INTROGRESSION BREEDING IN CULTIVATED PLANTS

EDITED BY: Jaime Prohens, Pietro Gramazio, Mariola Plazas and Laura Toppino

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-673-9
DOI 10.3389/978-2-88971-673-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INTROGRESSION BREEDING IN CULTIVATED PLANTS

Topic Editors:
Jaime Prohens, Universitat Politècnica de València, Spain
Pietro Gramazio, University of Tsukuba, Japan
Mariola Plazas, Polytechnic University of Valencia, Spain
Laura Toppino, Council for Research in Agriculture and the Analysis of Agrarian Economy (CREA), Italy

Citation: Prohens, J., Gramazio, P., Plazas, M., Toppino, L., eds. (2021). Introgression Breeding in Cultivated Plants. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-673-9

Table of Contents

05 ***Editorial: Introgression Breeding in Cultivated Plants***
Pietro Gramazio, Jaime Prohens, Laura Toppino and Mariola Plazas

08 ***Breeding Has Increased the Diversity of Cultivated Tomato in The Netherlands***
Henk J. Schouten, Yury Tikunov, Wouter Verkerke, Richard Finkers, Arnaud Bovy, Yuling Bai and Richard G. F. Visser

20 ***High Resolution Mapping of a* Hordeum bulbosum*-Derived Powdery Mildew Resistance Locus in Barley Using Distinct Homologous Introgression Lines***
Parastoo Hoseinzadeh, Brigitte Ruge-Wehling, Patrick Schweizer, Nils Stein and Hélène Pidon

33 ***The Resurgence of Introgression Breeding, as Exemplified in Wheat Improvement***
Ming Hao, Lianquan Zhang, Shunzong Ning, Lin Huang, Zhongwei Yuan, Bihua Wu, Zehong Yan, Shoufen Dai, Bo Jiang, Youliang Zheng and Dengcai Liu

44 ***Characterization of the Wheat*-Psathyrostachys huashania *Keng 2Ns/2D Substitution Line H139: A Novel Germplasm With Enhanced Resistance to Wheat Take-All***
Shengsheng Bai, Fengping Yuan, Hanbing Zhang, Zhenyue Zhang, Jixin Zhao, Qunhui Yang, Jun Wu and Xinhong Chen

54 ***Development of Wheat*-Aegilops caudata *Introgression Lines and Their Characterization Using Genome-Specific KASP Markers***
Surbhi Grewal, Manel Othmeni, Jack Walker, Stella Hubbart-Edwards, Cai-yun Yang, Duncan Scholefield, Stephen Ashling, Peter Isaac, Ian P. King and Julie King

66 ***Introgression of Two Broad-Spectrum Late Blight Resistance Genes,* Rpi-Blb1 *and* Rpi-Blb3*, From* Solanum bulbocastanum *Dun Plus Race-Specific* R *Genes Into Potato Pre-breeding Lines***
Elena Rakosy-Tican, Ramona Thieme, Janine König, Marion Nachtigall, Thilo Hammann, Tunde-Eva Denes, Klaudia Kruppa and Márta Molnár-Láng

83 ***Introgression Breeding in Barley: Perspectives and Case Studies***
Javier Hernandez, Brigid Meints and Patrick Hayes

98 ***Introgression Leads to Genomic Divergence and Responsible for Important Traits in Upland Cotton***
Shoupu He, Pengpeng Wang, Yuan-Ming Zhang, Panhong Dai, Mian Faisal Nazir, Yinhua Jia, Zhen Peng, Zhaoe Pan, Junling Sun, Liru Wang, Gaofei Sun and Xiongming Du

110 ***Diverse Wheat-Alien Introgression Lines as a Basis for Durable Resistance and Quality Characteristics in Bread Wheat***
Eva Johansson, Tina Henriksson, Maria Luisa Prieto-Linde, Staffan Andersson, Rimsha Ashraf and Mahbubjon Rahmatov

125 ***Potential Short-Term Memory Induction as a Promising Method for Increasing Drought Tolerance in Sweetpotato Crop Wild Relatives [*Ipomoea *series* Batatas *(Choisy) D. F. Austin]***
Fernando Guerrero-Zurita, David A. Ramírez, Javier Rinza, Johan Ninanya, Raúl Blas and Bettina Heider

140 ***Introgression Breeding in Cowpea [*Vigna unguiculata *(L.) Walp.]***
Ousmane Boukar, Michael Abberton, Olaniyi Oyatomi, Abou Togola, Leena Tripathi and Christian Fatokun

151 ***Spike-Stalk Injection Method Causes Extensive Phenotypic and Genotypic Variations for Rice Germplasm***
Yuanyi Hu, Bigang Mao, Yumei Xia, Yan Peng, Dan Zhang, Li Tang, Ye Shao, Yaokui Li and Bingran Zhao

161 ***Current Perspectives on Introgression Breeding in Food Legumes***
Aditya Pratap, Arpita Das, Shiv Kumar and Sanjeev Gupta

179 ***Dynamic Programming for Resource Allocation in Multi-Allelic Trait Introgression***
Ye Han, John N. Cameron, Lizhi Wang, Hieu Pham and William D. Beavis

frontiers in Plant Science

EDITORIAL
published: 28 September 2021
doi: 10.3389/fpls.2021.764533

Editorial: Introgression Breeding in Cultivated Plants

Pietro Gramazio[1,2]*, *Jaime Prohens*[3], *Laura Toppino*[4] and *Mariola Plazas*[2]

[1] Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan, [2] Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-Universitat Politècnica de València, Valencia, Spain, [3] Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Valencia, Spain, [4] CREA, Council for Agricultural Research and Economics, Research Centre for Genomics and Bioinformatics, Montanaso Lombardo, Italy

Keywords: wild relatives, pre-breeding, climate change, diversity, genetic base

Editorial on the Research Topic

Introgression Breeding in Cultivated Plants

The increasing demand for plant products in a climate change scenario is a major challenge for breeders, who are loaded with the pressing need to develop varieties that are more productive, resilient, and efficient in the use of resources (Langridge et al., 2021). Crop wild relatives (CWRs) are a largely unexplored source of genetic variation for many traits, particularly tolerance to biotic and abiotic stresses (Dempewolf et al., 2017), but also for other traits that can contribute to improving the yield, quality, and adaptability of our crops to adverse conditions along with broadening their genetic base (Warschefsky et al., 2014). However, the use of CWRs is not an easy task due to breeding barriers between CWRs and their cultivated relatives that often hamper or, in practice, prevent their use in breeding (Prohens et al., 2017). In addition, CWRs usually display a pattern of traits that are pivotal for survival in the wild, such as physical and chemical defenses, dormancy, or mechanisms for seeds dispersal (easy seed shattering) among others, that turn out undesirable in cultivated species (Meyer and Purugganan, 2013). Despite these difficulties, capturing the useful genetic diversity from CWRs through pre-breeding efforts, combining a broad array of conventional and new molecular techniques, provides new opportunities to develop a new generation of plant materials with improved features introgressed from CWRs (Prohens et al., 2017, Langridge et al., 2021). The urgency to make use of the diversity present in CWRs provided the impetus for this Research Topic. The 14 papers collected here report the development of new introgression materials and strategies, as well as new relevant information, for introgression breeding in a broad range of crops.

Overall, cereals are the primary staple crops in most regions of the world and is one of the crop groups where greater efforts were put into introgression breeding (Dempewolf et al., 2017). Two review papers address the strategies followed in wheat (Hao et al.) and barley (Hernandez et al.). Bread wheat (*Triticum aestivum* L.) is a hexaploid allopolyploid (Feldman and Levin, 2012) with many CWRs exhibiting different ploidy levels in the large *Triticeae* tribe within the *Poaceae* family (Lu and Ellstrand, 2014). Hao et al. presented a strategy based on the exploitation of synthetic hexaploid wheats by double top cross followed by two phases of selection, allowing the development of three new cultivars. This novel strategy could also be extended to other allopolyploid crops where the progenitor species are known. Barley (*Hordeum vulgare* L.) is the fourth most cultivated cereal with two major germplasm groups (two-row vs. six-row) encompassing a large genetic diversity of landraces (Milner et al., 2019). As shown by Hernandez et al., landraces can significantly contribute to barley breeding through introgression into elite material. In this case, the authors introduce past and current efforts for identifying genetic diversity in the barley genepool for improving multiple

Edited and reviewed by:
Petr Smýkal,
Palacký University, Olomouc, Czechia

***Correspondence:**
Pietro Gramazio
piegra@upv.es

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *25 August 2021*
Accepted: *06 September 2021*
Published: *28 September 2021*

Citation:
Gramazio P, Prohens J, Toppino L and Plazas M (2021) Editorial: Introgression Breeding in Cultivated Plants. Front. Plant Sci. 12:764533. doi: 10.3389/fpls.2021.764533

traits, as well as several case studies for introgression breeding from landraces for stripe and stem rust resistance.

The development of new sets of introgression materials in cereals and their genetic and phenotypic characterization contributes to broadening the genetic base and use in breeding (Lu and Ellstrand, 2014). New lines of wheat with introgression from *Aegilops caudata* L. have been obtained by Grewal et al. after crossing this species with the bread wheat cultivar Paragon, showing the *ph1*/*ph1* mutant genotype, to obtain a set of recombinant introgression lines (ILs) with potential materials carrying the enhanced disease resistance from *A. caudata*. Apart from methods based on sexual crossings, the so-called Exogenous DNA Transfer (EDT) methods allow circumventing hybridization and introgression from CWRs that are not sexually compatible with the crop (Ali et al., 2015; Jiang et al., 2021). One of these methods, the spike-stalk injection method, consisting of the injection of exogenous DNA close to the inflorescence a few days after meiosis (De la Pena et al., 1987), has been exploited by Hu et al. to develop a rice variant with introgressions from *Oryza eichingeri* characterized by higher plant height, panicle length and spikelet number. Using introgression materials developed using CWRs as donors, Bai et al. found resistance to the take-all disease caused by the fungus *Gauemannomyces tritici* in 2Ns/2D substitution line of bread wheat with *Psathyrostachys huashania* Keng. In barley, Hoseinzadeh et al. used a set of ILs with *H. bulbosum* L. to fine map a gene from this latter species conferring resistance to powdery mildew. Finally, Johansson et al. evaluated wheat materials with introgressions from rye, *Leymus* spp. and *Thinopyrum junceiforme* (Á. Löve & D. Löve) Á. Löve for tolerance to several diseases and pests, as well as for quality traits. They found materials with multiple resistances, including resistance to stripe rust, Russian wheat aphid, and Syrian Hessian fly.

Legumes are an important source of protein and oil and can make an effective contribution to sustainable agriculture due to their capacity of symbiotic associations with *Rhizobium* nitrogen-fixing bacteria (Soumare et al., 2020). Pratap et al. reviewed the strategies and achievements made in introgression breeding of legumes by developing different segregating populations for the dissection of traits of interest, mostly for tolerance to biotic and abiotic stresses, existing in the CWRs of eight legumes. Introgression breeding has also resulted in significant improvements in cowpea [*Vigna unguiculata* (L.) Walp.], a multipurpose crop grown in tropical and subtropical regions with high potential for improvement as highlighted by the review of Boukar et al.

Root crops such as potato (*Solanum tuberosum* L.) or sweet potato [*Ipomoea batatas* (L.) Lam.] are important polyploid staple crops for which introgression breeding can make significant contributions for their enhancement and adaptation to climate change (George et al., 2018; Bethke et al., 2019). In fact, potato is one of the crops in which a greater utilization has been made for introgression breeding (Dempewolf et al., 2017), with *Phytophthora infestans* resistance being one of the main targets in the use of CWRs (Su et al., 2020). Rakosy-Tican et al. used somatic hybrids of potato with the resistant species *S. bulbocastanum* Dunal, in which four resistance genes have been identified and cloned, to develop three potato cultivars carrying the resistance genes. Some of the selected clones gave yield and quality similar to the recurrent parent while being resistant. Drought represents a major constrain in some of the major areas of sweet potato cultivation and CWRs carrying tolerance traits to drought have been identified (Nhanala and Yencho, 2020). Guerrero-Zurita et al. subjected 53 accessions of 10 sweet potato CWRs to potential short-memory induction treatment consisting of priming the plants after flowering onset with different water restriction periods. By measuring different ecophysiological indicators they identified *I. triloba* L. and *I. trifida* (Kunth) G. Don as the most promising species for introgression breeding of sweet potato for drought tolerance. In addition, the methodology used can be very efficient for the identification of potential sources of tolerance among CWRs in other crops.

Tomato is one of the crops where most introgression work has been performed, resulting in most of the modern tomato varieties carrying introgressions from several CWRs (Causse et al., 2013). Schouten et al. found that, since the 1960's, the genetic diversity of commercial tomato varieties has continuously increased due to the introgressions from CWRs, mostly aimed at improving resistance to diseases but also fruit quality traits. This article shows how introgression breeding can recover the diversity lost in the different genetic bottlenecks occurring since the domestication of tomato.

Introgression breeding proved also efficient in industrial crops (Dempewolf et al., 2017) where CWRs can contribute to improving biotic and abiotic stress tolerance as well as quality. He et al. genotyped 582 tetraploid cotton (*Gossypium hirsutum* L.) accessions by RAD-seq and found a low diversity in the landraces, while modern cultivars generally contained introgressions from several CWRs which, depending on the introgressed region, contributed to several important traits as also to the genomic differentiation of modern groups.

Introgression breeding programmes may have different levels of complexity depending on the objectives and on the plant materials, genetic diversity, and resources available (Prohens et al., 2017). Therefore, designing an introgression programme is not a trivial task. In this respect, Han et al. have devised a Markov Decision Processes model for optimizing the strategies for the allocation of resources in introgression breeding programmes. The simulations performed show that the model improves the efficiency in the use of resources in the introgression process, making an effective contribution in optimizing the resources available and in consequence maximizing genetic gains.

Overall, the papers from the present Research Topic highlight the importance and efficiency of introgression breeding for broadening the genetic base of our crops and making available to breeders new sources of variation. Despite the effort performed in some major crops (Dempewolf et al., 2017), introgression breeding is still in its infancy in many crops, being challenging in those requiring long generation times, such as fruit trees, or when cross-incompatibility or sterility prevents introgression breeding. In these cases, "new plant breeding techniques" such as cis-genesis and gene editing can contribute to introgression breeding without the need for sexual hybridization (Prohens et al., 2017).

In any case, in order to promote introgression breeding among the public and private breeders community, more efforts should be devoted to collecting, conserving and enhancing the CWRs, which are underrepresented in the germplasm collections, as well as to proceed to its extensive high-throughput phenotyping and genotyping. In this way, we are confident that by using the largely unexplored genetic diversity of CWRs and landraces introgression breeding can make an effective contribution to addressing the challenges of improving agricultural productivity and sustainability in a climate change scenario.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Ali, A., Bang, S. W., Chung, S. M., and Staub, J. E. (2015). Plant transformation *via* pollen tube-mediated gene transfer. *Plant Mol. Biol. Rep.* 33, 742–747. doi: 10.1007/s11105-014-0839-5

Bethke, P. L., Halterman, D. A., and Jansky, S. H. (2019). Potato germplasm enhancement enters the genomics era. *Agronomy* 9:575. doi: 10.3390/agronomy9100575

Causse, M., Desplat, N., Pascual, L., Le Paslier, M. C., Sauvage, C., Bauchet, G., et al. (2013). Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genom.* 14:791. doi: 10.1186/1471-2164-14-791

De la Pena, A., Lörz, H., and Schell, J. (1987). Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nature* 325, 274–276. doi: 10.1038/325274a0

Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., and Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Sci.* 57, 1070–1082. doi: 10.2135/cropsci2016.10.0885

Feldman, M., and Levin, A. A. (2012). Genome evolution due to allopolyploidization in wheat. *Genetics* 192, 763–774. doi: 10.1534/genetics.112.146316

George, T. S., Taylor, M. A., Dodd, I. C., and White, P. J. (2018). Climate change and consequences for potato production: a review of tolerance to emerging abiotic stress. *Potato Res.* 60, 239–268. doi: 10.1007/s11540-018-9366-3

Jiang, D., Wang, F., Zhang, H., Gao, W., Tong, X. L. C., Chen, G., et al. (2021). Formation of proto-Kranz in C3 rice induced by spike-stalk injection method. *Int. J. Mol. Sci.* 22:4305. doi: 10.3390/ijms22094305

Langridge, P., Braun, H., Hulke, B., Ober, E., and Prassana, B. M. (2021). Breeding crops for climate resilience. *Theor. Appl. Genet.* 134, 1607–1611. doi: 10.1007/s00122-021-03854-7

Lu, B. R., and Ellstrand, N. (2014). World food security and the tribe Triticeae (Poaceae): genetic resources of cultivated, wild, and weedy taxa for crop improvement. *J. Syst. Evol.* 52, 661–666. doi: 10.1111/jse.12131

Meyer, R., and Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14, 840–852. doi: 10.1038/nrg3605

Milner, S. G., Jost, M., Taketa, S., Mazón, E. R., Himmelbach, A., Oppermann, M., et al. (2019). Genebank genomics highlights the diversity of a global barley collection. *Nat. Genet.* 51, 319–326. doi: 10.1038/s41588-018-0266-x

Nhanala, S. E. C., and Yencho, C. G. (2020). Assessment of the potential of wild *Ipomoea* spp. For the improvement of drought tolerance in cultivated sweet potato *Ipomoea batatas* (L.). *Lam. Crop Sci.* 61, 234–249. doi: 10.1002/csc2.20363

Prohens, J., Gramazio, P., Plazas, M., Dempewolf, H., Kilian, B., Díez, M. J., et al. (2017). Introgressiomics: a new approach for using crop wild relatives in breeding for adaptation to climate change. *Euphytica* 213:158. doi: 10.1007/s10681-017-1938-9

Soumare, A., Diedhiou, A. G., Thuita, M., Hafidi, M., Ouhdouch, Y., Gopalakrishnan, S., et al. (2020). Exploiting biological nitrogen fixation: a route towards sustainable agriculture. *Plants* 9:1011. doi: 10.3390/plants9081011

Su, Y., Viquez-Zamora, M., den Uil, D., Sinnige, J., Kruyt, H., Vossen, J., et al. (2020). Introgression of genes for resistance against *Phytophthora infestans* in diploid potato. *Am. J. Potato Res.* 97, 33–42. doi: 10.1007/s12230-019-09741-8

Warschefsky, E., Penmetsa, R. V., Cook, D. R., and von Wettberg, E. J. B. (2014). Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *Am. J. Bot.* 101, 1791–1800. doi: 10.3732/ajb.1400116

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 20 December 2019
doi: 10.3389/fpls.2019.01606

Breeding Has Increased the Diversity of Cultivated Tomato in The Netherlands

Henk J. Schouten[1]†, Yury Tikunov[1†], Wouter Verkerke[2], Richard Finkers[1], Arnaud Bovy[1], Yuling Bai[1] and Richard G.F. Visser[1]*

[1] Plant Breeding, Wageningen University & Research, Wageningen, Netherlands, [2] Business Unit Greenhouse Horticulture, Wageningen University & Research, Bleiswijk, Netherlands

It is generally believed that domestication and breeding of plants has led to genetic erosion, including loss of nutritional value and resistances to diseases, especially in tomato. We studied the diversity dynamics of greenhouse tomato varieties in NW Europe, especially The Netherlands, over the last seven decades. According to the used SNP array, the genetic diversity was indeed very low during the 1960s, but is now eight times higher when compared to that dip. The pressure since the 1970s to apply less pesticides led to the introgression of many disease resistances from wild relatives, representing the first boost of genetic diversity. In Europe a second boost ensued, largely driven by German popular media who named poor tasting tomatoes *Wasserbomben* (water bombs). The subsequent collapse of Dutch tomato exports to Germany fueled breeding for fruit flavor, further increasing diversity since the 1990s. The increased diversity in composition of aroma volatiles observed starting from 1990s may reflect the efforts of breeders to improve fruit quality. Specific groups of aroma compounds showed different quantitative trend over the decades studied. Our study provides compelling evidence that breeding has increased the diversity of tomato varieties considerably since the 1970s.

Keywords: tomato varieties, diversity, introgressions, metabolomics, breeding

OPEN ACCESS

Edited by:
Jaime Prohens, Polytechnic University of Valencia, Spain

Reviewed by:
Mathilde Causse, Genetics and Improvement of Fruit and Vegetables (INRA), France
Santiago García Martínez, Universidad Miguel Hernández de Elche, Spain

****Correspondence:***
Henk J. Schouten
henk.schouten@wur.nl

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *24 September 2019*
Accepted: *14 November 2019*
Published: *20 December 2019*

Citation:
Schouten HJ, Tikunov Y, Verkerke W, Finkers R, Bovy A, Bai Y and Visser RGF (2019) Breeding Has Increased the Diversity of Cultivated Tomato in The Netherlands. Front. Plant Sci. 10:1606. doi: 10.3389/fpls.2019.01606

INTRODUCTION

A recent paper in *Nature Biotechnology* on *de novo* domestication of tomato voiced the general belief that "breeding of crops over millennia for yield and productivity has led to reduced genetic diversity. As a result, beneficial traits of wild species, such as disease resistance and stress tolerance, have been lost (…). Despite the increases in yield conferred by domestication, the breeding focus on yield has been accompanied by a loss of genetic diversity and reduced nutritional value and taste" (Zsögön, 2018).

Reduction of diversity among crop varieties poses risks for cultivation, especially when most varieties carry the same genetic basis for resistance to diseases and pests. If a disease resistance is overcome in one variety, other varieties become susceptible too. This leads to agricultural vulnerability which can affect the entire chain, especially if there are no alternatives for disease control, such as appropriate, authorized pesticides. History has provided several examples, such as the Panama disease (Fusarium oxysporum f. sp. cubense) epidemic in banana (Ploetz, 2000; Garcia, 2013), or the southern corn leaf blight (*Helminthosporium maydis*) outbreak in maize (Horsfall, 1972; Kemble et al., 1982). Because the number of authorized pesticides has decreased and continues to decrease,

crop protection has to rely more and more on resistances that should have not a narrow genetic basis.

The loss of genetic variation in crops due to the modernization of agriculture has been denoted as genetic erosion (Tanksley and McCouch, 1997). During domestication preferred genotypes were selected, leading to loss of alleles and a decrease in genetic diversity of landraces compared to wild accessions (Bai and Lindhout, 2007; Blanca, 2015; Lin et al., 2014). Two principal occurrences affecting crop diversity have been identified: 1) the replacement of landraces by commercial varieties; and 2) more recent additional changes in the diversity of commercial varieties caused by plant breeding (van de Wouw et al., 2010a). Breeding can reduce genetic diversity by continued selection in the breeding germplasm, or may broaden genetic diversity through the introgression of alleles from wild relatives. The question remains whether the increase in diversity because of introgression has compensated the reduction of genetic diversity due to inbreeding and selection.

We have studied this for tomato, as particularly in this crop there have been indications of serious genetic erosion (Lin et al., 2014). Furthermore, Tieman and Klee (Tieman, 2017; Klee and Tieman, 2018) mentioned that "modern commercial varieties contain significantly lower amounts of many (…) important flavor chemicals than older varieties" as a result of intensive selection for production traits, such as yield and disease resistance, at the expense of flavor. We studied the evolution of diversity of commercial tomato varieties in NW Europe since the 1950s. To do this, we looked at both genetic variation at the DNA level, and phenotypic variation, including disease resistances, fruit size, and flavor components.

MATERIALS AND METHODS

Tomato Varieties and Growing Conditions

Ninety tomato varieties introduced in the Netherlands for commercial glasshouse fresh fruit production in a time period from 1950 till 2016 were selected by random picking without any prior knowledge about any of genetic or phenotypic parameters analyzed, to have about 12 varieties per each decade (**Table S1**). Although for recent decades far more varieties were available compared to the 1950s, we decided to have a balanced sampling with similar sampling sizes for the different decades, thus preventing changes in diversity due to differences in number of sampled varieties per decade. The varieties were grown in a glasshouse at standard commercial growing conditions in the summer of 2017. Three plants per variety were grown using a randomized block experimental design.

Genetic Diversity Analysis

Young leaf material was collected from the 90 tomato varieties, freeze dried, and sent to Trait Genetics, Germany, for genotyping by means of a Illumina® SolCAP SNP-array (Sim, 2012a). This yielded 7720 SNP-marker scores per variety. The SNP-scores were visualized in Excel, using a blanc cell in case the score resembled the reference genome. In addition to the SNP array genomic DNA from two recent varieties from this list, i.e. Merlice and Bambelo, was re-sequenced using Illumina HiSeq 150 paired ends sequencing, for detailed analysis of introgressions.

Genetic Diversity Index Calculation

Within each decade we calculated for each SNP the nucleotide frequencies among the varieties that were commercially introduced in that decade, and used these frequencies for calculating the genetic diversity index (H) of Nei (1973), according the equation:

$$H = 1 - \sum_i p_i^2$$

where: p = frequency of nucleotide i.

This is also named "expected heterozygosity." The H-values per decade were averaged among all markers, giving a measure for genetic diversity for each decade. The frequency p was also calculated at the diploid level, looking at frequencies of allelic combinations in commercial varieties.

Definition of Basal Genome and Introgressions

For finding introgression that were deliberately introduced by breeders, we needed an introgression free reference genome. The generally used reference genome of tomato refers to cv. Heinz. However, this variety may harbor introgressions. Therefore, we defined a "basal genome," which is the consensus genome of the sampled varieties from the 1950s and 1960s. For each SolCap array marker, we selected the most common nucleotide in these two decennia. Marker scores that deviated from this basal genome were highlighted in Excel, using conditional formatting. This revealed introgression haploblocks.

Analysis of the Basic Fruit Flavor Parameters

For the analysis of the basic flavor parameters: soluble solids content (Brix), titratable acidity, firmness, and juiciness from 20 to 30 fruit per variety were harvested per a variety at a mature ripe stage. Fruit ripening was judged using intensity of pigmentation and firmness. Also average fruit weight was recorded for every variety analyzed.

For soluble solids content and titratable acidity measurements, fruit quarters (for fruits >30 g per fruit) and whole fruits (<30 g per fruit) were homogenized for 15 s in a Vita-prep 3 blender. The soluble solids content was measured directly from the homogenized fruit sample as °Brix by means of a Refracto 30PX digital refractometer (Mettler Toledo). Titratable acidity (mmol H3O+/100 gram fresh weight) was determined by means of potentiometric endpoint titration with 0.1 mol/l NaOH till pH 8.2 by means of a T50 titrator (Mettler Toledo).

For texture measurements, from each fruit a 10-mm-diameter disk was excised from the fruit pericarp at the locular region by means of a cork borer. The disks were patted dry by rolling on filter paper and were weighed on an analytical balance (Mettler Toledo XA 204 DeltaRange). Five disks were enclosed between

sheets of screening cloth (Agratex, Ludvig Svensson), placed between two sheets of pre-weighed filter paper, having the skin down (Whatman 1003-917) and compressed by means of an Instron 3343 Universal Testing Machine at a speed of 60 mm/min with a flat plate plunger to 900 N. After this compressing, the filter paper was re-weighed. The juiciness (% Juice) was calculated as the weight increase of the filter papers divided by the fresh weight of the disks. Pericarp firmness was defined as the force [Newton (N)] at break of the force/deformation curve of the five simultaneously compressed disks. The data of the basic flavor parameters are present in **Supplemental Data (Data Sheet S7)**.

Analysis of Volatile Compounds

Nine mature ripe fruits per variety were pooled to have a representative sample. Fruits were cut, immediately frozen in liquid nitrogen and were ground to fine powder under liquid nitrogen using A11 analytical mill (IKA). Volatile organic compounds (VOCs) were analyzed, identified, and quantified using a Solid Phase Micro Extraction–Gas Chromatography–Mass Spectrometry (SPME-GC-MS) method previously described (Tikunov, 2005; Tikunov, 2013). Frozen fruit powder (1 g fresh weight) was weighed into a 5-ml screw-cap vial, closed, and incubated for 10 min at 30°C. An aqueous EDTA-sodium hydroxide (NaOH) solution was prepared by adjusting 100 mM EDTA to pH of 7.5 with NaOH. Then, 1 ml of the EDTA-NaOH solution was added to the sample to give a final EDTA concentration of 50 mM. Solid $CaCl_2$ was immediately added to give a final concentration of 5 M. The closed vials were sonicated for 5 min. A 1-ml aliquot of the pulp was transferred into a 10-ml crimp cap vial (Waters), capped, and used for SPME GC-MS analysis. Volatiles were automatically extracted from the vial headspace and injected into the GC-MS *via* a Combi PAL autosampler (CTC Analytics). Headspace volatiles were extracted by exposing a 65-µm polydimethylsiloxane-divenylbenzene SPME fiber (Supelco) to the vial headspace for 20 min under continuous agitation and heating at 50°C. The fiber was desorbed in the GC-MS injection port and compounds were separated on an HP-5 (50 m × 0.32 mm × 1.05 µm) column with helium as carrier gas (37 kPa). Mass spectra in the 35 to 400 m/z range were recorded by an MD800 electron impact MS (Fisons Instruments) at a scanning speed of 2.8 scans/s and an ionization energy of 70 eV. The chromatography and spectral data were evaluated using Xcalibur software (Thermo Scientific). MSClust software (Tikunov et al., 2011) was used to extract volatile compound mass spectral information from the chromatograms. Forty-six compounds were identified by comparison with authentic chemical standards (**Data Sheet S5**); the other were tentatively identified using MSSearch software (Thermo) and the NIST mass spectral library (www.nist.gov).

Analysis of Flavor Diversity

To estimate time trends of basic flavor components and individual volatile compounds, the Mann-Kendall trend test (Gilbert, 1987) implemented in Past3 software (https://folk.uio.no/ohammer/past/) was used. To discover trends in the diversity of volatile compound composition, a pairwise Euclidean distances matrix was calculated based on the quantitative profiles of 69 annotated volatiles (**Data Sheet S5**). Then a mean distance was calculated for each variety by averaging its distances to all other varieties registered within ±5 years. These mean distances of all 90 varieties were subjected to the Mann-Kendall trend test.

RESULTS AND DISCUSSION

Genetic Diversity of Tomato Varieties During the Previous Seven Decades

We collected leaf samples from 90 tomato varieties that were commercially released between 1950 and 2016 in NW Europe (**Table S1**). All varieties have been used in greenhouses for the production of tomatoes for the fresh market. We analyzed roughly equal numbers of varieties per decade from the 1950s till the 2010s. The varieties were genotyped using the SolCap SNP platform (Sim, 2012a). This yielded 7720 SNP-marker scores per variety (**Data Sheet S1**). SNP-markers containing missing values over the majority of varieties were removed, leaving 7,661 SNP-markers. We grouped the varieties per decade and calculated for each SNP the genetic diversity index (*H*) of Nei (Nei, 1973), also referred to as expected heterozygosity. These *H*-values per decade were averaged over all markers, providing a measure for the genetic diversity of each decade (**Figure 1**). This figure clearly shows that the genetic diversity among commercial tomato varieties was low during the 1950s, and even lower during the 1960s. However, from the 1970s onwards, the diversity increased up to eight-fold compared to the 1960s, according to the studied SNPs, and using the diversity index *H*. Apparently, the increase in diversity caused by introgressions far exceeded the decrease in diversity caused by selection.

This increase in genetic diversity in recent varieties compared to the low diversity in those from the 1950s and 1960s is not restricted to a few loci only, but has occurred across the whole genome (**Figures 1** and **2**). Chromosomes 4, 5, 6, 9, 11, and 12 show a particularly pronounced increase in diversity in modern varieties. Apparently, the increased genetic diversity has not been limited to a small number of genes or a few chromosome arms, but has encompassed the majority of all tomato genes. However, some regions have hardly changed since the 1950s, e.g., the upper half of Chr. 2, harboring repeats of 45S ribosomal DNA (The Tomato Genome Consortium, 2012), and the central parts of Chrs. 3, 7, and 10 (**Figure 2**), being the centromeric regions of these chromosomes (Víquez-Zamora, 2014).

As **Figure 1** shows, genetic diversity was very low in the 1950s and 1960s. Based on this observation, we defined a "base tomato genome," representing for each SNP-marker the most prevalent nucleotide during these two decades. The nucleotides deviating from this "base genome" were regarded as introgressions. **Figure 1** illustrates the gradual increase in the proportion of introgressed DNA. Currently, more than a quarter of the genome (28%) is composed of such introgressions.

Before 1970, nearly all tomato varieties were homozygous, but from 1980 onwards, nearly all new commercial varieties are hybrids (**Figure S1**), giving an extra layer of genetic diversity, i.e., within the varieties. In **Figures 1** and **2**, we did not separate heterozygous from homozygous introgressions, giving heterozygous introgressions

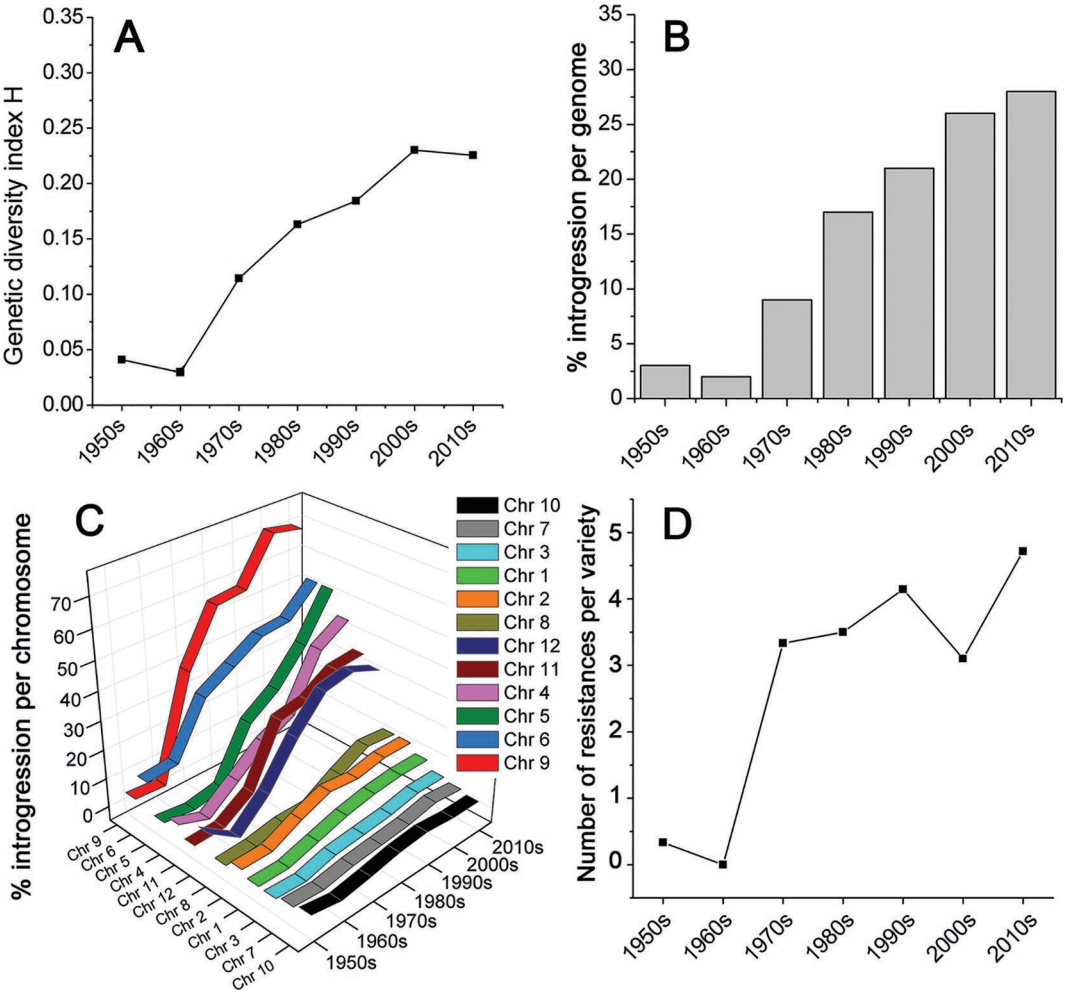

FIGURE 1 | Genetic diversity of tomato varieties, commercially introduced from 1950 till 2016. **(A)** The diversity within a decade is expressed as Nei's index, also referred to as expected heterozygosity. **(B)** The proportion of the genome and **(C)** individual chromosomes of commercial tomato varieties, consisting of introgressions compared to the prevailing genome of tomato varieties in the 1950s and 1960s. **(D)** Average number of diseases and pest to which the investigated varieties are resistant, according to the official NAKTuinbouw variety database [https://www.naktuinbouw.com/groente/variety-description/tomato-solanum-lycopersicum-l; (Banga and Aalbersberg, 1995)].

the same weight as homozygous ones. However, in **Data Sheet S3**, we show the separate introgressions in the individual varieties, distinguishing homozygous from heterozygous haploblocks. **Data Sheet S3** clearly show the increase in abundancy of haploblocks from the 1970s onwards. Several of these haploblocks are very large, reflecting linkage drag. The "basal genome" without deliberate introgressions is shown in **Data Sheet S3** too. This basal genome consists of the consensus marker scores of the varieties sampled in the 1950s and 1960s.

The First Diversity Boost: Introgressions for Resistances

The chromosomes differed considerably in their introgression composition (**Figure 1C**). Two groups of chromosomes can be distinguished. In the first group of chromosomes (Chr. 1, 2, 3, 7, 8, 10), only 5 to 15% of the chromosomal DNA has been altered since the 1960s. In the second group (Chr. 4, 5, 6, 9, 11, 12), between 30 and 70% of the chromosome has been replaced by introgressions. There is one chromosome (Chr. 9) that consists of approximately 70% introgressed DNA, compared to the 1960s.

This huge change in the composition of Chr. 9 (**Figure 1C**) was caused by a large introgression fragment from *Solanum peruvianum*. This introgression carries the tomato mosaic virus (ToMV) resistance gene *Tm2* (derived from *S. peruvianum* PI 126926) or its allele $Tm2^2$ (derived from *S. peruvianum* PI 18650) (Lanfermeijer et al., 2003; Lin et al., 2014). We re-sequenced the recently (2013) introduced variety "Merlice," being homozygous for $Tm2^2$ (**Data Sheet S2**). The exotic fragment encompasses 79% (53 Mb) of Chr 9 in this modern variety (**Figure S2**). Breeding companies started selling tomato varieties with this introgression in the 1970s, and the proportion of varieties carrying this introgression has increased ever since (58% in the 1970s to 93% in the 2010s; **Figure S3**). However, the introgression size has remained large, showing the co-introgression of the majority of Chr. 9 from the wild species [**Data Sheet S3**, (Víquez-Zamora, 2013; Lin et al., 2014)].

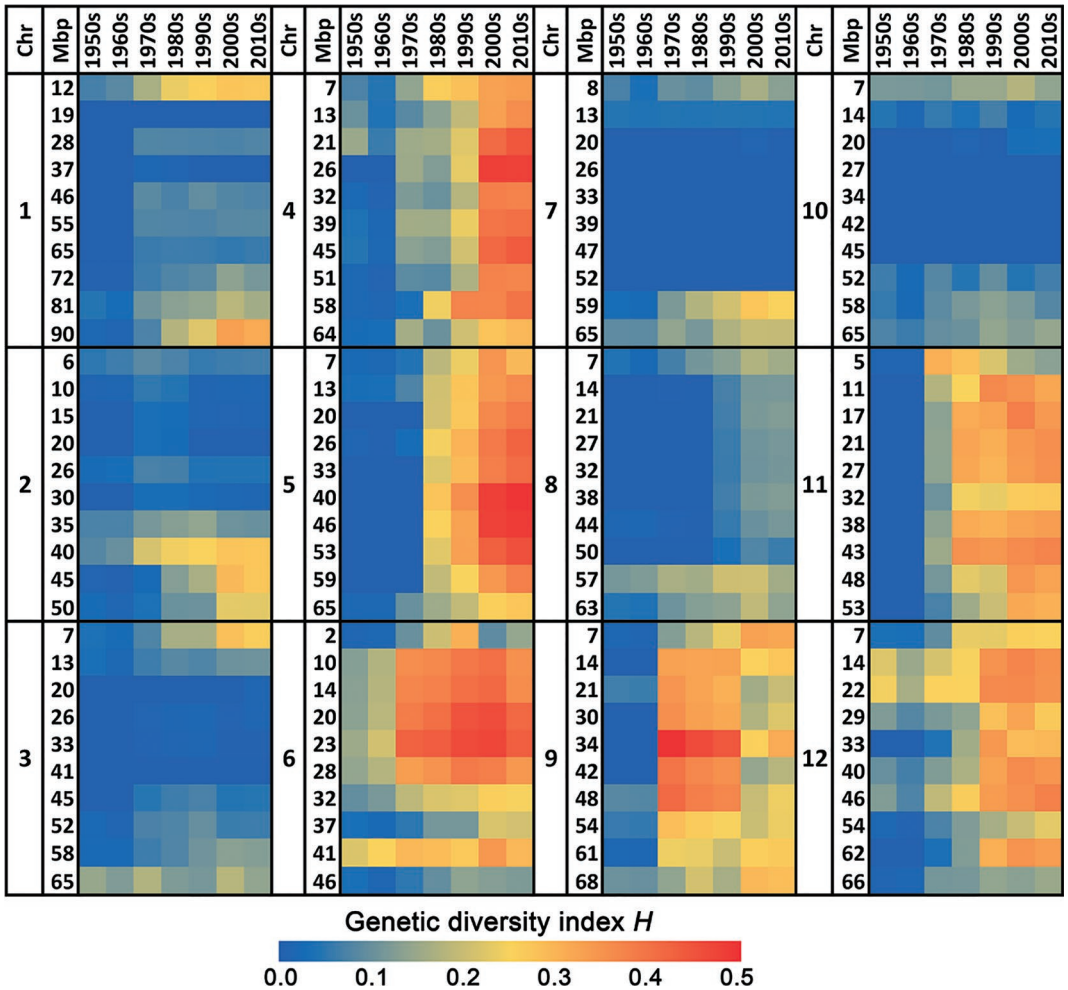

FIGURE 2 | Heat map of the genetic diversity in the 12 chromosomes of commercial tomato varieties in course of time. The blue color indicates a very low genetic diversity in that decade for the respective chromosomal fragment, whereas a red color represents a very high genetic diversity.

This linkage drag has likely been caused by recombination suppression during meiosis, possibly due to large structural rearrangements such as an inversion in this region (Bonierbale et al., 1988).

From the 1970s onwards, the genes *Cf-4* and *Cf-9* were introgressed at the top of chromosome 1, for providing resistance to leaf mold disease, caused by *Cladosporium fulvum* (**Data Sheet S6**). These genes descend from *S. pimpinellifolium* (**Data Sheet S6**). Two other resistance genes for controlling this pathogen, i.e., *Cf-2* and *Cf-5*, were introgressed at the top of chromosome 6. These genes were introgressed since the 1970s too.

Another resistance gene on chromosome 6, introgressed during the same period, is the *Mi-1* gene from *S. peruvianum,* conferring resistance to southern root-knot nematode (*Meloidogyne incognita*). This introgression has remained very large since its introgression, covering nearly 60% of the chromosome [**Data Sheet S6**; (Víquez-Zamora, 2013; Lin et al., 2014)]. Many more resistance genes have been introgressed since the 1970s, on nearly all chromosomes of tomato (**Data Sheet S6**).

Analysis of the phenotypic traits of the varieties since the 1960s, as described by descriptive variety lists (Banga and Aalbersberg, 1995), showed that also at the phenotypic level, the diversity among varieties was very small in the 1960s, but from the 1970s onwards, an increasing number of resistances to diseases and pests were introduced (**Figure 1D**, **Data Sheet S4**).

The Second Diversity Boost: Fruit Quality Traits; "Discharging Water Bombs"

Poor flavor is the most frequent reason for consumer dissatisfaction of tomatoes (Fernqvist and Hunter, 2012). In the late 1980s and early 1990s, German popular news media coined the phrase *Wasserbomben* (German for "water bombs") to describe poor, watery-tasting Dutch tomatoes (Hendriks, 2016). At that time, the vast majority of Dutch fresh tomatoes were being exported to Germany. Following the damage to their reputation, Dutch tomato exports to Germany collapsed shortly after (Hendriks, 2016). This became an important milestone in tomato breeding in NW Europe, marking the need for an adjustment of selection criteria towards balancing agronomical traits, such as high yield and resistances, with consumer quality traits, such as flavor, fruit size, shape, and

color. Also, research on better and quicker measurement techniques of flavor started. The *Wasserbomben* crisis fueled the second boost of diversity, namely a diversity in fruit types and improved flavors.

Fruit Size

The most obvious phenotypic diversification that occurred from the 1980s onwards has been fruit size (**Figure 3A**). While varieties that entered the market before 1980 showed little variation in average weight, ranging predominantly between 50 and 100 g per fruit, the introduction of new fruit types, such as cherry and cocktail tomatoes (<25 g per fruit) as well as large fruited varieties (100–300 g) led to an enormous diversification of fruit sizes from the 1980s onwards (**Figures 3A** and **S4C**). Víquez-Zamora et al. (2013) showed that the cherry-like fruit sizes were obtained by introgressions of large parts of Chrs. 4, 5, and 12

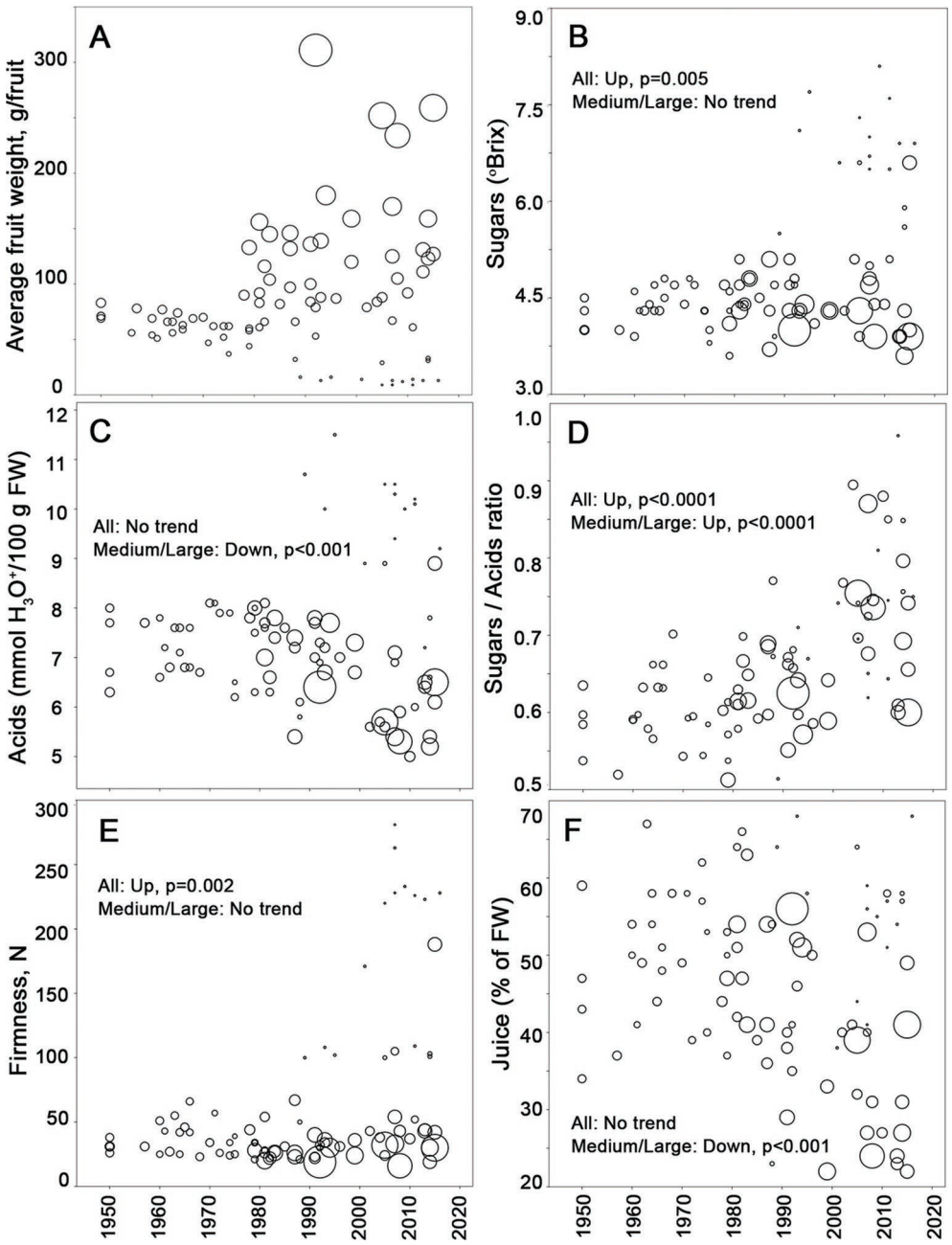

FIGURE 3 | The development of average fruit weight and the basic fruit flavor parameters of tomato varieties, commercially introduced between 1950 and 2016. Each variety is represented by a circle, whose size is proportional to its average fruit weight. Significance of quantitative trends is indicated with p-values of Mann-Kendall trend test for all varieties analyzed (All) and for medium-sized and large fruited tomatoes with average weight >25 g per fruit (Medium/Large). **(A)** Average fruit weight, g/fruit, **(B)** Sugar content in ripe fruit measured as refraction index, ºBrix, **(C)** Titratable acids in ripe fruit, mmol H_3O^+/100 g FW, **(D)** Sugar/Acid ratio in ripe fruit, **(E)** Firmness of ripe fruit, N, **(F)** Amount of liquid (juice) released by ripe fruit, % of FW.

from *S. pimpinellifolium*. This is consistent with the increased diversity of these chromosomes since the 1990s, apparent from the heat map in **Figure 2**.

Flavor

Tomato fruit flavor is determined by a combination of five essential chemical and textural components: 1) the concentration of sugars in ripe tomato, mainly fructose and glucose, which can be very well approximated by soluble solids content, measured as a refraction index (°Brix); 2) acidity, which is determined by the concentration of organic acids, mainly citrate; 3) firmness of the fruit pericarp; 4) fruit juiciness; and 5) aroma, caused by a complex combination of VOCs. The ratio of the first two components—sugar content and acidity—is one of the main parameters for the perception of sweetness (Tandon et al., 2003; Zanor, 2009). The sugar/acid ratio showed an overall significant increase in the last three decades (**Figure 3D**). On the one hand, cherry tomatoes with a higher sugar content were introduced in that period (**Figure 3B** and **Table S2**). Concurrently, there was a reduction in the acidity of medium-sized and large fruits (**Figure 3C**). Fruit pericarp firmness has increased mainly due to the introduction of cherry varieties (**Figure 3E**). Variation in juiciness appears to have been considerable in all decades since 1950, although interestingly, the proportion of varieties whose fruit tissue released less liquid has increased during the last three decades (**Figure 3F**). These may represent an increasing interest in "non-leaky" varieties for use in salads and on bread.

Most of the aroma active VOCs in tomato fruit can be classified into five distinct groups according to their biosynthetic origins (Rambla et al., 2014): 1) VOCs derived from fatty acids; 2) phenolic VOCs; 3) phenylpropanoid VOCs; 4) VOCs derived from sulfur-containing and branched chain amino acids; and 5) carotenoid breakdown products (Buttery et al., 1987; Buttery et al., 1988; Krumbein and Auerswald, 1998; Baldwin et al., 2000; Tandon et al., 2000; Tandon et al., 2001; Selli et al., 2014; Du et al., 2015; Tieman, 2017). Overall, the VOC composition of tomato fruits has diversified significantly, particularly over the last two decades (**Figure 4A**). However, different chemicals with distinct

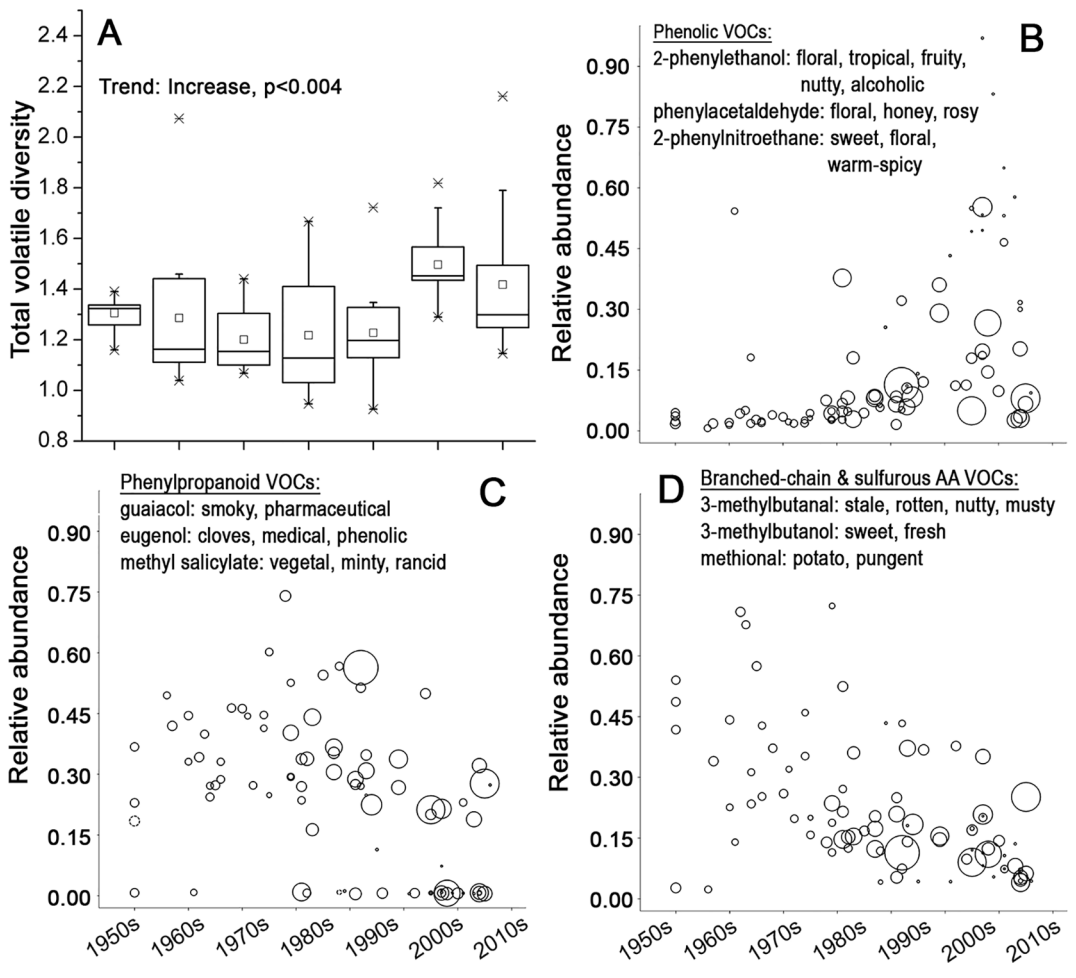

FIGURE 4 | The composition of volatiles in tomato fruits in course of time. **(A)** Boxplot of the quantitative diversity of volatiles of fruits of tomato varieties in the respective decades. The diversity is expressed as Euclidian distances between quantities of 60 identified volatile compounds (VOCs). **(B–D)** Relative average abundance of three VOC groups with their individual aroma characters. Each variety is represented by a circle, whose size is proportional to the average fruit weight. **(B)** Phenolic VOCs, **(C)** phenypropanoid VOCs, and **(D)** VOCs derived from branched-chain and sulfurous amino acids in fruits of 90 tomato varieties.

aroma characteristics showed different temporal trends (**Table S3**). The second group of phenolic VOCs associated with floral and sweet aroma (Tandon et al., 2000; Baldwin et al., 2008; Mayer, 2008; Selli et al., 2014; Du et al., 2015) increased dramatically in the fruits of more modern varieties when compared to earlier varieties (**Figures 4B** and **S5**). The third group of phenylpropanoid VOCs, which express a "smoky" (Tikunov, 2013) or "medical" aroma (Causse, 2002), and the fourth group of VOCs produced from branched-chain and sulfurous amino acids, associated with earthy/musty/pungent/medicinal types of aroma (Baldwin et al., 2004; Baldwin et al., 2008), showed a very significant reduction in fruits of the more recent varieties (**Figures 4C**, **D**, **Figure S5** and **Table S3**).

Tomato fruit flavor is a product of a complex interaction between all the mentioned factors (Baldwin et al., 2008). Our data indicate that breeding has led to a clear increase in variation of different flavor components from the beginning of 1990s. This can be regarded as the second boost of diversity, since it began two decades later than the general increase in genetic diversity. More specifically, there has been an increasing proportion of varieties with a higher potential to express sweet/fruity types of flavor along with reduced expression of potential off-flavors.

Some recent studies, such as a work of Tieman and Klee (Tieman, 2017; Klee and Tieman, 2018), comparing modern commercial varieties to old, non-commercial heirloom tomatoes, conclude that the modern varieties had a lower flavor quality compared to the old varieties due to breeding. This seeming discrepancy with our findings might be due to the difference in the definition of "old," which in Tieman et al. study refers to tomato varieties that were not subjected to breeding for traits relevant for industrial tomato production, and some of them may originate from long before 1950s. In our collection, the oldest varieties of 1950–1960s are all indeterminate greenhouse varieties that already went through the breeding for the production-related traits, which make their genetic diversity much narrower compared to the heirlooms. Secondly, our modern varieties have been bred for greenhouses in NW Europe, whereas Tieman and Klee studied open field varieties for N. America.

Considering the existence of other sources of biodiversity, e.g., landraces or heirloom varieties predating the 1950s and wild relatives, there may still be room for improvement of flavor of modern commercial tomatoes by further enrichment with alleles that can further improve fruit flavor (Tieman, 2017). However, our data indicate that modern breeding for quality for the consumer is on the right track. Pre-harvest, harvesting, and post-harvest practices are other important factors affecting variation in flavor quality. Therefore, breeding for and introduction of superior genotypes must be complemented by appropriate growing and post-harvest practices to ensure the maximum translation of genetics into flavor of desirable quality, and to deliver it to consumers at its best.

Diversity Evolution in a Wider Temporal Context

In our analysis, we looked at varieties from the 1950s till the present time, reflecting the period of modern commercial tomato breeding. In order to put the evolution of diversity in a wider temporal context, we calculated the diversity index among 385 ancestors (*S. lycopersicum* var. *lycopersicum*, *S. lycopersicum* var. *cerasiforme*, and *S. pimpinellifolium*), and 129 vintage accessions (including landraces and heirlooms). We used the SolCap array data and the classification of genotypes from Blanca et al. (2015), selecting the same SNP-markers as were used for **Figure 1**.

Figure 5 shows the relatively high diversity among ancestors belonging to *S. lycopersicum* and *S. pimpinellifolium*. Domestication of these species and genetic bottlenecks due to transport (Blanca, 2012; Blanca, 2015) led to vintage types. According to **Figure 5**, only one third of the variation in the ancestors was still present in the vintage accessions. Further selection and inbreeding led to the commercial varieties in the 1960s, representing only 10% of the diversity of the ancestors. However, introgressions of resistances to diseases and pests in the 1970s and 1980s increased the genetic diversity considerably, even above the level of the vintage types. The second boost of diversity, i.e., the breeding for fruit size diversity, color differences, improved taste, and additional resistances, further increased the diversity index, nearly to the level of the ancestors (**Figure 5**). We have to keep in mind here that the introgressions in the modern varieties not only descend from *S. lycopersicum* and *S. pimpinellifolium*, i.e., the species from which the domesticated tomatoes have been derived, but also from more distant relatives, including *S. peruvianum, S. pennellii, S. chilense*, and *S. habrochaites* (**Data Sheet S6**). This explains the high level of diversity of the modern varieties compared to the ancestors *S. lycopersicum* and *S. pimpinellifolium*. As the SolCap array is based on SNPs between accessions of *S. lycopersicum* and *S. pimpinellifolium* (Sim, 2012b), the genetic diversity of modern varieties harboring introgressions from other wild species might not have been fully captured by the array. It is therefore possible that the genetic diversity among the evaluated varieties from the 1970s till now might be even higher than we report.

Blanca et al. (2015) studied the domestication of tomato in Ecuador, Peru, and Mesoamerica, but also included contemporary genotypes in their analysis. These genotypes referred predominantly to genotypes from breeding programs in the USA for the fresh market and processing market. Blanca et al. concluded that genotypes for the processing and fresh market showed slightly higher levels of diversity when compared to vintage types. They mention that this is likely due to the effect of introgression during breeding, and differentiation into distinct market classes, which is consistent with our findings.

When focusing in rarefaction analyses on alleles that are present in specific (sub)groups only, so-called private alleles, Blanca et al. found that the processing contemporary types showed a higher frequency of private alleles compared to genotypes for the fresh market. The frequency of private alleles was lowest in vintage types.

Earlier work of Sim et al. (2012a) also showed that contemporary genotypes for the fresh market showed higher genetic diversity compared to the vintage sub-population. The same held for the processing group. The highest diversity was found in the cherry tomatoes, as appeared both from the genetic diversity indices as from the rarefaction analysis (Sim, 2012a).

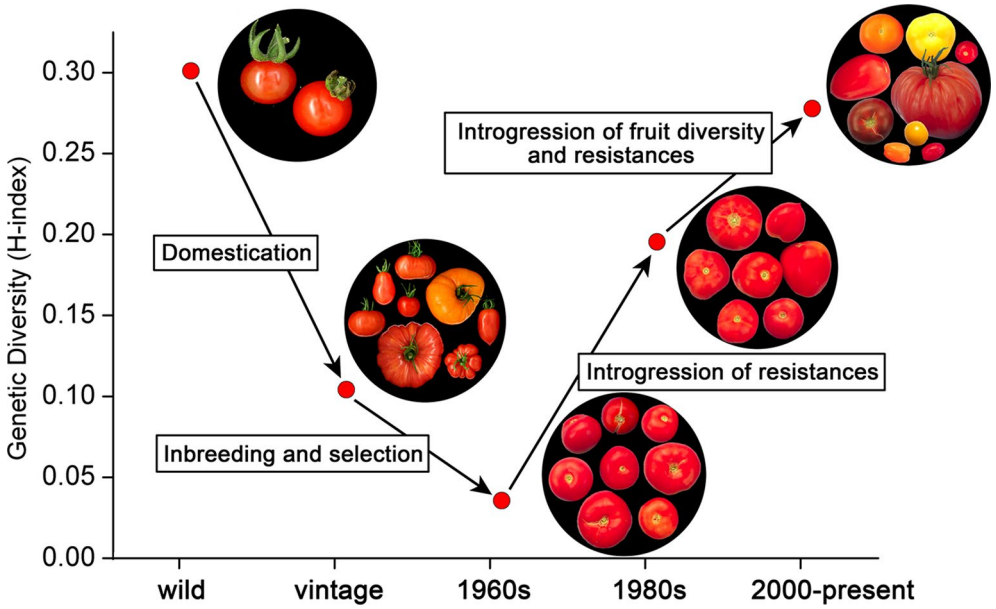

FIGURE 5 | The evolution of genetic diversity in tomato. The upper left group represents the ancestors (*S. lycopersicum* var. *lycopersicum*, *S. lycopersicum* var. *cerasiforme*, and *S. pimpinellifolium*), which gave rise to the vintage types including landraces through a process of domestication [data from (Blanca, 2015)]. Inbreeding and selection among these vintage tomatoes led to commercial varieties in the 1960s with a very low genetic and phenotypic diversity. From the 1970s onwards, resistances to diseases and pests were introgressed from distant species, including *S. peruvianum*, *S. pennellii*, *S. chilense*, and *S. habrochaites*, increasing genetic diversity among commercial tomato varieties considerably. After the 1980s, fruit size, color, and flavor started to vary substantially, further increasing the genetic diversity of modern varieties.

Suitability of the Marker Platform

The SolCap array has been based on 7,720 SNPs, discovered in RNAseq data from *S. lycopersicum* and *S. pimpinellifolium* accessions (Causse, 2002). We removed 59 markers that gave missing values for >80% of the varieties before 1970, leaving 7,661 SNP markers. These markers covered the whole tomato genome, although the number of markers per Mbp varied. After 1970, breeders started to introgress resistances from distant species, including *S. peruvianum, S. pennellii, S. chilense*, and *S. habrochaites* (**Data Sheet S6**). As the SolCap array was not designed to capture genetic variation from these species, we wondered whether the number of missing marker data would increase after 1970, due to poor hybridization of DNA sequences from these wild species to the oligos on the array. **Figure S6** shows that the number of missing data indeed increased in course of time. However, the percentage of missing marker values was still low (<0.5%). Therefore, we conclude that the number of missing SNP calls from the SolCap array was negligible and did not influence the conclusions.

However, we are aware that the SolCAP array may overlook quite some variation from wild species that were not included in the initial set of genotypes when selecting the SolCAP SNPs. Resequencing reveals far more SNPs, as we exemplify for two resequenced varieties (**Figure S2**). However, using, e.g., variant calling files (VCFs) based on resequencing data does have limitations too, as only reads that align to the reference genome are considered, disregarding many reads that are not aligned to the reference genome. Unmapped reads may harbor even more variation.

The selection of SNPs and their physical positions does influence the absolute values of the genetic diversity index H. Using another selection, or using resequencing data and haploblocks, rather than the SolCap array data, would influence the absolute levels of the diversity measure. However, we believe the trends will remain very comparable.

Putting Biodiversity in a Wider Perspective Regarding Other Crops

For agricultural field crops, several studies have been performed on the genetic diversity of varieties. Wouw et al. (van de Wouw et al., 2010b) performed a meta-analysis based on data from 44 published papers, addressing diversity trends in released crop varieties of eight different field crops in the twentieth century. The studies encompassed variety diversity in many countries in the world, not only in Europe and North America. Wheat was the most represented crop, with 26 of the 44 papers focused on wheat. For wheat the lowest diversity occurred in the period from the 1960s till the 1980s. This decrease was 6% compared to the first half of the century and was significant. However, a recovery of diversity was observed in the 1990s. For seven other major crops (barley, maize, oat, flax, soybean, pea, rice), a dip in diversity was also observed in the 1960s, but the recovery was earlier compared to wheat (van de Wouw et al., 2010b). Apparently, the trends we have described for tomato are similar for other crops. However, the changes in genetic diversity have been far more pronounced and clearer for tomato, not showing an increase of a few percent,

but a nine-fold increase since the 1960s. Therefore, we conclude that the concern about decreasing diversity among varieties, due to modern plant breeding, is not supported by our study. In contrast, we have observed a tremendous increase in diversity, both at the genotypic and phenotypic level.

DATA AVAILABILITY STATEMENT

Access to cultivar resequencing data can be obtained by contacting the corresponding author. All other datasets generated and analyzed for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

HS and RV set up the initial experiments for determining biodiversity levels of varieties and made the first outlines for a paper. HS, RV, AB, YT, and YB subsequently conceptualized the paper. HS coordinated the research, performed the genetic analyses, and wrote the original draft. YT, AB, and WV were responsible for the analyses of flavor components and provided text, figures, and tables for the manuscript. YT prepared the final figures. RF did bioinformatic analyses of re-sequencing data regarding introgressions. YB analyzed introgressions of resistance genes, and RV supervised the process. All authors reviewed and rewrote (parts of) the manuscript.

FUNDING

The research was financed by Plant Breeding (WUR) and by the Board for Plant Varieties, The Netherlands. The views expressed in the text are the authors' own and do not necessarily reflect the opinions of the Dutch Board for Plant Varieties.

ACKNOWLEDGMENTS

We acknowledge Tomato World and the Centre for Genetic Resources, the Netherlands, for providing tomato seeds. We are grateful to Maarten Peters (Unifarm, WUR) and Fien Meijer (Plant Breeding, WUR) for taking care of the tomato plants, Viola Ghio (RIKILT, WUR) for the collaboration during fruit harvesting and analysis of volatiles, Gerard Bijsterbosch and Yan Zhe (Plant Breeding, WUR) for DNA sampling and bioinformatic analyses, NAKTuinbouw for providing phenotypic evaluations of registered varieties, tomato breeders (Ruud Verhoef - Rijk Zwaan, Martijn van Stee–Enza Zaden, Lennard de Vos–Axia Seeds, Doug Heath–Bejo Seeds, Theo Schotte–Solynta, and Jaap Hoogstraten–Bayer) for giving information regarding tomato fruit sizes, and tomato grower Ton Janssen for giving historical information regarding the *Wasserbombe* crisis. We are also grateful to Martijn van Kaauwen (WUR) who supported the bioinformatic analysis, and Peter Bourke (WUR) for linguistic improvements. We further acknowledge the comments from reviewers after submission of the manuscript, leading to a significant improvement of the text.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01606/full#supplementary-material

FIGURE S1 | The percentage of studied tomato varieties, being hybrids. The horizontal axis shows the decade of commercial introduction.

FIGURE S2 | The frequency of SNPs, when aligning 150 bp reads of the modern cv. 'Merlice' to Chr. 9 of the reference tomato genome (cv. 'Heinz'). The figure illustrates the large introgression (yellow) harboring the gene *Tm-2²* (orange) from *S. peruvianum* P.I.18650. This gene provides resistance to tomato mosaic virus (ToMV). The distal parts of Chr. 9 that show high sequence similarity to the reference genome are represented by green blocks. Possible rearrangements of the introgression compared to the reference genome are not shown.

FIGURE S3 | The percentage of investigated tomato varieties being resistant to tomato mosaic virus.

FIGURE S4 | The average fruit size class and weight of 284 tomato varieties (including the 90 varieties evaluated in the present study) registered in the official Dutch cultivar registry (https://www.raadvoorplantenrassen.nl/nl/rassenregister/) from 1950 till 2016. **(A)** Average fruit weight of 90 tomato varieties determined experimentally in the present study; **(B)** Relationship between fruit size class (as assigned in the official variety registry) of the 90 studied varieties and the average fruit weight determined experimentally in the present study; **(C)** Distribution of fruit sizes and calculated fruit weight of all 284 commercial varieties registered in the period from 1950 till 2016, which is divided into three sub-periods to have a balanced number of cultivars per sub-period.

FIGURE S5 | Quantitative variation of aroma related volatile compounds in ripe fruits of 90 tomato varieties from 1950 till 2016. The five biosynthetic groups are highlighted by different colors: green – branched-chain and sulphur-containing amino acid derived volatiles, red – phenolic volatiles, blue – phenylpropanoid volatiles, yellow – lipid derived volatiles; purple – carotenoid derived volatiles. The vertical axis represents relative abundance of a compound, where the maximum abundance of a compound is set at 1.00. Each circle represents a tomato variety, and a size of a circle is proportional to the average fruit weight of that variety.

FIGURE S6 | The proportion of SolCap SNP markers that had missing values. We took the 7661 markers into account that we used for calculation of the Nei-index *H* (Fig. 1), disregarding the markers that failed for all varieties, or failed for >80% of the varieties before 1970.

DATA SHEET S1 | The SolCap array SNP-scores of the 90 tomato varieties.

DATA SHEET S2 | Alignment of the *TM-2* and *TM-2²* alleles for resistance to the Tomato Mosaic Virus (ToMV), the susceptible allele *tm-2*, and the sequence of cv. 'Merlice' at this locus. This alignment shows that 'Merlice' harbors *TM-2²* homozygously.

DATA SHEET S3 | The introgressions in 90 tomato varieties, commercially introduced since 1950. These varieties were genotyped, using the SolCap array (https://www.illumina.com/library-prep-array-kit-selector/kits-and-arrays/the-solcap-tomato-consortium.html). Column 1 to 3 provide the ColCap marker names and physical positions. A cell is blanc in case the marker scores for both alleles are identical to the prevalent score for that marker in the 1950s and 1960s. A cell is light blue, if only one marker score deviates from the 'standard tomato genome', therefore reflecting heterozygous introgressions. The dark blue cells represent homozygous introgressions.

DATA SHEET S4 | Disease resistances of the 90 varieties according the official tomato variety registration of NAKTuinbouw.

DATA SHEET S5 | GC-MS data of 66 annotated volatile compounds. Quantitative values of volatiles presented as counts (AU - arbitrary units) of corresponding selective ions.

DATA SHEET S6 | Reported resistance genes in tomato, and their positions on the tomato genome.

DATA SHEET S7 | Flavor components of the investigated tomato varieties. Sugar: estimated as refraction (°Brix); % juice: the amount of juice, squeezed from the pericarp relatively to the weight of the pericarp before squeezing; Acidity: the amount of titratable acid (mmol H3O+/100gr); Firmness: the power needed to break the pericarp (N). For some varieties, there were not sufficient numbers of fruits available for reliable measurements.

TABLE S1 | List of studied tomato varieties. The varieties are sorted according to their years of commercial introduction in The Netherlands according to the Dutch Variety Register of the Raad voor Plantenrassen (https://nederlandsrassenregister.nl/).

TABLE S2 | Trends for basic flavor parameters. Mann-Kendall trend statistics for all the varieties studied (n = 90), and for the non-cherry (medium-sized and large fruits, >30 g/fruit) varieties. Trend coefficient S estimates strength and directionality of a trend. The statistical significance is represented by a P-value. Trends are estimated for the entire collection of 90 varieties, and for a sub-set of medium-sized and large fruited varieties (Non-cherry). SSC – soluble solid content, equivalent of sugar content in tomato, TA – titratable acidity, Firmness – firmness of fruit pericarp, Juiciness – amount of liquid released from tomato fruit pericarp.

TABLE S3 | Volatile compounds which had a statistically significant quantitative trends in fruits of tomato varieties introduced from 1950 till 2016. Directionality and strength of trends is represented by Mann-Kendall trend coefficient S with positive and negative values indicating an increase and decrease, respectively, in concentration during the period of time studied. Trends are estimated in the entire collection of 90 varieties and in a sub-set of medium and large fruited varieties (Non-cherry). P-values represent significance of the trend

REFERENCES

Bai, Y., and Lindhout, P. (2007). Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? *Ann. Bot.* 100, 1085–1094. doi: 10.1093/aob/mcm150

Baldwin, E. A., Scott, J. W., and Shewmaker, C. K. (2000). Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. *Pl. Health Progr.* 35, 1013–1022. doi: 10.21273/HORTSCI.35.6.1013

Baldwin, E. A., Goodner, K., Plotto, A., Pritchett, K., and Einstein, M. (2004). Effect of volatiles and their concentration on perception of tomato descriptors. *J. Food Sci.* 69, S310–S318. doi: 10.1111/j.1750-3841.2004.tb18023.x

Baldwin, E. A., Goodner, K., and Plotto, A. (2008). Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J. Food Sci.* 73S, 294–307. doi: 10.1111/j.1750-3841.2008.00825.x

Banga, O., and Aalbersberg, W. (1995). *Beschrijvende rassenlijst voor groentegewassen,* Vol. 1 (Wageningen: Leiter-Nypels).

Blanca, J., Canizares, J., Cordero, L., Pascual, L., Díez, M. J., and Nuez, F. (2012). Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. *PloS One* 7, e48198. doi: 10.1371/journal.pone.0048198

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., et al. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16, 257. doi: 10.1186/s12864-015-1444-1

Bonierbale, M. W., Plaisted, R. L., and Tanksley, S. D. (1988). RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120, 1095–1103.

Buttery, R. G., Teranishi, R., and Ling, L. C. (1987). Fresh tomato aroma volatiles: a quantitative study. *J. Agric. Food Chem.* 35, 540–544. doi: 10.1021/jf00076a025

Buttery, R. G., Teranishi, R., Ling, L. C., Flath, R. A., and Stern, D. J. (1988). Quantitative studies on origins of fresh tomato aroma volatiles. *J. Agric. Food Chem.* 36, 1247–1250. doi: 10.1021/jf00084a030

Causse, M., Saliba-Colombani, V., Lecomte, L., Duffé, P., Rousselle, P., and Buret, M. (2002). QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. *J. Exp. Bot.* 53, 2089–2098. doi: 10.1093/jxb/erf058

Du, X., Song, M., Baldwin, E., and Rouseff, R. (2015). Identification of sulphur volatiles and GC-olfactometry aroma profiling in two fresh tomato cultivars. *Food Chem.* 171, 306–314. doi: 10.1016/j.foodchem.2014.09.013

Fernqvist, F., and Hunter, E. (2012). Who's to blame for tasteless tomatoes? The effect of tomato chilling on consumers' taste perceptions. *Europ. J. Hortic. Sci.* 77 (S), 193–198.

Garcia, F. A., Ordonez, N., Konkol, J., AlQasem, M., Naser, Z., Abdelwali, M., et al. (2013). First report of *Fusarium oxysporum f.* sp. *cubense* tropical race 4 associated with panama disease of banana outside Southeast Asia. *Plant Dis.* 98, 694. doi: 10.1094/PDIS-09-13-0954-PDN

Gilbert, R. O. (1987). *Statistical methods for environmental pollution monitoring* (New York: Wiley). doi: 10.2307/2531935

Hendriks, A. (2016). *De tomaat* (Amsterdam: Nieuw Amsterdam).

Horsfall, J. G., Brandow, G. E., Brown, W. L., Day, P. R., Gabelman, W. H., Hanson, J. B., et al. (1972). *Genetic vulnerability of major crops* (National Academies of Sciences).

Kemble, R. J., Flavell, R. B., and Brettell, R. I. S. (1982). Mitochondrial DNA analyses of fertile and sterile maize plants derived from tissue culture with the Texas male sterile cytoplasm. *Theor. Appl. Genet.* 62, 213–217. doi: 10.1007/BF00276241

Klee, H. J., and Tieman, D. M. (2018). The genetics of fruit flavour preferences. *Nat. Genet.* 19, 347–356. doi: 10.1038/s41576-018-0002-5

Krumbein, A., and Auerswald, H. (1998). Characterization of aroma volatiles in tomatoes by sensory analyses. *Nahrung* 42, 395–399. doi: 10.1002/(SICI)1521-3803(199812)42:06<395::AID-FOOD395>3.3.CO;2-Z

Lanfermeijer, F. C., Dijkhuis, J., Sturre, M. J. G., de Haan, P., and Hille, J. (2003). Cloning and characterization of the durable tomato mosaic virus resistance gene *Tm-2(2)* from *Lycopersicon esculentum*. *Plant Mol. Biol.* 52, 1037–1049. doi: 10.1023/A:1025434519282

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., et al. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* 46, 1220–1226. doi: 10.1038/ng.3117

Mayer, F., Takeoka, G. R., Buttery, R. G., Whitehand, L. C., Naim, M., and Rabinowitch, H. D. (2008). Studies on the aroma of five fresh tomato cultivars and the precursors of cis- and trans-4,5-epoxy-(E)-2-decenals and methional. *J. Agric. Food Chem.* 56, 3749–3757. doi: 10.1021/jf0732915

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *PNAS* 70, 3321–3323. doi: 10.1073/pnas.70.12.3321

Ploetz, R. C. (2000). Panama disease: a classic and destructive disease of banana. *Online Pl. Health Prog.* doi: 10.1094/PHP-2000-1204-01-HM

Rambla, J. L., Tikunov, Y. M., Monforte, A. J., Bovy, A. G., and Granell, A. (2014). The expanded tomato fruit volatile landscape. *J. Exp. Bot.* 65, 4613–4623. doi: 10.1093/jxb/eru128

Selli, S., Kelebek, H., Ayseli, M. T., and Tokbas, H. (2014). Characterization of the most aroma-active compounds in cherry tomato by application of the aroma extract dilution analysis. *Food Chem.* 165, 540–546. doi: 10.1016/j.foodchem.2014.05.147

Sim, S.-C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M. W., Van Deynze, A., et al. (2012a). High-density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding. *PloS One* 7, e45520–e45518. doi: 10.1371/journal.pone.0045520

Sim, S.-C., Van Deynze, A., Stoffel, K., Douches, D. S., Zarka, D., Ganal, M. W., et al. (2012b). Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PloS One* 7, e40563. doi: 10.1371/journal.pone.0040563

Tandon, K. S., Baldwin, E. A., and Shewfelt, R. L. (2000). Aroma perception of individual volatile compounds in fresh tomatoes (*Lycopersicon esculentum*, Mill.) as affected by the medium of evaluation. *Postharv. Biol. Tech.* 20, 261–268. doi: 10.1016/S0925-5214(00)00143-5

Tandon, K. S., Jordán, M., Baldwin, E. A., and Goodner, K. L. (2001). Characterization of fresh tomato aroma volatiles using GC-olfatometry. *Proc. Fla. State Hortic. Soc.* 114, 142–144.

Tandon, K. S., Baldwin, E. A., Scott, J. W., and Shewfelt, R. L. (2003). Linking sensory descriptors to volatile and nonvolatile components of fresh tomato flavor. *J. Food Sci.* 68, 2366–2371. doi: 10.1111/j.1365-2621.2003.tb05774.x

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

The Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/nature11119

Tieman, D., Zhu, G., Resende, M. F. R., Lin, T., Nguyen, C., Bies, D., et al. (2017). A chemical genetic roadmap to improved tomato flavor. *Science* 355, 391–394. doi: 10.1126/science.aal1556

Tikunov, Y. M., Laptenok, S., Hall, R. D., Bovy, A., and de Vos, R. C. H. (2011). MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. *Metabolomics* 8, 714–718. doi: 10.1007/s11306-011-0368-2

Tikunov, Y., Lommen, A., De Vos, C., Verhoeven, H., Bino, R., Hall, R., et al. (2005). A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiol.* 139, 1125–1137. doi: 10.1104/pp.105.068130

Tikunov, Y. M., Molthoff, J., De Vos, R. C. H., Beekwilder, J., van Houwelingen, A., van der Hooft, J. J. J., et al. (2013). Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. *Plant Cell.* 25, 3067–3078. doi: 10.1105/tpc.113.114231

Víquez-Zamora, M., Vosman, B., van de Geest, H., Bovy, A., Visser, R. G. F., Finkers, R. et al. (2013). Tomato breeding in the genomics era: insights from a SNP array. *BMC Genomics* 14, 354. doi: 10.1186/1471-2164-14-354

Víquez-Zamora, M., Caro, M., Finkers, R., Tikunov, Y., Bovy, A., Visser, R. G. F., et al. (2014). Mapping in the era of sequencing: high density genotyping and its application for mapping TYLCV resistance in *Solanum pimpinellifolium*. *BMC Genomics* 15, 1152. doi: 10.1186/1471-2164-15-1152

van de Wouw, M., Kik, C., van Hintum, T., van Treuren, R., and Visser, B. (2010a). Genetic erosion in crops: concept, research results and challenges. *Plant Genet. Resour.* 8, 1–15. doi: 10.1017/S1479262109990062

van de Wouw, M., van Hintum, T., Kik, C., van Treuren, R., and Visser, B. (2010b). Genetic diversity trends in twentieth century crop cultivars: a meta analysis. *Theor. Appl. Genet.* 120, 1241–1252. doi: 10.1007/s00122-009-1252-6

Zanor, M. I., Rambla, J. L., Chaib, J., Steppa, A., Medina, A., Granell, A., et al. (2009). Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J. Exp. Bot.* 60, 2139–2154. doi: 10.1093/jxb/erp086

Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., et al. (2018). *De novo* domestication of wild tomato using genome editing. *Nat. Biotechnol.* 36, 1211–1216. doi: 10.1038/nbt.4272

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 March 2020
doi: 10.3389/fpls.2020.00225

High Resolution Mapping of a *Hordeum bulbosum*-Derived Powdery Mildew Resistance Locus in Barley Using Distinct Homologous Introgression Lines

Parastoo Hoseinzadeh[1†], Brigitte Ruge-Wehling[2], Patrick Schweizer[3‡], Nils Stein[1,4] and Hélène Pidon[1]*

[1] Genomics of Genetic Resources, Department of Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany, [2] Institute for Breeding Research on Agricultural Crops, Julius Kühn Institute (JKI), Sanitz, Germany, [3] Pathogen-Stress Genomics, Department of Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany, [4] Department of Crop Sciences, Center for Integrated Breeding Research (CiBreed), Georg-August-University, Göttingen, Germany

Powdery mildew caused by *Blumeria graminis* f. sp. *hordei (Bgh)* is one of the main foliar diseases in barley (*Hordeum vulgare* L.; *Hv*). Naturally occurring resistance genes used in barley breeding are a cost effective and environmentally sustainable strategy to minimize the impact of pathogens, however, the primary gene pool of *H. vulgare* contains limited diversity owing to recent domestication bottlenecks. To ensure durable resistance against this pathogen, more genes are required that could be unraveled by investigation of secondary barley gene-pool. A large set of *Hordeum bulbosum* (*Hb*) introgression lines (ILs) harboring a diverse set of desirable resistance traits have been developed and are being routinely used as source of novel diversity in gene mapping studies. Nevertheless, this strategy is often compromised by a lack of recombination between the introgressed fragment and the orthologous chromosome of the barley genome. In this study, we fine-mapped a *Hb* gene conferring resistance to barley powdery mildew. The initial genotyping of two *Hb* ILs mapping populations with differently sized 2HS introgressions revealed severely reduced interspecific recombination in the region of the introgressed segment, preventing precise localization of the gene. To overcome this difficulty, we developed an alternative strategy, exploiting intraspecific recombination by crossing two *Hv/Hb* ILs with collinear *Hb* introgressions, one of which carries a powdery mildew resistance gene, while the other doesn't. The intraspecific recombination rate in the *Hb*-introgressed fragment of 2HS was approximately 20 times higher than it was in the initial simple ILs mapping populations. Using high-throughput genotyping-by-sequencing (GBS), we allocated the resistance gene to a 1.4 Mb interval, based on an estimate using the *Hv* genome as reference, in populations of only 103 and 146 individuals, respectively, similar to what is expected at this locus in barley. The most likely candidate resistance gene within this

OPEN ACCESS

Edited by:
Laura Toppino, Council for Research in Agriculture and the Analysis of Agrarian Economy (CREA), Italy

Reviewed by:
Ernesto Igartua, Spanish National Research Council, Spain
Antonin Dreiseitl, Agricultural Research Institute Kromeriz, Czechia

***Correspondence:**
Hélène Pidon
pidon@ipk-gatersleben.de

†Present address:
Parastoo Hoseinzadeh, KWS SAAT SE & Co., KGaA, Einbeck, Germany

‡*Deceased*

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *06 December 2019*
Accepted: *13 February 2020*
Published: *03 March 2020*

Citation:
Hoseinzadeh P, Ruge-Wehling B, Schweizer P, Stein N and Pidon H (2020) High Resolution Mapping of a Hordeum bulbosum-Derived Powdery Mildew Resistance Locus in Barley Using Distinct Homologous Introgression Lines. Front. Plant Sci. 11:225. doi: 10.3389/fpls.2020.00225

interval is part of the coiled-coil nucleotide-binding-site leucine-rich-repeat (CC-NBS-LLR) gene family, which is over-represented among genes conferring strong dominant resistance to pathogens. The reported strategy can be applied as a general strategic approach for identifying genes underlying traits of interest in crop wild relatives.

Keywords: introgression lines, *Hordeum bulbosum*, recombination, mapping, powdery mildew, resistance, crop wild relative

INTRODUCTION

The increased genetic uniformity of cultivated crops, makes them highly vulnerable to various biotic and abiotic stresses, leading to crop yield losses and serious food security issues (Hoisington et al., 1999). Disease resistance breeding is a cost effective and environmentally sustainable strategy for minimizing the damage caused by plant pathogens. Hence, plant breeders are continuously working to discover novel sources of genetic resistance. Crop wild relatives (CWRs) are a reservoir of genetic variation providing an important source of novel alleles for the genetic improvement of cultivated species. Crosses between cultivars and CWRs have been carried out in several crop species to unlock this favorable genetic diversity (Tanksley and McCouch, 1997; Feuillet et al., 2008). Some prominent examples of the introgression of favorable disease resistance alleles from CWRs are the introductions of late blight resistance into potato from the wild potato *Solanum demissum* (Rao, 1979; Prescott-Allen and Prescott-Allen, 1986), and of stem rust resistance genes *Sr21* (Chen et al., 2015) and *Sr39* (Kerber and Dyck, 1990), both effective against the race Ug99, into bread wheat from *Triticum monococcum* and *Aegilops speltoides*, respectively.

Barley (*Hordeum vulgare* L.), the fourth most important cereal crop in the world, is affected each year by up to 30% potential yield loss due to pests and diseases (Savary et al., 2012). Limitations in the availability of novel resistance genes or alleles in the primary gene pool of barley, comprising the cultivated barley *H. vulgare* spp. *vulgare* and its wild progenitor *H. vulgare* spp. *spontaneum*, has directed the focus of research toward other barley gene pools. Bulbous barley, *Hordeum bulbosum* (*Hb*), a wild self-incompatible species and the only member of the secondary gene pool of cultivated barley (von Bothmer et al., 1995) is resistant to many barley pathogens (Xu and Kasha, 1992; Walther et al., 2000). A large panel of *Hb* introgression lines (ILs) harboring a diverse spectrum of resistance traits has been developed during recent years (Pickering et al., 1995; Johnston et al., 2009). This resource comprises ILs carrying, among others, the barley leaf rust resistance gene *Rph26* on chromosome $1H^bL$ (Yu et al., 2018); barley leaf rust gene *Rph22* (Johnston et al., 2013) and barley mild mosaic virus gene $Rym16^{Hb}$ (Ruge-Wehling et al., 2006) both located on chromosome $2H^bL$; barley yellow dwarf virus resistance gene $Ryd4^{Hb}$ on chromosome $3H^bL$ (Scholz et al., 2009) as well as loci conferring powdery mildew resistance located on chromosome $2H^bS$, $2H^bL$ and $7H^bL$ in barley/*Hb* introgression lines (Xu and Kasha, 1992; Pickering et al., 1995; Shtaya et al., 2007). These ILs represent a unique genetic resource for improving barley resistance to pathogens and for scientific investigation of resistance mechanisms as they provide access to further genetic diversity out of the primary gene pool of barley (Tanksley and Nelson, 1996; Zamir, 2001; Johnston et al., 2009). Genotyping-by-sequencing (GBS) of 145 *Hv/Hb* introgression lines (Wendler et al., 2014, 2015) has provided an extensive pool of molecular markers, sequence resources and single-nucleotide polymorphisms (SNPs) information, greatly improving the efficiency of mapping *Hb* loci.

Since the identification and the extensive use of the durable and complete *mlo* resistance gene in European germplasm (Jørgensen, 1992), the threat of powdery mildew to barley has been largely mitigated. However, *mlo* is associated with yield penalties (Kjær et al., 1990) and increased susceptibility to some hemibiothrophic and necrotrophic fungi (Jarosch et al., 1999; Kumar et al., 2001; McGrann et al., 2014). Thus, the search for alternative sources of resistance to powdery mildew remains important for barley breeding (Czembor, 2002; Corrion and Day, 2015). The wild progenitor *H. vulgare* spp. *spontaneum* proved to be a great source of diversity of resistance genes (Fischbeck et al., 1976; Dreiseitl and Bockelman, 2003). However, these genes are mostly race-specific, major resistance genes. They are particularly effective, but prone to be rapidly overcome by emerging new pathotypes. Therefore, the search for non-host resistance, from plant species to which the pathogen is not coevolutionary adapted, represent a great hope to achieve a complete and durable resistance. Several *Hv*/*Hb* introgression lines carrying a locus conferring powdery mildew resistance have been described (Xu and Kasha, 1992; Pickering et al., 1995, 1998; Shtaya et al., 2007). The *Hb* accession A42 displays a dominant high resistance to powdery mildew that has been localized on the short arm of chromosome $2H^b$ in preliminary studies (Szigat and Szigat, 1991; Michel, 1996). Interestingly, several significant QTLs and major genes associated with powdery mildew resistance have repeatedly been reported in this region in cultivated barley (Backes et al., 2003; von Korff et al., 2005; Řepková et al., 2009). Moreover, a resistance gene to powdery mildew has been reported in this region in various *Hb* accessions (Pickering et al., 1995; Zhang et al., 2001; Shtaya et al., 2007).

While the potential value of untapped genetic diversity of CWR is immense, their application in breeding programs through the use of ILs is hampered by negative linkage drag, mainly caused by severely repressed genetic recombination (Wijnker and de Jong, 2008; Prohens et al., 2017), which can confer yield penalties or other unfavorable characteristics (Hospital, 2001). The degree of drag is correlated with the size of introgressed CWR chromatin segments, and thus can be mitigated by reducing the size of respective introgressions through recombination (Frisch and Melchinger, 2001). However, the efficient utilization of *Hb* germplasm in barley crop

improvement and the genetic mapping of loci contributed by *Hb* suffers from highly reduced frequency of recombination in introgressed intervals up to 14-fold compared to intraspecific barley crosses (Ruge et al., 2003; Ruge-Wehling et al., 2006; Kakeda et al., 2008; Johnston et al., 2013). Possible explanations for this phenomenon include excessive sequence diversity, structural variation among *Hordeum* genomes, and probably other unknown mechanisms (Pickering, 1991; Hoffmann et al., 2004; Wendler et al., 2017). To reduce the negative linkage drag, precise delimitation of the causal gene is required, which usually demands intensive screening of large segregating *Hv/Hb* ILs populations.

Canady et al. (2006) compared the recombination rate in crosses between cultivated tomato (*Solanum lycopersicum*) and ILs with *Solanum lycopersicoides* fragments in *S. lycopersicum* backgrounds with the recombination in crosses between ILs with *S. lycopersicoides* fragments and ILs with *Solanum pennellii* fragments in the same region. They showed that tomato ILs with overlapping fragments from closely related species exhibited increased recombination rates in those fragments. Similarly, in barley, Johnston et al. (2015) demonstrated the usefulness of intraspecific recombination between *Hb* ILs to overcome the negative linkage between genes conferring pathogen resistance and reduced yield. They crossed two recombinant ILs containing an *Hb* locus on chromosome 2HL comprising the genes *Rph22* and $Rym16^{Hb}$, together with the proximal region of the original introgression for one of them, and the distal region for the other. They obtained four lines with reduced introgressions around the locus of interest for which the yield was close to the one of the recurrent *Hv* genotype.

The current study aimed to map a dominant powdery mildew resistance locus on chromosome $2H^bS$ introgressed from the tetraploid A42 powdery mildew resistant *Hb* accession into a susceptible barley cultivar "Borwina." Mapping in populations of over 200 F7 and BC1F6 from crosses between a susceptible barley cultivar and an *Hb* IL showed severely repressed recombination between the introgressed segment and the *Hv* genome. To overcome this difficulty, we exploited intraspecific recombination instead of interspecific recombination by crossing the IL carrying the resistance locus with another *Hv/Hb* IL carrying a homologous *Hb* introgression without resistance loci.

MATERIALS AND METHODS

Fungal Isolate and Phenotyping Test

The swiss powdery mildew field isolate CH4.8 was chosen based on its hypersensitive response it triggers on resistant plants and its capacity to overcome the resistance gene *MlRu2* present in Borwina, allowing a clear discrimination between resistant and susceptible plants. The resistance test was carried out in two technical replicates by detached leaf assay on the second seedling leaf sampled 14 days after sowing, as described in Hoseinzadeh et al. (2019). The inoculated detached leaves were kept for 7 days in a growth cabinet (MLR-352H, Panasonic, Japan) with LED light sources (S2 20W matt nw, ARTEKO-LED, Germany) at 20°C with 60% humidity and a 16:8 h photoperiod. Powdery mildew resistance phenotype was scored macroscopically based on percentage infection area as described by Mains and Dietz (1930). Plants without mycelial growth and sporulation but presenting necrotic flecks (infection type 1)were classified as resistant, while those having infection types similar to the susceptible parent "Borwina," with sporulation of well-developed colonies (infection types 3 and 4) were classified as susceptible.

Plant Material and Population Development

Initial genetic mapping was performed on the F7 and BC1F6 populations "5216" and "4176," respectively, both derived from crossing a tetraploid derivative of the colchicine treated German winter barley cultivar "Borwina" and the tetraploid *Hb* accession A42, which is resistant to several European barley powdery mildew isolates. Both "5216" and "4176" populations were derived from selfing of a single plant from the previous generation established as both resistant and heterozygous at the resistance locus. The crossing scheme used to generate those populations is described in **Figure 1**.

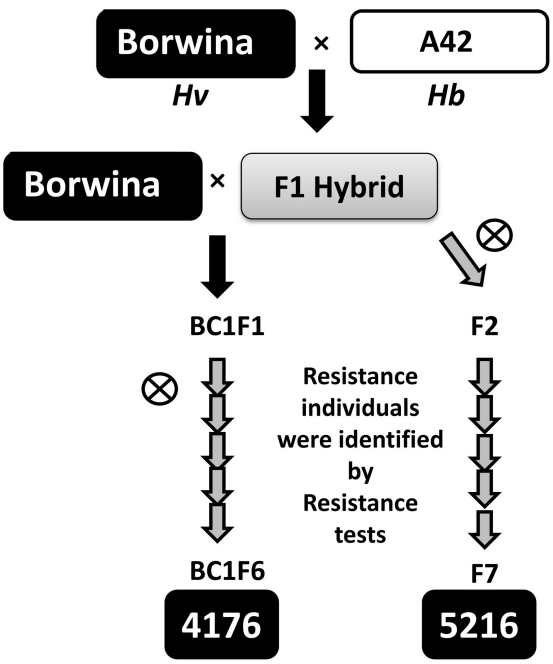

FIGURE 1 | Schematic outline of the mapping population design. The tetraploid *Hb* accession A42, resistant to the isolate CH4.8, was crossed to the tetraploid derivative of the susceptible barley cultivar "Borwina." From tetraploid F1 hybrids, two different introgression mapping populations were developed: the "4176" population was developed by backcrossing the F1 hybrid once to the parent "Borwina," followed by five generations of selfing. A gray arrow indicates a selfing generation. The "5216" population was derived from six generations of selfing from the F1 hybrid. In the course of development of populations "4176" and "5216," the generations were diploid from BC1F2 and F2, respectively. Through the population development, each new selfing generation was obtained by selecting and selfing a single resistant heterozygous plant identified by resistance test with the powdery mildew isolate CH4.8 and marker data analysis, in order to promote recombination.

We hypothesized that differences in the sequence or organization between *Hv* and *Hb* orthologous genome regions would severely reduce meiotic recombination. To test this hypothesis, we analyzed intraspecific *Hb*/*Hb* recombination in a *Hv* background. We generated two F2 populations by crossing two independent *Hv*/*Hb* ILs carrying independent but overlapping *Hb* introgressions at the terminal 2HS chromosome, thus representing different *Hb* genotypes at the resistance locus. Three introgression lines "IL 88," "IL 99," and "IL 116," developed in New Zealand Institute for Crop and Food Research, were selected that carry independent *Hb* introgressions at end of the short arm of barley chromosome 2H (Wendler et al., 2015; **Table 1**). These three ILs were phenotyped for resistance against isolate CH4.8 (**Figure 2**). Only "IL 88" displayed full susceptibility to CH4.8 and was used for further population development.

The homozygous resistant ILs 5216/4_40 and 4176/26_33 from the populations "5216" and "4176" were crossed to the susceptible "IL 88." From each cross, a single F1 plant was selfed, resulting in 103 and 146 F2 seeds, respectively. These two F2 populations were named "dIL_5216" and "dIL_4176," respectively.

Genomic DNA Extraction

Genomic DNA was extracted from third leaves of barley seedlings using a guanidine isothiocyanate-based DNA isolation method in 96-well plate format as described earlier (Milner et al., 2018). dsDNA concentration was measured by Qubit® 2.0 Fluorometer using the Qubit™ dsDNA BR (Broad Range) Assay Kit (Invitrogen, Carlsbad, CA, United States) following the manufacturer's protocol.

Marker Development

To screen the initial IL mapping populations "4176" and "5216" for recombinants, nine CAPS markers (**Supplementary Table 1**) were designed to be evenly distributed over the distal 20 cM of barley chromosome 2HS, based on the conserved interspecific SNPs identified by targeted enrichment re-sequencing of 145 *Hb* ILs (Wendler et al., 2015). This set of ILs included the F4 homozygous 2HS IL (3026) from the mapping population "5216" and its associated donor lines, plus four additional *Hv* cultivars and four *Hb* accessions (Wendler et al., 2015). Only conserved *Hv*/*Hb* SNPs with a minimum of six-fold read coverage and located in the target region on the barley draft genome (International Barley Genome Sequencing Consortium., 2012) were selected and converted into CAPS markers using SNP2CAPS software (Thiel et al., 2004). Primer design was carried out using the default settings of Primer3 v.0.4.01 (Koressaar and Remm, 2007; O'Halloran, 2015) with minor modifications: The primer length was set between 19–21 bp. Primer melting temperature (Tm) was set to minimum Tm = 58°C, optimum Tm = 59°C and maximum Tm = 60°C. The product size was defined to be between 700 and 1,000 bp and Guanine-Cytosine content (GC-content) was set within the range of 50–55%.

CAPS Genotyping

Genotyping of populations "4176" and "5216" with the described CAPS markers was performed in a 20 µl PCR reaction volume including 20 ng genomic DNA, 0.1 U of HotStarTaq DNA Polymerase (Qiagen, Hilden, Germany), 1x PCR reaction buffer containing 15 mM $MgCl_2$ (Qiagen, Hilden, Germany), 0.2 mM of each dNTP (Fermentas, Fermentas, St. Leon-Rot, Germany), and 0.5 mM of each primer. All fragments were amplified using the following touchdown PCR profile: an initial denaturing step of 15 min at 95°C was followed by four cycles with denaturation at 95°C for 30 s, annealing at 62°C for 30 s (decreasing by 1°C per cycle), and extension at 72°C for 30 s. A final extension step was performed at 72°C for 7 min. The enzymatic digestion of the amplicons was performed in a 10 µl volume containing 5 µl of PCR product, 1× of appropriate buffer (New England Biolabs, Hitchin, United Kingdom), 1 U of enzyme (New England Biolabs, Hitchin, United Kingdom) and adjusted to final volume by adding molecular biology grade pure water. The reaction mix was incubated for 1 h at recommended incubation temperature. The digested PCR products were resolved by 1.5–2.5% gel-electrophoresis depending on amplicon size.

Genotyping-by-Sequencing and Data Analysis

GBS was used, following published procedures (Mascher et al., 2013b), to check the genetic purity and state of heterozygosity of F1 hybrid seeds derived from crosses between overlapping 2HS introgression lines as well as to genotype the whole F2 populations derived from these crosses. DNA of the progeny and parental lines were pooled per Illumina HiSeq2500 lane in an equimolar manner and sequenced for 107 cycles, single read, using a custom sequencing primer as previously described (Hoseinzadeh et al., 2019). The reads were aligned to the TRITEX genome assembly of barley cultivar Morex (Monat et al., 2019) with BWA-MEM version 0.7.12a (Li and Durbin, 2009). Variants were filtered following the protocol of Milner et al. (2018) for a minimum depth of sequencing of four to accept a genotype call, a minimum

TABLE 1 | *Hb* introgression lines containing segments overlapping with IL "4176" and IL "5216" and their observed resistance phenotype to the *Bgh* isolate CH4.8.

IPK ID	IL code (New Zealand)	Crossing scheme	Introgression location GBS	Start (Mb)	End (Mb)	Phenotype of ILs to isolate CH4.8
88	200A3/7/M1	Emir x A17/1	2HS	0	19,4	Susceptible
99	213G3/2/2/2/1	Emir x (2920/4 x Tinos16/1)	2HS, 6HS	0	22,1	Resistant
116	230H24/5/M1/M1	Morex x 2032	2HS	0	19,4	Resistant

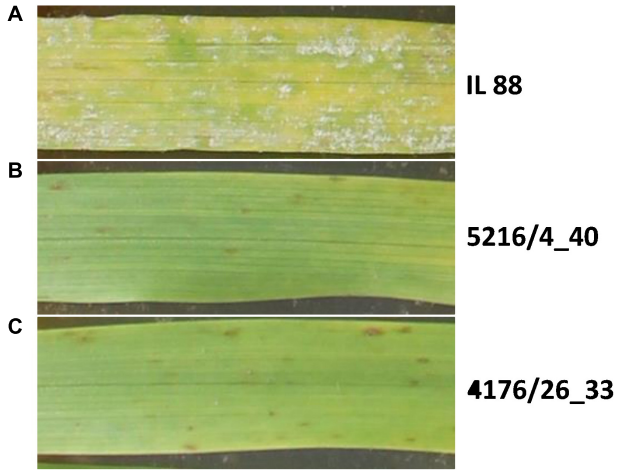

FIGURE 2 | Powdery mildew infection (isolate CH4.8) on the second leaf of seedling from the three independent *Hv*/*Hb* parental ILs, 7 days after inoculation. **(A)** Phenotype of the susceptible parental introgression line "IL 88" population "dIL_5216" and "dIL_4176." **(B)** Phenotype of the resistant parental introgression line 5216/4_40 from population "dIL_5216." **(C)** Phenotype of the resistant parental introgression line 4176/26_33 from population "dIL_4176."

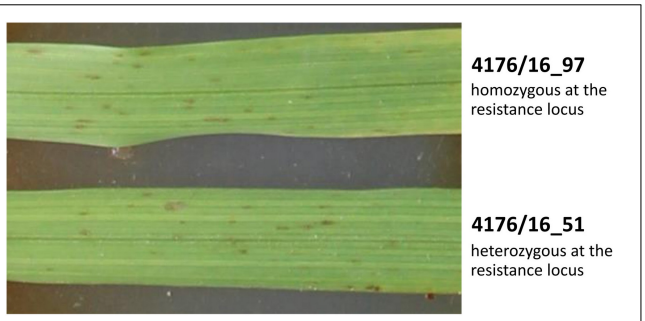

FIGURE 3 | Powdery mildew infection (isolate CH4.8) on the second leaf of seedling from two BC2F6 ILs from population "4176," 7 days after inoculation. 4176/16_97 is homozygous at the resistance locus whereas 4176/16_51 is heterozygous. Their resistance phenotype is identical and present necrotic lesions characteristic of HR.

mapping quality score of the SNPs (based on read depth ratios calculated from the total read depth and depth of the alternative allele) of six, a maximum fraction of heterozygous call of 70% and a maximum fraction of 25% of missing data. The resulting tables of polymorphisms are provided in **Supplementary Tables 2**, **3**.

RESULTS

Inheritance of the Resistance Contributed by *Hb*

The susceptible parental *Hv* genotype "Borwina" consistently displayed a leaf infection area of ≥80% in all experiments. The scoring rates of all susceptible individuals of both populations "4176" and "5216" was similar to the susceptible parent and did not significantly vary between phenotyping experiments, indicating high infection efficiency and reproducibility in all phenotyping experiments. The resistance phenotype to the CH4.8 powdery mildew isolate of plants carrying the *Hb* introgressed segment in a heterozygous state was identical to that of plants homozygous for the *Hb* fragment (**Figure 3**). The resistance phenotype was invariably accompanied with a hypersensitive response (HR) forming a necrotic lesion.

Phenotypic segregation for powdery mildew resistance against CH4.8 isolate was consistent with a 3:1 ratio (resistant/susceptible, R/S, $P < 0.05$) in all mapping populations, indicating the control of resistance by a single dominant resistance gene (**Table 2**). We propose the temporary name *Mlhb.A42* for this locus, based on previous naming of *Hb* powdery mildew resistance genes (Pickering et al., 1995; Steffenson, 1998).

TABLE 2 | Phenotypic segregation of powdery mildew resistance in each of 2HS IL mapping populations.

Mapping population	Number of resistant lines	Number of susceptible lines	χ^2
"4176" (BC1F6)	199	67	0,005**
"5216" (F7)	150	53	0,13**
"dIL_5216" (F2)	81	22	0.73**
"dIL_4176" (F2)	114	32	0.74*

χ^2 indicates the result of the Chi squared test performed to determine the goodness to fit a 3R:1S ratio, expected for a dominant monogenic inheritance. Asterisks indicate significance of the test with a p-value inferior to 1% (**) and 5% (*).

Recombination Frequency and Mapping of *Mlhb.A42* in Single-Introgression Line Mapping Populations

Genotyping of the mapping populations employed nine CAPS markers designed based on existing exome capture re-sequencing data (**Supplementary Table 1**). This showed that the BC1F6 population "4176" carried a longer introgressed segment compared to the F7 population "5612" (**Figure 4**). The genotyping results for population "5612" confirmed the extent of the introgressed segment, previously detected by exome capture in the F6 generation (Wendler et al., 2015).

Genotyping of 266 and 203 individuals in the initial IL mapping populations "4176" and "5612," uncovered only three and two recombinants within the introgressed segments, respectively (**Figure 4**). The introgressed segment therefore has a genetic length of approximately 1 centiMorgan (cM). Yet, the segment of population "5612" represents 10 cM on the barley POPSEQ map (Mascher et al., 2013a; Wendler et al., 2015). This confirms the anticipated reduced recombination between the *Hv* genome and the introgressed *Hb* segment. The resulting genetic interval for *Mlhb.A42* is flanked by markers CAPS02 and CAPS11, corresponding to a 9.5 Mbp interval between bp positions 2,269,761 and 11,819,688 on chromosome 2HS.

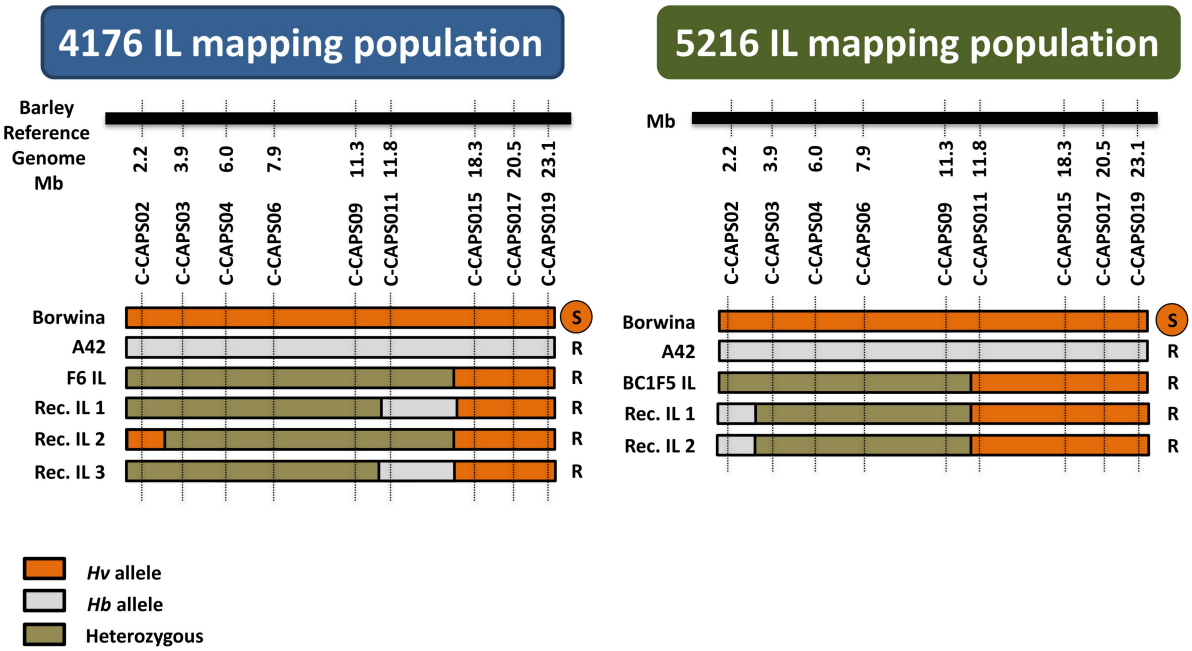

FIGURE 4 | Graphical genotypes and phenotypes of the recombinant *Hv/Hb* ILs from the two IL mapping populations. "Borwina" and "A42" are the original susceptible and resistant parents of the populations. "F6 IL" and "BC1F5 IL" represent the plant that was selfed in the previous generation to obtain populations "4176" and "5216," respectively. "Rec. IL" 1 to 3 are the identified recombinant plants. Nine CAPS markers named C-CAPS02 to C-CAPS019, were developed based on conserved *Hv/Hb* SNP loci described in Wendler et al. (2014), spanning the terminal 23 Mbp of barley chromosome 2HS. The black horizontal bars represent schematically the barley reference genome. The physical position of each selected SNPs on the barley reference genome is written below the black line. The homozygous state of the alleles from susceptible and resistant parents is shown as orange and gray colors, respectively, whereas heterozygous state is shown in green. The phenotype of each recombinant is indicated on the right of their genotype (R = resistant; S = susceptible). The number of recombinants identified through screening with the developed markers was only three and two in IL mapping population "4176" and "5216," respectively.

Recombination and Mapping of *Mlhb.A42* in Double Introgression Lines

The 103 and 146 F2 plants from the two double introgression populations "dIL_5216" and "dIL_4176," respectively, were genotyped by GBS and phenotyped for resistance against CH4.8 isolate. Based on obtained SNP variants across candidate interval for *Mlhb.A42* defined in the initial mapping populations, 14 and 36 recombinants were obtained in "dIL_5216" and "dIL_4176," respectively (**Figure 5**). The phenotypes of all non-recombinant plants corresponded to their genotype in this interval. Association between the phenotypes and genotypes of the recombinants defined overlapping 1.7 Mbp (between 7,943,409 and 9,595,691 bp) and 1.4 Mbp (between 8,193,151 and 9,595,691 bp) intervals on barley chromosome 2HS (Monat et al., 2019) in "dIL_5216" and "dIL_4176," respectively (**Figure 5**).

Candidate Genes Within the Target Interval

Based on the annotated barley reference genome sequence (Monat et al., 2019), 46 high confidence (HC) genes (**Table 3**) are located within the 1.4 Mbp *Mlhb.A42* interval as delimited in the "dIL_4176" population. Those genes with a putative functional annotation included a nucleotide-binding-site leucine-rich-repeat class of gene (NBS-LLR), an HR-like lesion-inducing protein-coding gene, a lectin receptor kinase (LecRK) and a Heat shock protein 90, all gene functions that could be directly or indirectly implied with plant resistance to biotic and/or abiotic stresses and therefore, qualify as candidate genes for *Mlhb.A42*. HORVU.MOREX.r2.2HG0082250 is annotated as a LecRK. However, analysis of its protein sequence with InterPro (Mitchell et al., 2018), showed that, like LecRKs, it is composed of an extracellular legume (L-type) lectin domain and a transmembrane domain. However, its cytoplasmic domain is only 23 amino-acid long and does not bear a kinase domain, making this gene a probable lectin receptor-like protein (LecRLP).

DISCUSSION

We report the fine mapping of *Mlhb.A42*, a dominant powdery mildew resistance locus originating from *Hb*, using mapping populations derived by crossing two independent, resistant and susceptible, *Hv*/*Hb* ILs. The strong suppression of recombination between homeologous genomic segments in *Hv*/*Hb* introgression lines, which typically results in severe linkage drag, previously represented a barrier to the efficient utilization of *Hb* germplasm in barley crop improvement, and to the isolation of disease resistance genes introgressed from *Hb*.

The genomic resource created by the GBS study of Wendler et al. (2015) on 145 *Hv/Hb* ILs proved to be a useful tool

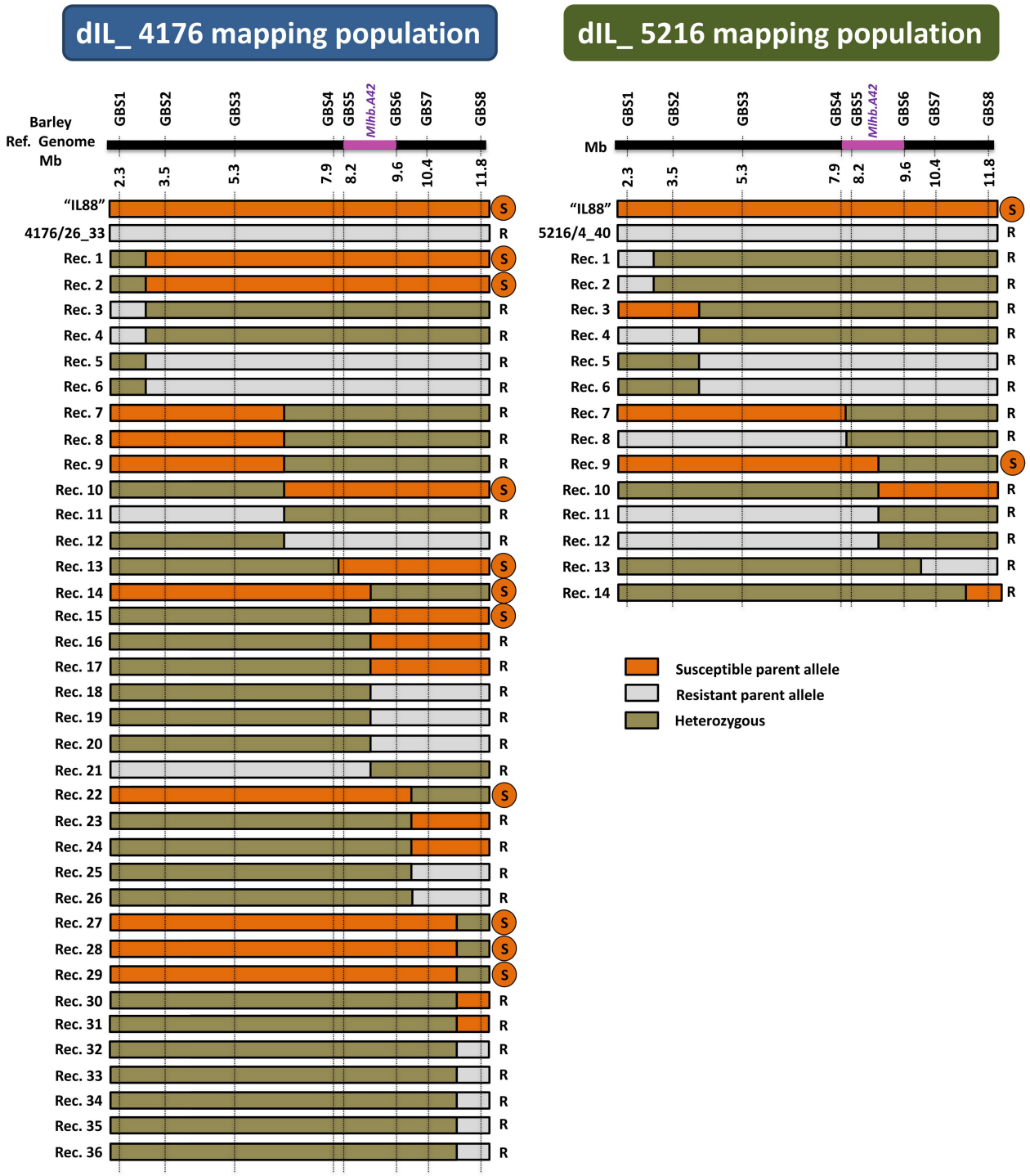

FIGURE 5 | Intraspecific mapping of the powdery mildew resistance locus *Mlhb.A42* in F2 populations derived from two different 2HS *Hv/Hb* introgression lines. The horizontal black bar schematically represents barley chromosome 2HS. The physical coordinates of each GBS marker are indicated below it. The genomic region containing the *Mlhb.A42* locus on barley chromosome 2HS is shown in purple. The graphical genotypes of recombinants are indicated as horizontal bars. The genotype of the susceptible parent "IL88" is represented in orange, the one of the resistant parents 4176/26_33 and 5216/4_40 is represented in gray, and heterozygous state is shown in olive green. The phenotype of each recombinant plant is indicated on the right of its genotype (R = resistant; S = susceptible). The recombination rates in "dIL_4176" and "dIL_5216" are 24.7 and 13.6%, respectively, corresponding to a 20-fold increase compared to single introgression lines. The *Mlhb.A42* gene was allocated to a 1.4 Mb interval, based on an estimate using the *Hv* genome as reference.

TABLE 3 | High confidence (HC) genes based on automated annotation of barley reference genome.

Gene name	Start[1]	End[1]	Annotation
HORVU.MOREX.r2.2HG0081650	8223803	8224706	Serine/threonine-protein phosphatase 7 long form-like protein
HORVU.MOREX.r2.2HG0081660	8227095	8233246	NBS-LRR-like resistance protein
HORVU.MOREX.r2.2HG0081670	8278638	8280547	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein, putative
HORVU.MOREX.r2.2HG0081680	8302525	8304486	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein, putative
HORVU.MOREX.r2.2HG0081690	8320263	8321485	12-oxophytodienoate reductase-like protein
HORVU.MOREX.r2.2HG0081700	8328715	8329968	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein, putative
HORVU.MOREX.r2.2HG0081720	8364720	8365206	NADH-ubiquinone oxidoreductase chain 1
HORVU.MOREX.r2.2HG0081740	8381983	8383512	ATP synthase subunit alpha
HORVU.MOREX.r2.2HG0081760	8470715	8473118	HR-like lesion-inducing protein-related protein
HORVU.MOREX.r2.2HG0081770	8474573	8487616	Actin-related protein
HORVU.MOREX.r2.2HG0081780	8511271	8512941	Cytochrome P450
HORVU.MOREX.r2.2HG0081810	8576680	8577147	Cytochrome P450
HORVU.MOREX.r2.2HG0081820	8607940	8608750	Cytochrome P450
HORVU.MOREX.r2.2HG0081830	8615323	8617046	Cytochrome P450
HORVU.MOREX.r2.2HG0081840	8664852	8679924	Cytochrome P450
HORVU.MOREX.r2.2HG0081850	8686994	8691803	Kaurene synthase
HORVU.MOREX.r2.2HG0081860	8729062	8731108	Cytochrome P450, putative
HORVU.MOREX.r2.2HG0081880	8896440	8898414	Copalyl diphosphate synthase
HORVU.MOREX.r2.2HG0081890	8905788	8907365	Cytochrome P450
HORVU.MOREX.r2.2HG0081900	8930477	8930938	Isoaspartyl peptidase/L-asparaginase
HORVU.MOREX.r2.2HG0081920	8961925	8964012	Copalyl diphosphate synthase
HORVU.MOREX.r2.2HG0081930	8965503	8967761	Cytochrome P450
HORVU.MOREX.r2.2HG0081980	9008320	9009897	Cytochrome P450
HORVU.MOREX.r2.2HG0082000	9064624	9068933	Copalyl diphosphate synthase
HORVU.MOREX.r2.2HG0082010	9095159	9096736	Cytochrome P450
HORVU.MOREX.r2.2HG0082020	9253167	9258032	Agenet domain, putative
HORVU.MOREX.r2.2HG0082050	9310583	9312095	Chalcone synthase
HORVU.MOREX.r2.2HG0082060	9337837	9339296	O-methyltransferase family protein
HORVU.MOREX.r2.2HG0082080	9357711	9359256	Glycosyltransferase
HORVU.MOREX.r2.2HG0082090	9360337	9378163	ABC transporter B family protein
HORVU.MOREX.r2.2HG0082100	9379916	9381655	Glycosyltransferase
HORVU.MOREX.r2.2HG0082140	9435844	9439807	Transcription factor
HORVU.MOREX.r2.2HG0082150	9441094	9442595	Serine/threonine-protein kinase
HORVU.MOREX.r2.2HG0082160	9462767	9463051	TTF-type zinc finger protein with HAT dimerization domain-containing protein
HORVU.MOREX.r2.2HG0082180	9468781	9469636	F-box protein PP2-A13
HORVU.MOREX.r2.2HG0082190	9472177	9476645	ATP sulfurylase (Sulfate adenylyltransferase)
HORVU.MOREX.r2.2HG0082200	9478069	9482961	Zinc finger family protein
HORVU.MOREX.r2.2HG0082210	9488137	9491244	carbohydrate esterase, putative (DUF303)
HORVU.MOREX.r2.2HG0082220	9498236	9499725	Serpin
HORVU.MOREX.r2.2HG0082240	9505214	9505813	Maternal effect embryo arrest protein
HORVU.MOREX.r2.2HG0082250	9538406	9539521	Lectin receptor kinase
HORVU.MOREX.r2.2HG0082260	9543886	9546930	Carboxymethylenebutenolidase-like protein
HORVU.MOREX.r2.2HG0082270	9550542	9550993	NAD(P)-binding rossmann-fold protein
HORVU.MOREX.r2.2HG0082280	9555561	9557307	Nicotianamine synthase
HORVU.MOREX.r2.2HG0082310	9591121	9593401	Heat shock protein 90
HORVU.MOREX.r2.2HG0082320	9595584	9597283	caspase-6 protein

[1] *Coordinates based on TRITEX Morex assembly (Monat et al., 2019).*

to identify suitable partners for the development of double ILs populations. The exploitation of intraspecific recombination allowed us to overcome the barrier to recombination usually observed in IL populations. Recombination rates in the region carrying the introgressed fragment were estimated to be 24.7% ("dIL_4176") and 13.6% ("dIL_5216"), comparable to the corresponding 10% rate observed in pure *Hv/Hv* mapping populations [e.g. POPSEQ map (Mascher et al., 2013a)]. The rate of recombination was exceeded by a factor of 20-fold as compared to the F7 and BC1F6 single IL populations "5216" and "4176"

(approximately 1%). The polymorphisms and markers identified in this study can be converted into KASP assays which would enable for rapid and high-throughput screening of large breeding population for the purpose of introgression of *Mlhb.A42* into barley cultivars.

The 1.4 Mbp identified target interval is containing 46 annotated HC genes on the most recent chromosome-scale genome assembly of cultivar "Morex" (Monat et al., 2019). The powdery mildew resistance conferred by this *Hb* locus is dominantly inherited, displaying chlorotic/necrotic flecks of HR with collapsed hyphae at inoculation sites, suggesting a salicylic acid (SA)-independent resistance pathway. Genes from the NBS-LLR or the receptor-like-kinase (RLK) families are over-represented among genes conferring this type of strong dominant resistance to pathogens (Kourelis and van der Hoorn, 2018), making genes from these families likely candidates in the context of this study. HORVU.MOREX.r2.2HG0081660, one of the annotated genes in the interval, is a coiled-coil (CC)-NBS-LRR gene and therefore the most likely candidate gene. However, HORVU.MOREX.r2.2HG0082250 is annotated as a LecRK and could also be a good candidate. LecRKs are a type of RLK characterized by an extracellular lectin motif (Wu and Zhou, 2013). They have been described as implicated in biotic stress resistance, mostly to bacteria and fungi (Singh and Zimmerli, 2013). This type of gene has been identified in resistance to oomycetes (Wang et al., 2015a,b; Balagué et al., 2017) and fungi (Huang et al., 2013; Wang et al., 2014a) in *Arabidopsis thaliana* and to wheat powdery mildew in *Haynaldia villosa* (Wang et al., 2018). According to InterPro (Mitchell et al., 2018), HORVU.MOREX.r2.2HG0082250 does not bear a kinase domain and is likely to be a L-type LecRLP. So far, the only LecRLPs described have a lectin-like Lysin-motif (LysM)-type lectin domain. Two LysM-LecRLPs from *A. thaliana* and three from rice have been identified reported in context of disease resistance through interaction with the LysM-LecRK CERK1. The rice LysM-LecRLP CEBiP recognizes chitin and, through a direct interaction with CERK1, confers pattern-triggered immunity against fungi (Shimizu et al., 2010). Similarly, LYP4 and LYP6 both perceive peptidoglycan and chitin and interact with CERK1 (Liu et al., 2012). In *A. thaliana*, LYM1 and LYM3 sense pectidoglycan and trigger immunity, through CERK1, to bacterial infection (Willmann et al., 2011). *A. thaliana* contains only four L-type LecRLPs (Bellande et al., 2017) but, so far, no L-type LecRLPs have been functionally described. However, a mode of action of HORVU.MOREX.r2.2HG0082250 similar to the one of LysM-LecRLPs is a possibility for further investigation.

Almost all the genes annotated in the delimited target interval might be directly or indirectly involved in resistance to biotic and abiotic stresses and should be considered tentative candidates. NADH-ubiquinone oxidoreductase is involved in intracellular ROS production that can prevent pathogen infection (van der Merwe and Dubery, 2006). Cytochrome P450 and O-methyltransferase proteins are responsible for production of several molecules that can play a role in resistance to pathogens or pests (He and Dixon, 2000; Dixon, 2001; Noordermeer et al., 2001). Some glycosyl-transferase genes have been identified as necessary for the HR (Langlois-Meurinne et al., 2005). The rice gene *OsHRL* encodes for a HR-like lesion inducing protein and has been shown to be associated with resistance to bacterial blight (Park et al., 2010). Copalyl-diphosphate synthases are implicated in the biosynthesis of phytohormones including gibberellic acid or phytoalexins, which contribute to defensive secondary metabolism (Prisic et al., 2004; Harris et al., 2005). Finally, HORVU.MOREX.r2.2HG0082310 is annotated as a Heat shock protein 90 (Hsp90), which are involved in stress resistance, in particular disease resistance, mediating signal transduction for HR (Xu et al., 2012). Notably, it has been shown by virus-induced gene silencing that an *Hsp90* gene is required for *Mla13* resistance of barley to powdery mildew (Hein et al., 2005).

In addition to the annotated genes, other genes, e.g. members of NBS-LRR and RLK families, might be present in the interval of the resistant *Hb* parent but missing in the respective interval of the "Morex" reference sequence. Indeed, *Hv* and *Hb* diverged 6 million years ago, accumulating structural variations since then (Jakob and Blattner, 2006). Therefore, this *Hb* resistance to powdery mildew could be due to a gene absent from the barley reference genome. In particular, NBS-LRR genes are subject to frequent duplication (Flagel and Wendel, 2009), and resistances conferred by NBS-LRRs genes are frequently due to presence/absence variation of such genes (Grant et al., 1998; Griffiths et al., 1999; Henk et al., 1999; Stahl et al., 1999; Tian et al., 2002). The wheat powdery mildew-resistance gene *Pm21* (Xing et al., 2018) originates from the wheat/*Dasypyrum villosum* translocation line T6AL.6VS and is localized in a region presenting a high level of synteny with the *Mlhb.A42* locus. This gene confers broad spectrum dominant resistance against wheat powdery mildew isolates and encodes a RPP13-like NBS-LRR gene (He et al., 2017; Xing et al., 2018). The protein sequence of *Pm21* only shares 34% identity with the translated nucleotide sequence of HORVU.MOREX.r2.2HG0081660. However, its localization in a syntenic region to the resistance conferring *Mlhb.A42* locus provides a tempting working hypothesis that *Pm21* and *Mlhb.A42* could represent orthologous genes or members of the same locally evolved gene family.

A resequencing strategy of the *Mlhb.A42* locus in both resistant and susceptible *Hb* will be necessary to ascertain the structure of the locus and the number and nature of the candidate genes. With the development of new sequencing methods this could be achieved by Cas9-guided enrichment of the locus (Wang et al., 2014b) or sorting and sequencing for assembly of the chromosome 2H in one of the two introgression lines (Thind et al., 2017). Moreover, a locus of resistance to powdery mildew has been identified on chromosome 2HS in the *Hb* accessions S1 – where it was named *Mlhb1.a* – (Pickering et al., 1995; Shtaya et al., 2007), 2032 (Zhang et al., 2001; Shtaya et al., 2007), and A17 (Shtaya et al., 2007). Allelism tests are required to check whether the same locus is involved in the resistance from those four accessions or not. Xu and Kasha (1992) identified another *Hb* powdery mildew resistance gene in the accession GBC141. However, the causal dominant gene, designated as *Mlhb2.b*, was

located by Kasha et al. (1996) on chromosome 2HL and therefore not allelic to *Mlhb1.a* or *Mlhb.A42*.

Durability and spectrum of a resistance gene are two major criteria to assess its application potential (Mundt, 2014). Durability of resistance genes can become a major concern as deployed resistance genes are experiencing a boom-and-bust phenomenon (Gladieux et al., 2015). The spectrum of *Mlhb.A42* was not evaluated in this study and should be tested in order to ascertain its potential for field resistance. Durability is difficult to estimate under laboratory scenarios. The durability of orthologs genes can be used as a proxy, yet quite imperfect. As discussed earlier, the wheat resistance gene *Pm21*could be orthologous to *Mlhb.A42*. Varieties carrying *Pm21* have increasingly been cultivated in China in the recent years (Bie et al., 2015) and its durability can therefore be evaluated in real conditions. Unfortunately, in some wheat fields, new *Bgt* isolates, virulent against *Pm21* have been identified (Shi et al., 2009; Yang et al., 2009). However, resistance based on this gene persisted close to 40 years (Tang et al., 2018) and is still effective against more than a thousand of field isolates in China (Zeng et al., 2014) and Poland (Czembor et al., 2014). To counteract the risk of isolates breaking the resistance provided by a single locus it is of ongoing importance to identify new resistance genes and to pyramid new loci with existing sources of resistance to increase the durability of resistance in the system (Wu et al., 2019). Moreover, exploiting CWR resistances could be a way to unravel even more effective resistance genes. Indeed, Wang et al. (2019) showed that the *Hb* LecRLK gene of resistance to leaf rust *Rph22* confers a stronger resistance to leaf rust adapted to *Hv* than its *Hv* ortholog *Rphq2*. The hypothesis is that crop receptors have a lower recognition of crop-adapted pathogens than the CWR receptors because of adaptation of the pathogens during centuries of coevolution with their host plant.

In the current study we report genetic mapping of *Mlhb.A42*, a dominant resistance locus introgressed to cultivated barley from *Hb*. This work is a proof-of-concept study for establishing the basic steps of map-based cloning of genes present in *Hv/Hb* IL collections by exploiting double ILs mapping populations. Using this strategy, we circumvented the limitation of repressed meiotic recombination which was frequently observed in attempts of genetic mapping employing populations derived between *Hv*/*Hb* introgression lines and pure barley cultivars. Here, we observed similar or even higher recombination rates as expected in *Hv* and thus providing a major step toward facilitated exploitation of secondary gene pool-derived resistance genes in barley crop improvement.

DATA AVAILABILITY STATEMENT

The GBS datasets generated and analyzed in this study are deposited at e!DAL PGP repository (Arend et al., 2016) under the doi: 10.5447/IPK/2020/2.

AUTHOR CONTRIBUTIONS

PH performed the experimental work. PH and HP performed the data analysis and wrote the manuscript. BR-W provided seed material and sample information for two IL mapping populations "4176" and "5216." PS supervised the phenotyping. NS designed the study, supervised the experimental work, and contributed to the writing of the manuscript. All authors read, corrected, and approved the manuscript.

FUNDING

This manuscript is part of a Ph.D. study of PH, which was financially supported by a grant from the German Research Foundation (DFG) to PS and NS ("DURESTrit," STE 1102/5-1/608699) in frame of the ERACAPS initiative. This study was further supported by German Ministry of Food and Agriculture (BMEL, FKZ:2818201615) by reason to the decision of the German Bundestag.

ACKNOWLEDGMENTS

We gratefully appreciate the excellent technical support by Mary Ziems in performing crossings and Susanne Koenig in GBS library preparation. We kindly acknowledge Dr. Axel Himmelbach for his valuable support in the sequencing, Dr. Martin Mascher for his invaluable guidance in data processing and analysis, Dr. Dimitar Douchkov for his input on the phenotyping test and fungal material, and Dr. Timothy Rabanus-Wallace for language editing. We are grateful to Dr. Neele Wendler for her support in the employment of the previously developed GBS and exome capture data presented in the current study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00225/full#supplementary-material

REFERENCES

Arend, D., Junker, A., Scholz, U., Schüler, D., Wylie, J., and Lange, M. (2016). PGP repository: a plant phenomics and genomics data publication infrastructure. *Database* 2016:baw033. doi: 10.1093/database/baw033

Backes, G., Madsen, L., Jaiser, H., Stougaard, J., Herz, M., Mohler, V., et al. (2003). Localisation of genes for resistance against *Blumeria graminis* f. sp. hordei and *Puccinia graminis* in a cross between a barley cultivar and a wild barley (*Hordeum vulgare* ssp. spontaneum) line. *Theor. Appl. Genet.* 106, 353–362. doi: 10.1007/s00122-002-1148-1

Balagué, C., Gouget, A., Bouchez, O., Souriac, C., Haget, N., Boutet-Mercey, S., et al. (2017). The *Arabidopsis thaliana* lectin receptor kinase LecRK-I.9 is required for full resistance to *Pseudomonas syringae* and affects jasmonate signalling. *Mol. Plant Pathol.* 18, 937–948. doi: 10.1111/mpp.12457

Bellande, K., Bono, J.-J., Savelli, B., Jamet, E., and Canut, H. (2017). Plant lectins and lectin receptor-like kinases: how do they sense the outside? *Int. J. Mol. Sci.* 18:1164. doi: 10.3390/ijms18061164

Bie, T., Zhao, R., Jiang, Z., Gao, D., Zhang, B., and He, H. (2015). Efficient marker-assisted screening of structural changes involving *Haynaldia villosa* chromosome 6V using a double-distal-marker strategy. *Mol. Breed.* 35:34. doi: 10.1007/s11032-015-0211-y

Canady, M. A., Ji, Y., and Chetelat, R. T. (2006). Homeologous recombination in *Solanum lycopersicoides* introgression lines of cultivated tomato. *Genetics* 174, 1775–1788. doi: 10.1534/genetics.106.065144

Chen, S., Rouse, M. N., Zhang, W., Jin, Y., Akhunov, E., Wei, Y., et al. (2015). Fine mapping and characterization of *Sr21*, a temperature-sensitive diploid wheat resistance gene effective against the *Puccinia graminis* f. sp. tritici Ug99 race group. *Theor. Appl. Genet.* 128, 645–656. doi: 10.1007/s00122-015-2460-x

Corrion, A., and Day, B. (2015). *"Pathogen Resistance Signalling in Plants," in eLS*. New York, NY: American Cancer Society, 1–14. doi: 10.1002/9780470015902.a0020119.pub2

Czembor, H. J., Domeradzka, O., Czembor, J. H., and Mańkowski, D. R. (2014). Virulence structure of the powdery mildew (*Blumeria graminis*) population occurring on triticale (x Triticosecale) in poland. *J. Phytopathol.* 162, 499–512. doi: 10.1111/jph.12225

Czembor, J. H. (2002). Resistance to powdery mildew in selections from Moroccan barley landraces. *Euphytica* 125, 397–409. doi: 10.1023/A:1016061508160

Dixon, R. A. (2001). Natural products and plant disease resistance. *Nature* 411, 843–847. doi: 10.1038/35081178

Dreiseitl, A., and Bockelman, H. E. (2003). Sources of powdery mildew resistance in a wild barley collection. *Genet. Resour. Crop. Evol.* 50, 345–350. doi: 10.1023/A:1023953819787

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. *Trends Genet.* 24, 24–32. doi: 10.1016/j.tig.2007.11.001

Fischbeck, G., Schwarzbach, E., Sobel, Z., and Wahl, I. (1976). Mildew resistance in Israeli populations of 2-rowed wild barley (*Hordeum spontaneum*). *Z. Pflanzenzüchtung* 76, 163–166.

Flagel, L. E., and Wendel, J. F. (2009). Gene duplication and evolutionary novelty in plants. *New Phytol.* 183, 557–564. doi: 10.1111/j.1469-8137.2009.02923.x

Frisch, M., and Melchinger, A. E. (2001). The length of the intact donor chromosome segment around a target gene in marker-assisted backcrossing. *Genetics* 157, 1343–1356.

Gladieux, P., Feurtey, A., Hood, M. E., Snirc, A., Clavel, J., Dutech, C., et al. (2015). The population biology of fungal invasions. *Mol. Ecol.* 24, 1969–1986. doi: 10.1111/mec.13028

Grant, M. R., McDowell, J. M., Sharpe, A. G., de Torres Zabala, M., Lydiate, D. J., and Dangl, J. L. (1998). Independent deletions of a pathogen-resistance gene in *Brassica* and *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15843–15848. doi: 10.1073/pnas.95.26.15843

Griffiths, A. J. F., Gelbart, W. M., Miller, J. H., and Lewontin, R. C. (eds). (1999). "Chromosomal rearrangements," in *Modern Genetic Analysis* (New York, NY: W. H. Freeman).

Harris, L. J., Saparno, A., Johnston, A., Prisic, S., Xu, M., Allard, S., et al. (2005). The Maize *An2* gene is induced by *Fusarium* attack and encodesan ent-copalyl diphosphate synthase. *Plant Mol. Biol.* 59, 881–894. doi: 10.1007/s11103-005-1674-8

He, H., Zhu, S., Ji, Y., Jiang, Z., Zhao, R., and Bie, T. (2017). Map-based cloning of the gene *Pm21* that confers broad spectrum resistance to wheat powdery mildew. *bioRxiv* [preprint]. doi: 10.1101/177857

He, X.-Z., and Dixon, R. A. (2000). Genetic manipulation of isoflavone 7-O-methyltransferase enhances biosynthesis of 4'-O-Methylated isoflavonoid phytoalexins and disease resistance in alfalfa. *Plant Cell* 12, 1689–1702. doi: 10.1105/tpc.12.9.1689

Hein, I., Barciszewska-Pacak, M., Hrubikova, K., Williamson, S., Dinesen, M., Soenderby, I. E., et al. (2005). Virus-induced gene silencing-based functional characterization of genes associated with powdery mildew resistance in barley. *Plant Physiol.* 138, 2155–2164. doi: 10.1104/pp.105.062810

Henk, A. D., Warren, R. F., and Innes, R. W. (1999). A new ac-like transposon of *Arabidopsis* Is associated with a deletion of the *RPS5* disease resistance gene. *Genetics* 151, 1581–1589.

Hoffmann, A. A., Sgrò, C. M., and Weeks, A. R. (2004). Chromosomal inversion polymorphisms and adaptation. *Trends Ecol. Evol.* 19, 482–488. doi: 10.1016/j.tree.2004.06.013

Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J.-M., Skovmand, B., Taba, S., et al. (1999). Plant genetic resources: what can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U.S.A.* 96, 5937–5943. doi: 10.1073/pnas.96.11.5937

Hoseinzadeh, P., Zhou, R., Mascher, M., Himmelbach, A., Niks, R. E., Schweizer, P., et al. (2019). High resolution genetic and physical mapping of a major powdery mildew resistance locus in barley. *Front. Plant Sci.* 10:146. doi: 10.3389/fpls.2019.00146

Hospital, F. (2001). Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. *Genetics* 158, 1363–1379.

Huang, P., Ju, H.-W., Min, J.-H., Zhang, X., Kim, S.-H., Yang, K.-Y., et al. (2013). Overexpression of L-type lectin-like protein kinase 1 confers pathogen resistance and regulates salinity response in *Arabidopsis thaliana*. *Plant Sci.* 20, 98–106. doi: 10.1016/j.plantsci.2012.12.019

International Barley Genome Sequencing Consortium. (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543

Jakob, S. S., and Blattner, F. R. (2006). A chloroplast genealogy of *Hordeum* (Poaceae): long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Mol. Biol. Evol.* 23, 1602–1612. doi: 10.1093/molbev/msl018

Jarosch, B., Kogel, K.-H., and Schaffrath, U. (1999). The Ambivalence of the barley *Mlo* locus: mutations conferring resistance against powdery mildew (*Blumeria graminis* f. sp. hordei) Enhance Susceptibility to the Rice Blast Fungus Magnaporthe grisea. *Mol. Plant-Microbe Interact.* 12, 508–514. doi: 10.1094/MPMI.1999.12.6.508

Johnston, P. A., Meiyalaghan, V., Forbes, M. E., Habekuß, A., Butler, R. C., and Pickering, R. (2015). Marker assisted separation of resistance genes *Rph22* and *Rym16Hb* from an associated yield penalty in a barley: *Hordeum bulbosum* introgression line. *Theor. Appl. Genet.* 128, 1137–1149. doi: 10.1007/s00122-015-2495-z

Johnston, P. A., Niks, R. E., Meiyalaghan, V., Blanchet, E., and Pickering, R. (2013). Rph22: mapping of a novel leaf rust resistance gene introgressed from the non-host *Hordeum bulbosum* L. into cultivated barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 126, 1613–1625. doi: 10.1007/s00122-013-2078-9

Johnston, P. A., Timmerman-Vaughan, G. M., Farnden, K. J. F., and Pickering, R. (2009). Marker development and characterisation of *Hordeum bulbosum* introgression lines: a resource for barley improvement. *Theor. Appl. Genet.* 118, 1429–1437. doi: 10.1007/s00122-009-0992-7

Jørgensen, I. H. (1992). Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. *Euphytica* 63, 141–152. doi: 10.1007/BF00023919

Kakeda, K., Ibuki, T., Suzuki, J., Tadano, H., Kurita, Y., Hanai, Y., et al. (2008). Molecular and genetic characterization of the S locus in *Hordeum bulbosum* L., a wild self-incompatible species related to cultivated barley. *Mol. Genet. Genomics* 280, 509–519. doi: 10.1007/s00438-008-0383-9

Kasha, K. J., Pickering, R. A., William, H. M., Hill, M. A., Oro, R., Reader, S., et al. (1996). "GISH and RFLP facilitated identification of a barley chromosome carrying powdery mildew resistance from Hordeum bulbosum," in *Proceedings 7th International Barley Genetics Symposium, Saskatoon*, Canada.

Kerber, E. R., and Dyck, P. L. (1990). Transfer to hexaploid wheat of linked genes for adult-plant leaf rust and seedling stem rust resistance from an amphiploid of *Aegilops speltoides* × *Triticum monococcum*. *Genome* 33, 530–537. doi: 10.1139/g90-079

Kjær, B., Jensen, H. P., Jensen, J., and Jørgensen, J. H. (1990). Associations between three *ml-o* powdery mildew resistance genes and agronomic traits in barley. *Euphytica* 46, 185–193. doi: 10.1007/BF00027217

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23, 1289–1291. doi: 10.1093/bioinformatics/btm091

Kourelis, J., and van der Hoorn, R. A. L. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell* 30, 285–299. doi: 10.1105/tpc.17.00579

Kumar, J., Hückelhoven, R., Beckhove, U., Nagarajan, S., and Kogel, K.-H. (2001). A compromised mlo pathway affects the response of barley to the necrotrophic fungus *Bipolaris sorokiniana* (*Teleomorph*: *Cochliobolus sativus*) and its toxins. *Phytopathology* 91, 127–133. doi: 10.1094/PHYTO.2001.91.2.127

Langlois-Meurinne, M., Gachon, C. M. M., and Saindrenan, P. (2005). Pathogen-responsive expression of glycosyltransferase genes *UGT73B3* and *UGT73B5* is necessary for resistance to *Pseudomonas syringae* pv tomato in *Arabidopsis*. *Plant Physiol.* 139, 1890–1901. doi: 10.1104/pp.105.067223

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Liu, B., Li, J.-F., Ao, Y., Qu, J., Li, Z., Su, J., et al. (2012). Lysin Motif6Containing Proteins LYP4 and LYP6 play dual roles in peptidoglycan and chitin perception in rice innate immunity. *Plant Cell* 24, 3406–3419. doi: 10.1105/tpc.112.102475

Mains, E. B., and Dietz, S. M. (1930). Physiologic forms of barley mildew. *Erysiphe graminis hordei*. *Phytopathology* 20, 229–239.

Mascher, M., Muehlbauer, G. J., Rokhsar, D. S., Chapman, J., Schmutz, J., Barry, K., et al. (2013a). Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). *Plant J.* 76, 718–727. doi: 10.1111/tpj.12319

Mascher, M., Wu, S., Amand, P. St, Stein, N., and Poland, J. (2013b). Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PLoS One* 8:e076925. doi: 10.1371/journal.pone.0076925

McGrann, G. R. D., Stavrinides, A., Russell, J., Corbitt, M. M., Booth, A., Chartrain, L., et al. (2014). A trade off between mlo resistance to powdery mildew and increased susceptibility of barley to a newly important disease. *Ramularia leaf spot*. *J. Exp. Bot.* 65, 1025–1037. doi: 10.1093/jxb/ert452

Michel, M. (1996). *Untersuchungen zur Übertragung von Resistenzgenen aus der Wildart Hordeum bulbosum L. in die Kulturgerste Hordeum vulgare L. Diss.* Munich: TU München.

Milner, S., Jost, M., Taketa, S., Mazón, E., Himmelbach, A., Oppermann, M., et al. (2018). Genebank genomics reveals the diversity of a global barley collection. *Nat. Genet. Forthcomin.* 51, 319–326. doi: 10.1038/s41588-018-0266-x

Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., et al. (2018). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 47, D351–D360. doi: 10.1093/nar/gky1100

Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., et al. (2019). TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *bioRxiv* [preprint]. doi: 10.1101/631648

Mundt, C. C. (2014). Durable resistance: a key to sustainable management of pathogens and pests. *Infect. Genet. Evol.* 27, 446–455. doi: 10.1016/j.meegid.2014.01.011

Noordermeer, M. A., Veldink, G. A., and Vliegenthart, J. F. G. (2001). Fatty acid hydroperoxide lyase: a plant cytochrome P450 enzyme involved in wound healing and pest resistance. *ChemBioChem* 2, 494–504. doi: 10.1002/1439-7633(20010803)2:7/8<494::aid-cbic494>3.0.co;2-1

O'Halloran, D. M. (2015). PrimerView: high-throughput primer design and visualization. *Source Code Biol. Med.* 10:8. doi: 10.1186/s13029-015-0038-2

Park, S.-R., Moon, S.-J., Shin, D.-J., Kim, M.-G., Hwang, D.-J., Bae, S.-C., et al. (2010). Isolation and characterization of rice *OsHRL* gene related to bacterial blight resistance. *Plant Pathol. J.* 26, 417–420. doi: 10.5423/PPJ.2010.26.4.417

Pickering, R. A. (1991). Comparison of crossover frequencies in barley (*Hordeum vulgare*) and *H. vulgare* × *H. bulbosum hybrids* using a paracentric inversion. *Genome* 34, 666–673. doi: 10.1139/g91-102

Pickering, R. A., Hill, A. M., Michel, M., and Timmerman-Vaughan, G. M. (1995). The transfer of a powdery mildew resistance gene from *Hordeum bulbosum* L to barley (*H. vulgare* L.) chromosome 2 (2I). *Theor. Appl. Genet.* 91, 1288–1292. doi: 10.1007/BF00220943

Pickering, R. A., Steeffenson, B. J., Hill, A. M., and Borovkova, I. (1998). Association of leaf rust and powdery mildew resistance in a recombinant derived from a *Hordeum vulgare* × *Hordeum bulbosum* hybrid. *Plant Breed* 117, 83–84. doi: 10.1111/j.1439-0523.1998.tb01453.x

Prescott-Allen, C., and Prescott-Allen, R. (1986). *The First Resource.* New Haven, CT: Yale University Press.

Prisic, S., Xu, M., Wilderman, P. R., and Peters, R. J. (2004). Rice contains two disparate ent-copalyl diphosphate synthases with distinct metabolic functions. *Plant Physiol.* 136, 4228–4236. doi: 10.1104/pp.104.050567

Prohens, J., Gramazio, P., Plazas, M., Dempewolf, H., Kilian, B., Díez, M. J., et al. (2017). Introgressiomics: a new approach for using crop wild relatives in breeding for adaptation to climate change. *Euphytica* 213:158. doi: 10.1007/s10681-017-1938-9

Rao, M. V. P. (1979). *The Transfer of Alien Genes for Stem rust Resistance to Durum Wheat*, ed. S. Ramanujam (New Delhi: Indian Society of Genetics & Plant Breeding), 338–341.

Řepková, J., Teturová, K., Dreiseitl, A., and Soldánová, M. (2009). Characterization and chromosomal location of powdery mildew resistance genes from wild barley PI282605. *J. Plant Dis. Prot.* 116, 257–259. doi: 10.1007/BF03356319

Ruge, B., Linz, A., Pickering, R., Proeseler, G., Greif, P., and Wehling, P. (2003). Mapping of *Rym14Hb*, a gene introgressed from *Hordeum bulbosum* and conferring resistance to BaMMV and BaYMV in barley. *Theor. Appl. Genet.* 107, 965–971. doi: 10.1007/s00122-003-1339-4

Ruge-Wehling, B., Linz, A., Habekuß, A., and Wehling, P. (2006). Mapping of *Rym16Hb*, the second soil-borne virus-resistance gene introgressed from *Hordeum bulbosum*. *Theor. Appl. Genet.* 113, 867–873. doi: 10.1007/s00122-006-0345-8

Savary, S., Ficke, A., Aubertot, J.-N., and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Secur.* 4, 519–537. doi: 10.1007/s12571-012-0200-5

Scholz, M., Ruge-Wehling, B., Habekuß, A., Schrader, O., Pendinen, G., Fischer, K., et al. (2009). *Ryd4Hb*: a novel resistance gene introgressed from Hordeum bulbosum into barley and conferring complete and dominant resistance to the barley yellow dwarf virus. *Theor. Appl. Genet.* 119, 837–849. doi: 10.1007/s00122-009-1093-3

Shi, Y., Wang, B., Li, Q., Wu, X., Wang, F., Liu, H., et al. (2009). Analysis of the virulent genes of *Erysiphe graminis* f. sp. tritici and the resistance genes of wheat commercial cultivars in Shaanxi Province. *J. Triticeae Crop.* 29, 706–711.

Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., et al. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J.* 64, 204–214. doi: 10.1111/j.1365-313X.2010.04324.x

Shtaya, M. J. Y., Sillero, J. C., Flath, K., Pickering, R., and Rubiales, D. (2007). The resistance to leaf rust and powdery mildew of recombinant lines of barley (*Hordeum vulgare* L.) derived from *H. vulgare* × *H. bulbosum* crosses. *Plant Breed* 126, 259–267. doi: 10.1111/j.1439-0523.2007.01328.x

Singh, P., and Zimmerli, L. (2013). Lectin receptor kinases in plant innate immunity. *Front. Plant Sci.* 4:124. doi: 10.3389/fpls.2013.00124

Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M., and Bergelson, J. (1999). Dynamics of disease resistance polymorphism at the Rpm1 locus of *Arabidopsis*. *Nature* 400, 667–671. doi: 10.1038/23260

Steffenson, B. J. (1998). Coordinators reports: disease and pest resistance genes. *Barley Genet. newslette*

Szigat, G., and Szigat, G. (1991). Amphidiploid hybrids between *Hordeum vulgare* and *H. bulbosum*-basis for the development of new initial material for winter barley breeding. *Vor. Pflanzenzüchtg* 20, 34–39.

Tang, S., Hu, Y., Zhong, S., and Luo, P. (2018). The potential role of powdery mildew-resistance Gene *Pm40* in Chinese wheat-breeding programs in the Post-*Pm21* Era. *Engineering* 4, 500–506. doi: 10.1016/j.eng.2018.06.004

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

Tanksley, S. D., and Nelson, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203. doi: 10.1007/BF00223376

Thiel, T., Kota, R., Grosse, I., Stein, N., and Graner, A. (2004). SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Res.* 32, e5–e5. doi: 10.1093/nar/gnh006

Thind, A. K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., et al. (2017). Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nat. Biotechnol.* 35, 793–796. doi: 10.1038/nbt.3877

Tian, D., Araki, H., Stahl, E., Bergelson, J., and Kreitman, M. (2002). Signature of balancing selection in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11525–11530. doi: 10.1073/pnas.172203599

van der Merwe, J. A., and Dubery, I. A. (2006). Benzothiadiazole inhibits mitochondrial NADH: ubiquinone oxidoreductase in tobacco. *J. Plant Physiol.* 163, 877–882. doi: 10.1016/j.jplph.2005.08.016

von Bothmer, R., Jacobsen, N., Baden, C., Jorgensen, R. B., and Linde-Laursen, I. (1995). "An ecogeographical study of the genus Hor-deum," in *Systematic and Ecogeographic Studies on Crop Gene Pools 7*, (Rome: International Plant Genetic Resources Institute).

von Korff, M., Wang, H., Léon, J., and Pillen, K. (2005). AB-QTL analysis in spring barley. I. Detection of resistance genes against powdery mildew, leaf rust and scald introgressed from wild barley. *Theor. Appl. Genet.* 111, 583–590. doi: 10.1007/s00122-005-2049-x

Walther, U., Rapke, H., Proeseler, G., and Szigat, G. (2000). *Hordeum bulbosum*-a new source of disease resistance-transfer of resistance to leaf rust and mosaic viruses from *H. bulbosum* into winter barley. *Plant Breed* 119, 215–218. doi: 10.1046/j.1439-0523.2000.00475.x

Wang, Y., Bouwmeester, K., Beseh, P., Shan, W., and Govers, F. (2014a). Phenotypic analyses of *Arabidopsis* T-DNA insertion lines and expression profiling reveal that multiple L-Type lectin receptor kinases are involved in plant immunity. *Mol. Plant-Microbe Interact.* 27, 1390–1402. doi: 10.1094/MPMI-06-14-0191-R

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., et al. (2014b). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951. doi: 10.1038/nbt.2969

Wang, Y., Cordewener, J. H. G., America, A. H. P., Shan, W., Bouwmeester, K., and Govers, F. (2015a). *Arabidopsis* lectin receptor kinases LecRK-IX.1 and *LecRK-IX*.2 Are functional analogs in regulating phytophthora resistance and plant cell death. *Mol. Plant-Microbe Interact* 28, 1032–1048. doi: 10.1094/MPMI-02-15-0025-R

Wang, Y., Subedi, S., de Vries, H., Doornenbal, P., Vels, A., Hensel, G., et al. (2019). Orthologous receptor kinases quantitatively affect the host status of barley to leaf rust fungi. *Nat. Plants* 5, 1129–1135. doi: 10.1038/s41477-019-0545-2

Wang, Y., Weide, R., Govers, F., and Bouwmeester, K. (2015b). L-type lectin receptor kinases in *Nicotiana benthamiana* and tomato and their role in *Phytophthora* resistance. *J. Exp. Bot.* 66, 6731–6743. doi: 10.1093/jxb/erv379

Wang, Z., Cheng, J., Fan, A., Zhao, J., Yu, Z., Li, Y., et al. (2018). LecRK-V, an L-type lectin receptor kinase in *Haynaldia villosa*, plays positive role in resistance to wheat powdery mildew. *Plant Biotechnol. J.* 16, 50–62. doi: 10.1111/pbi.12748

Wendler, N., Mascher, M., Himmelbach, A., Johnston, P., Pickering, R., and Stein, N. (2015). Bulbosum to Go: a toolbox to utilize *Hordeum vulgare*/*bulbosum* Introgressions for Breeding and Beyond. *Mol. Plant* 8, 1507–1519. doi: 10.1016/j.molp.2015.05.004

Wendler, N., Mascher, M., Nöh, C., Himmelbach, A., Scholz, U., Ruge-Wehling, B., et al. (2014). Unlocking the secondary gene-pool of barley with next-generation sequencing. *Plant Biotechnol. J.* 12, 1122–1131. doi: 10.1111/pbi.12219

Wendler, N., Mascher, M., Himmelbach, A., Bini, F., Kumlehn, J., and Stein, N. (2017). A High-density, sequence-enriched genetic map of *Hordeum bulbosum* and Its collinearity to *H. vulgare*. *Plant Genome* 10, 1–11. doi: 10.3835/plantgenome2017.06.0049

Wijnker, E., and de Jong, H. (2008). Managing meiotic recombination in plant breeding. *Trends Plant Sci.* 13, 640–646. doi: 10.1016/j.tplants.2008.09.004

Willmann, R., Lajunen, H. M., Erbs, G., Newman, M.-A., Kolb, D., Tsuda, K., et al. (2011). *Arabidopsis* lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19824–19829. doi: 10.1073/pnas.1112862108

Wu, X. X., Xu, X. F., Ma, D. X., Chen, R. Z., Li, T. Y., and Cao, Y. Y. (2019). Virulence structure and its genetic diversity analyses of *Blumeria graminis* f. sp. tritici isolates in China. *BMC Evol. Biol.* 19:183. doi: 10.1186/s12862-019-1511-3

Wu, Y., and Zhou, J.-M. (2013). Receptor-Like Kinases in Plant Innate Immunity. *J. Integr. Plant Biol.* 55, 1271–1286. doi: 10.1111/jipb.12123

Xing, L., Hu, P., Liu, J., Witek, K., Zhou, S., Xu, J., et al. (2018). *Pm21* from *Haynaldia villosa* Encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Mol. Plant* 11, 874–878. doi: 10.1016/j.molp.2018.02.013

Xu, J., and Kasha, K. J. (1992). Transfer of a dominant gene for powdery mildew resistance and DNA from *Hordeum bulbosum* into cultivated barley (*H. vulgare*). *Theor. Appl. Genet.* 84, 771–777. doi: 10.1007/BF00227383

Xu, Z.-S., Li, Z.-Y., Chen, Y., Chen, M., Li, L.-C., and Ma, Y.-Z. (2012). Heat shock protein 90 in plants: molecular mechanisms and roles in stress responses. *Int. J. Mol. Sci.* 13, 15706–15723. doi: 10.3390/ijms131215706

Yang, L., Xiang, L., Zeng, F., Wang, H., Shi, W., and Yu, D. (2009). Virulence gene structure analysis of *Blumeria graminis* f. sp. tritici in Hubei. *Plant Prot* 35, 76–79.

Yu, X., Kong, H. Y., Meiyalaghan, V., Casonato, S., Chng, S., Jones, E. E., et al. (2018). Genetic mapping of a barley leaf rust resistance gene *Rph26* introgressed from *Hordeum bulbosum*. *Theor. Appl. Genet.* 131, 2567–2580. doi: 10.1007/s00122-018-3173-8

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.* 2, 983–989. doi: 10.1038/35103590

Zeng, F., Yang, L., Gong, S., Shi, W., Zhang, X., Wang, H., et al. (2014). Virulence and Diversity of *Blumeria graminis* f. sp. tritici Populations in China. *J. Integr. Agric.* 13, 2424–2437. doi: 10.1016/S2095-3119(13)60669-3

Zhang, L., Pickering, R. A., and Murray, B. G. (2001). Hordeum vulgare × *H. bulbosum tetraploid* hybrid provides useful agronomic introgression lines for breeders. *New Zeal. J. Crop Hortic. Sci.* 29, 239–246. doi: 10.1080/01140671.2001.9514185

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 06 March 2020
doi: 10.3389/fpls.2020.00252

The Resurgence of Introgression Breeding, as Exemplified in Wheat Improvement

*Ming Hao[1,2], Lianquan Zhang[1,2], Shunzong Ning[2], Lin Huang[2], Zhongwei Yuan[2], Bihua Wu[2], Zehong Yan[2], Shoufen Dai[2], Bo Jiang[2], Youliang Zheng[2] and Dengcai Liu[1,2]**

[1] State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Sichuan Agricultural University, Ya'an, China, [2] Triticeae Research Institute, Sichuan Agricultural University, Ya'an, China

Breeding progress in most crops has relied heavily on the exploitation of variation within the species' primary gene pool, a process which is destined to fail once the supply of novel variants has been exhausted. Accessing a crop's secondary gene pool, as represented by its wild relatives, has the potential to greatly expand the supply of usable genetic variation. The crop in which this approach has been most strongly championed is bread wheat (*Triticum aestivum*), a species which is particularly tolerant of the introduction of chromosomal segments of exotic origin thanks to the genetic buffering afforded by its polyploid status. While the process of introgression can be in itself cumbersome, a larger problem is that linkage drag and/or imperfect complementation frequently impose a yield and/or quality penalty, which explains the reluctance of breeders to introduce such materials into their breeding populations. Thanks to the development of novel strategies to induce introgression and of genomic tools to facilitate the selection of desirable genotypes, introgression breeding is returning as a mainstream activity, at least in wheat. Accessing variation present in progenitor species has even been able to drive genetic advance in grain yield. The current resurgence of interest in introgression breeding can be expected to result in an increased deployment of exotic genes in commercial wheat cultivars.

Keywords: alien introgression, marker, synthetic wheat, *Triticum aestivum*, Triticeae

OPEN ACCESS

Edited by:
Pietro Gramazio, University of Tsukuba, Japan

Reviewed by:
Mark Owen Winfield, University of Bristol, United Kingdom
Jaroslav Doležel, Institute of Experimental Botany CAS, Czechia

***Correspondence:**
Dengcai Liu
dcliu7@sicau.edu.cn;
dcliu7@yahoo.com

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *06 January 2020*
Accepted: *18 February 2020*
Published: *06 March 2020*

Citation:
Hao M, Zhang L, Ning S, Huang L, Yuan Z, Wu B, Yan Z, Dai S, Jiang B, Zheng Y and Liu D (2020) The Resurgence of Introgression Breeding, as Exemplified in Wheat Improvement. Front. Plant Sci. 11:252. doi: 10.3389/fpls.2020.00252

INTRODUCTION

The process of crop domestication, in which selection is applied to a population of wild plants, inevitably introduces a major genetic bottleneck (Simmonds, 1976). After many centuries of farmer-based selection, the late 19th century ushered in the hugely successful era of scientific plant breeding, a process which is predicated on creating novel combinations of the genetic variants represented in the crop's primary gene pool (Borlaug, 1983). While continuing to deliver genetic advances in many crops, the evidence now suggests that a yield plateau has been reached in certain crops, leading to concerns that the still growing global demand for plant-based products (food, feed, fiber, and industrial products such as starch and oil) will not be easily met without innovations being made in breeding technology (Tester and Langridge, 2010). First and foremost is the need therefore to expand the genetic base of the most intensely bred crop species, notably maize, rice, and wheat.

The secondary gene pool of a crop species is populated by its wild (as well as, in some cases, its cultivated) relatives. Accessing this gene pool typically requires *ab initio* the creation of a wide hybrid between the domesticate and one of its related species. Reducing the representation of the exotic parent's genome is most straight-forwardly achieved by backcrossing the hybrid to the domesticate (Anderson and Hubricht, 1938; Anderson, 1949), resulting in so-called "introgression" materials. While creating a viable wide hybrid is generally considered to require human intervention, wide hybrids clearly occur in the wild, as witnessed by the evolution of many natural allopolyploids, such as tetraploid and hexaploid wheat, oat, cotton, sugarcane, canola, coffee, and tobacco (Ribeiro-Carvalho et al., 2004; Dvorak et al., 2006; Udall and Wendel, 2006; Soltis et al., 2009). Introgression has made some significant contributions to crop improvement: an example is the introduction into rice of a gene conditioning male sterility gene from its wild relative *Oryza rufipogon*, which has facilitated the production of plentiful and affordable F_1 hybrid grain, and thereby led to the widespread exploitation of heterosis (Yuan, 1966). In another example, a chromosome comprising a segment of a wheat chromosome fused to one from rye (*Secale cereale*) by the early 1990s was present in almost one half of the bread wheat cultivars bred in 17 countries (Rabinovich, 1998). Polyploid species such as wheat tend to be better able than diploid ones to tolerate the presence of exotic chromatin due to the buffering provided by the presence of multiple homeologs (Sears, 1956; Dubcovsky and Dvorak, 2007).

Bread wheat represents a case model for introgression breeding. The species belongs to the Triticeae tribe which houses over 500 species, including several other important crop species [macaroni wheat (*T. durum*), barley (*Hordeum vulgare*), and rye] and fodder species (Wang et al., 1994; Yen et al., 2005). The genome of each members of the tribe comprises a combination of 26 related, but distinct genomes (Liu et al., 2017a). The first reported wide hybrids made between Triticeae species date back to an attempt to cross bread wheat with rye by Wilson (1876). The literature describing so-called "chromosome engineering" (Sears, 1972) is dominated by experiments conducted using bread wheat as the recipient. Wide hybrids involving Triticeae species have been used to understand a number of basic phenomena, such as the loss of centromeric histone H3 which accompanies chromosome elimination (Sanei et al., 2011), the formation of unreduced gametes underlying spontaneous chromosome doubling (Hao et al., 2014) and the restriction of chromosome pairing in allopolypoids to homologs (Sears, 1976).

Many papers have been published on the theme of introgression in wheat, and some useful reviews have been assembled by Sears (1981), Friebe et al. (1996), Kole (2011), and Mason (2017). The aim here was therefore not to provide another review of wheat introgression, but rather to discuss how some recently developed breeding strategies are contributing to a resurgence of interest in introgression breeding. Although the focus of the review is very much concentrated on wheat, many of these issues will be equally applicable to other crop species.

THE BREAD WHEAT GENE POOL

The gene pool concept proposed by Harlan and de Wet (1971) rests on a classification of the ease/difficulty of generating viable sexual hybrids, which is the *sine qua non* of introgression. In their scheme, a crop species' near relatives fall into either the primary (GP-1), secondary (GP-2) or tertiary (GP-3) gene pools. Membership of the three gene pools in Triticeae is determined largely by the species' genomic constitution (Jiang et al., 1993; Feuillet et al., 2008). In the case of common wheat as the recipient species, we here re-consider Triticeae species as one of four types based on the genome constitution and the ease/difficulty of introgression breeding.

GP-1 species share the same three genomes (B, A, and D) which make up the bread wheat genome. These species, such as *T. spelta* and *T. macha*, are fully sexually compatible with bread wheat. As a result, introgression is achievable through conventional crossing and selection, since recombination between the chromosomes of the recipient bread wheat and the donor relative is effectively unrestricted. Some back-crossing is usually required to restore the bread wheat genetic background.

GP-2 species share only some of the bread wheat genomes. Two prominent examples are *T. turgidum,* a group of tetraploid wheats of genomic constitution BA and *Aegilops tauschii* (goatgrass), the donor of bread wheat's D genome. Like GP-1 species, these are generally readily crossable with wheat, although because of their unbalanced chromosome constitution, the resulting hybrids are typically only poorly fertile. Nevertheless, introgression is relatively straight-forward as it can be affected by homologous recombination. These species are of particular value as a genetic resource for bread wheat improvement. As a result, there has been a concerted, worldwide effort to mine variation from these GP-2 species (particularly *Ae. tauschii*) by creating synthetic hexaploid wheats (SHWs), which replicate the natural wide cross responsible for the speciation of hexaploid wheat (Mujeeb-Kazi et al., 1996).

GP-3 species share no homologous genomes with bread wheat; they thus include the majority of the Triticeae species, including the two domesticates rye (R genome) and barley (H genome). For these species, not only are wide crosses less easy to develop (generally requiring embryo rescue because the hybrid endosperm fails to develop), but also introgression has to rely on inducing either homeologous recombination or a chromosome breakage-fusion event.

GP-2/3 species feature at least one genome present in bread wheat, alongside at least one which is homeologous. Examples include *Ae. cylindrica* (CD) and *Ae. ventricosa* (DN) and the large number of synthetic wheat × rye amphiploids (BAR and BADR), referred to as triticales. Here, introgression is possible via homologous recombination, provided that the target gene resides within a chromosome belonging to the homologous genome. Otherwise, as for the GP-3 species, introgression has to rely on inducing either homeologous recombination or a chromosome breakage-fusion event.

In addition, it has been proposed to add a fourth gene pool (GP-4) to acknowledge the potential of transgenesis (Suslow et al., 2002; Kumar and Rustgi, 2014) and somatic hybridization

(Xia et al., 2003) to introduce genes without any requirement for the prior formation of a sexual hybrid. While wheat's GP-4 in principle harbors every living organism, from microbe to mammal, it also includes a number of related species, notably sorghum (*Sorghum bicolor*), Job's tears (*Coix lacryma-jobi*), and Cogon grass (*Imperata cylindrica*) (Liu et al., 2014); while it is possible to culture *in vitro* immature hybrid embryos formed when wheat is pollinated by these species and to regenerate viable plants, the non-wheat chromosomes are rapidly eliminated during the zygote's early cell divisions, so that the regenerants are effectively wheat haploids. Note in passing that access to transgenic crops is limited in many countries, meaning that GP-4 represents at best a theoretical resource, at least for the present.

BARRIERS TO EXPLOITING INTROGRESSION MATERIALS IN WHEAT IMPROVEMENT

While a substantial research investment has been made into creating introgression materials, their impact on wheat improvement has been relatively modest. The major reason for this outcome is that many of genotypes have proven to be defective in terms of plant type, grain yield and/or grain quality, reflecting a combination of linkage drag and an inadequate level of genetic complementation.

When a target gene is introduced, whether this is achieved using recombination or chromosome breakage/fusion, it will inevitably be accompanied by other genes linked to it on the introgressed segment; some of these genes may have deleterious consequences on the plant's agronomy, productivity or grain quality. The ideal therefore is to engineer a transfer which involves as short an introgressed segment as possible. Linkage drag can potentially be overcome where the introgression has been achieved by homologous recombination, since the length of the segment can in principle simply be reduced by inducing further rounds of recombination enabled by a program of backcrossing. Linkage drag is more difficult to negate in materials which have been generated as a result of a homeologous recombination event, because no further recombination will occur once the wild type restriction over the pairing of homeologs has been restored (Sears, 1976).

The potential for inadequate genetic complementation becomes increasingly likely where the donor species is only distantly related to the recipient. In this case, the donor and recipient species have been isolated from one another over such a long period that they will have diverged substantially at the genetic level. As a result, the gene content of the introgressed exotic segment and the wheat segment which has been replaced may not be the same, leading to a progeny in which the introgressed genes cannot fully complement those present on the replaced wheat chromosome segment. This situation can lead to deleterious effects on the plant's agronomic performance (Birchler and Veitia, 2012). While genetically unbalanced genotypes can be informative for the purpose of genetic analysis, they are seldom of value in the context of varietal improvement.

The genetic background in which an exotic transfer has been engineered can also discourage the breeder take-up of introgression materials. Much of the research effort in bread wheat has been focused on the Sichuan province landrace Chinese Spring (CS), because of its choice for the development of the aneuploid and mutant stocks required for chromosome engineering (Sears and Miller, 1985). This choice has had some unfortunate consequences, since the phenotype of CS is deficient with respect to several key quantitative traits, and has proven difficult to correct (Liu et al., 2018). Nevertheless, a substantial body of germplasm harboring introgression products has been developed by various programs, although so far their impact on wheat improvement has been minimal, given that breeders are unwilling to break up the constellation of favorable alleles which they so laboriously assembled through many years of crossing and selection. Some recent advances in introgression methodology and the development of genomic resources have the potential to overcome some of the problems associated with the exploitation of introgression materials.

THE USE OF DOUBLE MONOSOMICS TO GENERATE ROBERTSONIAN TRANSLOCATIONS

Unpaired meiotic chromosomes (univalents) have a tendency to spontaneously break at their centromere during anaphase I to form two fragments. This process can give rise to a Robertsonian translocation (RobT) where two different chromosomes simultaneously break in the same cell, since chromosomal fragments readily fuse with one another (Robertson, 1916; Sears, 1952). Balancing or compensating RobTs (cRobTs), in which a chromosome arm becomes fused to the opposite arm of its homeolog, represent a key intermediate in the process of inducing introgression in the situation where homologous recombination is not feasible. A small number of cRobTs have indeed had a considerable impact on wheat improvement; the most notable examples are the wheat–rye translocation 1BL.1RS, in which the short arm of wheat chromosome 1B has been replaced by the short arm of rye chromosome 1R (Mettin et al., 1973; Zeller, 1973), and the 6AL.6VS translocation, in which the long arm of wheat chromosome 6A has become fused to the short arm of *Dasypyrum villosum* chromosome 6V. The former cRobT was ubiquitous among high yielding CIMMYT and European wheats during the closing years of the 20th century, while the latter one, which carries a gene determining resistance to the foliar pathogen *Blumeria graminis* (powdery mildew), is present in many current Chinese cultivars (Cao et al., 2011).

Constructing a wheat plant in which two chromosomes are present as monosomics is a relatively straight-forward procedure given the range of cytogenetic stocks which have been assembled (Davies et al., 1985; Lukaszewski, 1993, 1997; Friebe et al., 2005). A list of some successful inductions is presented in **Table 1**. The principle of the double monosomic method is based on the certainty that, thanks to the suppression of homeolog pairing, the two monosomes will remain unpaired at meiotic metaphase I (giving rise to a meiotic constitution of $20'' + 2'$); selfed

TABLE 1 | Examples of compensating Robertsonian translocations generated by double monosomic method.

Donor species	Target traits and genes	Translocation chromosome	References
Haynaldia villosa	Resistance to wheat spindle streak mosaic virus	4VS·4DL	Zhang Q. et al., 2005*
H. villosa	Increasing gluten strength	1V#3S·1DL	Zhao et al., 2010**
H. villosa	Stem rust *Sr52*	T6AS·6V#3L	Qi et al., 2011**
H. villosa	/	A set of RobTs involving 12 *D. villosum* chromosomes	Liu et al., 2011a**
Aegilops searsii	Stem rust resistance	$3S^{s}S{\cdot}3AL$, $3S^{s}S{\cdot}3BL$, $3S^{s}S{\cdot}3DL$	Liu et al., 2011b*
Thinopyrum intermedium	Resistance to streak mosaic virus	7BS·7S#3L	Liu et al., 2011c**
Th. intermedium	Stem rust *Sr44*	7J#1S·7DL	Liu et al., 2013**
Th. bessarabicum	High Fe and Zn contents	$6E^{b}S{\cdot}6DL$	Ardalani et al., 2016*
Th. elongatum	Improving flour quality	1AS.1EL	Tanaka et al., 2017*
Rye	Stem rust *Sr59*	2DS·2RL	Rahmatov et al., 2016*
Rye	Powdery mildew *Pm56*	6RS·6AL	Hao et al., 2018*
Barley	β-glucan CslF6; long spikes	A complete set of six RobTs	Danilova et al., 2018**
Barley	Salt tolerance and β-glucan content	7BS·7HL	Türkösi et al., 2018**

**The production of double monosomic plants by crossing between substitution and normal lines; **The production of double monosomic plants by crossing between monosomic and addition lines.*

progeny of such a plant are subsequently screened, using a combination of genetic markers and karyotyping, to detect *de novo* cRobT products. Both marker and karyotyping technology have benefited from progress in DNA analysis, the former taking the form of DNA-based markers such as microsatellites and single nucleotide polymorphisms (SNPs), and the latter exploiting either genomic or fluorescence *in situ* hybridization (GISH and FISH). Double monosomics are most easily generated by crossing a whole chromosome substitution with a euploid plant (**Figure 1**): for instance, Ardalani et al. (2016) crossed a 6Eb(6D) substitution with the cultivar "Roushan," and among 80 F_2 segregants of the resulting $20'' + 6D' + 6Eb'$ plant, it was possible to identify, using a PCR-based screen, a plant carrying a 6DL.6EbS cRobT. Similarly, Hao et al. (2018) only needed to screen 69 segregants to uncover a 6AL.6RS cRobT. A second approach begins with a cross between an established monosomic line ($20'' + 1'$) and an addition line ($21'' + 1''$), as exemplified by Danilova et al. (2018); in this case, each of the three stocks monosomic for a group 7 homeolog ($20'' + 7A'$, $20'' + 7B'$ and $20'' + 7D'$) were crossed with a line carrying a disomic dose of barley chromosome 7H ($21'' + 7H''$); selection among the resulting F_1s was then made for plants having a somatic chromosome number of 42, which were expected to be of meiotic constitution $20'' + 7A/B/D' + 7H'$. A marker-based screen of the progeny of these selections was successful in identifying all six potential cRobTs involving each arm of the barley chromosome and its opposite wheat arm, and these were confirmed using GISH analysis.

EXPLOITING THE *Ph* SYSTEM TO GENERATE SUB-CHROMOSOMAL ARM INTROGRESSIONS

While Sears (1956) showed that it was possible to transfer a segment of an *Ae. umbellulata* chromosome carrying a gene conditioning leaf rust resistance into wheat by irradiation with X-rays, chromosome engineering was greatly accelerated following the discovery that the suppression of meiotic pairing between wheat homeologs is under relatively simple genetic control. The most important locus is *Ph1*, mapping to chromosome arm 5BL (Okamoto, 1957; Riley and Chapman, 1958), and now known to have formed as a result of the insertion of a segment of subtelomeric heterochromatin within a cluster of *cdc2*-related genes (Griffiths et al., 2006). Deleting the *Ph2* locus (mapping to chromosome arm 3DS) has an intermediate effect on pairing suppression (Mello-Sampayo, 1971; Sutton et al., 2003), while a large effect locus has recently been mapped to chromosome arm 3AL by Fan et al. (2019). A number of minor effect loci were identified by Sears (1976). Importantly for chromosome engineering technology, the suppression of homoelog pairing imposed by both *Ph1* and *Ph2* extends beyond wheat itself, and thus can be exploited to induce recombination between a number of wheat chromosomes and their homelogs from related species (Riley et al., 1968; Koebner and Shepherd, 1985, 1986, 1987; Ceoloni et al., 1992).

Since large translocations are prone to be deleterious because of linkage drag and/or inadequate genetic complementation, reducing their size, while maintaining the presence of the gene(s) targeted for introgression, is a desirable goal (Moore et al., 1993). Given that the success rate in inducing introgressions is quite low, and that conventional cytological methods are not capable of distinguishing between large and small introgression segments, pioneering attempts to achieve this goal were largely unsuccessful. Recent advances with respect to both chromosome manipulation and particularly in genetic marker technology, however, are changing this picture (Winfield et al., 2016; King et al., 2017, 2019; Koo et al., 2017). A list of successful examples is presented in **Table 2**. In a recent case example, an attempt to shorten the 6AL.6VS translocation has been described by Lukaszewski and Cowger (2017). The target gene on the *Ha. villosa* chromosome arm is *Pm21*, which protects against infection by *B. graminis* (Cao et al., 2011). To reduce the length of 6VS chromatin present, the cRobT was introduced into a background deficient for *Ph1* to generate plants of constitution $19'' + 5B''$ (*ph1b*) + 6AL.6VS′ + 6A′. A marker-based

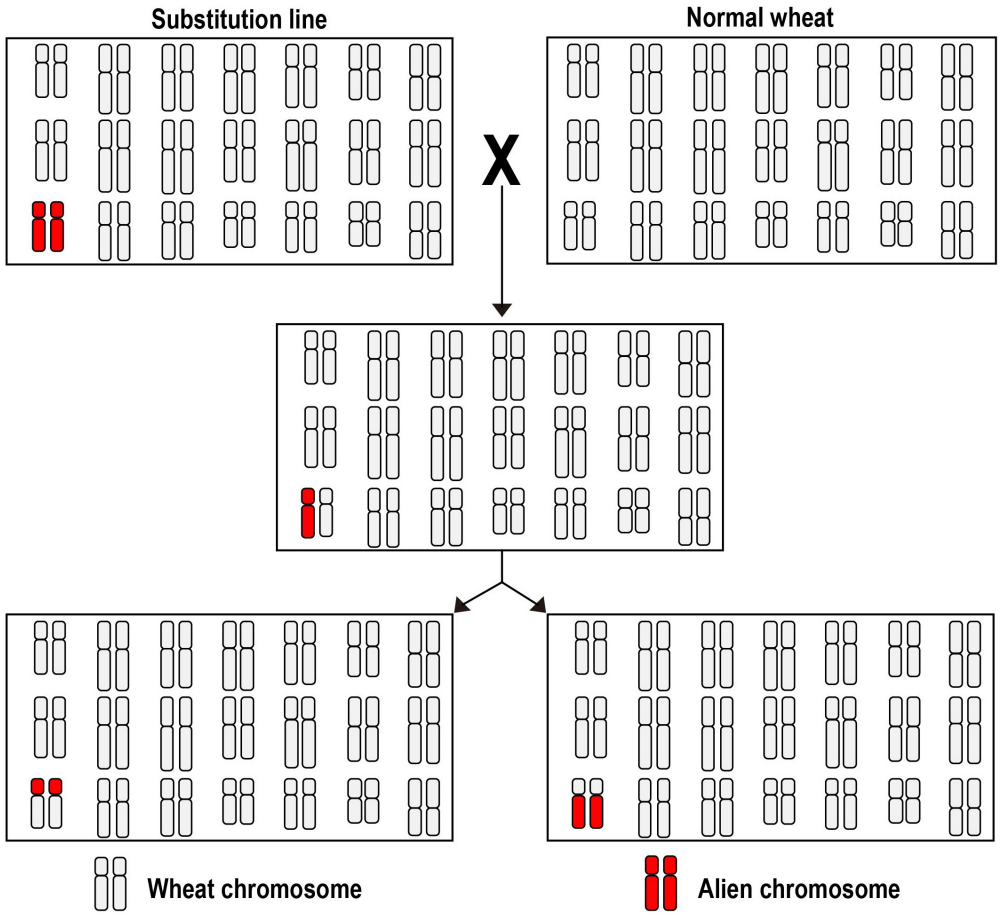

FIGURE 1 | The double monosomic method use to induce cRobTs. A euploid cultivar is crossed with a whole chromosome substitution line in which an exotic chromosome has replaced its wheat homeolog. The resulting F_1 hybrid carries two chromosomes which remain as univalents at meiosis metaphase I, giving the opportunity for a joint breakage/fusion event at anaphase I to create a cRobT.

screen of 997 progeny allowed for the identification of 29 new translocations involving combinations of segments from chromosome arms 6AS and 6VS. The critical step was then to cross a plant containing a translocation comprising the terminal segment of 6VS (including *Pm21*) with one (also including *Pm21*) which carried a translocation comprising the proximal end of 6VS; since the two translocations shared a small fragment 6VS, it was then possible to exploit homologous recombination to generate a chromosome harboring a small interstitial segment of 6VS (including *Pm21*) inserted into the sub-terminal region of 6AS (**Figure 2**).

HOMOLOGY-DIRECTED INTROGRESSION AIMED AT ENHANCING COMPLEX TRAITS

Most of the introgression experiments reported in the literature have sought to transfer a single gene trait, most prominently resistance to disease. Most breeders' traits – notably grain yield – are, however, under polygenic control. The "omnigenic model" has been developed as a means of accounting for the observation that despite the inheritance of certain human diseases being due to the effect of many genomically dispersed genes, some of these genes encode products having no obvious connection to the expression of the disease (Boyle et al., 2017). The model implies that gene regulatory networks are sufficiently interconnected with one another such that all genes expressed in disease-relevant cells are liable to affect the functions of core disease-related genes and that most of the heritability can be explained by effects on genes outside core pathways. An equivalent situation probably exists with respect to the genetic determination of yield, which is clearly complex given the large number of quantitative trait loci which have been identified as contributing to the trait. A consequence of the hexaploid status of bread wheat is that interactions between homeologous genes are likely an important component of the overall genetic determination of yield-related traits (Flood et al., 1992; Yang et al., 2017). A further consequence is that the triplication of many genes in bread wheat, arising from the presence of three homeologs, provides a buffering effect, such that allelic variation at one of the copies may result in just

TABLE 2 | Examples of small translocations generated by manipulation with *ph1b*.

Alien donor species	Target traits and genes	Translocation chromosomes	References
Aegilops umbellulata	/	1U/1B; 1U/1D	Koebner and Shepherd, 1987
Aegilops speltoides	Leaf rust *Lr47*; green bug *Gb5*	7S/7A	Dubcovsky et al., 1998
Ae. speltoides	Stem rust/*Sr39*	2S/2B	Niu et al., 2011
Ae. speltoides	/	2S/2B	Zhang et al., 2017, 2018
Ae. speltoides	Tan spot *TsrAes1*; *Septoria nodorum* blotch *SnbAes1*	2S/2B	Zhang et al., 2019b
Aegilops sharonensis	*Lr56/Yr38*	T6AS·6AL-6S^{sh}/6A	Marais et al., 2010
Aegilops searsii	Powdery mildew *Pm57*	2S^{s}#1/2B	Liu et al., 2017b
Secale cereale	Removing *secalin* of T1RS.1BL	T1RS·1BL/1B; T1RS·1BL/1D	Koebner and Shepherd, 1985, 1986
S. cereale	Removing *secalin* of T1RS.1BL	T1RS·1BL/1B	Lukaszewski, 2000
S. cereale	/	T2RS·2BL/2B; T2BS·2RL/2B	Lukaszewski et al., 2004
S. cereale	Removing *secalin* of T1RS.1BL	T1DS-1RS-1DS·1DL/1D	Anugrahwati, 2009
Thinopyrum ponticum	Leaf rust *Lr19*; yellow pigment	T7DS·7DL-7EL/7A; T7DS·7DL-7EL/7D	Zhang W. et al., 2005
Th. ponticum Th. intermedium	Barley yellow dwarf virus *Bdv2*; Rust resistance *Lr19* and *Sr25*	T7DS·7J?/T7DS·7DL-7S?	Ayala-Navarrete et al., 2007
Th. intermedium	Wheat streak mosaic virus *Wsm3*	T7BS·7S#3L/7B	Danilova et al., 2017
Th. elongatum	/	2E/2B	Zhang et al., 2017, 2018
Hordeum vulgare	/	4H^{ch}/4D	Rey et al., 2015
H. vulgare	β-glucan synthesis *HvCslF6*	T7HL·7AS/7A; T7HL·7BS/7B; T7HL·7DS/7D	Danilova et al., 2019
Elymus tsukushiensis	*Fusarium* head blight *Fhb6*	1Ets#1S/1A	Cainong et al., 2015
Haynaldia villosa	*Powdery mildew Pm21*	T6VS·6AL/6A	Lukaszewski and Cowger, 2017
Ha. villosa	Wheat yellow mosaic virus *Wss1*	T4VS·4DL/4D	Zhao et al., 2013; Dai et al., 2019
Ha. villosa	Stem rust *Sr52*	T6AS-6V#3L/6A	Li et al., 2019

a minor phenotypic effect (Dubcovsky and Dvorak, 2007). Thus, although variation at a single locus may well have an insubstantial effect on grain yield, combining a number of novel introgression segments into a single genotype has the potential to generate a significant positive effect. This possibility has been validated by the experience with SHWs, pursued originally by CIMMYT in the last decade of the 20th century, and later much expanded, both at CIMMYT and elsewhere. It has become clear that introgressing multiple segments from an SHW parent can significantly enhance the yield potential of wheat (Hoisington et al., 1999; Coghlan, 2006; Warburton et al., 2006; Dreisigacker et al., 2008; Trethowan and Mujeeb-Kazi, 2008; Yang et al., 2009; Li et al., 2018; Hao et al., 2019; Zhang et al., 2019a). The Chinese cultivar Chuanmai 42, for example, which has as one of its parents the CIMMYT SHW Syn769 (*T. durum* cv. Decoy 1/*Ae. tauschii* 188), recorded a yield of more than 6 t/ha in Sichuan regional trials, out-performing the commercial check variety by 20% over two consecutive years (Yang et al., 2009). It has been estimated that farmers using this cultivar would gain an extra 0.5–0.8 t/ha of grain (Li et al., 2014b). The recently released cultivars Shumai 580, Shumai 969, and Shumai 830 (Hao et al., 2019) were all bred from an SHW derived from the cross *T. turgidum* AS2255/*Ae. tauschii* AS60. Shumai 580 out-performed the best local check cultivar by 56% (equivalent to ∼2 t/ha) in a yield trial in Yunnan province, while Shumai 969 has remained the highest yielding cultivar in Sichuan trials over the past decade. The yield per spike achieved in Sichuan province by Shumai 830 has been consistently higher than any other commercial cultivar's.

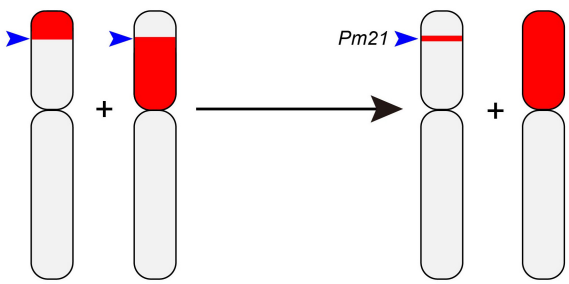

FIGURE 2 | The introgression of a small fragment of *Ha. villosa* chromosome 6V into a distal site of wheat chromosome arm 6AS. Two independent introgression lines were inter-crossed: one carries the terminal portion of chromosome arm 6VS and the other a proximal segment, while both carry *Pm21*, which lies in the short common interstitial segment. As a result of a homologous recombination event in the F_1 hybrid, segregants are generated which carry only the short interstitial segment [Figure redrawn from Lukaszewski and Cowger (2017)].

SEQUENCE-BASED ANALYSIS OF THE GENOMES OF SHW DERIVED CULTIVARS

A characterization of the introgressions carried by the three Shumai cultivars, as assessed at the sequence level by Hao et al. (2019), has shown that the ratio of the SHW parent's DNA retained lay in the range 12.4–15.0%. This outcome coincides well with the predicted ratio of 12.5% based on the assumption

that the breeding strategy employed did not discriminate against SHW alleles (except for those responsible for visible deleterious traits such as shattering, lateness and tallness). The genomes of the three Shumai cultivars featured a unique set of introgressions from SHW-L1: alleles from the B genome of the SHW genome were prominent in Shumai 580, those from the D genome in Shumai 969 and those from the A genome in Shumai 830. Although SHWs themselves tend to be poor agronomic performers, an analysis of a recombinant inbred population formed from a cross between SHW-L1 and an elite cultivar has shown that nearly a half (40/86) of the positive alleles at loci affecting yield were inherited from the SHW parent (Hao et al., 2019). Similar phenomenon has also been observed in other plant systems such as tomato (de Vicente and Tanksley, 1993). In addition to these quantitative trait loci, the Shumai cultivars inherited SHW alleles at *Ppd-A1*, *Vrn-A1*, *Vrn-B1*, *TaTEF*, *GPC-2*, *TaGASR7*, *TaGW2*, *TaCKX6*, and *TaSus1*, all of which are genes having a known impact on productivity (Nadolska-Orczyk et al., 2017). A similar representation of SHW alleles in materials selected from advanced backcross populations has been reported by Huang et al. (2003, 2004). The take home message is that while SHWs in themselves cannot compete with elite cultivars, they do represent a valuable source of beneficial alleles, so there is every reason to include them in bread wheat improvement programs.

In addition, SHW is different from bread wheat since its newly synthetic process may introduce transcriptome shock (Li et al., 2014a, 2015). Some marked differences have been noted between the transcriptomes of SHWs and conventional cultivars (Hao et al., 2017; Ramírez-González et al., 2018), resulting in the proposition that aspects of regulatory control differ (Li et al., 2015). Thus, besides the clear contribution of DNA sequence polymorphism, SHWs probably also deliver variation at the RNA level.

A STRATEGY FOR BREEDING BASED ON SHWS

The successful use of unadapted germplasm requires the introduction of genes determining the target trait(s) and selection against those determining deleterious traits. It has been conventionally held for a long time that using unadapted germplasm as a parent in a breeding program is inherently risky, since the size of breeding population required to breed these deleterious traits out may be uneconomically large. The indications are, however, that the efficiency of introgression breeding can be improved by adopting more focused strategies for both crossing and selection. The advanced backcross quantitative trait locus approach proposed by Tanksley and Nelson (1996) provides a method to retain favorable alleles inherited from the exotic donor parent while returning the background to that of the elite recipient. The concept has been applied to introgression programs in tomato (Tanksley et al., 1996), rice (Xiao et al., 1998), and bread wheat (Huang et al., 2003, 2004; Liu et al., 2006). However, in commercial practice, it has become evident that the necessity for extensive backcrossing has been exaggerated. Of 46 SHW-based cultivars released in 15 countries in the period 2003–2017, only four have been based

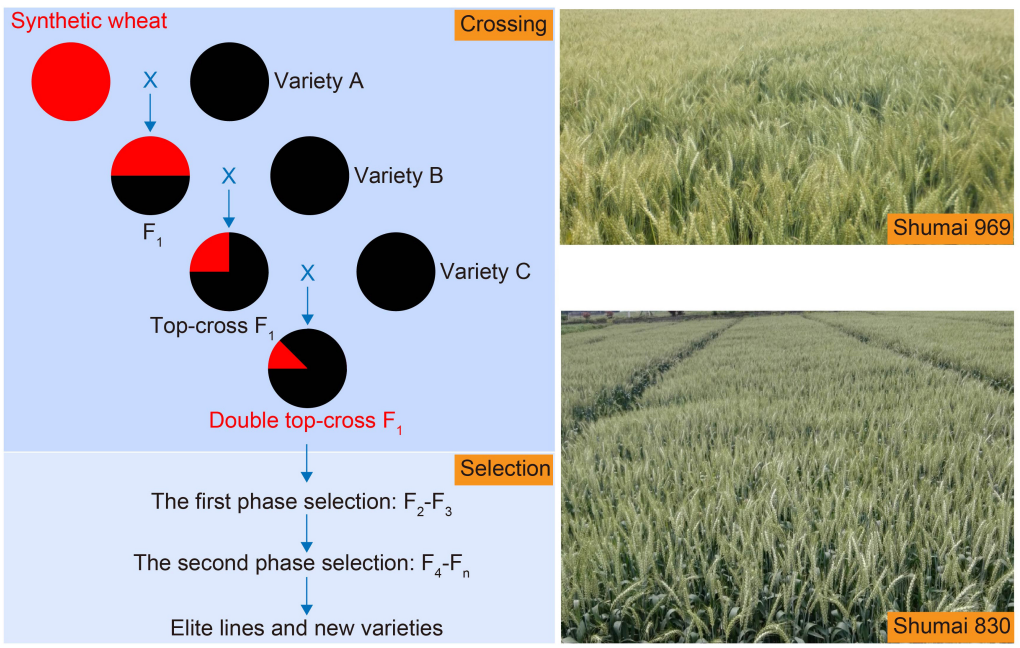

FIGURE 3 | The double top-cross (DTC) and two phase selection (2PS) strategies deployed to manage introgression into elite germplasm from an SHW. The DTC populations retain on average 12.5% of the nuclear genome of the SHW parent. The aim of the first selection phase, carried out in the F_2 and F_3 generations, is to eliminate serious agronomic defects, while that of the second phase (applied from the F_4 generation onward) is to improve yield. Shumai 969 and Shumai 830 were both bred using this strategy.

on backcrossing to an elite cultivar (Li et al., 2018). Rather, the experience has been that top-crossing is a more effective approach, perhaps because of the unlikelihood that any single elite cultivar has the genetic content to neutralize the many defects present in an SHW, whereas including two or more elite parents increases the chance of success.

In the authors' hands, an effective strategy has proven to be one based on a double top cross (DTC), followed by two phases of selection (2PS) (Hao et al., 2019; **Figure 3**). The scheme, summarized as SHW-L1/B//C///D (where B through D indicate three independent elite cultivars), implies that at its end the population's predicted mean proportion of nuclear DNA inherited from SHW-L1 will be 12.5%. During the F_2 and F_3 generations, the population size was reduced by selecting against tough glumes, late maturity, tall stature and yellow rust susceptibility, while selection from the F_4 onward was directed at yield. The program has already led to the release of three cultivars: Shumai 580 and Shumai 969 both emerged from rather small populations (∼100–200 F_2 and ∼100–200 F_3 plants), whereas the selection of Shumai 830 required a more conventional population size, because one of its parental lines was a less well established cultivar. A fourth release (Shumai 114) is imminent. Overall, the indication is that the DTC-2PS strategy can be highly effective for introgressing material from an SHW.

The effectiveness of the DTC-2PS strategy relies on a number of factors. Firstly, it assumes that the SHW donor harbors plentiful genetic variation, allowing introgression to potentially contribute to the improvement of extant traits or even the creation of novel ones. It has been suggested that inheritance of the rapid grain filling trait exhibited by SHW-L1 partially underlies the improved yield potential of the three Shumai cultivars (Hao et al., 2019). Secondly, the DTC scheme combines allelic variation from more than one elite cultivar, providing a wider pool of alleles for generating novel combinations: each of the elite cultivars has its specific constellation of favorable alleles available to interact with the donor's genome. Thirdly, the early generations are used to select for major genes controlling the important domestication traits, leaving behind a much smaller population for the breeder to handle the issue of yield advance.

As demonstrated by Li et al. (2020), the DTC-2PS strategy has also been successfully deployed to improve the genetic background of a line harboring a 6AL.6RS wheat/rye translocation, valued for the presence of *Pm56* (Hao et al., 2018). Two genetic backgrounds provided the starting material: one was CS and the other the Sichuan province landrace Kaixian-luohanmai. The strategy is quite general, so could readily be used to widen the genetic base of any crop species. It is of particular relevance to allopolyploid species where the progenitor parents are known (for example in cotton, 4x *Brassica* spp. and groundnut), since the domesticate has become genetically isolated from the wild progenitor species and its genetic base has been narrowed as a result of domestication and subsequent breeding (Stebbins, 1950; Buckler et al., 2001). The expectation is that introgression from exotic germplasm will feature strongly in future crop improvement programs.

AUTHOR CONTRIBUTIONS

MH and DL contributed to the conception and design of the study and wrote the first draft of the manuscript. All authors contributed to the manuscript revision, read, and approved the submitted version.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (91935303, 31671689, and 31671682), the Chinese Government National Key Research and Development Program (2016YFD0102000), the Sichuan Provincial Agricultural Department Innovative Research Team (wheat-10), and the Sichuan Province Science and Technology Department Crops Breeding Project (2016NYZ0030).

REFERENCES

Anderson, E. (1949). *Introgressive Hybridization*. New York, NY: Wiley & Sons.

Anderson, E., and Hubricht, L. (1938). Hybridization in *Tradescantia*. III. The evidence for introgressive hybridization. *Am. J. Bot.* 25, 396–402. doi: 10.1002/j.1537-2197.1938.tb09237.x

Anugrahwati, D. R. (2009). Isolation of new 1DS/1RS wheat–rye recombinations. *J. Crop Agro.* 2, 60–69. doi: 10.1007/BF00226788

Ardalani, S., Mirzaghaderi, G., and Badakhshan, H. (2016). A robertsonian translocation from *Thinopyrum bessarabicum* into bread wheat confers high iron and zinc contents. *Plant Breed.* 135, 286–290. doi: 10.1111/pbr.12359

Ayala-Navarrete, L., Bariana, H. S., Singh, R. P., Gibson, J. M., Mechanicos, A. A., and Larkin, P. J. (2007). Trigenomic chromosomes by recombination of *Thinopyrum intermedium* and *Th. ponticum* translocations in wheat. *Theor. Appl. Genet.* 116, 63–75. doi: 10.1007/s00122-007-0647-5

Birchler, J. A., and Veitia, R. A. (2012). Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc. Natl Acad Sci. U.S.A.* 109, 14746–14753. doi: 10.1073/pnas.1207726109

Borlaug, N. E. (1983). Contributions of conventional plant breeding to food production. *Science* 219, 689–693. doi: 10.1126/science.219.4585.689

Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169, 1177–1186. doi: 10.1016/j.cell.2017.05.038

Buckler, E., Thornsberry, J., and Kresovich, S. (2001). Molecular diversity, structure and domestication of grasses. *Genet. Res.* 77, 213–218.

Cainong, J. C., Bockus, W. W., Feng, Y., Chen, P., Qi, L., Sehgal, S. K., et al. (2015). Chromosome engineering, mapping, and transferring of resistance to *Fusarium* head blight disease from *Elymus tsukushiensis* into wheat. *Theor. Appl. Genet.* 128, 1019–1027. doi: 10.1007/s00122-015-2485-1

Cao, A., Xing, L., Wang, X., Yang, X., Wang, W., Sun, Y., et al. (2011). Serine/threonine kinase gene *Stpk-V*, a key member of powdery mildew resistance gene *Pm21*, confers powdery mildew resistance in wheat. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7727–7732. doi: 10.1073/pnas.1016981108

Ceoloni, C., Signore, G. D., Ercoli, L., and Donini, P. (1992). Locating the alien chromatin segment in common wheat-*Aegilops longissima* mildew resistant transfers. *Hereditas* 116, 239–245. doi: 10.1111/j.1601-5223.1992.tb00148.x

Coghlan, A. (2006). *Synthetic Wheat Offers Hope to the World*. London: John Murray Learning.

Dai, K., Zhao, R., Shi, M., Xiao, J., Yu, Z., Jia, Q., et al. (2019). Dissection and cytological mapping of chromosome arm 4VS by the development of

wheat-*Haynaldia villosa* structural aberration library. *Theor. Appl. Genet.* 133, 217–226. doi: 10.1007/s00122-019-03452-8

Danilova, T. V., Friebe, B., Gill, B. S., Poland, J., and Jackson, E. (2018). Development of a complete set of wheat–barley group-7 Robertsonian translocation chromosomes conferring an increased content of β-glucan. *Theor. Appl. Genet.* 131, 377–388. doi: 10.1007/s00122-017-3008-z

Danilova, T. V., Poland, J., and Friebe, B. (2019). Production of a complete set of wheat–barley group-7 chromosome recombinants with increased grain β-glucan content. *Theor. Appl. Genet.* 132, 3129–3141. doi: 10.1007/s00122-019-03411-3

Danilova, T. V., Zhang, G., Liu, W., Friebe, B., and Gill, B. S. (2017). Homoeologous recombination-based transfer and molecular cytogenetic mapping of a wheat streak mosaic virus and *Triticum* mosaic virus resistance gene *Wsm3* from *Thinopyrum intermedium* to wheat. *Theor. Appl. Genet.* 130, 549–556. doi: 10.1007/s00122-016-2834-8

Davies, P. A., Pallotta, M. A., and Driscoll, C. J. (1985). Centric fusion between nonhomologous rye chromosomes in wheat. *Can. J. Genet.Cytol.* 27, 627–632. doi: 10.1139/g85-094

de Vicente, M. C., and Tanksley, S. D. (1993). QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134, 585–596.

Dreisigacker, S., Kishii, M., Lage, J., and Warburton, M. (2008). Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement. *Aust. J. Agric. Res.* 59, 413–420.

Dubcovsky, J., and Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316, 1862–1866. doi: 10.1126/science.1143986

Dubcovsky, J., Lukaszewski, A. J., Echaide, M., Antonelli, E. F., and Porter, D. R. (1998). Molecular characterization of two Triticum speltoides interstitial translocations carrying leaf rust and greenbug resistance genes. *Crop Sci.* 38, 1655–1660. doi: 10.2135/cropsci1998.0011183X003800060040x

Dvorak, J., Akhunov, E. D., Akhunov, A. R., Deal, K. R., and Luo, M. C. (2006). Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Mol. Biol. Evol.* 23, 1386–1396. doi: 10.1093/molbev/msl004

Fan, C., Luo, J., Zhang, S., Liu, M., Li, Q., Li, Y., et al. (2019). Genetic mapping of a major QTL promoting homoeologous chromosome pairing in a wheat landrace. *Theor. Appl. Genet.* 132, 2155–2166. doi: 10.1007/s00122-019-03344-x

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. *Trends Genet.* 24, 24–32. doi: 10.1016/j.tig.2007.11.001

Flood, R. G., Lagudah, E. S., and Halloran, G. M. (1992). Expression of vernalization requirement and spikelet number in synthetic hexaploid wheats and their *Triticum tauschii* and tetraploid wheat parents. *Ann. Bot.* 69, 213–217. doi: 10.1093/oxfordjournals.aob.a088332

Friebe, B., Jiang, J., Raupp, W. J., McIntosh, R. A., and Gill, B. S. (1996). Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91, 59–87. doi: 10.1007/bf00035277

Friebe, B., Zhang, P., Linc, G., and Gill, B. S. (2005). Robertsonian translocations in wheat arise by centric misdivision of univalents at anaphase I and rejoining of broken centromeres during interkinesis of meiosis II. *Cytogenet. Genome Res.* 109, 293–297. doi: 10.1159/000082412

Griffiths, S., Sharp, R., Foote, T. N., Bertin, I., Wanous, M., Reader, S., et al. (2006). Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* 439, 749–752. doi: 10.1038/nature04434

Hao, M., Li, A., Shi, T., Luo, J., Zhang, L., Zhang, X., et al. (2017). The abundance of homoeologue transcripts is disrupted by hybridization and is partially restored by genome doubling in synthetic hexaploid wheat. *BMC Genomics* 18:149. doi: 10.1186/s12864-017-3558-0

Hao, M., Liu, M., Luo, J., Fan, C., Yi, Y., Zhang, L., et al. (2018). Introgression of powdery mildew resistance gene *Pm56* on rye chromosome arm 6RS into wheat. *Front. Plant Sci.* 9:1040. doi: 10.3389/fpls.2018.01040

Hao, M., Luo, J., Zeng, D., Zhang, L., Ning, S., Yuan, Z., et al. (2014). *QTug. sau-3B* is a major quantitative trait locus for wheat hexaploidization. *G3: Genes Genom. Genet* 4, 1943–1953. doi: 10.1534/g3.114.013078

Hao, M., Zhang, L., Zhao, L., Dai, S., Li, A., Yang, W., et al. (2019). A breeding strategy targeting the secondary gene pool of bread wheat: introgression from a synthetic hexaploid wheat. *Theor. Appl. Genet.* 132, 2285–2294. doi: 10.1007/s00122-019-03354-9

Harlan, J. R., and de Wet, J. M. J. (1971). Towards a rational classification of cultivated plants. *Taxon* 20, 509–517. doi: 10.2307/1218252

Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J. M., Skovmand, B., Taba, S., et al. (1999). Plant genetic resources: what can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U.S.A.* 96, 5937–5943. doi: 10.1073/pnas.96.11.5937

Huang, X. Q., Coster, H., Ganal, M. W., and Roder, M. S. (2003). Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 106, 1379–1389. doi: 10.1007/s00122-002-1179-7

Huang, X. Q., Kempf, H., Ganal, M. W., and Roder, M. S. (2004). Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and a synthetic wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 933–943. doi: 10.1007/s00122-004-1708-7

Jiang, J., Friebe, B., and Gill, B. S. (1993). Recent advances in alien gene transfer in wheat. *Euphytica* 73, 199–212. doi: 10.1038/hdy.2012.116

King, J., Grewal, S., Yang, C., Hubbart, S., Scholefield, D., Ashling, S., et al. (2017). A step change in the transfer of interspecific variation into wheat from *Amblyopyrum muticum*. *Plant Biot. J.* 15, 217–226. doi: 10.1111/pbi.12606

King, J., Newell, C., Grewal, S., Hubbart-Edwards, S., Yang, C., Scholefield, D., et al. (2019). Development of stable homozygous wheat/*Amblyopyrum muticum* (*Aegilops mutica*) introgression lines and their cytogenetic and molecular characterization. *Front. Plant Sci.* 10:34. doi: 10.3389/fpls.2019.00034

Koebner, R. M. D., and Shepherd, K. W. (1985). Induction of recombination between rye chromosome 1RL and wheat chromosomes. *Theor. Appl. Genet.* 71, 208–215. doi: 10.1007/BF00252057

Koebner, R. M. D., and Shepherd, K. W. (1986). Controlled introgression to wheat of genes from rye chromosome arm 1RS by induction of allosyndesis. *Theor. Appl. Genet.* 73, 197–208. doi: 10.1007/BF00289275

Koebner, R. M. D., and Shepherd, K. W. (1987). Allosyndetic recombination between a chromosome of *Aegilops umbellulata* and wheat chromosomes. *Heredity* 59, 33–45. doi: 10.1038/hdy.1987.94

Kole, C. (2011). *Wild crop Relatives: Genomic and Breeding Resources. Cereals*. Berlin: Springer-Verlag.

Koo, D. H., Liu, W., Friebe, B., and Gill, B. S. (2017). Homoeologous recombination in the presence of *Ph1* gene in wheat. *Chromosoma* 126, 531–540. doi: 10.1007/s00412-016-0622-5

Kumar, N., and Rustgi, S. (2014). "Agronomically relevant traits transferred to major crop plants by alien introgressions," in *Alien Gene Transfer in Crop Plants*, Vol. 1, eds A. Pratap and J. Kumar (New York, NY: Springer), 211–245. doi: 10.1007/978-1-4614-8585-8_9

Li, A., Liu, D., Wu, J., Zhao, X., Hao, M., Geng, S., et al. (2014a). mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* 26, 1878–1900. doi: 10.1105/tpc.114.124388

Li, A., Liu, D., Yang, W., Kishii, M., and Mao, L. (2018). Synthetic hexaploid wheat: yesterday, today, and tomorrow. *Engineering* 4, 552–558. doi: 10.1016/j.eng.2018.07.001

Li, A. L., Geng, S. F., Zhang, L. Q., Liu, D. C., and Mao, L. (2015). Making the bread: insights from newly synthesized allohexaploid wheat. *Mol. Plant* 8, 847–859. doi: 10.1016/j.molp.2015.02.016

Li, H., Dong, Z., Ma, C., Tian, X., Qi, Z., Wu, N., et al. (2019). Physical mapping of stem rust resistance gene *Sr52* from *Dasypyrum villosum* based on *ph1b*-Induced homoeologous recombination. *Int. J. Mol. Sci.* 20:4887. doi: 10.3390/ijms20194887

Li, J., Wan, H. S., and Yang, W. Y. (2014b). Synthetic hexaploid wheat enhances variation and adaptive evolution of bread wheat in breeding processes. *J. Syst. Evol.* 52, 735–742. doi: 10.1111/jse.12110

Li, Q. C., Huang, L., Li, Y. Z., Fan, C. L., Xie, D., Zhao, L. B., et al. (2020). Genetic stability of 6RS/6AL translocation chromosome and its transmission through gametes. *Acta Agronomica Sinica* 46, 513–519. doi: 10.3724/SP.J.1006.2020.91051

Liu, C., Qi, L., Liu, W., Zhao, W., Wilson, J., Friebe, B., et al. (2011a). Development of a set of compensating *Triticum aestivum–Dasypyrum villosum* robertsonian translocation lines. *Genome* 54, 836–844. doi: 10.1139/g11-051

Liu, D., Zhang, H., Zhang, L., Yuan, Z., Hao, M., and Zheng, Y. (2014). "Distant hybridization: a tool for interspecific manipulation of chromosomes," in *Alien*

Gene Transfer in Crop Plants, Vol. 1, eds A. Pratap and J. Kumar (New York, NY: Springer), 25–42. doi: 10.1007/978-1-4614-8585-8_2

Liu, D., Zhang, L., Hao, M., Ning, S., Yuan, Z., Dai, S., et al. (2018). Wheat breeding in the hometown of Chinese Spring. *Crop J.* 6, 82–90. doi: 10.1016/j.cj.2017.08.009

Liu, D. C., Hao, M., Li, A. L., Zhang, L. Q., Zheng, Y. L., and Mao, L. (2017a). "Allopolyploidy and interspecific hybridization for wheat improvement," in *Polyploidy and Hybridization for Crop Improvement*, ed. A. S. Mason (Cambridge: CRC Press), 27–52.

Liu, S., Zhou, R., Dong, Y., Li, P., and Jia, J. (2006). Development, utilization of introgression lines using a synthetic wheat as donor. *Theor. Appl. Genet.* 112, 1360–1373. doi: 10.1007/s00122-006-0238-x

Liu, W., Danilova, T. V., Rouse, M. N., Bowden, R. L., Friebe, B., Gill, B. S., et al. (2013). Development and characterization of a compensating wheat-*Thinopyrum intermedium* robertsonian translocation with *Sr44* resistance to stem rust (Ug99). *Theor. Appl. Genet.* 126, 1167–1177. doi: 10.1007/s00122-013-2044-6

Liu, W., Jin, Y., Rouse, M., Friebe, B., Gill, B., and Pumphrey, M. O. (2011b). Development and characterization of wheat-*Ae. searsii* robertsonian translocations and a recombinant chromosome conferring resistance to stem rust. *Theor. Appl. Genet.* 122, 1537–1545. doi: 10.1007/s00122-011-1553-4

Liu, W., Koo, D. H., Xia, Q., Li, C., Bai, F., Song, Y., et al. (2017b). Homoeologous recombination-based transfer and molecular cytogenetic mapping of powdery mildew-resistant gene *Pm57* from *Aegilops searsii* into wheat. *Theor. Appl. Genet.* 130, 841–848. doi: 10.1007/s00122-017-2855-y

Liu, W., Seifers, D. L., Qi, L. L., Friebe, B., and Gill, B. S. (2011c). A compensating wheat-*Thinopyrum intermedium* robertsonian translocation conferring resistance to wheat streak mosaic virus and *Triticum* mosaic virus. *Crop Sci.* 51, 2382–2390. doi: 10.1007/s00122-016-2834-8

Lukaszewski, A. J. (1993). Reconstruction in wheat of complete chromosomes 1B and 1R from the 1RS.1BL translocation of 'Kavkaz' origin. *Genome* 36, 821–824. doi: 10.1139/g93-109

Lukaszewski, A. J. (1997). Further manipulation by centric misdivision of the 1RS.1BL translocation in wheat. *Euphytica* 94, 257–261.

Lukaszewski, A. J. (2000). Manipulation of the 1RS.1BL translocation in wheat by induced homoeologous recombination. *Crop Sci.* 40, 216–225. doi: 10.2135/cropsci2000.401216x

Lukaszewski, A. J., and Cowger, C. (2017). Re-Engineering of the *Pm21* transfer from *Haynaldia villosa* to bread wheat by induced homoeologous recombination. *Crop Sci.* 57, 2590–2594. doi: 10.2135/cropsci2017.03.0192

Lukaszewski, A. J., Rybka, K., Korzun, V., Malyshev, S. V., Lapinski, B., and Whitkus, R. (2004). Genetic and physical mapping of homoeologous recombination points involving wheat chromosome 2B and rye chromosome 2R. *Genome* 47, 36–45. doi: 10.1139/g03-089

Marais, G. F., Badenhorst, P. E., Eksteen, A., and Pretorius, Z. A. (2010). Reduction of *Aegilops sharonensis* chromatin associated with resistance genes *Lr56* and *Yr38* in wheat. *Euphytica* 171, 15–22. doi: 10.1007/s10681-009-9973-9

Mason, A. S. (2017). *Polyploidy and Hybridization for Crop Improvement*. Boca Raton, FL: CRC Press.

Mello-Sampayo, T. (1971). Genetic regulation of meiotic chromosome pairing by chromosome 3D of *Triticum aestivum*. *Nat. New Biol.* 230, 22–23. doi: 10.1038/newbio230022a0

Mettin, D., Bluthner, W. D., and Schlegel, G. (1973). "Additional evidence of spontaneous 1B/1R wheat-rye substitutions and translocations," in *Proceedings of the 4th International Wheat Genetetics Symposium* (Columbia, MO: Agricultural Experiment Station), 179–184.

Moore, G., Abbo, S., Cheung, W., Foote, T., Gale, M., Koebner, R., et al. (1993). Key features of cereal genome organization as revealed by the use of cytosine methylation-sensitive restriction endonucleases. *Genomics* 15, 472–482. doi: 10.1006/geno.1993.1097

Mujeeb-Kazi, A., Rosas, V., and Roldan, S. (1996). Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. *(Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats *(T. turgidum* L. s. lat. × *T. tauschii*; 2n = 6x = 42, AABBDD) and its potential utilization for wheat improvement. *Genet. Resour. Crop Evolut.* 43, 129–134. doi: 10.1007/bf00126756

Nadolska-Orczyk, A., Rajchel, I. K., Orczyk, W., and Gasparis, S. (2017). Major genes determining yield-related traits in wheat and barley. *Theor. Appl. Genet.* 130, 1081–1098. doi: 10.1007/s00122-017-2880-x

Niu, Z., Klindworth, D. L., Friesen, T. L., Chao, S., Jin, Y., Cai, X., et al. (2011). Targeted introgression of a wheat stem rust resistance gene by DNA marker-assisted chromosome engineering. *Genetics* 187, 1011–1021. doi: 10.1534/genetics.110.123588

Okamoto, M. (1957). Asynapsis effect of chromosome V. *Wheat Inf. Serv.* 5:6.

Qi, L. L., Pumphrey, M. O., Friebe, B., Zhang, P., Qian, C., Bowden, R. L., et al. (2011). A novel Robertsonian translocation event leads to transfer of a stem rust resistance gene (*Sr52*) effective against race Ug99 from *Dasypyrum villosum* into bread wheat. *Theor. Appl. Genet.* 123, 159–167. doi: 10.1007/s00122-011-1574-z

Rabinovich, S. V. (1998). "Important of wheat-rye translocation for breeding modern cultivars of *Triticum aestivum* L," in *Wheat: Prospects for Global Improvement*, eds H. J. Braun, F. Altay, W. E. Kronstad, S. P. S. Beniwal, and A. McNAB (Dordrecht: Kluwer Academic), 401–418. doi: 10.1007/978-94-011-4896-2_55

Rahmatov, M., Rouse, M. N., Nirmala, J., Danilova, T., Friebe, B., Steffenson, B. J., et al. (2016). A new 2DS.2RL Robertsonian translocation transfers stem rust resistance gene *Sr59* into wheat. *Theor. Appl. Genet.* 129, 1383–1392. doi: 10.1007/s00122-016-2710-6

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361:eaar6089. doi: 10.1126/science.aar6089

Rey, M. D., Calderón, M. C., and Prieto, P. (2015). The use of the *ph1b* mutant to induce recombination between the chromosomes of wheat and barley. *Front. Plant Sci.* 6:160. doi: 10.3389/fpls.2015.00160

Ribeiro-Carvalho, C., Guedes-Pinto, H., Igrejas, G., Stephenson, P., Schwarzacher, T., and Heslop-Harrison, J. S. (2004). High levels of genetic diversity throughout the range of the Portuguese wheat landrace 'Barbela'. *Ann. Bot.* 94, 699–705. doi: 10.1093/aob/mch194

Riley, R., and Chapman, V. (1958). Genetic control of the cytologically diploid behaviour of hexaploid wheat. *Nature* 182, 713–715. doi: 10.1038/182713a0

Riley, R., Chapman, V., and Johnson, R. (1968). Introduction of yellow rust resistance of *Aegilops comosa* into wheat by genetically induced homoeologous recombination. *Nature* 217, 383–384. doi: 10.1038/217383a0

Robertson, W. M. R. B. (1916). Chromosome studies. I. Taxonomic relationships shown in the chromosomes of *Tettegidae* and *Acrididiae*: V-shaped chromosomes and their significance in *Acrididiae*, *Locustidae* and *Grillidae*: chromosomes and variation. *J. Morphol.* 27, 179–331.

Sanei, M., Pickering, R., Kumke, K., Nasuda, S., and Houben, A. (2011). Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13373–13374. doi: 10.1073/pnas.1103190108

Sears, E. R. (1952). Misdivision of univalents in common wheat. *Chromosoma* 4, 535–550. doi: 10.1007/bf00325789

Sears, E. R. (1956). The transfer of leaf rust resistance from *Aegilops umbellulata* to wheat. *Brookhaven Symp. Biol.* 9, 1–22.

Sears, E. R. (1972). "Chromosome engineering in wheat," in *Proceedings of the 4th Stadler Genetics Symposium* (Columbia, MO: University of Missouri), 23–38.

Sears, E. R. (1976). Genetic control of chromosome pairing in wheat. *Ann. Rev. Genet.* 10, 31–51. doi: 10.1146/annurev.ge.10.120176.000335

Sears, E. R. (1981). "Transfer of alien genetic material to wheat," in *Wheat science-today and tomorrow*, eds L. T. Evans and W. J. Peacock (Cambridge: Cambridge University Press), 75–89.

Sears, E. R., and Miller, T. E. (1985). The history of Chinese Spring wheat. *Cer. Res. Commun.* 13, 261–263.

Simmonds, N. W. (1976). *Evolution of Crop Plants*. New York, NY: Longman.

Soltis, D. E., Albert, V. A., Leebens-Mack, J., Bell, C. D., Paterson, A. H., Zheng, C., et al. (2009). Polyploidy and angiosperm diversification. *Am. J. Bot.* 96, 336–348. doi: 10.3732/ajb.0800079

Stebbins, G. L. Jr. (1950). *Variation and Evolution in Plants*. New York, NY: Columbia University Press.

Suslow, T., Thomas, B., and Bradford, K. (2002). *Biotechnology Provides New Tools for Plant Breeding*. Oakland, CA: UCANR Publications.

Sutton, T., Whitford, R., Baumann, U., Dong, C. M., Able, J. A., and Langridge, P. (2003). The *Ph2* pairing homoeologous locus of wheat (*Triticum aestivum*): identification of candidate meiotic genes using a comparative genetics approach. *Plant J.* 36, 443–456. doi: 10.1046/j.1365-313x.2003.01891.x

Tanaka, H., Nabeuchi, C., Kurogaki, M., Garg, M., Saito, M., Ishikawa, G., et al. (2017). A novel compensating wheat–*Thinopyrum elongatum* robertsonian translocation line with a positive effect on flour quality. *Breed. Sci.* 67, 509–517. doi: 10.1270/jsbbs.17058

Tanksley, S. D., Grandillo, S., Fulton, T. M., Zamir, D., Eshed, Y., Petiard, V., et al. (1996). Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theor. Appl. Genet.* 92, 213–224. doi: 10.1007/BF00223378

Tanksley, S. D., and Nelson, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203. doi: 10.1007/BF00223376

Tester, M., and Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science* 327, 818–822. doi: 10.1126/science.1183700

Trethowan, R. M., and Mujeeb-Kazi, A. (2008). A Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. *Crop Sci.* 48, 1255–1265. doi: 10.2135/cropsci2007.08.0477

Türkösi, E., Darko, E., Rakszegi, M., Molnár, I., Molnár-Láng, M., and Cseh, A. (2018). Development of a new 7BS.7HL winter wheat-winter barley robertsonian translocation line conferring increased salt tolerance and (1,3; 1,4)-β-D-glucan content. *PLoS One* 13:e0206248. doi: 10.1371/journal.pone.0206248

Udall, J. A., and Wendel, F. F. (2006). Polyploidy and crop improvement. *Crop Sci.* 46, 3–14.

Wang, R. R. C., Bothmer, R. V., Dvorak, J., Fedak, G., Linde-Laursen, I., and Muramatsu, M. (1994). "Genome symbols in the Triticeae (Poaceae)," in *Proceedings of the 2nd International Triticeae Symposium*, eds R. R. C. Wang, K. B. Jensen, and C. Janssi (Logan, UT: Utah State University), 29–34.

Warburton, M. L., Crossa, J., Franco, J., Kazi, M., Trethowan, R., Rajaram, S., et al. (2006). Bringing wild relatives back into the family: recovering genetic diversity in CIMMYT improved wheat germplasm. *Euphytica* 149, 289–301. doi: 10.1007/s10681-005-9077-0

Wilson, A. S. (1876). On wheat and rye hybrids. *Trans. Proc. Bot. Soc. Edinburgh* 12, 286–288.

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L., Benbow, H. R., Wilkinson, P. A., et al. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biot. J.* 14, 1195–1206. doi: 10.1111/pbi.12485

Xia, G., Xiang, F., Zhou, A., Wang, H., and Chen, H. (2003). Asymmetric somatic hybridization between wheat (*Triticum aestivum* L.) and *Agropyron elongatum* (Host) Nevishi. *Theor. Appl. Genet.* 107, 299–305. doi: 10.1007/s00122-003-1247-7

Xiao, J., Li, J., Grandillo, S., Ahn, S. N., Yuan, L., Tanksley, S. D., et al. (1998). Identification of trait-improving quantitative trait loci alleles from a wild rice relative. *Oryza Rufipogon Genet.* 150, 899–909.

Yang, W., Liu, D., Li, j, Zhang, L., Wei, H., Hu, X., et al. (2009). Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China. *J. Genet. Genomics* 36, 539–546. doi: 10.1016/s1673-8527(08)60145-9

Yang, W., Lou, X., Li, J., Pu, M., Mirbahar, A. A., Liu, D., et al. (2017). Cloning and functional analysis of MADS-box genes, *TaAG-A* and *TaAG-B*, from a wheat K-type cytoplasmic male sterile line. *Front. Plant Sci.* 8:1081. doi: 10.3389/fpls.2017.01081

Yen, C., Yang, J. L., and Yen, Y. (2005). Hitoshi Kihara, Askell Löve, and the modern genetic concept of the genera in the tribe Triticeae (Poaceae). *Acta Phytotaxon. Sin.* 43, 82–93.

Yuan, L. P. (1966). A preliminary report on male sterility in rice, *Oryza sativa* L. *Sci. Bull.* 4, 32–34.

Zeller, F. J. (1973). "1B/1R wheat-rye chromosome substitutions and translocations," in *Proceedings of the 4th International Wheat Genetics Symposium* (Columbia, MO: Agricultural Experiment Station), 209–221.

Zhang, C., Huang, L., Zhang, H., Hao, Q., Lyu, B., Wang, M., et al. (2019a). An ancestral NB-LRR with duplicated 3′UTRs confers stripe rust resistance in wheat and barley. *Nat. Commun.* 10:4023. doi: 10.1038/s41467-019-11872-9

Zhang, Q., Li, Q., Wang, X., Wang, H., Lang, S., Wang, Y., et al. (2005). Development and characterization of a *Triticum aestivum-Haynaldia villosa* translocation line T4VS·4DL conferring resistance to wheat spindle streak mosaic virus. *Euphytica* 145, 317–320. doi: 10.1007/s10681-005-1743-8

Zhang, W., Cao, Y., Zhang, M., Zhu, X., Ren, S., Long, Y., et al. (2017). Meiotic homoeologous recombination-based alien gene introgression in the genomics era of wheat. *Crop Sci.* 57, 1189–1198. doi: 10.2135/cropsci2016.09.0819

Zhang, W., Lukaszewski, A. J., Kolmer, J., Soria, M. A., Goyal, S., and Dubcovsky, J. (2005). Molecular characterization of durum and common wheat recombinant lines carrying leaf rust resistance (*Lr19*) and yellow pigment (*Y*) genes from *Lophopyrum ponticum*. *Theor. Appl. Genet.* 111, 573–582. doi: 10.1007/s00122-005-2048-y

Zhang, W., Zhu, X., Zhang, M., Chao, S., Xu, S., and Cai, X. (2018). Meiotic homoeologous recombination-based mapping of wheat chromosome 2B and its homoeologues in *Aegilops speltoides* and *Thinopyrum elongatum*. *Theor. Appl. Genet.* 131, 2381–2395. doi: 10.1007/s00122-018-3160-0

Zhang, W., Zhu, X., Zhang, M., Shi, G., Liu, Z., and Cai, X. (2019b). Chromosome engineering-mediated introgression and molecular mapping of novel *Aegilops speltoides*-derived resistance genes for tan spot and Septoria nodorum blotch diseases in wheat. *Theor. Appl. Genet.* 132, 2605. doi: 10.1007/s00122-019-03374-5

Zhao, R., Wang, H., Xiao, J., Bie, T., Cheng, S., Jia, Q., et al. (2013). Induction of 4VS chromosome recombinants using the CS *ph1b* mutant and mapping of the wheat yellow mosaic virus resistance gene from *Haynaldia villosa*. *Theor. Appl. Genet.* 126, 2921–2930. doi: 10.1007/s00122-013-2181-y

Zhao, W., Qi, L., Gao, X., Zhang, G., Dong, J., Chen, Q., et al. (2010). Development and characterization of two new *Triticum aestivum–Dasypyrum villosum* robertsonian translocation lines T1DS·1V#3L and T1DL·1V#3S and their effect on grain quality. *Euphytica* 175, 343–350. doi: 10.1007/s10681-010-0177-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 March 2020
doi: 10.3389/fpls.2020.00233

Characterization of the Wheat-*Psathyrostachys huashania* Keng 2Ns/2D Substitution Line H139: A Novel Germplasm With Enhanced Resistance to Wheat Take-All

Shengsheng Bai[1], Fengping Yuan[1], Hanbing Zhang[1], Zhenyue Zhang[1], Jixin Zhao[2], Qunhui Yang[1], Jun Wu[1] and Xinhong Chen[1]**

[1] Shaanxi Key Laboratory of Plant Genetic Engineering Breeding, College of Agronomy, Northwest A&F University, Yangling, China, [2] Shaanxi Research Station of Crop Gene Resources and Germplasm Enhancement, Ministry of Agriculture, Yangling, China

OPEN ACCESS

Edited by:
Jaime Prohens,
Universitat Politècnica de València, Spain

Reviewed by:
Eva Darko,
Hungarian Academy of Sciences (MTA), Hungary
Michał Tomasz Kwiatek,
Poznań University of Life Sciences, Poland

***Correspondence:**
Jun Wu
13572016162@163.com
Xinhong Chen
cxh2089@126.com

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 19 November 2019
Accepted: 14 February 2020
Published: 10 March 2020

Citation:
Bai S, Yuan F, Zhang H, Zhang Z, Zhao J, Yang Q, Wu J and Chen X (2020) Characterization of the Wheat-Psathyrostachys huashania Keng 2Ns/2D Substitution Line H139: A Novel Germplasm With Enhanced Resistance to Wheat Take-All. Front. Plant Sci. 11:233. doi: 10.3389/fpls.2020.00233

Take-all is a devastating soil-borne disease that affects wheat production. The continuous generation of disease-resistance germplasm is an important aspect of the management of this pathogen. In this study, we characterized the wheat-*Psathyrostachys huashania* Keng (*P. huashania*)-derived progeny H139 that exhibits significantly improved resistance to wheat take-all disease compared with its susceptible parent 7182. Sequential genomic *in situ* hybridization (GISH) and multicolor fluorescence *in situ* hybridization (mc-FISH) analyses revealed that H139 is a stable wheat-*P. huashania* disomic substitution line lacking wheat chromosome 2D. Expressed sequence tag-sequence tagged site (EST-STS) marker and Wheat Axiom 660K Genotyping Array analysis further revealed that H139 was a novel wheat-*P. huashania* 2Ns/2D substitution line. In addition, the H139 line was shown to be cytologically stable with a dwarf phenotype and increased spikelet number. These results indicate that H139, with its enhanced wheat take-all disease resistance and desirable agronomic traits, provides valuable genetic resources for wheat chromosome engineering breeding.

Keywords: take-all, *Psathyrostachys huashania*, *Triticum aestivum*, novel germplasm, H139

INTRODUCTION

Wheat (*Triticum aestivum L.*) is one of the most important global food crops, contributing nearly a fifth of the total calories consumed by humans (International Wheat Genome Sequencing Consortium (IWGSC), 2018). Wheat production is of great significance to global food security and quality of life. Wheat take-all is a destructive root disease caused by the recently reclassified soil-borne ascomycete fungus *Gauemannomyces tritici* (*G. tritici*), previously referred to as *Gauemannomyces graminis var. tritici* (Walker, 1981; Hernández-Restrepo et al., 2016). This pathogen is widely distributed in wheat-producing areas worldwide (Cook, 2003; Youn-Sig and Weller, 2013). Generally, wheat roots with take-all disease exhibit black spots followed by gradual necrosis of vasculature tissues, resulting in hindered plant growth and development (Youn-Sig and Weller, 2013). In severe infections, take-all disease causes seedling death or, upon extensive invasion of the crown, the death

of mature plants, resulting in reduced yield or crop failure (Penrose, 1985). Take-all disease has a considerable impact on the security of wheat production.

The use of disease-resistant varieties remains an economically viable and environmentally friendly measure to control wheat diseases (McMillan et al., 2014; Wu et al., 2017). Penrose (1985) described a laboratory method for the uniform inoculation of wheat seedlings with take-all pathogen in order to evaluate disease resistance in wheat cultivars. Jia et al. (1982) and Scott et al. (1989) confirmed the variation in take-all resistance in wheat cultivars, and described how cultivars with stronger tillering ability and well-developed roots usually exhibit greater disease resistance. Wang et al. (2012) adopted an optimized disease resistance screening method, based on fungal-colonized agar plugs inoculated 2 cm below wheat seeds in combination with a 0–6 disease grading standard, to identify two varieties with moderate resistance against *G. tritici* (G1037) among 69 Chinese wheat cultivars. McMillan et al. (2014) identified seven *T. monococcum* accessions with moderate resistance to take-all disease in naturally inoculated fields; however, the susceptibility of *T. monococcum* to take-all disease in a seedling pot test does not accurately reflect its disease resistance under field conditions.

The utilization of genetic resources from disease-resistant wild relatives had been proven valuable for wheat improvement (Fu et al., 2003; An et al., 2013; Li et al., 2018). An important aspect of take-all disease resistance research is to identify and develop new resistant germplasm. Oat (*Avena sativa* L.) displays near-complete disease resistance to wheat take-all (*G. tritici*), owing to its root system that can produce the antifungal compound avenacin (Papadopoulou et al., 1999; Qi et al., 2004). However, oats are susceptible to *G. graminis var. avenae*, the causal pathogen of oats take-all (Crombie et al., 1986). Rye (*Secale cereale* L.) is considered more resistant to take-all than common wheat, with a disease resistance ranging from moderate to high (Cook, 2003). *Psathyrostachys huashania* Keng (*P. huashania*) (2n = 14, NsNs) is regarded as a wild relative with high resistance to wheat take-all (Wang and Shang, 2000). *P. huashania* is primarily distributed in the Huashan section of the Qinling Mountains, China (Kuo, 1987; Baden, 1991). This species is favored by researchers because of its cold and drought tolerance, early maturity, perennial traits, dwarf stature, and resistance to wheat take-all (*G. tritici*), stripe rust (*Puccinia striiformis* f. sp. *tritici*), and powdery mildew (*Blumeria graminis* f. sp. *tritici*) (Chen et al., 1996; Fu et al., 2003; Du et al., 2013a).

A previous genetically distant hybridization between wheat line 7182 and *P. huashania* was performed by our research team in 1989, from which the haploid line H881 was successfully isolated (Chen et al., 1991). Subsequently, a series of wheat-*P. huashania* addition lines were developed and characterized (Chen et al., 1996; Wu et al., 2007). Initially, Wang and Shang (2000) used a series of wheat-*P. huashania*-derived offspring to screen for resistance to take-all disease at the seedling stage. This revealed that *P. huashania* displayed high disease resistance, whereas the wheat parental line 7182 was susceptible. In the derived offspring, disease resistance of the addition line H1 was similar to that of *P. huashania*, six addition lines and three substitution lines displayed a moderate level of disease resistance, and the remaining lines were susceptible. More recently, synthetic approaches including cytological observations, GISH, EST-STS marker analyses, morphological analysis, and agronomic trait evaluation have led to the successful identification of a series of wheat-*P. huashania* disomic addition lines, including 1Ns (Du et al., 2014d), 2Ns (Du et al., 2014c), 3Ns (Du et al., 2014b), 4Ns (Du et al., 2014a), 5Ns (Du et al., 2013a), 6Ns (Du et al., 2013b), and 7Ns (Du et al., 2013c), and one disomic substitution line, namely 16-6 (Du et al., 2015). However, the use of wheat-*P. huashanica*-derived offspring for enhanced resistance against wheat take-all disease remains underexplored. It is worth mentioning that: both 16-6 and H139 in this study were derived from the distant hybridization of common wheat 7182 × *P. huashania*. The F_1 generation H881 (2n = 28, ABDNs) was self sterile. To restore the fertility of the hybrid, 7182 was used as the male parent to back-cross with H811 for two times, and then self-crossed for multiple generations. 16-6 and H139 were selected from BC_2F_6 and BC_2F_8 generations, respectively (**Supplementary Figure S1**).

With the development of high-throughput sequencing technology, the number of available SNP in common wheat research has been greatly enriched. At present, the high-density Wheat SNP Arrays provides a powerful tool for the genotype of wheat and its wild relatives as well as their introgression lines. Zhou et al. (2018) used the wheat 660K single nucleotide polymorphism (SNP) array to characterize the complete set of wheat-*A. cristatum* addition/substitution lines according to their homoeologous relationships. This wheat 660K SNP array is highly efficient with a wide range of potential applications. The common wheat 7182 is a self-pollinating plant, whereas *P. huashania* is a cross-pollinating plant. Theoretically, in the homozygous wheat genome background, the wheat chromosomes with homologous *P. huashania* chromosomes should have a higher percentage of heterozygous genotypes than wheat chromosomes for addition/substitution lines, and the substituted wheat chromosomes should have more missed genotype markers and ratio than wheat chromosomes.

In this study, we characterized a wheat-*P. huashanica* derived line H139 that displays wheat take-all disease resistance at both the seedlings and mature stages. Molecular cytological and Wheat 660K Genotyping Array were performed to determine the chromosome composition of H139. Moreover, the morphological traits of H139 were evaluated. Our results indicate that the H139 germplasm resource can serve as bridge material for wheat improvement.

MATERIALS AND METHODS

Experimental Materials

In the wheat take-all disease susceptibility seedling pot bioassay, oat (*Avena sativa* cv. Bayou8, resistant to take-all) and hexaploid wheat (*Triticum aestivum* L. cv. Yangmai158, susceptible to take-all) were used as controls to evaluate the disease resistance of line H139, *P. huashania*, and wheat parent 7182.

In the wheat take-all disease susceptibility mature plant test, wild oat (*Avena fatua* L., resistant), rye (cv. Dongmu70, resistant),

TABLE 1 | List of experimental materials used in various tests.

Methods	H139	7182	*P. HS*	YM158	16-6	BY8	Wild Oat	DM70	CS
Take-all test on seedling	+	+	+	+	−	+	−	−	−
Take-all test on mature plant	+	+	+	+	−	−	+	+	−
Sequential GISH and mc-FISH	+	−	−	−	−	−	−	−	−
EST-STS makers analysis	+	+	+	−	−	−	−	−	+
Wheat 660K SNP array	+	+	−	−	+	−	−	−	−
Morphological characterization	+	+	−	−	+	−	−	−	−

Both H139 and 16-6 were wheat-Psathyrostachys huashania 2Ns/2D substitution lines; 7182 was wheat parent of H139 and 16-6; P. HS: Psathyrostachys huashania Keng or P. huashania; YM158: Yangmai158, a Chinese wheat cultivar that was as susceptible control to take-all; BY8: Bayou8 (Avena sativa L.) was as resistant control to take-all in this study; Wild Oat: Avena fatua L., was as resistant control to take-all from field of Yangling, China; DM70: Dongmu70 (Secale cereale L.), is a forage rye variety; CS; Chinese Spring. +: the experimental material has been applied to the test, −: no applied.

and wheat Yangmai158 (susceptible) were used as controls. See **Table 1** for details of experimental materials used in this study.

G1280, an isolate of *G. tritici* with strong pathogenicity (Feng et al., 2013), was used as the infectious material. These plant materials are preserved at the Shaanxi Key Laboratory of Genetic Engineering for Plant Breeding, College of Agronomy, Northwest A&F University, Shaanxi, China.

Take-All Disease Susceptibility Test

According to the methods of Wang et al. (2012) with some modification. Specific practices were as follows: equal size seeds were disinfected for 5 min with 70% alcohol, washed three times with sterilized water, then sprouted under 24°C. Fourteen-millimeter diameter flower pots were prepared by filling half full with autoclaved vermiculite. Ten 5-mm diameter agar disks containing take-all infectious material were placed evenly on the vermiculite surface (hyphae down), and then covered with a further 2-cm thickness of vermiculite. Ten sprouted seeds were then sown evenly and covered with thin layer of vermiculite. Three replicates were set up per treatment. All plots were placed in a manual climatic cabinet in a randomized design (16-h day length, temperature 15°C, twice weekly watering). After 22 days of growth, the roots were rinsed gently with running water and disease assessment was performed.

For take-all disease assessment in seedlings, we used the disease level standard 0–6 scale as reported by Wang et al. (2012). Grade 0: no symptoms; Grade 1: 1–10% roots infected; Grade 2: 10–50% roots infected; Grade 3: 50–100% roots infected; Grade 4: seedling stem base appears brown or with black surface covering; Grade 5: yellowing of seedling and dwarfing; Grade 6: seedling death. Take-all disease index was calculated for each pot as follows:

$$\text{Index} = \frac{0 * X_0 + 1 * X_1 + 2 * X_2 + 3 * X_3 + 4 * X_4 + 5 * X_5 + 6 * X_6}{(X_0 + X_1 + X_2 + X_3 + X_4 + X_5 + X_6) * 6} * 100\%$$

where X_1, X_2, X_3, X_4, X_5, X_6 denotes the number of seedlings at Grades 1–6 take-all disease level.

For the take-all disease susceptibility test on mature plants, 30-cm diameter pots were prepared by filling half full with field soil and smoothing the soil surface. Then, 5-cm diameter disk containing take-all infectious material was placed under wheat seedlings, and the seedlings were transplanted to the prepared pots. Three replicates of five seedlings per pot were created, which were watered twice a week until the grain filling stage. Plants were taken out and the soil was gently washed away from roots. Each root was separated and arranged on a white plate. The total number of roots and the number of infected roots were counted for each plant. The proportion of roots infected in each whole-plant root system was estimated and graded into six categories according to McMillan et al. (2014): no symptoms, slight 1 (1–10% roots infected), slight 2 (11–25%), moderate 1 (26–50%), moderate 2 (51–75%), and sever (more than 75%).

Sequential GISH and mc-FISH

GISH was conducted to detect *P. huashanica* chromatin of H139. Seeds of H139 were germinated at 24°C on moistened filter paper. Actively growing roots were cut, placed into ice-cold water for 24 h, and then stored in Carnoy's solution (95% ethanol-acetic acid 3:1, v/v). Chromosome preparation of root tip cells of H139 was performed as Han et al. (2006). The genomic DNA of *P. huashanica* was extracted from fresh leaves using the improved CTAB method (Cota-Sánchez et al., 2006). *P. huashanica* DNA probe was labeled with digoxigenin-11-dUTP using the nick translation method (An et al., 2013). The amount of probe DNA was 2.5 ng, and the ratio of probe/blocking DNA was 1:50 for one slide. Chromosomes were counterstained with 4, 6-diamidino-2-phenylindole (DAPI). Mitotic chromosomes number counts and fluorescent signals were viewed and photographed using a microscope (Olympus BX60) with a Photometrics SenSys CCD camera.

To further distinguish the substituted wheat chromosomes in H139, the slide was eluted in 2 × X_1 SSC for 2 min, and transferred into 70% alcohol for 20 min, transferred into 100% alcohol for 1 h, and dried in air, and then was prepared to complete mc-FISH. The mc-FISH was performed using pAs1 (Rayburn and Gill, 1986) and pSc119.2 (Mcintyre et al., 1990) as probes. The two probes were mixed according to the proportion of 1:1 with chromosome hybridization. Detailed information about the hybridization patterns of pAs1 and pSc119.2 karyotype

data were reported by previous studies (Mukai et al., 1993; Schneider et al., 2003; Tang et al., 2014). The counterstain, detection, and visualization of chromosomes were performed as described above.

EST-STS Marker Analyses

EST-STS marker analyses was conducted to determine the homologous group of *P. huashania* chromosomes contained in H139 to the 7182 counterpart. Genomic DNA from H139, 7182 and *P. huashanica* were isolated as described previously (Cota-Sánchez et al., 2006). To characterize the genomic composition of H139, 75 pairs of EST-STS markers with polymorphisms between *P. huashania* and 7182 were used (**Supplementary Table S1**). PCR reaction conditions and procedures were performed as the previously published method (Du et al., 2013a). And the PCR products were separated in 8% non-denaturing polyacrylamide gels and then silver-stained and photographed.

Wheat 660K Genotyping Array Analyses

The Wheat Axiom 660K Genotyping Array has proven to be a powerful tool for identifying homologous group relationships of alien chromosomes in addition or substitution lines (Zhou et al., 2018). Wheat axiom 660K genotyping array assays were performed by Beijing CapitalBio Technology, Co., Ltd.[1] H139, 16-6 and 7182 were genotyped using the Affymetrix GeneTitan System. SNP genotype calling and clustering was performed with the Polyploidy version in Affymetrix Genotyping ConsoleTM (GTC) software (Wu et al., 2018). The statistics of heterozygous and missed SNP, as well as the screening and distribution of differences SNP between H139 and 16-6, were completed by Office Excel 2010. Compared with the parent 7182, the Ns chromosome that was homologous to wheat chromosome with the high ratio of heterozygous genotypes was considered additional alien chromosomes. And chromosome with the high proportion of missing genotype data was considered replaced wheat chromosomes in the substitution lines.

Morphological Trait Evaluation

The morphological traits of H139 and 7182 and 16-6 were evaluated from 2016 to 2018. A randomized complete block design was arranged with three replications. At maturity, 30 plants of each line were harvested to evaluate their morphological traits, including plant height, spike length, spikelets per spike, kernels per spikelet, kernels per spike, and 1000-grain weight. The Duncan's multiple range procedure was used to compare significant differences between H139 and wheat parent 7182 for morphological traits. Comparisons were made using the analysis of variance procedure in SPSS Statistics v22.

[1]http://www.capitalbiotech.com

RESULTS

Evaluation of H139 Take-All Disease Resistance

The wheat take-all disease susceptibility test on seedlings was completed in an artificial climate cabinet with controlled temperature, humidity, and light conditions. The disease-resistant control oat (Bayou 8) displayed complete immunity to wheat take-all disease, whereas the susceptible control Yangmai 158 was sufficiently infected with an average disease index (ADI) of 59.4%. The wheat parent 7182 displayed severe disease (ADI 63.0%); however, *P. huashania* showed high take-all disease resistance, with an ADI of only 5.6%, and the wheat-*P. huashania*-derived progeny line H139 was moderately resistant (**Figure 1A**), with a disease index of 34.9%. The Duncan's multiple range test showed that the disease resistance of H139 was significantly higher ($P < 0.01$) than that of 7182.

The wheat take-all disease susceptibility test on mature plants was completed in the Northwest Institute of Botany field. The wild oat disease-resistant control displayed the least infection, with an average infected root rate of 9.40% and a disease index of 22.2%, whereas rye (Dongmu 70) displayed moderate infection, with an average infected root rate of 16.5% and a disease index of 42.2%. The susceptible cultivar Yangmai 158 displayed considerable infection, with an average infected root rate of 62.7% and a disease index of 77.3%. The disease index of *P. huashania* was 33.3%, which was between the disease indices of wild oat and rye, and its average infected root rate of 11.3% was close to that of rye (**Figure 1B**). The wheat parent 7182 exhibited less disease susceptibility than Yangmai158, with an average infected root rate of 44.3% and a disease index of 63.3%. Finally, the disease index of H139 was 43.6% and the average infected root rate was 22.4%, thus indicating that the H139 line has enhanced wheat take-all disease resistance compared with the susceptible parent 7182.

Sequential GISH and mc-FISH Analyses

Using the genomic DNA of *P. huashania* as probe, GISH analysis of root tip cells of H139 was performed to identify the presence of *P. huashania* chromatin. The result demonstrated that there were 42 chromosomes in H139, of which two showed

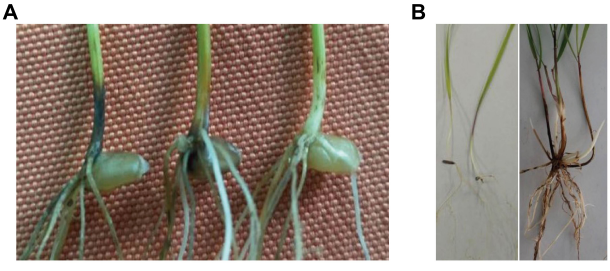

FIGURE 1 | Evaluation of wheat take-all disease susceptibility test. **(A)** Seedlings test of Yangmai 158 (left), 7182 (middle), and H139 (right). **(B)** *P. huashania* performance in the wheat take-all disease susceptibility test performed on seedlings (left) and mature plants (right).

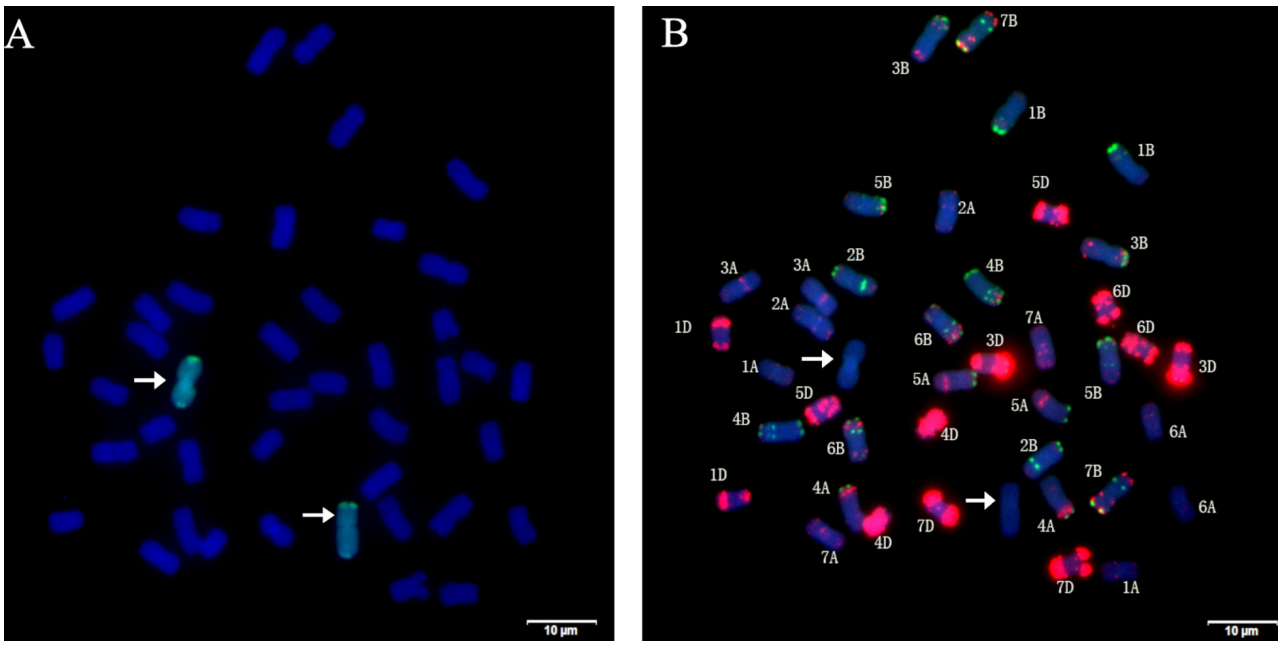

FIGURE 2 | Sequential GISH and mc-FISH analysis of root tip cells of H139. **(A)** GISH analysis of H139 shows two green hybridization signals by using *P. huashania* genomic DNA as probe, and the wheat chromosomes were counterstained with DAPI (blue). **(B)** mc-FISH on the same metaphase after GISH analysis of H139 by using pAs1 (red) and pSc119.2 (green) as probe simultaneously. Arrows indicate the pair of *P. huashania* chromosomes.

green fluorescence signals (**Figure 2A**), thus indicating that H139 contains a pair of *P. huashania* chromosomes. For three consecutive years, H139 was self-pollinated and its offspring were analyzed by GISH, through which no chromosome segregation was observed. Therefore, H139 is considered a genetically stable disomic substitution line.

In order to determine the substituted wheat chromosomes, mc-FISH with probes pAs1 and pSc119.2 was performed in H139. Compared with the karyotype data of pAs1 and pSc119.2, 20 pairs of wheat chromosomes were successfully identified, but the 2D chromosome was not found (**Figure 2B**). Therefore, we conclude that the wheat 2D chromosome was replaced by a pair of *P. huashania* chromosomes in the H139 line.

EST-STS Marker Analysis

Seventy-five pairs of EST-STS primers, each polymorphic between *P. huashania* and 7182, were used to identify homologous group of *P. huashania* chromosomes in H139. Among these, 10 pairs of EST-STS primers amplified the same specific band in H139 and *P. huashania* that was absent in 7182 (**Figure 3**). These markers were mainly distributed in the homologous group 2 of common wheat. Therefore, we deduced that the *P. huashania* chromosomes introduced into H139 belonged to homologous group 2. Based on the above results, H139 is considered a wheat-*P. huashania* 2Ns/2D substitution line.

Wheat 660K SNP Array Analyses

Compared with EST-STS markers, Wheat Axiom 660K Genotyping Array has a greater number of markers that are distributed more evenly across wheat chromosomes. To further clarify the chromosome composition of H139, the Wheat Axiom 660K Genotyping Array was used for genotyping. The ratio of heterozygous genotypes on each of the wheat chromosomes was determined in H139 and 7182. The results showed that the ratio of heterozygous genotypes of H139 on wheat chromosome 2D was significantly higher than 2D of 7182, and was also the highest ratio (31.73) compared with other wheat chromosomes in H139 and 7182 (**Figure 4A**). The ratio of heterozygous genotypes of other wheat chromosomes ranged from 3.6 to 10.3%, which was much lower than that of 2D. Therefore, we can confirm that the *P. huashania* chromosome in H139 was 2Ns. This result further validates the results of EST-STS marker analysis.

Moreover, the ratio of missed markers on each chromosome was also calculated in H139 and 7182. The data showed that the missed marker rate of H139 on wheat chromosome 2D was significantly higher than 2D of 7182, and the missed marker rate of other wheat chromosomes ranged from 2.6 to 6.29%, which was much lower than that of 2D (**Figure 4B**). Therefore, it can be concluded that wheat chromosome 2D is replaced by *P. huashania* chromosome 2Ns in H139, which is consistent with the results of GISH and mc-FISH, and further support the notion that H139 is a wheat-*P. huashania* 2Ns/2D substitution line.

In addition, 16-6, a wheat-*P. huashania* 2Ns/2D substitution line previously reported by Du et al. (2015), had also been re-validated as wheat-*P. huashania* 2Ns/2D substitution line based on the wheat 660K Genotyping Array data (**Figure 4**). However, 16-6 and H139 were further analyzed and showed that there were 53 4564 (91.5%) same SNP, and 49 909 (8.5%) differential SNP between H139 and 16-6. These

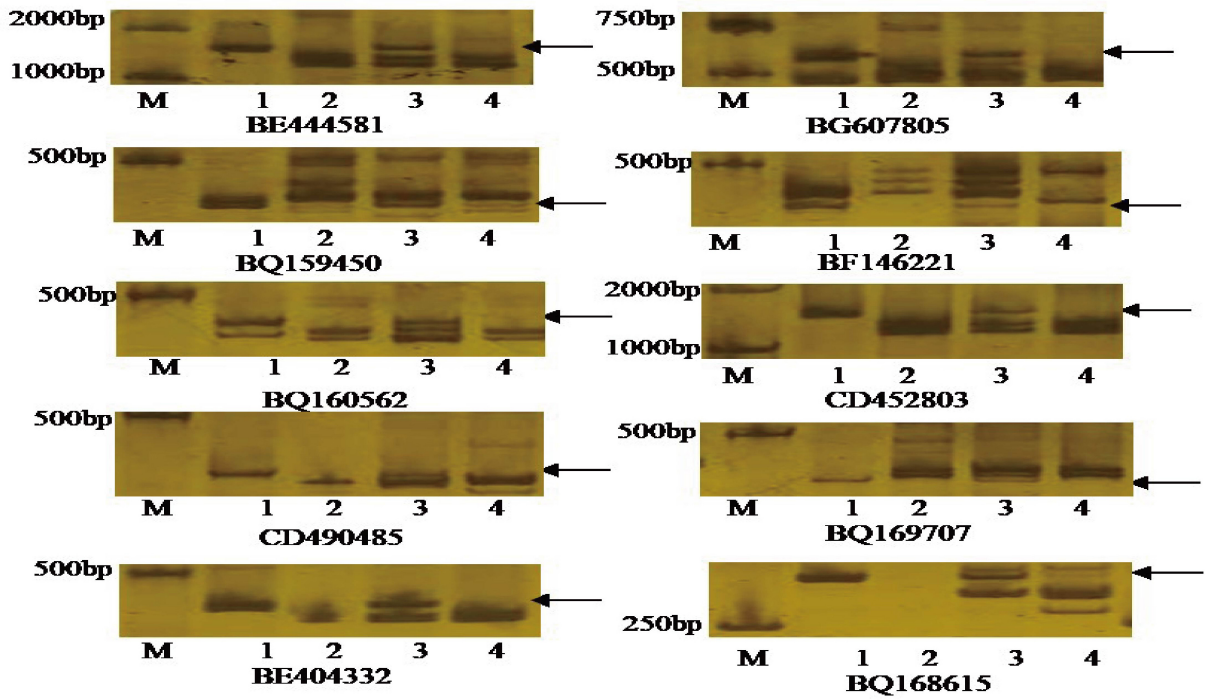

FIGURE 3 | Amplification patterns of the specific EST-STS markers with distributed in wheat homologous group 2. These EST-STS makers amplified the same specific band in H139 and *P. huashania* that was absent in 7182, respectively. M: DNA marker (DL2000); (1) *P. huashania*; (2) 7182; (3) H139; (4) Chinese Spring.

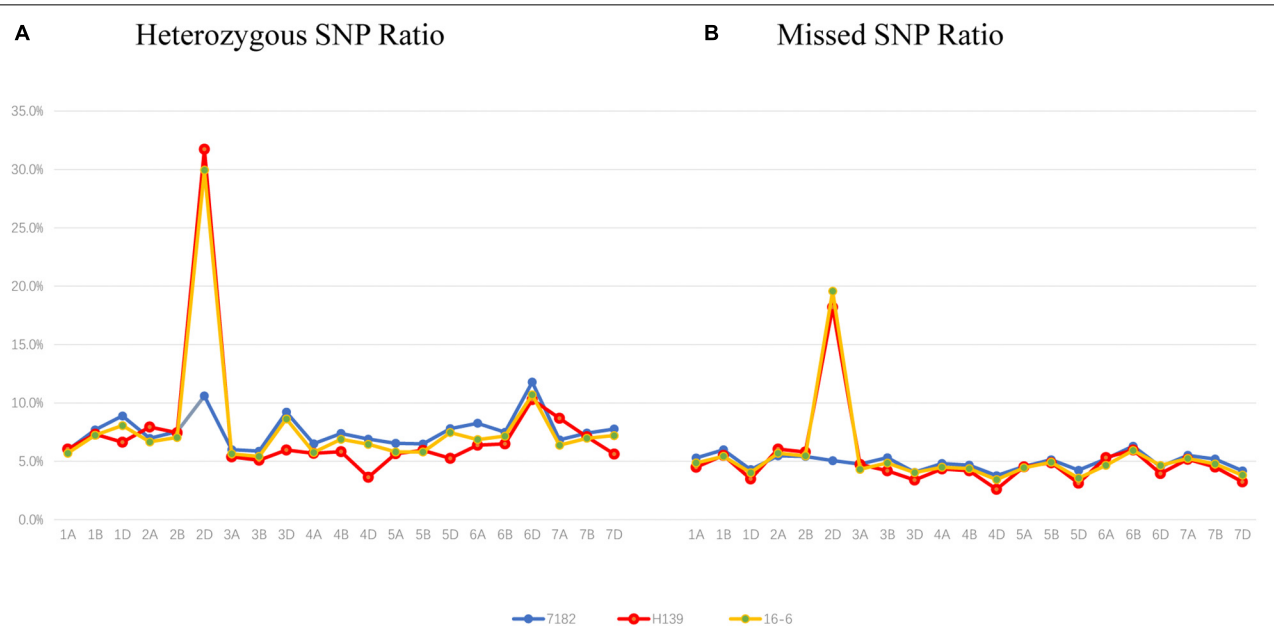

FIGURE 4 | Wheat 660K SNP Array data analysis of 7182 and H139 and wheat-*P. huashania* 2Ns/2D substitution line 16-6. **(A)** The ratio of heterozygous genotypes on each of the wheat chromosomes showed that the ratio of wheat chromosome 2D in H139 and 16-6 was significantly higher than 2D of 7182, and was also the highest ratio compared with other wheat chromosomes. This indicates that *P. huashania* chromosome was 2Ns in H139 and 16-6. **(B)** The missed markers rate of each wheat chromosomes showed that the ratio of wheat chromosome 2D in H139 and 16-6was significantly higher than 2D of 7182, and the missed marker rate of other wheat chromosomes were much lower than that of 2D. This indicates that wheat chromosome 2D was replaced by *P. huashania* chromosome 2Ns in H139 and 16-6.

differences were widely distributed on 21 pairs of chromosomes (**Supplementary Table S2**).

Among these, 2785 SNP were distributed on the 2Ns(2D) chromosome, and the distribution ratio of these SNP in 2Ns(2D) was relatively high compared with other chromosomes (**Supplementary Table S2**). Both H139 and 16-6 are wheat-*P. huashania* 2Ns/2D substitution lines, and their second homologous groups include 2A, 2B, and 2Ns chromosomes without 2D. Therefore, the 2785 SNP likely originated from the 2Ns chromosome. According to the physical position of these SNP on 2D chromosome, the distribution proportion of these differential SNP within 1 Mb was calculated on, the results showed that there was a peak between 200 and 300 Mb (**Figure 5**). These results demonstrate that there are partial differences in genome between them. And the 2Ns chromosome in H139 is partial differences than 16-6.

Morphological Trait Evaluation

Morphological trait surveys of three consecutive years showed that H139 performed almost identically to wheat parent 7182 regarding seedling growth habit, heading time, spike type, and spike color. Compared with its wheat parent 7182, however, H139 was shorter in plant stature and greater spikelets per spike (**Figure 6**). Although H139 also exhibited greater spike length, kenels per spike, and 1000-kernel weight than 7182, these differences were not statistically significant (**Table 2**). Compared with 16-6, there were significant differences between H139 and 16-6 regarding plant height, spike length, spikelets per spike, kernels per spikelet, kernels per spike, and 1000-kernel weight (**Table 2**).

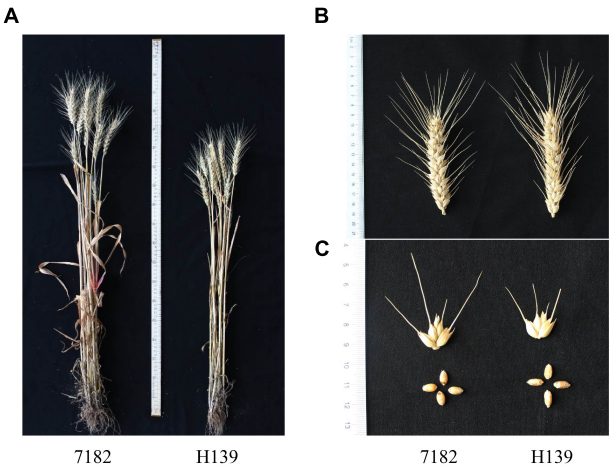

FIGURE 6 | Plant **(A)**, spike **(B)**, spikelet and grain **(C)** morphology of H139 and the parent 7182.

DISCUSSION

Take-all is a typical soil-borne disease more pervasive with continuous cropping that seriously threatens wheat yield and grain quality. Limited germplasm resistant to take-all disease has been identified in common wheat (McMillan et al., 2014). Chromosome engineering is important for crop breeders to develop new disease-resistant germplasm and broaden the genetic base of wheat (Jiang et al., 1993; Zhuang et al., 2015; An et al., 2019). Hollins et al. (1986) found that the

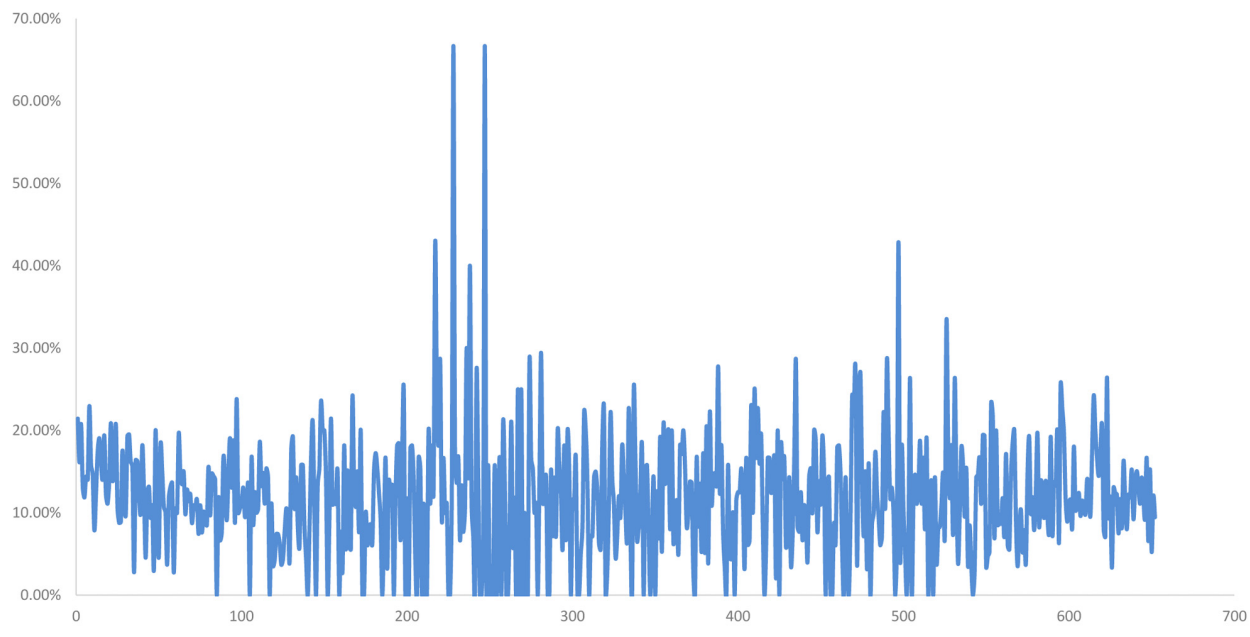

FIGURE 5 | The distribution ratio of differential SNP per Mb on 2Ns(2D) chromosome between in H139 and 16-6. Abscissa is the physical position of SNP on chromosome, and the unit is Mb. The ordinate is the ratio of the number of differential SNP to the total SNP in 1Mb.

TABLE 2 | Data analysis of agronomic traits of 7182 and H139 and 16-6.

	7182		H139		16-6	
Characters	**Mean**	**Range**	**Mean**	**Range**	**Mean**	**Range**
Plant height(cm)	75.3	(65–82)	60**	(51–71)	75.9	(71–80)
Spike length(cm)	8.8	(6.5–10.5)	8.9	(6–11)	12.2**	(9.5–14)
Spikelets per spike	19.9	(16–24)	22.1*	(17–26)	25.9**	(24–27)
Kernels per spikelet	3	(2–4)	3	(2–4)	4	(2–5)
Kenels per spike	55.8	(32–85)	56	(42–84)	75.7**	(66–85)
Thousand-grain weight(g)	46.2	(45.6–47.2)	47.1	(46.7–47.5)	50.5**	(48–53.3)

*There were significant differences in means at the **$P < 0.01$ level and *$P < 0.05$ level, based on Duncan's multiple range tests.*

triticale was moderate resistance to take-all compared with wheat and rye. And the application of triticale is considered to be an immediately available method to improve wheat take-all resistance. But the introduction of individual pairs of rye chromosomes into wheat genetic background did not significantly slight take-all susceptible. Wang et al. (2003) reported 9 lines with relatively high resistance to take-all in seedling from 19 lines [wheat-*P. huashania* intermedium derivative lines and wheat-*Haynaldia vilosa* L. Schur (2n = 14, VV) intermedium derivatives lines], and then in further field test they found H922-9-12, V2 and V9129-1 expressed high resistance in wheat anthesis period, and H922-9-12 and V2 were high resistance in harvest period. And V2 was a wheat-*Haynaldia vilosa* substitution line, but there was no chromosome location information of take-all disease resistance. *P. huashania* was shown to be a precious germplasm resource with high resistance to take-all disease (Wang and Shang, 2000; Fu et al., 2003). Since the 1980s, our research team has been committed to introducing individual desirable characteristics of *P. huashania* into common wheat by chromosome engineering (Chen et al., 1991). Wang and Shang (2000) reported that *P. huashania* was high resistant to wheat take-all disease and represents a new disease-resistant germplasm resource. Their study reported that seven addition lines, three substitution lines, and two translocation lines with moderate resistance to *G. tritici* through a seedlings susceptibility test. However, there was also no chromosome location information about take-all disease resistance. And the take-all disease resistance of *P. huashanica* and its derivatives in mature plants remains poorly understood.

In this study, we demonstrated through take-all disease resistance seedling assays the disease immunity of control oat, the high-level disease resistance of parent *P. huashanica*, the disease susceptibility of both control Yangmai158 and wheat parent 7182, and the enhanced disease-resistance of line H139. Furthermore, the disease resistance of *P. huashanica* and H139 was analyzed in mature plants, for which we used equal pathogen inoculation. Interestingly, wild oat as a disease-resistant control was infected by *G. tritici*, but the infection was slight. Our results slightly differ from those of McMillan et al. (2014), whereby oat plants were immune to wheat take-all disease in field tests. This discrepancy may be due to differences in experimental oat varieties or inoculation methods. The method used by McMillan et al. (2014) involved natural field pathogen inoculation; however, we used a manual inoculation method that may have resulted in more uniform, high-level pathogen exposure than that occurs in natural field conditions. This greater pathogen inoculation may have resulted in a relatively higher take-all disease index, which is supported by the higher disease indices observed for the susceptible control Yangmai 158 and disease-resistant control rye. Although take-all disease was more prevalent in field trials, H139 maintained a moderate disease index and a lower average disease root rate than the wheat parent 7182. This indicates that H139 plants are resistant to take-all disease at the seedling and mature stage.

Furthermore, H139, which represents a new wheat-*P. huashania* 2Ns/2D substitution line resistant to take-all disease, was characterized by GISH, mc-FISH, molecular marker assays, Wheat 660K Genotyping Array analysis, and morphological traits evaluation. In an earlier study by Du et al. (2014c), the wheat-*P. huashanica* 2Ns disomic addition line 3-6-4-1 was shown to be generally resistant to stripe rust in both seedlings and mature plants. Following this, Du et al. (2015) described the wheat-*P. huashania* 2Ns/2D substitution line 16-6, which exhibited resistance to mixed races of stripe rust (CYR31, CYR32, and SY11-14) in mature plants. The disease resistance observed in these studies was attributed to the *P. huashania* parent, thus demonstrating that the introduction of *P. huashania* chromosome 2Ns could enhance the disease resistance of common wheat. Interestingly, in the pre-experiment of take-all identification, the disease index of 16-6 was greater than 50%, which was considered as a susceptible line. In addition, some morphological characters of H139 and 16-6 are also significantly different. To clarify the reason of their differences, we analyzed the distribution of differential SNP between 16-6 and H139, and found that there could be an obvious peak in 200–300 Mb region between the 2Ns chromosomes of the H139 and 16-6. The difference of SNP in the specific chromosome region of 2Ns may be the cause of the difference of resistance and morphology traits between the two substitution lines. And the uneven distribution of the polymorphic SNP may indicate a deletion or a unexplored chromosome variation on the 2Ns chromosome in one of the substitution lines. The wheat genetic background of the substitution lines 16-6 and H139 are the same as that of 7182 (Du et al., 2015), which provide an opportunity to develop a genetic mapping population for the 2Ns chromosome using the two 2Ns(2D) substitution lines as crossing partners. Producing knock-out mutants by EMS mutagenesis and the application of MutChromSeq approach together with the results of genetic mapping could be a promising way to identify resistance gene against take-all disease. Another important application of line H139 is to hybridize it with other intermediate materials such as hexaploid *triticale* (AABBRR), or directly with rye (Dongmu 70) that resistant to wheat take-all, to develop genetic offspring of multiple genera, and aggregate multiple resistance genes, so as to further improve the resistance to wheat take-all. Meanwhile, hybridization among different genera will further broaden the genetic variation of wheat and enrich the diversity of wheat germplasm materials.

CONCLUSION

Through genetically distant hybridization and chromosome engineering, the desirable disease resistance characteristics of *P. huashanica* were introduced into the genetic background of common wheat, which significantly improved take-all disease resistance. Wheat-*Psathyrostachys huashania* Keng 2Ns/2D substitution line H139, as a new germplasm with resistance to wheat take-all and desirable agronomic traits, can be employed as valuable genetic resources for wheat chromosome engineering breeding.

DATA AVAILABILITY STATEMENT

The wheat 660k array data of H139 has been successfully uploaded to GEO, The accession number is GSE143188. We may view the GSE143188 study at: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143188.

AUTHOR CONTRIBUTIONS

SS and XC designed and directed the study. SS, QY, JZ, and JW selected the wheat-*P. huashanica* 2Ns/2D substitution line H139. SS, FY, HZ, and ZZ performed the experiments. SS, HZ, and JW analyzed the data. JW and XC contributed reagents, materials, and analysis tools. SS wrote the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 31571650 and 31771785), the National Key Research and Development Program of China (No. 2017YFD0100701), the Key Projects in Shaanxi Provincial Agricultural Field (No. 2018ZDXM-NY-006), and the Key Research and Development Project of Shaanxi Province (No. 2019ZDLNY04-05).

ACKNOWLEDGMENTS

We thank Pro. Huaigu Chen, the Institute of Plant Protection, Jiangsu Academy of Agricultural Science, Nanjing, China, for *G. tritici* material G1280. We thank Dr. Xiaolong Wang from the College of Food Engineering and Nutritional Science of Shaanxi Normal University, Shaanxi, China, for oat material Beiyou 8. We also thank Dr. Jianhui Wu and TopEdit UC for useful advice and linguistic editing of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00233/full#supplementary-material

FIGURE S1 | The development process of wheat-*Psathyrostachys huashania* 2Ns/2D substitution line H139 and 16-6.

TABLE S1 | 75 EST-STS markers with polymorphisms between *P. huashania* and 7182 were used in this study.

TABLE S2 | The distribution proportion of differential SNP on chromosomes between H139 and 16-6.

REFERENCES

An, D. G., Ma, P. T., Zheng, Q., Fu, S. L., Li, L. H., Han, F. P., et al. (2019). Development and molecular cytogenetic identification of a new wheat-rye 4R chromosome disomic addition line with resistances to powdery mildew, stripe rust and sharp eyespot. *Theor. Appl. Genet.* 132, 257–272. doi: 10.1007/s00122-018-3214-3

An, D. G., Zheng, Q., Zhou, Y. L., Ma, P. T., Lv, Z. L., Li, L. H., et al. (2013). Molecular cytogenetic characterization of a new wheat–rye 4R chromosome translocation line resistant to powdery mildew. *Chromosome Res.* 21, 419–432. doi: 10.1007/s10577-013-9366-8

Baden, C. (1991). A taxonomic revision of *Psathyrostachys* (Poaceae). *Nord. J. Bot.* 11, 3–26. doi: 10.1111/j.1756-1051.1991.tb01790.x

Chen, S. Y., Hou, W. S., Zhang, A. J., Fu, J., and Yang, Q. H. (1996). Breeding and cytogenetic study of Triticum aestivum-*Psathyrostachys huashanica* alien addition lines. *Acta Genet. Sin.* 23, 447–452.

Chen, S. Y., Zhang, A. J., and Fu, J. (1991). The hybridization between Triticum aestivum and *Psathyrotachys huashanica*. *Acta Genet Sin.* 18, 508–512.

Cook, R. J. (2003). Take-all of wheat. *Physiol. Mo. Plant Pathol.* 62, 73–86. doi: 10.1016/s0885-5765(03)00042-0

Cota-Sánchez, J. H., Remarchuk, K., and Ubayasena, K. (2006). Ready-touse DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. *Plant Mol. Biol. Rep.* 24, 161–167. doi: 10.1007/bf02914055

Crombie, W. M. L., Crombie, L., Green, J. B., and Lucas, J. A. (1986). Pathogenicity of 'take-all' fungus to oats: its relationship to the concentration and detoxification of the four avenacins. *Phytochemistry.* 25, 2075–2083. doi: 10.1016/0031-9422(86)80069-3

Du, W. L., Wang, J., Lu, M., Sun, S. G., Chen, X. H., Zhao, J. X., et al. (2013a). Molecular cytogenetic identification of a wheat-*Psathyrostachys huashanica* Keng 5Ns disomic addition line with stripe rust resistance. *Mol. Breed* 31, 879–888. doi: 10.1007/s11032-013-9841-0

Du, W. L., Wang, J., Lu, M., Sun, S. G., Chen, X. H., Zhao, J. X., et al. (2014a). Characterization of a whea-*Psathyrostachys huashanica* Keng 4Ns disomic addition line for enhanced tiller numbers and stripe rust resistance. *Planta* 239, 97–105. doi: 10.1007/s00425-013-1957-2

Du, W. L., Wang, J., Pang, Y. H., Li, Y. L., Chen, X. H., Zhao, J. X., et al. (2013b). Isolation and characterization of a *Psathyrostachys huashanica* Keng 6Ns chromosome addition in common wheat. *PLoS One* 8:e53921. doi: 10.1371/journal.pone.0053921

Du, W. L., Wang, J., Pang, Y. H., Wang, L. M., Wu, J., Zhao, J. X., et al. (2014b). Isolation and characterization of a wheat-*Psathyrostachys huashanica* Keng 3Ns disomic addition line with resistance to stripe rust. *Genome* 57, 37–44. doi: 10.1139/gen-2013-0199

Du, W. L., Wang, J., Pang, Y. H., Wu, J., Zhao, J. X., Liu, S. H., et al. (2014c). Development and application of PCR markers specific to the 1Ns chromosome of *Psathyrostachys huashanica* Keng with leaf rust resistance. *Euphytica* 200, 207–220. doi: 10.1007/s10681-014-1145-x

Du, W. L., Wang, J., Wang, L. M., Wu, J., Zhao, J. X., Liu, S. H., et al. (2014d). Molecular characterization of a wheat-*Psathyrostachys huashanica* Keng 2Ns disomic addition line with resistance to stripe rust. *Mol. Genet. Genomics* 289, 735–743. doi: 10.1007/s00438-014-0844-2

Du, W. L., Wang, J., Wang, L. M., Zhang, J., Chen, X. H., Zhao, J. X., et al. (2013c). Development and characterization of a *Psathyrostachys huashanica* Keng 7Ns chromosome addition line with leaf rust resistance. *PLoS One* 8:e70879. doi: 10.1371/journal.pone.0070879

Du, W. L., Zhao, J. X., Wang, J., Wang, L. M., Wu, J., Yang, Q. H., et al. (2015). Cytogenetic and molecular marker-based characterization of a wheat-*Psathyrostachys huashanica* Keng 2Ns(2D) substitution line. *Plant Mol. Biol. Rep.* 33, 414–423. doi: 10.1007/s11105-014-0761-x

Feng, Y. X., Li, W., Sun, H. Y., Deng, Y. Y., Yu, H. S., and Chen, H. G. (2013). Genetic diversity of Gaeumannomyces graminis var. tritici populations in Huang-Huai winter wheat production region of China. ACTA. Phytophylacica. *SINICA* 40, 495–501.

Fu, J., Wang, M. N., Zhao, J. X., Chen, S. Y., Hou, W. S., and Yang, Q. H. (2003). Studies on cytogenetics and utilization of wheat-Psathyrostachys huashanica medium material H8911 with resistance to wheat take-all fungus. *Acta Bot. Boreal. Occident. Sin.* 23, 2157–2162.

Han, F. P., Lamb, J. C., and Birchler, J. A. (2006). High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. *Proc. Nat. Acad. Sci. U.S.A* 103, 3238–3243. doi: 10.1073/pnas.0509650103

Hernández-Restrepo, M., Groenewald, J. Z., Elliott, M. L., Canning, G., Mcmillan, V. E., and Crous, P. W. (2016). Take-all or nothing. *Stud. Mycol.*

Hollins, T. W., Scott, P. R., and Gregory, R. S. (1986). The relative resistance of wheat, rye and triticale to take-all caused by *Gaeumannomyces graminis*. *Plant Pathol.* 35, 93–100. doi: 10.1111/j.1365-3059.1986.tb01986.x

International Wheat Genome Sequencing Consortium (IWGSC) (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:6403. doi: 10.1126/science.aar7191

Jia, T. Y., Wu, Q. B., Ye, X. C., Zang, F. C., Zhuo, Y. N., and Gong, B. Y. (1982). A preliminary study on the take-all of wheat. *Sci. Agric. Sin.* 15, 65–75.

Jiang, J., Friebe, B., and Gill, B. S. (1993). Recent advances in alien gene transfer in wheat. *Euphytica* 73, 199–212. doi: 10.1038/hdy.2012.116

Kuo, P. C. (1987). *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press, 51–104.

Li, F., Li, Y. H., Cao, L. R., Liu, , P. Y., Geng, M. M., Zhang, Q., et al. (2018). Simultaneous transfer of leaf rust and powdery mildew resistance genes from hexaploid triticale cultivar sorento into bread wheat. *Front. Plant Sci.* 9:85. doi: 10.3389/fpls.2018.00085

Mcintyre, C. L., Pereira, S., Moran, L. B., and Appels, R. (1990). New Secale cereale (Rye) DNA derivatives for the detection of rye chromosome segments in wheat. *Genome* 33, 635–640. doi: 10.1139/g90-094

McMillan, V. E., Gutteridge, R. J., and Hammond-Kosack, K. E. (2014). Identifying variation in resistance to the take-all fungus, Gaeumannomyces graminisvar.tritici, between different ancestral and modern wheat species. *BMC Plant Biol.* 14:212. doi: 10.1186/s12870-014-0212-8

Mukai, Y., Nakahara, Y., and Yamamoto, M. (1993). Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fuorescence in situ hybridization using total genomic and highly repeated DNA probes. *Genome* 36, 489–494. doi: 10.1139/g93-067

Papadopoulou, K., Melton, R. E., Leggett, M., and Osbourn, M. J. D. E. (1999). Compromised disease resistance in saponin-deficient plants. *Proc. Natl. Acad. Sci. U.S.A.* 96, 12923–12928. doi: 10.1073/pnas.96.22.12923

Penrose, L. (1985). Evidence for resistance in wheat cultivars grown in sand culture to the take-all pathogen, Gaeumannomyces graminis var. tritici. *Ann. Appl. Biol.* 107, 105–108. doi: 10.1111/j.1744-7348.1985.tb01552.x

Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R., and Osbourn, A. (2004). A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8233–8238. doi: 10.1073/pnas.0401301101

Rayburn, A. L., and Gill, B. S. (1986). Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. *Plant Mo. Biol. Rep.* 4, 102–109. doi: 10.1007/bf02732107

Schneider, A., Linc, G., and Molnár-Láng, M. (2003). Fluorescence in situ hybridization polymorphism using two repetitive DNA clones in different cultivars of wheat. *Plant Breed* 122, 396–400. doi: 10.1046/j.1439-0523.2003.00891.x

Scott, P. R., Hollins, T. H., and Summers, R. W. (1989). Breeding for resistance to two soil-borne diseases of cereals. *Vortraege Fuer Pflanzenzuechtung* 16, 217–230.

Tang, Z. X., Yang, Z. J., and Fu, S. L. (2014). Oligonucleotides replacing the roles of repetitive sequences pAs1, pSc119.2, *pTa*-535, pTa71, CCS1, and pAWRC.1 for FISH analysis. *J. Appl. Genet.* 55, 313–318. doi: 10.1007/s13353-014-0215-z

Walker, J. (1981). "Taxonomy of take-all fungi and related genera and species," in *Biology and control of take-all*, eds M. J. C. Asher, and P. J. Shipton (London: Academic Press), 15–74.

Wang, D. B., Wang, M. N., Jing, J. X., Shang, H. S., and Li, F. X. (2003). Study of resistance in distant hybridigation descendant of wheat to the take-all. *Acta Bot. Boreal.Occident. Sin.* 23, 1617–1620.

Wang, M. N., and Shang, H. S. (2000). Evalution of resistance in *Psathyrostachys huashaica* to wheat take-all fungus. *Acta Uninv. Agric. Boreali Occidentalis* 28, 69–71.

Wang, N., Feng, Y. X., Du, W. Z., Wang, Y., and Chen, H. G. (2012). Virulence of wheat Take-all pathogen and disease resistance of different wheat cultivars. *J. Plant Genet. Resour.* 13, 478–483.

Wu, J., Zhao, J. X., Chen, X. H., Liu, S. H., Yang, Q. H., Liu, W. X., et al. (2007). Cytology characteristic and GISH analysis on the progenies derived from common wheat (*T. aestivum* L.) ×*Psathyrostachys huashanica*. *J. Triticeae Crops* 27, 772–775.

Wu, J. H., Wang, Q. L., Liu, S. J., Huang, S., Mu, J. M., Zeng, Q. D., et al. (2017). Saturation mapping of a major effect QTL for stripe rust resistance on wheat chromosome 2B in cultivar Napo 63 using SNP genotyping arrays. *Front. Plant Sci.* 8:653. doi: 10.3389/fpls.2017.00653

Wu, J. H., Zeng, Q. D., Wang, Q. L., Liu, S. J., Yu, S. Z., Mu, J. M., et al. (2018). SNP-based pool genotyping and haplotype analysis accelerate fine-mapping of the wheat genomic region containing stripe rust resistance gene *Yr26*. *Theor. Appl. Genet.* 131, 1–16. doi: 10.1007/s00122-018-3092-8

Youn-Sig, K., and Weller, D. M. (2013). Take-all of wheat and natural disease suppression: a review. *Plant Pathol. J.* 29, 125–135. doi: 10.5423/ppj.si.07.2012.0112

Zhou, S. H., Zhang, J. P., Che, Y. H., Liu, W. H., Lu, Y. Q., Yang, X. M., et al. (2018). Construction of Agropyron, Gaertn. genetic linkage maps using a wheat 660k SNP array reveals a homoeologous relationship with the wheat genome. *Plant Biotechnol. J.* 16, 818–827. doi: 10.1111/pbi.12831

Zhuang, L. F., Liu, P., Liu, Z. Q., Chen, T. T., Wu, N., Sun, L., et al. (2015). Multiple structural aberrations and physical mapping of rye chromosome 2R introgressed into wheat. *Mol. Breed.* 35, 133.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 May 2020
doi: 10.3389/fpls.2020.00606

Development of Wheat-*Aegilops caudata* Introgression Lines and Their Characterization Using Genome-Specific KASP Markers

Surbhi Grewal[1]†, Manel Othmeni[1†], Jack Walker[1], Stella Hubbart-Edwards[1], Cai-yun Yang[1], Duncan Scholefield[1], Stephen Ashling[1], Peter Isaac[2], Ian P. King[1] and Julie King[1]*

[1] Division of Plant and Cop Sciences, Nottingham BBSRC Wheat Research Centre, University of Nottingham, Nottingham, United Kingdom, [2] IDna Genetics Ltd., Norwich Research Park, Norwich, United Kingdom

OPEN ACCESS

Edited by:
Laura Toppino,
Research Unit for Horticulture (ENT), Council for Research in Agriculture and the Analysis of Agrarian Economy (CREA), Italy

Reviewed by:
Filippo Maria Bassi,
International Center for Agricultural Research in the Dry Areas (ICARDA), Morocco
István Molnár,
Centre for Agricultural Research, Hungary

***Correspondence:**
Surbhi Grewal
surbhi.grewal@nottingham.ac.uk

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 02 March 2020
Accepted: 21 April 2020
Published: 14 May 2020

Citation:
Grewal S, Othmeni M, Walker J, Hubbart-Edwards S, Yang C, Scholefield D, Ashling S, Isaac P, King IP and King J (2020) Development of Wheat-Aegilops caudata Introgression Lines and Their Characterization Using Genome-Specific KASP Markers. Front. Plant Sci. 11:606. doi: 10.3389/fpls.2020.00606

Aegilops caudata L. [syn. *Ae. markgrafii* (Greuter) Hammer], is a diploid wild relative of wheat (2n = 2x = 14, CC) and a valuable source for new genetic diversity for wheat improvement. It has a variety of disease resistance factors along with tolerance for various abiotic stresses and can be used for wheat improvement through the generation of genome-wide introgressions resulting in different wheat–*Ae. caudata* recombinant lines. Here, we report the generation of nine such wheat–*Ae. caudata* recombinant lines which were characterized using wheat genome-specific KASP (Kompetitive Allele Specific PCR) markers and multi-color genomic *in situ* hybridization (mcGISH). Of these, six lines have stable homozygous introgressions from *Ae. caudata* and will be used for future trait analysis. Using cytological techniques and molecular marker analysis of the recombinant lines, 182 KASP markers were physically mapped onto the seven *Ae. caudata* chromosomes, of which 155 were polymorphic specifically with only one wheat subgenome. Comparative analysis of the physical positions of these markers in the *Ae. caudata* and wheat genomes confirmed that the former had chromosomal rearrangements with respect to wheat, as previously reported. These wheat–*Ae. caudata* recombinant lines and KASP markers are useful resources that can be used in breeding programs worldwide for wheat improvement. Additionally, the genome-specific KASP markers could prove to be a valuable tool for the rapid detection and marker-assisted selection of other Aegilops species in a wheat background.

Keywords: wheat, *Aegilops caudata*, introgression, recombinant chromosomes, KASP, GISH

INTRODUCTION

Wheat is one of the most widely cultivated crops worldwide, contributing about a fifth of the total calories and protein consumed by humans. Although wheat breeders report annual wheat yield gains, they are not sufficient to meet the needs of an ever-growing population. This is due to a lack of genetic diversity present in the gene pool of wheat that can be utilized for the development of higher yielding wheat varieties adapted to climate change. However, wild relatives of wheat possess

genetic diversity that could be exploited in wheat breeding programs through introgression of small segments of their genome, carrying desirable traits, into wheat (Feuillet et al., 2008).

Aegilops caudata L. [syn. *Ae. markgrafii* (Greuter) Hammer] is a diploid wild relative (2n = 2x = 14, CC) of hexaploid wheat, *Triticum aestivum* (2n = 6x = 42, AABBDD). It is native to the north-eastern Mediterranean basin with the main distribution from Greece to northern Iraq (Ohta and Yasukawa, 2015). It contributed the C-genome to allotetraploids *Aegilops cylindrica* (2n = 4x = 28; DDCC) and *Aegilops truncialis* (2n = 4x = 28; UUCC). *Ae. caudata* carries resistance genes for several diseases (Makkouk et al., 1994), especially against stripe rust (Valkoun et al., 1985; Baldauf et al., 1992; Toor et al., 2016), leaf rust (Iqbal et al., 2007; Riar et al., 2012; Gong et al., 2017), stem rust (Dyck et al., 1990; Gong et al., 2017) and powdery mildew (Gill et al., 1985; Valkoun et al., 1985; Baldauf et al., 1992; Weidner et al., 2012) and resistance against pests such as greenbug and hessian fly (Gill et al., 1985; Baldauf et al., 1992). Moreover, some accessions of *Ae. caudata* have shown tolerance to abiotic stresses such as freezing (Barashkova and Migushova, 1984; Iriki et al., 2001) and salinity (Gorham, 1990) and seed quality traits, such as increased grain mineral content (Wang et al., 2011). Thus *Ae. caudata* provides an important source of genetic variation for agronomically important traits that can be transferred into wheat as has been previously done for other Aegilops species (King et al., 2017, 2018).

In the past, *Ae. caudata* was not widely used in wheat pre-breeding programs due to the inability to identify the C-genome chromosomes in a wheat background. In addition, poor knowledge of the C-genome organization and syntenic relationships between wheat and *Ae. caudata* chromosomes has hampered the use of its genetic potential in wheat breeding. Over the past decades, however, efforts have been made to study the molecular organization of the *Ae. caudata genome* and its homology with wheat homoeologous groups. Due to the presence of metacentric and submetacentric chromosomes, *Ae. caudata* has been shown to have a highly asymmetric karyotype (Friebe et al., 1992). This is distinct from most of the Triticeae species and would suggest that chromosome collinearity in *Ae. caudata* is distorted compared to wheat. Fluorescence *in situ* hybridization (FISH) and molecular marker analysis of flow-sorted C-genome chromosomes have also confirmed genome rearrangements in *Ae. caudata* (Molnár et al., 2015, 2016). A set of wheat (cv. Alcedo)-*Ae. caudata* addition lines B-G (Schubert and Blüthner, 1992, 1995) have also been characterized extensively in previous studies using cytogenetic markers (Friebe et al., 1992), isozyme analysis (Schmidt et al., 1993), Simple Sequence Repeat (SSR) markers (Peil et al., 1998; Gong et al., 2017; Niu et al., 2018), FISH with cDNA probes (Danilova et al., 2017), Conserved Orthologous Sequence (COS) and PCR-based Landmark Unique Gene (PLUG) markers (Gong et al., 2017), and sequential FISH and genomic *in situ* hybridization (GISH) (Niu et al., 2018). Even though most of these studies found that the set of addition lines carried chromosomal rearrangements compared to wheat, the assignment of *Ae. caudata* chromosomes to corresponding wheat homoeologous groups was not consistent between them.

Despite numerous studies using wheat-*Ae. caudata* addition lines and the development of molecular markers capable of detecting *Ae. caudata* chromosomes in a wheat background, its utilization for wheat improvement through recombination with wheat chromosomes has rarely been reported (Dyck et al., 1990; Iqbal et al., 2007; Riar et al., 2012; Toor et al., 2016; Weidner et al., 2012). In this work, we report the generation of six stable wheat-*Ae. caudata* recombinant lines and the physical location of 182 KASP markers on *Ae. caudata* chromosomes, that can be used to detect the presence of *Ae. caudata* introgressions in a wheat background.

MATERIALS AND METHODS

Plant Material

Bread wheat cv. Paragon *ph1/ph1* mutant was pollinated with *Ae. caudata* accession 2090001 (obtained from Germplasm Resource Unit, GRU at the John Innes Centre, Norwich, United Kingdom) to produce F_1 hybrids. The origin of accession 2090001, according to the GRU database Seedstor, is unknown. There is no trait data available for this accession and it was thus, chosen at random. *T. aestivum* cv. Alcedo- *Ae. caudata* disomic addition lines B-G (TA3558-TA3563) carrying *Ae. caudata* chromosomes 2C, 5C, 6C, 7C, 3C, and 4C respectively, and the Alcedo cultivar were obtained from Wheat Genetics Resource Center (WGRC) at Kansas State University, United States.

The crossing program is as described in Grewal et al. (2018). In summary, the F_1 interspecific hybrids were grown to maturity and used as the female parent while backcrossed with Paragon wheat, *Ph1/Ph1*, to generate a BC_1 population. The BC_1 plants were then recurrently pollinated with Paragon *Ph1/Ph1* to produce BC_2, BC_3, BC_4 and BC_5 lines. 439 crossed seed were germinated and 278 plants were grown to maturity. Self-fertilization was facilitated by bagging three heads in each plant from the BC_2 generation onward. Cross fertility was calculated as the number of crosses producing seed.

Genotyping With KASP Markers

In this work, we used a set of previously produced 183 KASP markers polymorphic between hexaploid wheat and *Ae. caudata* accession 2090001 (Grewal et al., 2020a). Genomic DNA was isolated from leaf tissue of 10-day old seedlings in a 96-well plate as described by Thomson and Henry (1995).

Back-crossed and self-fertilized BC_xF_y (where x = 1–5 and y = 0–2) lines were genotyped alongside five wheat genotypes (Chinese Spring, Paragon, Pavon, Highbury, and Alcedo) and the *Ae. caudata* accession as controls. Even though 278 plants were grown to maturity, only those plants in the BC_1, BC_2, and BC_3 generations that produced either crossed or self-fertilized seed were genotyped. The wheat (Alcedo)-*Ae. caudata* B-G addition were also used as control lines to verify the specificity of the KASP markers (Grewal et al., 2020a) to each of the *Ae. caudata* chromosomes.

For each KASP^{TM} assay, two allele-specific primers and one common primer were used (**Supplementary Table S1**). Genotyping reactions were performed as described in

Othmeni et al. (2019) using a ProFlex PCR system (Applied Biosystems by Life Technology). A final reaction volume of 5 μl was used which included 1 ng genomic DNA, 2.5 μl KASP reaction mix (ROX), 0.068 μl primer mix and 2.43 μl nuclease free water. PCR conditions were set as 15 min at 94°C; 10 touchdown cycles of 10 s at 94°C, 1 min at 65–57°C (dropping 0.8°C per cycle); and 35 cycles of 10 s at 94°C, 1 min at 57°C. QuantStudio 5 (Applied Biosystems) was used for fluorescence detection of the reactions and the data was analyzed using the QuantStudio™ Design and Analysis Software V1.5.0 (Applied Biosystems).

Multi-Color Genomic *in situ* Hybridization (mcGISH)

Preparation of the root-tip metaphase chromosome spreads, the protocol for mcGISH and the image capture was as described in Grewal et al. (2020a). In summary, genomic DNA from *Triticum urartu* (to detect the A-genome), *Aegilops speltoides* (to detect the B-genome), and *Aegilops tauschii* (to detect the D-genome) and *Ae. caudata* were isolated using extraction buffer [0.1M Tris-HCl (pH 7.5), 0.05M EDTA (pH 8.0), 1.25% SDS]. Samples were incubated at 65°C for 1 h before being placed on ice and mixed with ice cold 6 M $NH_4C_2H_3O$ for 15 min. The samples were then spun down, the supernatant mixed with isopropanol to pellet the DNA and the isolated DNA further purified with phenol/chloroform. The genomic DNA of (1) *T. urartu* was labeled by nick translation with ChromaTide™ Alexa Fluor™ 488-5-dUTP (Invitrogen; C11397; colored green), (2) *Ae. speltoides* was labeled by nick translation with DEAC-dUTP (Jena Bioscience; NU-803-DEAC; colored blueish purple), (3) *Ae. tauschii* was labeled with ChromaTide™ Alexa Fluor™ 594-5-dUTP (Invitrogen; C11400; colored red) and 4) *Ae. caudata* was labeled by nick translation with ChromaTide™ Alexa Fluor™ 546-14-dUTP (Invitrogen; C11401; colored yellow). Slides were probed using 150 ng of *T. urartu*, 150 ng of *Ae. speltoides,* 300 ng of *Ae. tauschii* and 50 ng of *Ae. caudata* labeled genomic DNAs, in the ratio 3:3:6:1 (green: blue: red: yellow). No blocking DNA was used. DAPI was used for counterstaining all slides. Metaphases were detected using a high-throughput, fully automated Zeiss Axio ImagerZ2 upright epifluorescence microscope (Carl Zeiss Ltd., Oberkochen, Germany) with filters for DAPI (blue), Alexa Fluor 488 (green), Alexa Fluor 594 (red), Alexa Fluor 546 (yellow) and DEAC (aqua). Image capture was performed using a MetaSystems Coolcube 1m CCD camera and image analysis was carried out using Metafer4 (automated metaphase image capture) and ISIS (image processing) software (Metasystems GmbH, Altlussheim, Germany).

Fluorescence *in situ* Hybridization (FISH)

Root-tip metaphase chromosome spreads for Fluorescence *in situ* hybridization (FISH) were prepared as described above. Two repetitive DNA sequences, pSc119.2 (McIntyre et al., 1990), and pAs.1 (Rayburn and Gill, 1986) were used as probes. They were labeled with Alexa Fluor 488-5-dUTP (green) and Alexa Fluor 594-5-dUTP (red), respectively, and hybridized to the slide in the ratio 2:1 (green: red). Subsequent counterstaining and image capture were performed as described for GISH. The wheat FISH karyotype used for assigning chromosomes was that established by Tang et al. (2014).

RESULTS

Assignment of KASP Markers to *Ae. caudata* Chromosomes

In a previous study, we reported the generation of 183 KASP markers polymorphic between hexaploid wheat and *Ae. caudata* accession 2090001 (Grewal et al., 2020a). Although, the physical location of these markers on the wheat genome was reported, the study did not address the physical location of these markers on the *Ae. caudata* genome.

In this work, the KASP markers were tested on the Alcedo-*Ae. caudata* addition lines carrying chromosomes B-G. It was their physical location and distribution on these C-genome chromosomes, among the addition lines, that determined the assignment of the *Ae. caudata* chromosomes to the wheat homoeologous groups. When a KASP marker detected the presence of *Ae. caudata* in any one of the addition lines, the marker was assigned to the corresponding C-genome chromosome. Of the 183 KASP markers tested, 156 were individually assigned to an *Ae. caudata* chromosome present in the disomic addition lines (**Table 1**). This set is missing an addition line for chromosome A, but reports have suggested it to be homoeologous to group 1 of wheat (Friebe et al., 1992; Danilova et al., 2017; Niu et al., 2018). Thus, of the remaining 27 markers, one marker failed and the other 26 markers that did not detect *Ae. caudata* in the addition lines B-G but amplified the *Ae. caudata* accession were assigned to chromosome 1C.

There were 155 KASP markers specific to a wheat subgenome and 27 which had more than one corresponding sequence in the wheat genome and hence, were genome-non-specific for wheat (**Supplementary Table S2**). In total, 40 KASP markers were developed for the C-genome of *Ae. caudata* that were polymorphic with the A-genome of wheat, 53 with the B-genome and 62 with the D-genome of wheat. Chromosome 5C had the maximum number of KASP markers assigned to it (33) whereas the rest of the C-genome chromosomes had between 22 and 28 KASP markers assigned to them (**Table 1**). However, some C-genome chromosomes such as 2C and 4C, had markers assigned from non-homoeologous groups as shown in **Table 1**.

Generation and Identification of Wheat-*Ae. caudata* Recombinant Lines

To generate introgressions from *Ae. caudata* into wheat, a crossing program was initiated between the two species that resulted in 165 F_1 seeds from 49 crosses. A subsequent back-crossing program starting with 36 F_1 plants and involving 439 back-crosses over five generations (F_1-BC_4) led to the production of 1,700 back-crossed seed and 3,496 self-fertilized seed. The number of seeds sown, germination rate, cross fertility and seed set, etc., are summarized in **Table 2**.

TABLE 1 | Assignment of genome-specific and genome-non-specific KASP markers, and consequently homoeologous groups, to *Ae. caudata* chromosomes derived from Alcedo-*Ae. caudata* disomic addition lines and the distribution of the genome-specific KASP markers on the wheat A-, B-, and D-genomes.

Ae. caudata chromosome	Alcedo-*Ae. caudata* addition line	Genome-specific KASP markers			Genome-non-specific markers	Total number of KASP markers	Homoeologous group(s)
		A	B	D			
1C	–	7	7	9	3	26	1
2C	B	10	6	10	1	27	2/4
3C	F	2	6	7	7	22	3
4C	G	4	9	6	4	23	4/2/3/7
5C	C	8	14	10	1	33	5
6C	D	3	7	5	8	23	6
7C	E	6	4	15	3	28	7
Total		40	53	62	27	182	

TABLE 2 | Summary of number of seeds germinated and produced along with the number of crosses carried out and cross fertility in relation to the number of plants genotyped for each generation of the introgression program for *Ae. caudata* into wheat.

	Seeds sown	Plants grown to maturity	Germination rate (%)	Crosses made	Cross fertility (%)	Crossed seeds produced	Seeds/ Cross	Self-fertilized seeds produced	Plants genotyped	Plants with *Ae. caudata* chromatin
Wheat × *Ae. Caudata*	–	–	–	49	43	165	3.4	–	–	–
F_1	48	36	75	161	5	9	0.06	–	0	–
BC_1	9	7	33	76	29	42	0.6	–	3	3
BC_2	14	11	50	44	65	187	4.3	3	7	7
BC_3	38	28	74	98	87	842	8.6	296	24	21
BC_3F_1	45	30	67	–	–	–	–	1217	30	15
BC_3F_2	18	11	61	–	–	–	–	*	11	7
BC_4	81	50	62	60	100	620	10.3	1980	50	42
BC_4F_1	94	51	54	–	–	–	–	*	51	47
BC_5	92	54	59	–	–	–	–	*	54	39
Total	439	278		488		1865		3496	230	181

**Plants are currently undergoing self-fertilization.*

The KASP markers were also used to genotype a set of 230 wheat-*Ae. caudata* BC_xF_y lines to detect the presence of *Ae. caudata* introgressions in wheat. Of the 230 plants genotyped, 181 plants were found to have *Ae. caudata* chromatin (recombinant chromosomes and/or whole chromosomes) which included 3 BC_1, 7 BC_2, 21 BC_3, 15 BC_3F_1, 7 BC_3F_2, 42 BC_4, 47 BC_4F_1, and 39 BC_5 plants (**Table 2**). Analysis of the genotypes showed nine advanced back-crossed lines that had one or two recombinant chromosomes without the presence of any whole *Ae. caudata* chromosome (**Table 3**). Two recombinants were obtained from chromosome 1C, one from chromosome 4C, four from chromosome 5C and one from chromosome 7C. One introgression line had two of the above-mentioned recombinant chromosomes, one from 1C and one from 5C. Of the 7 BC_1 plants grown to maturity, only 3 BC_1 plants produced BC_2 seed. All of the eight recombinant chromosomes were first observed in these 3 BC_1 plants.

After self-fertilization, six of these BC_xF_y lines showed stable homozygous introgressions which were validated by mcGISH analysis (**Figures 1a–f**). Five out of these six lines had 42 chromosomes with the sixth line having 41 chromosomes including the recombinant chromosomes (**Table 3**). However, prior to mcGISH analysis, the genome-specific KASP markers enabled the identification of the homozygous lines and the wheat chromosome that had recombined with the *Ae. caudata* segment in the homozygous recombinant lines (**Figures 2a–h**).

Figure 2 shows the mcGISH image of the recombinant chromosome(s) (top) and the KASP marker analysis (bottom) of eight wheat-*Ae. caudata* introgression lines. The genotyping is displayed with the markers physically located on the wheat chromosomes (for e.g., chromosomes 1A, 1B, and 1D in **Figure 2a**) and the same set of markers on the *Ae. caudata* chromosomes (for e.g., chromosome 1C in **Figure 2a**). When a line had a homozygous introgression, the KASP markers on the wheat chromosome involved in the recombination showed homozygous green calls as shown in **Figures 2a,d,e,f,h**. This is due to the genome-specificity of the KASP markers in wheat. Thus, when two copies of the *Ae. caudata* segment replaced both copies of the homoeologous wheat regions in the homozygous recombinant chromosomes, it resulted in no wheat allele for the genome-specific KASP markers in those regions. All other KASP markers detecting the *Ae. caudata* segment but not physically located on the recombinant chromosome show a heterozygous call to indicate the presence of *Ae. caudata* in a wheat

TABLE 3 | List of wheat-*Ae. caudata* introgression lines containing recombinant chromosomes (RC) obtained in this study, the copy number of the RC(s) and the total number of chromosomes (2n) in each line.

Introgression line	Recombinant chromosome(s) (RC)	Copies of RC(s)	Total chromosome number 2n
BC_4F_1-157-9	T1AL-1CL.1CS	2	42
BC_4F_1-157-2	T5AL-5CL.5CS#1	2	42
BC_4F_1-158-12	T5AL-5CL.5CS#2	2	42
BC_4F_1-162-6	T5DL-5CL.5CS	2	42
BC_3F_2-160-4	T7DL-7CL.7CS	2	41
BC_4F_1-156-3	T1AL-1CL.1CS, T5AL-5CL.5CS#2	2	42
BC_4-284-4	T1DS.1DL-1CL	1	42
BC_4F_1-164-2	T3DS.3DL-4CL	1	41
BC_3F_1-401-8	T5BL-5CL.5CS	1	42

background. The marker analysis in **Figures 2a,e,h** indicated the C-genome segment had recombined with the A-genome of wheat whereas in **Figures 2b,f**, the markers indicated the C-genome segment had recombined with the D-genome of wheat. The mcGISH analysis, above, confirmed these results.

When a line had a heterozygous introgression, as shown in **Figures 2b,c,g**, all the KASP markers that were located within the introgressed *Ae. caudata* segment gave a heterozygous call and thus, were not able to indicate which wheat chromosome the segment had recombined with. The mcGISH analysis (top), however, indicated the wheat genome, B (**Figure 2g**) or D (**Figures 2b,c**), present in these recombinant chromosomes.

Physical Ordering of KASP Markers on C-Genome Chromosomes and Comparative Analysis With Wheat

Based on the molecular marker analysis of the Alcedo-*Ae. caudata* addition lines (**Table 1**) and the wheat-*Ae. caudata* back-crossed population (**Figure 2**; bottom), the mcGISH analysis of the recombinant chromosomes (**Figure 2**; top) and previous reports on C-genome chromosomal rearrangements (Danilova et al., 2017), the 182 KASP makers were tentatively ordered onto the seven *Ae. caudata* C-genome chromosomes (**Figure 3**). A comparative analysis of the markers on the C-genome chromosomes and their physical positions on the wheat A, B, and D-genome chromosomes (**Supplementary Table S2**), as indicated by the wheat reference genome assembly RefSeqv1 (International Wheat Genome Sequencing Consortium [IWGSC]et al., 2018), is shown in **Figure 4**.

KASP marker analysis of the recombinant lines showed that markers WRC0002-0077 on chromosome 1C (**Figure 3**) corresponding to wheat homoeologous group 1 markers, spanning the short arm and approximately half of the proximal region of the long arm (**Figure 4A**), detected the *Ae. caudata* segment present in the recombinant chromosome T1AL-1CL.1CS. The spread of these markers across homoeologous group 1 also matched the size of the segment subsequently shown by mcGISH (**Figure 2a**). KASP markers WRC0095-WRC0142 on chromosome 1C (**Figure 3**), corresponding to the distal end of the long arm of homoeologous group 1 (**Figure 4A**), detected the *Ae. caudata* introgression present in T1DS.1DL-1CL subsequently validated by mcGISH (**Figure 2b**). This suggested that the marker sequences on chromosome 1C were potentially collinear with their counterparts on wheat chromosomes 1A, 1B, and 1D.

Chromosome 4C of *Ae. caudata* had KASP markers from wheat homoeologous groups 2/3/4/7 assigned to it (**Table 1**). KASP markers WRC0387-0875 on chromosome 4C (**Figure 3**) detected the presence of an *Ae. caudata* segment in recombinant chromosome T3DS.3DL-4CL. However, **Figure 4F** shows that these markers were physically located on distal end of the long arm of wheat homoeologous group 3 (WRC0387, WRC0390 and WRC0335) with one marker near the centromere on homoeologous group 7 (WRC0875), indicating genomic rearrangements between homoeologous groups 3/4/7 in *Ae. caudata*. Due to the recombinant chromosome being present as a single copy in the introgression line, the KASP markers were not indicative of the wheat chromosome that had recombined with the *Ae. caudata* segment (**Figure 2c**). McGISH indicated it to be a D-genome chromosome which was subsequently confirmed to be chromosome 3D by FISH analysis (**Figures 5a,b**). The FISH analysis also showed that one copy of chromosome 1D was missing and hence, the total chromosome number for this line was 41 as mentioned in **Table 3**.

Four recombinants were obtained from chromosome 5C of *Ae. caudata*. KASP marker analysis showed that 21 markers, from WRC0668 to WRC0558 on chromosome 5C (**Figure 3**) detected all four recombinants. These markers correspond to wheat homoeologous group 5 spanning the small arm and approximately half of the proximal region of the long arm (**Figure 4C**). However, the four recombinants differed in their size and the wheat genome they had recombined with. For example, T5AL-5CL.5CS#1 and T5DL-5CL.5CS had the same 21 markers detecting the *Ae. caudata* segment but the former had recombined with chromosome 5A of wheat and the latter with chromosome 5D of wheat (**Figures 2e,f**, respectively). Recombinant chromosome T5BL-5CL.5CS had another 3 distal markers (WRC0703, WRC0595 and WRC0662; **Figure 3**), in addition to the 21, that detected the *Ae. caudata* segment in it and mcGISH analysis showed that it had recombined with the B-genome of wheat (**Figure 2g**). T5AL-5CL.5CS#2 had the longest segment from chromosome 5C of *Ae. caudata* and had another distal marker (WRC0645; **Figure 3**), making it a total of 25 markers, that detected this introgression. These markers indicated that recombination happened with chromosome 5A of wheat in this line and was also validated by mcGISH (**Figure 2h**). These segments assisted in the ordering of markers on chromosome 5C and comparative analysis showed that they were potentially collinear to corresponding sequences on wheat homoeologous group 5 (**Figure 4C**).

One large *Ae. caudata* segment was detected by 18 KASP markers, from WRC0908 to WRC0937 present on chromosome 7C (**Figure 3**). These markers indicated the segment to be to be homozygous and recombined with wheat chromosome 7D in the introgression line (**Figure 2d**). McGISH analysis of the recombinant chromosome T7DL-7CL.7CS showed that the

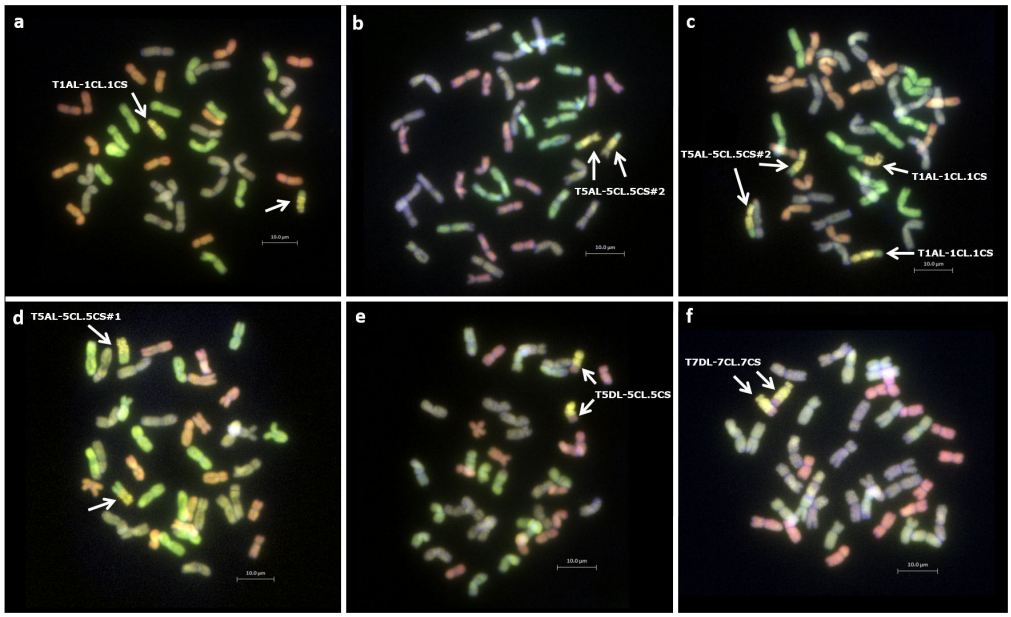

FIGURE 1 | Identification of wheat-*Ae. caudata* recombinant chromosomes in homozygous lines through mcGISH analysis of root-tip metaphase chromosomes spreads. Arrows indicate recombinant chromosomes. **(a)** BC_4F_1-157-9, **(b)** BC_4F_1-158-12, **(c)** BC_4F_1-156-3, **(d)** BC_4F_1-157-2, **(e)** BC_4F_1-162-6, and **(f)** BC_3F_2-160-4. *Ae. caudata* segments (yellow) recombined with the A- (green) and D- (red) genome chromosomes of wheat. Wheat B-genome is represented as grayish/purple chromosomes.

introgressed *Ae. caudata* segment included the short arm and a large part of the proximal region of the long arm of chromosome 7C, with the distal end of the recombinant chromosome being replaced with the distal end of chromosome 7DL (**Figure 2d**). However, the markers found in this segment were found to be present on the proximal region of the short arm and the whole of the long arm of wheat homoeologous group 7 (**Figure 4G**). The markers corresponding to the short arm of wheat homoeologous group 7 were not detecting this recombinant chromosome and were thus, physically ordered onto the distal end of chromosome 7C (**Figure 3**) indicating an inversion in *Ae. caudata* group 7 compared to wheat.

Where markers could not be ordered due to absence of any recombination with wheat, such as for homoeologous group 2, 3, 4, and 6, the markers on chromosome 2C, 3C, 4C (except the distal end of the long arm) and 6C were ordered based on a previous study by Danilova et al. (2017).

DISCUSSION

Assignment of KASP Markers to C-Genome Chromosomes Confirms Genomic Rearrangements

Ae. caudata has been shown to carry useful genetic diversity that can be used for wheat breeding. Its utilization for wheat improvement relies on the development of suitable and reliable molecular markers that can be used for the identification of *Ae. caudata* segments in introgression lines. In this research, a total of 182 KASP markers, developed in a previous study (Grewal et al., 2020a), were assigned to specific *Ae. caudata* chromosomes (**Table 1**), thereby, providing a rapid detection approach for the marker-assisted breeding of wheat-*Ae. caudata* introgression lines.

A number of studies have attempted assignment of various types of molecular markers to specific *Ae. caudata* chromosomes by testing them on the Alcedo-*Ae. caudata* disomic addition lines (Schubert and Blüthner, 1992, 1995). The variation observed in the results indicates the presence of genomic rearrangements. Recent studies using the Alcedo-*Ae. caudata* additions lines found that the C-genome chromosomes carried several inversions and translocations (Danilova et al., 2017; Gong et al., 2017; Niu et al., 2018). However, the deductions on the rearrangements carried by each chromosome were inconsistent among them. In this study, we used the same set of addition lines for the assignment of the KASP markers to the C-genome chromosomes and verified that *Ae. caudata* chromosomes B, C, D, E, F, and G were indeed 2C, 5C, 6C, 7C, 3C, and 4C respectively. Genotyping of these addition lines with the KASP markers showed that chromosome 2C had markers from wheat homoeologous groups 2/4 assigned to it (**Table 1**) which agreed with the work done by Danilova et al. (2017). Chromosome 4C had KASP markers assigned from homoeologous groups 4/2/3/7. This partially matched the results by Danilova et al. (2017) and Niu et al. (2018) who found that chromosome G (4C) had markers from homoeologous groups 4/2/3. Our studies found one marker from homoeologous group 7 assigned to chromosome 4C of *Ae. caudata*. Chromosome A is absent from this set of disomic addition lines. We assigned KASP markers

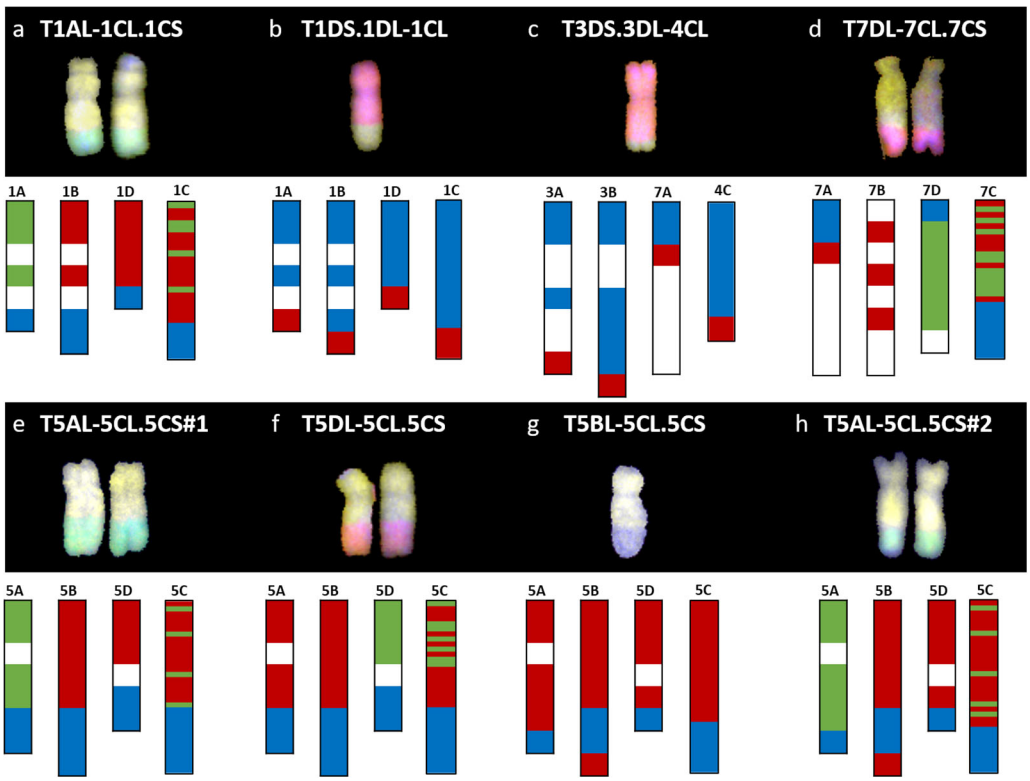

FIGURE 2 | Molecular characterization of wheat-*Ae. caudata* introgression lines using KASP markers and mcGISH. Top: the mcGISH image of the recombinant chromosome(s) and bottom: the genotyping data with KASP markers of 8 wheat-*Ae. caudata* introgression lines. **(a)** BC_4F_1-157-9, **(b)** BC_4-284-4, **(c)** BC_4F_1-164-2, **(d)** BC_3F_2-160-4, **(e)** BC_4F_1-157-2, **(f)** BC_4F_1-162-6, **(g)** BC3F1-401-8, and **(h)** BC_4F_1-158-12. C-genome segment is represented in yellow, A-genome in green, B-genome in blue, and D-genome in red. The genotyping is displayed with the markers distributed on the wheat chromosomes and the same set of markers ordered on the *Ae. caudata* chromosomes. White areas on the wheat chromosomes indicate the absence of KASP markers. Colored regions represent different genotype calls. Blue represents KASP markers homozygous for the wheat allele, red represents a heterozygous call and green represents a homozygous call for the *Ae. caudata* allele.

that were not assigned to chromosomes B-G to chromosome A and found that they were all homoeologous to the group 1 chromosomes of wheat (**Figure 4A**) which is consistent with previous reports that used FISH with cDNA probes and determined that chromosome A of *Ae. caudata* is homoeologous to wheat group 1 chromosomes (Danilova et al., 2014, 2017).

Generation of Wheat-*Ae. caudata* Recombinant Lines

It is to be noted that even though a considerable crossing program was undertaken, the eight wheat-*Ae. caudata* recombinant chromosomes were obtained from three BC_1 parents. In the inter-specific F_1 hybrids, which were haploid for the A, B, D, and C genomes, it is likely that the frequency of recombination between the chromosomes from different genomes would be low, in spite of the absence of the *Ph1* gene, leading to unviable gametes. This is exasperated by the genomic rearrangements in the C-genome possibly leading to inhibition of homoeologous chromosome pairing during meiosis. Hence, the cross fertility is lowest in the F_1 hybrids at 5% but increases substantially in the back-cross generations reaching 100% in the BC_4 population (**Table 2**). However, recombination events were first observed in the BC_1 plants, having occurred in the F_1 gametes, and no further homoeologous recombination was observed in subsequent back-cross generations.

Genotyping With Genome-Specific KASP Markers Allows Identification of Homozygous Recombinant Lines and Their Physical Mapping on *Ae. caudata* Chromosomes

Genotyping of the back-crossed population with the KASP markers showed that the markers were not only able to detect *Ae. caudata* in the wheat background but also indicated whether a line had a homozygous segment and which wheat chromosome had recombined with *Ae. caudata* in a homozygous line. Eight out of nine introgression lines had a single type of recombinant chromosome (**Table 3**). The ninth line was homozygous for two of these recombinant chromosomes. As shown in **Figure 2** (bottom), the KASP markers were able to detect whether a line was homozygous (red and green marker segments) or heterozygous (all red marker segments). A wheat chromosome that had green marker segments, representing

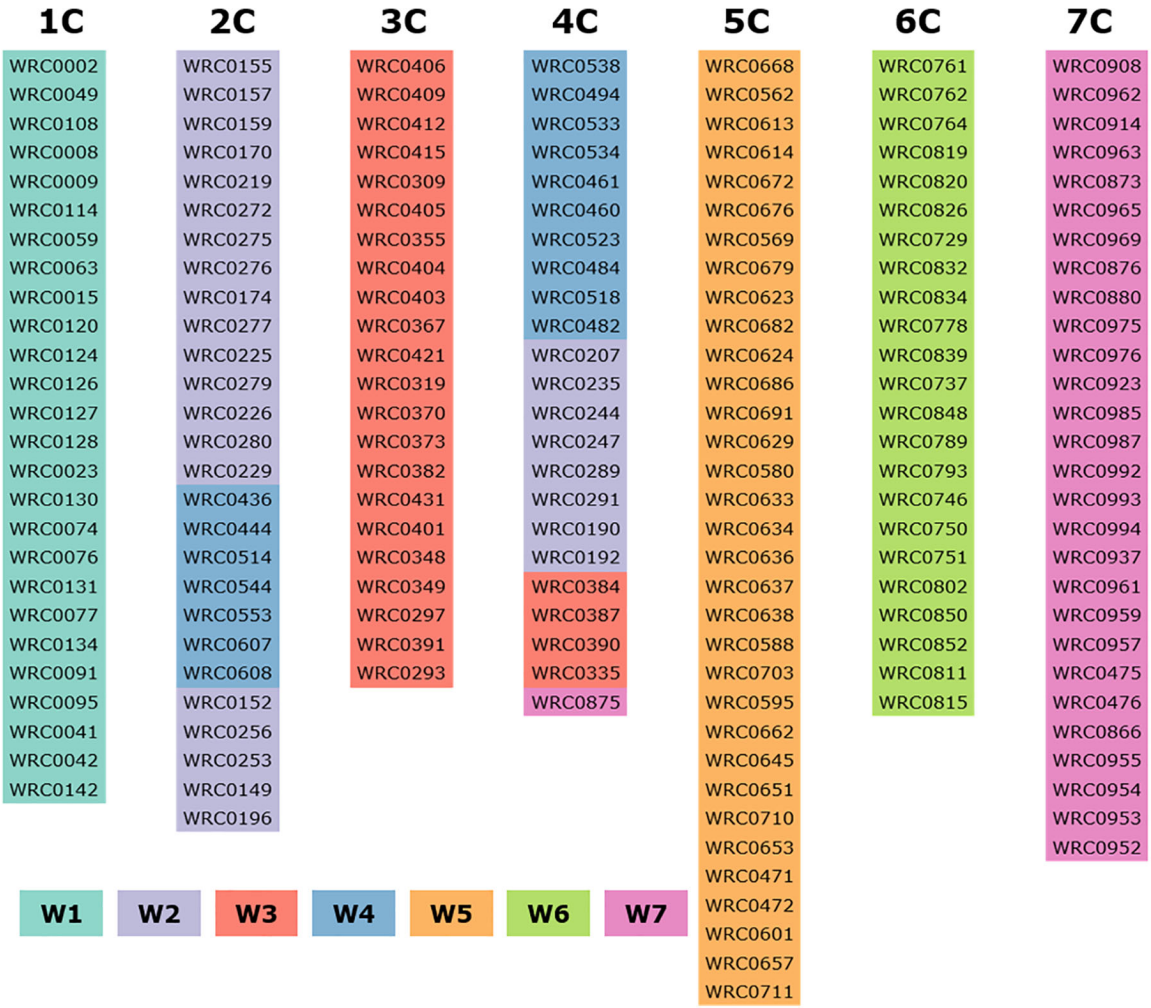

FIGURE 3 | Physical order of the 182 KASP markers assigned specifically to each *Ae. caudata* chromosome. W means wheat and W1–W7 represents wheat homoeologous group 1–7, respectively. Chromosomes 2C showing markers from homoeologous group 2/4 and chromosome 4C showing markers from homoeologous groups 4/2/3/7.

homozygous *Ae. caudata* alleles, was shown to be involved in the recombinant chromosome. The genotyping results were also validated by mcGISH analysis of the recombinant lines (**Figure 2**; top).

Genotyping of these recombinant lines helped to physically order some of the KASP markers along the *Ae. caudata* chromosomes (**Figure 3**). For example, the recombinant line with chromosome 7C showed that markers homoeologous to the distal end of the short arm of wheat homoeologous group 7 were potentially present on the distal end of the long arm of chromosome 7C in *Ae. caudata* (**Figure 4G**). This agrees with a previous report by Danilova et al. (2017). Taking this into account and the other marker assignments discussed above, we concluded that our results are most concurrent with the conclusions by Danilova et al. (2017). Thus, for the regions of the C-genome where no recombination was obtained, the markers were ordered according to Danilova et al. (2017) (**Figure 3**); and compared to the wheat genome (**Figure 4**) which showed all the potential inversions and translocations found in the C-genome of *Ae. caudata*.

Macrostructure of *Ae. caudata* Chromosomes

Chromosomes 1C, 5C, 6C, and 7C had maintained macro-synteny with wheat while chromosomes 1C and 5C were collinear with wheat and chromosomes 6C and 7C had inversions (**Figure 4**). In contrast, chromosomes 2C, 3C, and 4C had inter-chromosomal translocations and inversions (**Figure 4**). Due to the 4/5 translocation in wheat's A-genome, **Figure 4** showed markers corresponding to wheat chromosome 4A for *Ae. caudata* chromosome 5C (**Figure 4C**) and markers corresponding to wheat chromosome 5A for *Ae. caudata* chromosome 2C (that had a 2/4 translocation; **Figure 4E**). The macro-synteny and macro-collinearity with wheat could potentially be the reason we were able to obtain six recombinant chromosomes from *Ae. caudata*

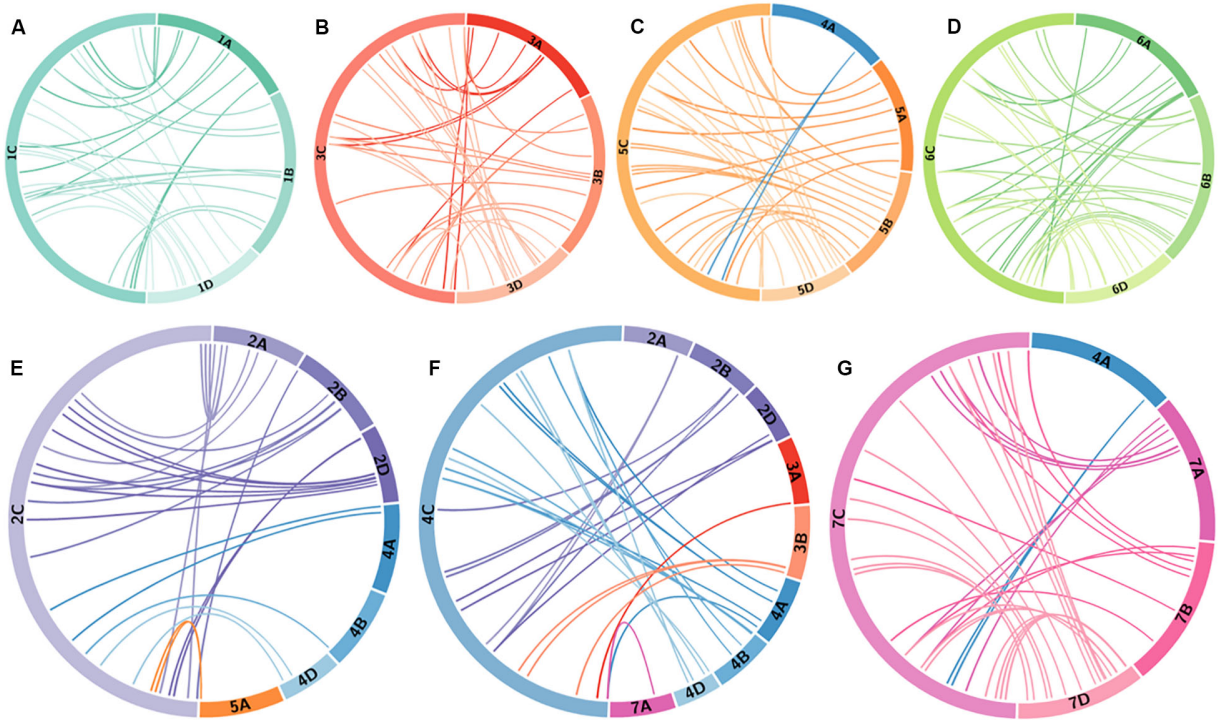

FIGURE 4 | Comparative analysis of *Ae. caudata* chromosome rearrangement compared to wheat chromosomes revealed by physically mapped KASP markers. Circos diagrams of *Ae. caudata* chromosome **(A)** 1C, **(B)** 3C, **(C)** 5C, **(D)** 6C, **(E)** 2C, **(F)** 4C, and **(G)** 7C. Corresponding markers are linked between their tentative physical order on the *Ae. caudata* C-genome chromosomes and their physical location (Mb) on the orthologous regions on the wheat genome.

chromosomes 1C and 5C (**Table 3** and **Figure 2**). Even though the majority of our results agreed with those of Danilova et al. (2017), we did not find any markers from wheat homoeologous group 7 translocated onto chromosome 6C (**Table 1** and **Figure 4F**). Our work also found one marker from wheat homoeologous group 7 assigned to chromosome 4C (not found by Danilova et al., 2017) along with three markers from wheat homoeologous 3. This may be due to differences in the *Ae. caudata* accessions used in the two studies.

In contrast to most diploid and allopolyploid Triticeae, that have preserved chromosome macrostructure, rye (*Secale cereale*), *Aegilops umbellulata* and *Ae. caudata* genomes are highly rearranged (Devos et al., 1993; Zhang et al., 1998; Danilova et al., 2017). As a consequence of evolution, chromosome rearrangements in the genomes of wild relative species can disrupt the collinearity between the wild relative and the homoeologous wheat chromosomes (Devos et al., 1993; Zhang et al., 1998). This can result in reduction or absence of meiotic chromosome pairing. However, the recombinant chromosome, T3DS.3DL.4CL, along with a recombinant from chromosome 7C, T7DL-7CL.7CS (which has an inversion from the short arm onto the long arm compared to wheat) developed in this work show that it is possible to get wheat-wild relative recombination from wild relative chromosomes that are highly rearranged with respect to wheat.

Genomic rearrangements can be caused in Triticeae interspecific crosses due to gametocidal (Gc) chromosomes, also known as cuckoo chromosomes, which cause chromosome breakage, in the gametes that lack them (Endo, 1978; Finch et al., 1984; Nasuda et al., 1998; Friebe et al., 2000). Gc chromosomes have been found in many Aegilops species (Endo and Katayama, 1978; Endo, 1982, 2007) and in C-genome-derived allotetraploids where Gc action was found on chromosome 2C of *Ae. cylindrica* (Endo, 1996) and on chromosome 3C of *Ae. truncialis* (Endo and Tsunewaki, 1975). Gc action can range from lethal to semi-lethal depending on the wheat genotypes in which the Gc chromosomes are present (Endo, 1990). Some Gc chromosomes, such as $4S^{sh}$ from *Aegilops sharonensis*, can cause complete sterility in gametes that lack them in Chinese Spring wheat and are thus, preferentially transmitted to the offspring (Miller et al., 1982; Grewal et al., 2017). The Gc action of chromosome 2C is semi-lethal in Chinese Spring wheat, allowing minor structural rearrangements in the wheat chromosomes to be retained. However, chromosome 3C has a severe Gc action in Chinese Spring but mild or semi-lethal in other cultivars resulting in only slight chromosome damage. This suggests that there might be an inhibitor gene in those cultivars that prevents the Gc chromosomes from being preferentially transmitted (Endo and Gill, 1996). Our work does not show any preferential transmission of chromosome 2C or 3C in the back-crossed population (data not shown) suggesting a possible mild Gc action or the complete absence of it. In this work, Paragon wheat was used as the wheat background for the *Ae. caudata* introgressions. Thus, apart from the *Ph1* mutation in the F_1 interspecific

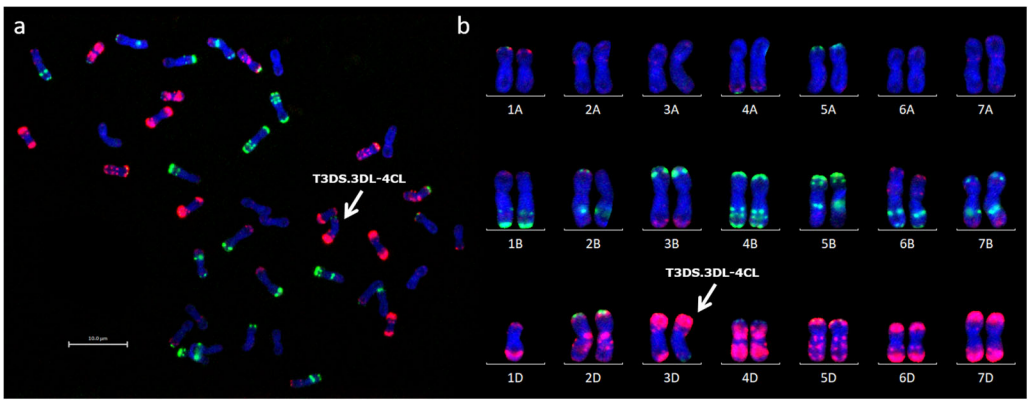

FIGURE 5 | FISH analysis of wheat-*Ae. caudata* introgression line carrying recombinant chromosome T3DS. 3DL-4CL. **(a)** FISH image of a root-tip metaphase spread of chromosomes from line BC_4F_1-164-2. Arrow indicates recombinant chromosome T3DS. 3DL-4CL. **(b)** FISH karyogram of the same spread from line BC_4F_1-164-2 showing 41 chromosomes including one copy of recombinant chromosome T3DS. 3DL-4CL and one copy of chromosome 1D. Sites of hybridization are shown with fluorescence-labeled probes, pSc119.2 (green) and pAs.1 (red). Wheat chromosomes were assigned to the karyogram according to the FISH karyotype established by Tang et al. (2014).

hybrids, a mild Gc action from C-genome chromosomes could potentially be responsible for the translocations between wheat and *Ae. caudata* chromosomes shown in this work. However, the results indicate that all recombinant chromosomes were obtained due to homoeologous recombination during meiosis. Six out of the eight recombinant chromosomes obtained in this work were between homoeologous linkage groups, i.e., chromosomes 1C of *Ae. caudata* with chromosomes 1A and 1D of wheat (**Table 3**) whereas any Gc action would have potentially resulted in random translocations. Confirmation of Gc action would need further cytological investigation of the gametophytes but it is also possible that the *Ae. caudata* accession used in this study does not possess any Gc action.

Potential Benefits for Breeding Increased Disease Resistance

To exploit the potential of the C-genome, *Ae. caudata*-derived wheat introgression lines were previously developed and characterized, primarily for disease resistance genes. Dyck et al. (1990) found stem rust resistance in crosses between *Ae. caudata* and wheat monosomic 5B lines. Another study mapped *Ae. caudata*-derived leaf rust resistance on chromosome 2A of wheat (Iqbal et al., 2007) while two powdery mildew resistant quantitative trait loci (QTL) were mapped onto chromosomes 1A and 7A of wheat (Weidner et al., 2012). Riar et al. (2012) and Toor et al. (2016) mapped leaf rust and stripe rust resistance, respectively, onto chromosome 5DS of wheat in one wheat-*Ae. caudata* introgression line. Here, we report the generation of nine wheat-*Ae. caudata* introgression lines (**Table 3**), of which six had stable homozygous recombinant wheat-*Ae. caudata* chromosomes (**Figure 1**). These lines will be made available for further trait characterization studies.

Previous work investigating wheat-*Ae. caudata* introgression lines did not attribute the disease resistance QTLs they found in these lines to a specific *Ae. caudata* chromosome. Gong et al. (2017) and Niu et al. (2018) evaluated the Alcedo-*Ae. caudata* addition lines for agronomic traits such as resistance to rust and powdery mildew diseases. Addition lines D (6C), F (3C), and G (4C) showed some resistance to powdery mildew isolates (Niu et al., 2018) while addition line E (7C) showed a higher resistance to powdery mildew (Gong et al., 2017; Niu et al., 2018). Gong et al. (2017) suggested chromosome D (6C) carried leaf rust resistance while Niu et al. (2018), suggested that chromosome B (2C) exhibited resistance to leaf rust which was comparable to the *Ae. caudata* parent. While Gong et al. (2017) suggested that chromosome E (7C) might carry some stem rust resistance, Niu et al. (2018) found no stem rust resistance in the addition lines. In contrast, Xu et al. (2009) had shown that addition lines C (5C) and D (6C) were resistant to the Ug99 race group of the stem rust pathogen (*Puccinia graminis* f. sp. *tritici*). Consequently, a new gene for Ug99 resistance was introgressed from chromosome D (6C) into wheat (Xu et al., 2017). Investigations into grain quality traits showed that chromosome E (7C) or F (3C) could increase protein and wet gluten content when introduced into wheat (Gong et al., 2017). In this study, we have produced four wheat-*Ae. caudata* recombinant lines with wheat from chromosome C (5C) and one from chromosome E (7C) (**Table 3**), among others, which could potentially carry disease resistance and grain quality QTLs.

CONCLUSION

Wheat-wild relative introgressions play an important role in wheat improvement and have been particularly exploited for disease resistance. However, a lack of compensation of the wheat genes by the wild relative chromosome segments and negative linkage drag can have a detrimental effect on agronomic performance of the wheat–wild relative introgression lines (Sears and Gustafson, 1993; Friebe et al., 1996). Therefore, better knowledge of the genome organization of the wild relative species is important before starting a new pre-breeding programs which

can be laborious and costly. In this work, we produced KASP markers that proved to be valuable tools to detect *Ae. caudata* chromatin in a wheat background. In addition, these KASP markers have shed further light on the structural rearrangements present in *Ae. caudata*. We have also produced nine wheat-*Ae. caudata* recombinant lines of which six have stable homozygous introgressions that can be used for further trait analysis studies.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

SG, MO, JW, CY, SH-E, DS, SA, IK, and JK carried out the crossing program. PI set up the high-throughput genotyping platform and developed the protocol. MO and JW performed the *in situ* hybridization experiments and the genotyping of wheat-*Ae. caudata* lines. SG analyzed the genotyping data, assigned the markers to the *Ae. caudata* chromosomes, and performed the comparative studies. SG, MO, IK, and JK conceived and designed the experiments. SG wrote the manuscript with assistance from MO, JK, and IK. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Biotechnology and Biological Sciences Research Council (Grant No. BB/P016855/1) as part of the Developing Future Wheat (DFW) programme. The funding body played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

ACKNOWLEDGMENTS

This manuscript has been released as a pre-print at bioRxiv (Grewal et al., 2020b).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00606/full#supplementary-material

TABLE S1 | Primer details for KASP assays.

TABLE S2 | Physical location and allelic information of SNPs in the KASP assays.

REFERENCES

Baldauf, F., Schubert, V., and Metzlaff, M. (1992). Repeated DNA sequences of *Aegilops markgrafii* (Greuter) Hammer var. markgrafii'. Cloning, sequencing and analysis of distribution in Poaceae species. *Hereditas* 116, 71–78. doi: 10.1111/j.1601-5223.1992.tb00802.x

Barashkova, E. A., and Migushova, E. F. (1984). Genotypic differences between *Aegilops* in resistance to low temperature. *Sel' skokhozyaistevennaya Biol.* 997, 41–46. doi: 10.1007/s00425-010-1184-z

Danilova, T. V., Akhunova, A. R., Akhunov, E. D., Friebe, B., and Gill, B. S. (2017). Major structural genomic alterations can be associated with hybrid speciation in *Aegilops markgrafii* (Triticeae). *Plant J.* 92, 317–330. doi: 10.1111/tpj.13657

Danilova, T. V., Friebe, B., and Gill, B. S. (2014). Development of a wheat single gene FISH map for analyzing homoeologous relationship and chromosomal rearrangements within the *Triticeae*. *Theor. Appl. Genet.* 127, 715–730. doi: 10.1007/s00122-013-2253-z

Devos, K. M., Atkinson, M. D., Chinoy, C. N., Francis, H. A., Harcourt, R. L., Koebner, R. M., et al. (1993). Chromosomal rearrangements in the rye genome relative to that of wheat. *Theor. Appl. Genet. Theor. Angew. Genet.* 85, 673–680. doi: 10.1007/bf00225004

Dyck, P. L., Kerber, E. R., and Martens, J. W. (1990). Transfer of a gene for stem rust resistance from *Aegilops caudata* to common wheat. *Can. J. Plant Sci.* 70, 931–934. doi: 10.4141/cjps90-114

Endo, T. R. (1978). "On the *Aegilops* chromosomes having gametocidal action on common wheat," in *Proceedings of the 5th International Wheat Genetics Symposium* (New Delhi: Indian Agricultural Research Institute), 306–314. doi: 10.1266/ggs.85.43

Endo, T. R. (1982). Gametocidal chromosomes of three *Aegilops* species in common wheat. *Can. J. Genet. Cytol.* 24, 201–206. doi: 10.1139/g82-020

Endo, T. R. (1990). Gametocidal chromosomes and their induction of chromosome mutations in wheat. *Jpn. J. Genet.* 65, 135–152. doi: 10.1266/jjg.65.135

Endo, T. R. (1996). Allocation of a gametocidal chromosome of *Aegilops* cylindrica to wheat homoeologous group 2. *Genes Genet. Syst.* 71, 243–246. doi: 10.1266/ggs.71.243

Endo, T. R. (2007). The gametocidal chromosome as a tool for chromosome manipulation in wheat. *Chromosome Res.* 15, 67–75. doi: 10.1007/s10577-006-1100-3

Endo, T. R., and Gill, B. S. (1996). The deletion stocks of common wheat. *J. Hered.* 87, 295–307. doi: 10.1093/oxfordjournals.jhered.a023003

Endo, T. R., and Katayama, Y. (1978). Finding of a selectively retained chromosome of *Aegilops caudata* L. in common wheat. *Wheat Inf. Serv.* 47-48, 32–35.

Endo, T. R., and Tsunewaki, K. (1975). Sterility of common wheat with *Aegilops triuncialis* cytoplasm. *J. Hered.* 66, 13–18.

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. *Trends Genet.* 24, 24–32. doi: 10.1016/j.tig.2007.11.001

Finch, R. A., Miller, T. E., and Bennett, M. D. (1984). "Cuckoo" *Aegilops* addition chromosome in wheat ensures its transmission by causing chromosome breaks in meiospores lacking it. *Chromosoma* 90, 84–88. doi: 10.1007/bf00352282

Friebe, B., Jiang, J., Raupp, W. J., McIntosh, R. A., and Gill, B. S. (1996). Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91, 59–87. doi: 10.1007/BF00035277

Friebe, B., Kynast, R. G., and Gill, B. S. (2000). Gametocidal factor-induced structural rearrangements in rye chromosomes added to common wheat. *Chromosome Res.* 8, 501–511. doi: 10.1023/a:1009219722418

Friebe, B., Schubert, V., Blüthner, W. D., and Hammer, K. (1992). C-banding pattern and polymorphism of *Aegilops caudata* and chromosomal constitutions of the amphiploid *T. aestivum* — Ae. caudata and six derived chromosome addition lines. *Theor. Appl. Genet.* 83, 589–596. doi: 10.1007/BF00226902

Gill, B. S., Sharma, H. C., Raupp, W. J., Browder, L. E., Hatchett, J. H., Harvey, T. L., et al. (1985). Evaluation of *Aegilops* species for resistance to wheat powdery mildew, wheat leaf rust, hessian fly, and greenbug. *Plant Dis.* 69, 314–316.

Gong, W., Han, R., Li, H., Song, J., Yan, H., Li, G., et al. (2017). Agronomic traits and molecular marker identification of wheat–*Aegilops caudata* addition lines. *Front. Plant Sci.* 8:1743. doi: 10.3389/fpls.2017.01743

Gorham, J. (1990). Salt tolerance in the triticeae: K/Na discrimination in *Aegilops* species. *J. Exp. Bot.* 41, 615–621. doi: 10.1093/jxb/41.5.615

Grewal, S., Gardiner, L. J., Ndreca, B., Knight, E., Moore, G., King, I. P., et al. (2017). Comparative mapping and targeted-capture sequencing of the gametocidal loci

in *Aegilops sharonensis*. *Plant Genome* 10, 1–16. doi: 10.3835/plantgenome2016.09.0090

Grewal, S., Hubbart-Edwards, S., Yang, C., Devi, U., Baker, L., Heath, J., et al. (2020a). Rapid identification of homozygosity and site of wild relative introgressions in wheat through chromosome-specific KASP genotyping assays. *Plant Biotechnol. J.* 18, 743–755. doi: 10.1111/pbi.13241

Grewal, S., Hubbart-Edwards, S., Yang, C., Scholefield, D., Ashling, S., Burridge, A., et al. (2018). Detection of *T. urartu* introgressions in wheat and development of a panel of interspecific introgression lines. *Front. Plant Sci.* 9:1565. doi: 10.3389/fpls.2018.01565

Grewal, S., Othmeni, M., Walker, J., Edwards, S. H., Yang, C.-Y., Scholefield, D., et al. (2020b). Development of wheat-*Aegilops caudata* introgression lines and their characterisation using genome-specific KASP markers. *bioRxiv* [preprint]. doi: 10.1101/2020.03.03.975326

Iqbal, N., Eticha, F., Khlestkina, E. K., Weidner, A., Röder, M. S., and Börner, A. (2007). The use of simple sequence repeat (SSR) markers to identify and map alien segments carrying genes for effective resistance to leaf rust in bread wheat. *Plant Genet. Resour.* 5, 100–103. doi: 10.1017/S1479262107672311

Iriki, N., Kawakami, A., Takata, K., Kuwabara, T., and Ban, T. (2001). Screening relatives of wheat for snow mold resistance and freezing tolerance. *Euphytica* 122, 335–341.

International Wheat Genome Sequencing Consortium [IWGSC], Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

King, J., Grewal, S., Yang, C. Y., Hubbart, S., Scholefield, D., Ashling, S., et al. (2017). A step change in the transfer of interspecific variation into wheat from *Ambylopyrum muticum*. *Plant Biotechnol. J.* 15, 217–226. doi: 10.1111/pbi.12606

King, J., Grewal, S., Yang, C.-Y., Hubbart Edwards, S., Scholefield, D., Ashling, S., et al. (2018). Introgression of *Aegilops speltoides* segments in *Triticum aestivum* and the effect of the gametocidal genes. *Ann. Bot.* 121, 229–240. doi: 10.1093/aob/mcx149

Makkouk, K. M., Ghulam, W., and Comeau, A. (1994). Resistance to barley yellow dwarf luteovirus in *Aegilops* species. *Can. J. Plant Sci.* 74, 631–634. doi: 10.4141/cjps94-113

McIntyre, C. L., Pereira, S., Moran, L. B., and Appels, R. (1990). New secale cereale (rye) DNA derivatives for the detection of rye chromosome segments in wheat. *Genome* 33, 635–640. doi: 10.1139/g90-094

Miller, T. E., Hutchinson, J., and Chapman, V. (1982). Investigation of a preferentially transmitted *Aegilops sharonensis* chromosome in wheat. *Theor. Appl. Genet.* 61, 27–33. doi: 10.1007/bf00261506

Molnár, I., Vrána, J., Burešová, V., Cápal, P., Farkas, A., Darkó, É, et al. (2016). Dissecting the U, M, S and C genomes of wild relatives of bread wheat (*Aegilops* spp.) into chromosomes and exploring their synteny with wheat. *Plant J.* 88, 452–467. doi: 10.1111/tpj.13266

Molnár, I., Vrána, J., Farkas, A., Kubaláková, M., Cseh, A., Molnár-Láng, M., et al. (2015). Flow sorting of C-genome chromosomes from wild relatives of wheat *Aegilops markgrafii*, *Ae. triuncialis* and *Ae. cylindrica*, and their molecular organization. *Ann. Bot.* 116, 189–200. doi: 10.1093/aob/mcv073

Nasuda, S., Friebe, B., and Gill, B. S. (1998). Gametocidal genes induce chromosome breakage in the interphase prior to the first mitotic cell division of the male gametophyte in wheat. *Genetics* 149, 1115–1124.

Niu, Z., Chao, S., Cai, X., Whetten, R. B., Breiland, M., Cowger, C., et al. (2018). Molecular and cytogenetic characterization of six wheat-*Aegilops markgrafii* disomic addition lines and their resistance to rusts and powdery mildew. *Front. Plant Sci.* 9:1616. doi: 10.3389/fpls.2018.01616

Ohta, S., and Yasukawa, N. (2015). "Genetic variation and its geographical distribution in *Aegilops caudata* L.: morphology, hybrid sterility and gametocidal genes," in *Advances in Wheat Genetics: From Genome to Field*, eds Y. Ogihara, S. Takumi, and H. Handa (Tokyo: Springer), 53–61.

Othmeni, M., Grewal, S., Hubbart-Edwards, S., Yang, C., Scholefield, D., Ashling, S., et al. (2019). The use of pentaploid crosses for the introgression of *Amblyopyrum muticum* and D-genome chromosome segments into durum wheat. *Front. Plant Sci.* 10:1110. doi: 10.3389/fpls.2019.01110

Peil, A., Korzun, V., Schubert, V., Schumann, E., Weber, W. E., and Röder, M. S. (1998). The application of wheat microsatellites to identify disomic *Triticum aestivum-Aegilops markgrafii* addition lines. *Theor. Appl. Genet.* 96, 138–146. doi: 10.1007/s001220050720

Rayburn, A. L., and Gill, B. S. (1986). Molecular identification of the D-genome chromosomes of wheat. *J. Hered.* 77, 253–255.

Riar, A. K., Kaur, S., Dhaliwal, H. S., Singh, K., and Chhuneja, P. (2012). Introgression of a leaf rust resistance gene from *Aegilops caudata* to bread wheat. *J. Genet.* 91, 155–161. doi: 10.1007/s12041-012-0161-7

Schmidt, J. C., Schubert, V., and Blüthner, W. D. (1993). Use of isozymes to characterize *Triticum aestivum — Aegilops markgrafii* addition lines. *Biochem. Physiol. Pflanzen* 188, 385–392. doi: 10.1016/S0015-3796(11)80141-1

Schubert, V., and Blüthner, W. D. (1992). Zerlegung des genoms von *Aegilops markgrafii* mit Hilfe von chromosomalen Additions-linien. *Kühn Archiv.* 86, 38–46.

Schubert, V., and Blüthner, W. D. (1995). "*Triticum aestivum-Aegilops markgrafii* addition lines: production and morphology," in *Proceedings of the 8th International Wheat Genetic Symposium* (Beijing: China Agricultural Scientech Press), 421–425.

Sears, E. R., and Gustafson, J. P. (1993). Use of radiation to transfer alien chromosome segments to wheat. *Crop Science* 33, 897–901. doi: 10.2135/cropsci1993.0011183X003300050004x

Tang, Z., Yang, Z., and Fu, S. (2014). Oligonucleotides replacing the roles of repetitive sequences pAs1, pSc119.2, pTa-535, pTa71, CCS1, and pAWRC.1 for FISH analysis. *J. Appl. Genet.* 55, 313–318. doi: 10.1007/s13353-014-0215-z

Thomson, D., and Henry, R. (1995). Single-step protocol for preparation of plant tissue for analysis by PCR. *BioTechniques* 19, 394–397.

Toor, P. I., Kaur, S., Bansal, M., Yadav, B., and Chhuneja, P. (2016). Mapping of stripe rust resistance gene in an *Aegilops caudata* introgression line in wheat and its genetic association with leaf rust resistance. *J. Genet.* 95, 933–938. doi: 10.1007/s12041-016-0718-y

Valkoun, J., Hammer, K., Kučerová, D., and Bartoš, P. (1985). Disease resistance in the genus *Aegilops* L. — stem rust, leaf rust, stripe rust, and powdery mildew. *Die Kulturpflanze* 33, 133–153. doi: 10.1007/bf01997267

Wang, S., Yin, L., Tanaka, H., Tanaka, K., and Tsujimoto, H. (2011). Wheat-*Aegilops* chromosome addition lines showing high iron and zinc contents in grains. *Breed. Sci.* 61, 189–195. doi: 10.1270/jsbbs.61.189

Weidner, A., Röder, M. S., and Börner, A. (2012). Mapping wheat powdery mildew resistance derived from *Aegilops markgrafii*. *Plant Genet. Resour.* 10, 137–140. doi: 10.1017/S1479262112000123

Xu, S. S., Jin, Y., Klindworth, D. L., Wang, R. R. C., and Cai, X. (2009). Evaluation and characterization of seedling resistances to stem rust ug99 races in wheat–alien species derivatives. *Crop Sci.* 49, 2167–2175. doi: 10.2135/cropsci2009.02.0074

Xu, S. S., Klindworth, D. L., Niu, Z., Zhang, Q., Chao, S., Friesen, T., et al. (2017). "Introgression of a new stem rust resistance gene from *Aegilops markgrafii* into wheat," in *Proceedings of 13th International Wheat Genetics Symposium*, eds H. Buertmayr, C. Lang-Mladek, B. Steiner, S. Michel, M. Buertmayr, and M. Lemmens (Vienna: University of Natural Resources and Life Sciences), 38–39.

Zhang, H., Jia, J., Gale, M. D., and Devos, K. M. (1998). Relationships between the chromosomes of *Aegilops umbellulata* and wheat. *Theor. Appl. Genet.* 96, 69–75. doi: 10.1007/s001220050710

Conflict of Interest: PI was employed by the company IDna Genetics Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 June 2020
doi: 10.3389/fpls.2020.00699

Introgression of Two Broad-Spectrum Late Blight Resistance Genes, *Rpi-Blb1* and *Rpi-Blb3*, From *Solanum bulbocastanum* Dun Plus Race-Specific *R* Genes Into Potato Pre-breeding Lines

Elena Rakosy-Tican[1], Ramona Thieme[2]*, Janine König[3], Marion Nachtigall[2], Thilo Hammann[2], Tunde-Eva Denes[1,4], Klaudia Kruppa[5] and Márta Molnár-Láng[5]*

[1] Plant Genetic Engineering Group, Department of Molecular Biology and Biotechnology, Babeş-Bolyai University, Cluj-Napoca, Romania, [2] Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Agricultural Crops, Julius Kühn-Institut, Quedlinburg, Germany, [3] Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Horticultural Crops, Julius Kühn-Institut, Quedlinburg, Germany, [4] Biological Research Centre, Jibou, Romania, [5] Centre for Agricultural Research, Hungarian Academy of Sciences, Martonvásár, Hungary

There is a wealth of resistance genes in the Mexican wild relative of cultivated *Solanum*, but very few of these species are sexually compatible with cultivated *Solanum tuberosum.* The most devastating disease of potato is late blight caused by the oomycete *Phytophthora infestans* (*Pi*). The wild hexaploid species *S. demissum,* which it is able to cross with potato, was used to transfer eleven race-specific genes by introgressive hybridization that were subsequently widely used in potato breeding. However, there are now more virulent races of *Pi* that can overcome all of these genes. The most sustainable strategy for protecting potatoes from late blight is to pyramid or stack broad-spectrum resistance genes into the cultivars. Recently four broad-spectrum genes (*Rpi*) conferring resistance to *Pi* were identified and cloned from the sexually incompatible species *S. bulbocastanum*: *Rpi-blb1* (RB), *Rpi-blb2*, *Rpi-blb3*, and *Rpi-bt1*. For this research, a resistant *S. bulbocastanum* accession was selected carrying the genes *Rpi-blb1* and *Rpi-blb3* together with race-specific *R3a* and *R3b* genes. This accession was previously used to produce a large number of somatic hybrids (SHs) with five commercial potato cultivars using protoplast electrofusion. In this study, three SHs with cv. 'Delikat' were selected and backcross generations (i.e., BC_1 and BC_2) were obtained using cvs. 'Baltica', 'Quarta', 'Romanze', and 'Sarpo Mira'. Their assessment using gene-specific markers demonstrates that these genes are present in the SHs and their BC progenies. We identified plants carrying all four genes that were resistant to foliage blight in greenhouse and field trials. Functionality of the genes was shown by using agro-infiltration with the effectors of corresponding *Avr* genes. For a number of hybrids and BC clones yield and tuber number were not significantly different from that

OPEN ACCESS

Edited by:
Mariola Plazas, Polytechnic University of Valencia, Spain

Reviewed by:
Hannele Lindqvist-Kreuze, International Potato Center, Peru
Weixing Shan, Northwest A&F University, China
Iga Tomczynska, Oregon State University, United States

***Correspondence:**
Elena Rakosy-Tican
elena.rakosy@ubbcluj.ro;
rakosy.elena@gmail.com
Ramona Thieme
ramona.thieme@julius-kuehn.de

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *27 September 2019*
Accepted: *04 May 2020*
Published: *03 June 2020*

Citation:
Rakosy-Tican E, Thieme R, König J, Nachtigall M, Hammann T, Denes T-E, Kruppa K and Molnár-Láng M (2020) Introgression of Two Broad-Spectrum Late Blight Resistance Genes, Rpi-Blb1 and Rpi-Blb3, From Solanum bulbocastanum Dun Plus Race-Specific R Genes Into Potato Pre-breeding Lines. Front. Plant Sci. 11:699. doi: 10.3389/fpls.2020.00699

of the parent cultivar 'Delikat' in field trials. The evaluation of agronomic traits of selected BC_2 clones and of their processing qualities revealed valuable material for breeding late blight durable resistant potato. We show that the combination of somatic hybridization with the additional use of gene specific markers and corresponding *Avr* effectors is an efficient approach for the successful identification and introgression of late blight resistance genes into the potato gene pool.

Keywords: effectors, detached leaf assay, introgression breeding, late blight field resistance, gene specific markers, processing qualities, tuber yield

INTRODUCTION

Late blight caused by *Phytophthora infestans* (Mont.) de Bary (*Pi*) is one of the most damaging diseases of potatoes as it does not only affect the above ground parts of the plants but also the tubers (Colton et al., 2006). Although the *Pi* genome has been sequenced (Haas et al., 2009), a deep understanding of the complex interaction and co-evolution of this pathogen and its host is lacking (Kamoun and Smart, 2005; Haverkort et al., 2009; Aguilera-Galvez et al., 2018). However, its potential in causing an epidemic coupled with its great adaptability led to many outbreaks and destructions of potato harvests during the course of human history (Vreugdenhil, 2007; Chakrabarti et al., 2017). Even today, the cost of control and lost production are estimated six billion euros per year, with half of this cost incurred in Europe (Haverkort et al., 2016). Current disease control relies on ten up to 16 fungicide applications per season (Hanson et al., 2007; Haverkort et al., 2009). Fungicides are not only costly but they pollute the environment and affect human health (Paro et al., 2012). Moreover, late blight resistance introduced into potato cultivars by classical breeding, based on eleven race-specific resistance genes from *S. demissum* was quickly overcome by new races of *Pi*.

Solanum bulbocastanum (*blb*), a wild diploid species of Mexican potato ($2n = 2x = 24$), is highly resistant to all known races of *Pi*, even when exposed to high concentrations of its spores. This species is a typical example of a 1EBN (endosperm balance number) species and hence cannot be conventionally crossed with cultivated potato (Johnston et al., 1980). To date, four of its Nucleotide Binding Sites–Leucine Rich Repeats (NBS-LRR) resistance genes have been identified and cloned: *Rpi-blb1* (Van der Vossen et al., 2003), also known as *RB* and *Rpi-bt1*–both loci located on chromosome VIII (Song et al., 2003; Oosumi et al., 2009), *Rpi-blb2*–on chromosome VI (Van der Vossen et al., 2005) and *Rpi-blb3* on chromosome IV (Lokossou et al., 2009). Nevertheless, it is likely there are other factors involved in its late blight resistance (Hein et al., 2009; Lokossou et al., 2010; Rietman, 2011). Horizontal or durable resistance is thought to be polygenic and hence a quantitative trait, with complex pathogen and host interaction mechanisms that are still poorly understood (Colton et al., 2006). For instance, it was suggested that the pathogenesis related protein StPRp27 contributes to a race-nonspecific resistance against *Pi* by inhibiting the development of the disease and could potentially be used in breeding for durable resistance to late blight (Shi et al., 2012).

Many of the *Rpi* genes are located in meta-QTLs, like *Rpi-blb2* on chromosome VI and *Rpi-blb3* on chromosome IV (Danan et al., 2011). The *Rpi* genes from *S. bulbocastanum* are considered to be broad spectrum resistance genes, as already shown for *Rpi-blb1* (Song et al., 2003), *Rpi-blb2* (Van der Vossen et al., 2005; Orbegozo et al., 2016), and *Rpi-blb3* (Zhu et al., 2012). The race-specific genes used in this study, originally identified in *S. demissum,* are *R3a* and *R3b,* which are members of the *R3* complex locus on chromosome XI. The *R2* gene recognizes the same effector as *Rpi-blb3*, coded by *Avr2* gene (Aguilera-Galvez et al., 2018). Huang et al. (2004) showed that the *R3* locus harbours two functionally distinct genes: *R3a* and *R3b*. The *R3a* gene is classified in the fast evolving class I (Huang et al., 2005). The *R3* locus in potato is an example of natural stacking of resistance genes. It contains the two closely linked *R3a* and *R3b* genes, which are 0.4 cM apart, with distinctly different resistances to particular types of *Pi*, *R3a* recognizes *Avr3a*, while *R3b* despite its sequence being similar to that of *R3a*, recognizes another effector encoded by the *Avr3b* gene (Li et al., 2011).

It has been shown that by combining broad-spectrum resistance genes by means of genetic engineering (cisgenesis), two or more gene stacks can confer durable late blight resistance in the field (Zhu et al., 2012; Haesaert et al., 2015). Unfortunately, genetic engineering, which uses cloned resistance genes from wild species of potato, is still not considered as non-genetically modified organism (GMO) and its use is restricted in Europe (directive 2001/18/EC), and globally by skeptical consumers. Despite considerable progress in the genetic analysis of quantitative resistance to late blight, under long day conditions, based on molecular markers (Gebhardt and Valkonen, 2001; Simko, 2002), breeders have made little progress in breeding resistant cultivars using marker-assisted selection (MAS). The major drawbacks are the tetrasomic inheritance of potato and the strong linkage between foliage resistance and late maturity (Bradshaw et al., 2004). On the other hand, for more than 100 years there was little progress in obtaining resistant cultivars by introducing single genes from *S. demissum* and conventional breeding using bridging species is very time consuming. This was the case for the commercial potato varieties 'Bionica' and 'Toluca' (Haverkort et al., 2009) in which the single broad-spectrum resistance gene *Rpi-blb2* was incorporated. It took more than 46 years of conventional breeding in Netherlands that started at the end of the fifties and used the late blight resistant *blb* as the source of the resistance and *S. acaule* as the bridging species (Hermsen and Ramanna, 1973).

Finding methods of deploying major resistance genes to *Pi* remains an important goal of potato breeding. Currently breeders isolate variants of *R* genes and deploy them in pyramids or stacks. It is expected that this will result in broad-spectrum recognition of *Pi* isolates and might provide a more durable resistance in the field (Jo et al., 2016). Recent screening of germplasm has revealed *Rpi* genes in wild species of *Solanum*, particularly broad-spectrum resistance genes in *blb* species (Wang et al., 2008, 2013; Hein et al., 2009). The current challenge is to select, judiciously combine and deploy sets of different resistance genes that confer durable late blight resistance on modern potato cultivars.

Previously, a large number of somatic hybrids (SHs) between potato cultivars and *S. bulbocastanum* were regenerated and analyzed for hybridity, ploidy and fertility (Rakosy-Tican et al., 2015). The goal of this study was to transfer broad-spectrum resistance genes *Rpi-blb1* and *Rpi-blb3* from *S. bulbocastanum* together with race-specific *R* genes (*R3a* and *R3b*), into the potato gene pool by somatic hybridization. A functional profiling of resistance genes by effectoromics, phenotypic expression of resistance in detached leaf assay and quantitative resistance assessment in the field were used to evaluate the resistance of the somatic hybrids (SHs) and derived back-cross progenies (BC_1 and BC_2). We tested the hypothesis if the four resistance genes introgressed by somatic hybridization present in the back-cross progenies confer resistance that can be confirmed in a greenhouse and field trials. Thus, the pre-breeding lines produced will be a valuable material for breeding durable late blight resistance potato cultivars.

MATERIALS AND METHODS

Plant Material

The morphology and nuclear genetic constitution of the SHs and backcross (BC) progenies, BC_1 and BC_2, between five commercial potato cultivars ('Baltica', 'Delikat', 'Quarta', 'Romanze', and 'Sarpo Mira') and *S. bulbocastanum* accession GLKS–31741 (*blb*41) (Groß Lüsewitz Potato Collections (GLKS) of the IPK Genebank, Leibniz-Institute of Plant Genetics and Crop Plant Research, Germany) have been analyzed (Rakosy-Tican et al., 2015). Only fertile SHs were chosen for further resistance assays. These included three hybrids with potato cv. 'Delikat' carrying all four resistance genes and their derived BC clones (see **Tables 1**–**3**).

Assessment of Resistance Genes

Screening for Rpi-blb Resistance Genes

Fusion parents, SHs and their BC progenies were screened for the resistance genes *Rpi-blb1* using specific markers (Wang et al., 2008), *Rpi-blb3* (Lokossou et al., 2009; Zhu et al., 2012), *R3a* (Huang et al., 2005), and *R3b* (Rietman, 2011). Amplifications were carried out in a 20 μL reaction mixture containing 1.5 mM $MgCl_2$, 200 μM of each dNTP, 0.5 μM primer, approximately 10 ng template DNA and 0.5 Unit InviTaq DNA Polymerase (Invitek). PCR was conducted using a standard set of conditions on a MJ Thermal cycler (PTC 200): initial denaturizing at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, $T_{annealing}$ for 45 s, 72°C for 45 s and a final extension step of 72°C for 5 min. Amplification products were size-separated on a 1.5% agarose gel stained with ethidium bromide and visualized using a UV transilluminator. The expected sizes of amplified fragments were 821 bp for *Rpi-blb1*–marker Blb1 F/R (Wang et al., 2008), and 757 bp for *Rpi-blb3*–marker Blb3 F/R (Lokossou et al., 2009). For the race-specific genes previously characterized in *S. demissum,* the sizes of amplified fragments were 982 bp for *R3a* (Huang et al., 2005) and 378 bp for *R3b* (Rietman, 2011).

Agro-Infiltration

Agro-infiltration was done using a culture of the *A. tumefaciens* strain AGL1 + pVirg transformed with a pK7WG2 plasmid containing the avirulence factors for: *Rpi-blb1* (*Avrblb1*), *Rpi-blb3* (*Avr2* - recognized also by *R2*), *R3a* and *R3b* (*Avr3a* and *Avr3b*), obtained from the University of Wageningen (V. Vleeshouwers). The PITG codes are: *Avr2* = PITG_22870; *Avr3a* = PITG_14371; *Avr3b* = PITG_18215; and *Avrblb1* = PITG_21388.

Cultivation and infiltration was performed as described by Du et al. (2014). The macroscopic response was scored 4 days after infiltration (dpi), when the dead cells (hypersensitive reaction, HR) were visible. As a negative control, the pK7WG2 empty vector was used (Vleeshouwers et al., 2008, 2011), which had the advantage that the cells killed by the infiltration (background necrosis) can be detected and hence, were not scored as a false positive. The parents were used as positive controls. The agro-infiltration and evaluation was done on three leaflets of one plant and the final appreciation of HR and 0 was based on three different plants of the same genotype.

Assessment of Ploidy and Genome Composition in SHs and Derived BCs

Determination of Ploidy Level by Classical Cytological Methods

Flow cytometry

The first pair of leaves of *in vitro* plants that were 8 weeks old were chopped with a razor blade in 800 μL of LB buffer, which was then filtered through a 25 μm mesh. Subsequently, 12.5 μL of buffer containing 1 μg/mg propidium iodide (PI) were added (Doležel et al., 2007). The nuclear DNA content was determined using a Becton Dickinson FacScan flow cytometer and the data analyzed using CellQuest software.

Determination of the ploidy level using cytological methods

Root tips of 1–2 cm in length were harvested from young plants grown *in vitro* on MS medium (Murashige and Skoog, 1962) containing 1 mg L^{-1} NAA, fixed in ethanol/acetic acid (3:1) and stored in 70% ethanol at 4°C. The root tips were washed twice in distilled water (10 min each), digested with a 1% Cellulase Onozuka R-10, 0.4% Cytohelicase and 0.4% Pectolyase in citrate buffer at pH = 4.8 (30 min) and squashed in 45% acetic acid. After the cover slips were removed in liquid nitrogen, slides were dehydrated by placing them sequentially in 70, 90, and finally 100% ethanol. Chromosomes were stained using 4′,6–diamidine-2′-phenylindolehydrochloride (DAPI) and washed with 2xSSC and distilled water. Probes were mounted in Vectashield (Vectashield, Vector Laboratories) to preserve the

TABLE 1 | Assessment of the somatic hybrid (SH) 82/4, between potato cv. 'Delikat' and *Solanum bulbocastanum* (*blb*41) and derived back-crossed progenies BC_1 and BC_2 (the cultivar used for crossing is indicated), for the presence of the resistance genes as revealed by gene specific markers, functional profiling of *Avr* genes by agro-infiltration, resistance to late blight in a detached leaf assay (DLA) and in the field (Δ-rAUDPC); standard varieties 'Adretta' (susceptible) and 'Sarpo Mira' (resistant to late blight) were used; the year of the field assays is given as a subscript to each value of Δ-rAUDPC; from each genotype three (*two) plants and three leaflets were infiltrated, hypersensitive reaction is indicated by HR and lack of reaction by 0: the numbers in brackets indicate the number of repetitions with the same reaction.

Parent/ SH/BC_1	Genotype ID	Resistance genes	Agro-infiltration with effectors:				DLA (mean ± SD)	*Phytophthora* attack Δ-rAUDPC (2009–2018)
			Avrblb1	*Avr2*	*Avr3a*	*Avr3b*		
Wild parent	*Blb*41	*Rpi-blb1, Rpi-blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.0 ± 0.00	-0.124_{13}; -0.405_{14}
Cultivar	Delikat	none	0(9)	0(9)	0(9)	0(9)	5.73 ± 1.17	0.311_{12}; 0.164_{15}; 0.477_{16}; 0.491_{18}
Cultivar	Quarta	*R3a, R3b*	0(9)	0(9)	HR(9)	HR(9)	4.80 ± 1.16	0.159_{09}; -0.008_{10}
Standardcultivar	Sarpo Mira	*R3a, R3b*	0(9)	0(9)	0(9)	HR(9)	2.11 ± 0.81	-0.573_{11}; -0.495_{12}; -0.332_{13}; $-0,567_{14}$; -0.286_{15}; -0.199_{16}; -0.128_{17}; -0.084_{18}
Standardcultivar	Adretta	nd	nd	nd	nd	nd	6.48 ± 1.57	0.246_{11}; 0.499_{12}; 0.255_{13}; 0.308_{14}; 0.416_{15}; 0.414_{16}; 0.364_{17}; 0.635_{18}
SH (Delikat + *blb*41)	82/4	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(8)	1.38 ± 0.36	-0.352_{12}; -0.075_{13}
BC_1 (SH x Quarta)	82/4/8	*Rpi-blb1*	HR(9)	HR(9)	0(9)	HR(8)	1.94 ± 0.26	0.096_{14}; 0.119_{15}; -0.012_{16}; 0.253_{17}; -0.072_{18}
	82/4/13	*Rpi-blb1, -blb3, R3a, R3b*	HR(8)	HR(9)	HR(8)	HR(9)	1.14 ± 0.30	-0.221_{14}; -0.355_{15}
	82/4/23	*Rpi-blb1, R3a, R3b*	HR(9)	0(9)	HR(9)	HR(9)	1.25 ± 0.45	-0.056_{13}; 0.094_{14}
	82/4/27	*Rpi-blb1*	nd	nd	nd	nd	1.43 ± 0.34	0.009_{14}; -0.055_{15}; -0.035_{16}
	82/4/30	*Rpi-blb1, R3a, R3b*	HR(9)	nd	HR(9)	HR(9)	1.53 ± 0.79	-0.122_{14}; -0.122_{15}; 0.039_{16}
	82/4/38	*Rpi-blb1, -blb3,*	HR(9)	HR(7)	0(9)	0(9)	1.45 ± 0.26	-0.282_{12}; 0.040_{13}
	82/4/43	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.52 ± 0.42	-0.031_{13}; 0.000_{16}
	82/4/46	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(8)	HR(9)	HR(9)	1.63 ± 0.53	-0.283_{14}; -0.402_{15}; 0.085_{16}
BC_2 (BC_1 x Sarpo Mira)	82/4/46/6	*R3a, R3b*	HR(6)*	HR(6)*	HR(6)*	HR(6)*	1.90 ± 0.19	0.152_{17}; -0.103_{18}
	82/4/46/14	*Rpi-blb1, -blb3, R3a, R3b*	HR(8)	HR(9)	HR(9)	0(9)	1.53 ± 0.25	-0.202_{17}; 0.044_{18}
BC_1 (SH x Quarta)	82/4/57	*Rpi-blb1*	HR(9)	HR(9)	0(9)	0(9)	2.60 ± 0.66	-0.159_{13}; 0.203_{14}
	82/4/68	*Rpi-blb1, -blb3, R3b*	HR(9)	HR(9)	0(9)	HR(9)	3.36 ± 1.58	-0.145_{13}; 0.164_{14}; 0.022_{15}; 0.174_{16}
BC_2 (BC_1 x Quarta)	82/4/68/3	*Rpi-blb1, -blb3*	HR(9)	0(9)	0(9)	0(9)	4.13 ± 0.81	0.508_{14}
	82/4/68/29	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(8)	HR(9)	HR(9)	2.36 ± 0.84	0.154_{14}
	82/4/68/45	*Rpi-blb1, -blb3, R3a*	HR(9)	HR(9)	HR(9)	0(9)	1.88 ± 0.24	0.985_{14}
BC_1 (SH x Quarta)	82/4/69	*Rpi-blb1, -blb3*	HR(9)	HR(9)	0(9)	0(9)	1.75 ± 0.62	-0.074_{12}; 0.113_{14}; 0.017_{15}; 0.003_{16}
BC_2 (BC_1 x Sarpo Mira)	82/4/69/7	*? R3b*	HR(9)	0(9)	0(9)	HR(9)	2.88 ± 1.26	-0.537_{17}; -0.130_{18}
BC_1 (SH x Quarta)	82/4/76	*Rpi-blb1, -blb3, R3a*	HR(6)*	HR(6)*	HR(6)*	0(6)*	1.53 ± 0.25	-0.041_{13}

Progenies with four stacked resistance genes, showing HR reaction and improved resistance to Pi in DLA and the field are highlighted in red; nd-not determined.

TABLE 2 | Assessment of the somatic hybrid (SH) 83/9, between potato cv. 'Delikat' and *Solanum bulbocastanum* (*blb*41) and derived back-crossed progenies BC_1 (the cultivar used for crossing is indicated), for the presence of the resistance genes as revealed by gene specific markers, functional profiling of *Avr* genes by agro-infiltration, resistance to late blight in a detached leaf assay (DLA) and in the field (Δ-rAUDPC); standard varieties 'Adretta' (susceptible) and 'Sarpo Mira' (resistant to late blight) were used; the year of the field assays is given as a subscript to each value of Δ-rAUDPC; from each genotype three plants and three leaflets were infiltrated, hypersensitive reaction is indicated by HR and lack of reaction by 0: the numbers in brackets indicate the number of repetitions with the same reaction.

Parents/ SH/BC_1	**Genotype ID**	**Resistance genes**	**Agro-infiltration with effectors**				**DLA (mean ± SD)**	***Phytophthora* attack Δ-rAUDPC (2009–2018)**
			Avrblb1	***Avr2***	***Avr3a***	***Avr3b***		
Wild parent	*Blb*41	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.00 ± 0.00	-0.124_{13}; -0.405_{14}
Cultivar	Delikat	none	0(9)	0(9)	0(9)	0(9)	5.73 ± 1.17	0.311_{12}; 0.164_{15}; $0.477_{16;}$ 0.491_{18}
Cultivar	Quarta	*R3a, R3b*	0(9)	0(9)	HR(9)	HR(9)	4.80 ± 1.16	0.159_{09}; -0.008_{10}
Standard cultivar	Sarpo Mira	*R3a, R3b*	0(9)	0(9)	0(9)	HR(9)	2.11 ± 0.81	-0.573_{11}; -0.495_{12}; -0.332_{13}; $-0,567_{14}$; -0.286_{15}; -0.199_{16}; -0.128_{17}; -0.084_{18}
Standardcultivar	Adretta	nd	nd	nd	nd	nd	6.48 ± 1.57	0.246_{11}; 0.499_{12}; 0.255_{13}; 0.308_{14}; 0.416_{15}; 0.414_{16}; 0.364_{17}; 0.635_{18}
SH (Delikat + *blb*41)	83/9	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(7)	1.34 ± 0.37	-0.324_{12}; 0.130_{13}
BC_1 (SH x Quarta)	83/9/3	*Rpi-blb1, R3a, R3b*	HR(9)	0(9)	HR(9)	HR(9)	1.40 ± 0.28	0.223_{14}
	83/9/8	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	0(9)	0(9)	HR(9)	3.06 ± 0.57	-0.115_{14};-0.130_{15}; 0.045_{16}
	83/9/11	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.10 ± 0.10	-0.049_{14};-0.331_{15}; -0.032_{16}; -0.307_{17}
	83/9/12	*Rpi-blb1, -blb3*	nd	nd	nd	nd	1.14 ± 0.15	-0.196_{14}; -0.234_{15}
	83/9/35	*Rpi-blb1, -blb3, R3a, R3b*	HR(8)	HR(9)	HR(9)	HR(8)	1.94 ± 0.44	-0.160_{14}; 0.011_{15}; 0.119_{16}
	83/9/39	*Rpi-blb3, R3a, R3b*	0(9)	HR(9)	HR(8)	HR(9)	1.60 ± 0.59	0.133_{14}
	83/9/44	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.30 ± 0.36	-0.326_{14}; -0.215_{15}
	83/9/47	*Rpi-blb1, -blb3, R3a, R3b*	HR(8)	HR(8)	HR(9)	HR(9)	2.06 ± 0.44	-0.101_{14}; -0.112_{15}; -0.005_{16}
	83/9/55	*Rpi-blb1, R3a, R3b*	HR(9)	0(9)	0(8)	HR(9)	1.80 ± 0.33	-0.230_{14}; -0.119_{15}; -0.292_{16}; 0.089_{17}; -0.198_{18}
	83/9/63	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	nd	HR(9)	HR(9)	1.87 ± 0.38	-0.103_{14}; -0.065_{15}; -0.170_{16}; 0.022_{17}; 0.047_{18}
	83/9/70	*Rpi-blb1, R3a, R3b*	HR(9)	0(9)	0(9)	HR(9)	2.25 ± 0.68	0.039_{14}
	83/9/73	*R3a, R3b*	0(9)	0(9)	HR(9)	HR(9)	2.25 ± 1.29	-0.088_{15}

Progenies with four stacked resistance genes, showing HR reaction and improved resistance to Pi in DLA and the field are highlighted in red; nd-not determined.

TABLE 3 | Assessment of the somatic hybrid (SH) 95/1, between potato cv. 'Delikat' and *Solanum bulbocastanum* (*blb*41) and derived back-crossed progenies BC_1 and BC_2 (the cultivar used for crossing is indicated), for the presence of the resistance genes as revealed by gene specific markers, functional profiling of *Avr* genes by agro-infiltration, resistance to late blight in a detached leaf assay (DLA) and in the field (Δ-rAUDPC); standard varieties 'Adretta' (susceptible) and 'Sarpo Mira' (resistant to late blight) were used; the year of the field assays is given as a subscript to each value of Δ-rAUDPC; from each genotype three plants and three leaflets were infiltrated, hypersensitive reaction is indicated by HR and lack of reaction by 0: the numbers in brackets indicate the number of repetitions with the same reaction.

Parents/ SH/BCs	Genotype ID	Resistance genes	Agro-infiltration with effectors				DLA (mean ± SD)	*Phytophthora* attack Δ-rAUDPC (2012–2018)
			Avrblb1	***Avr2***	***Avr3a***	***Avr3b***		
Wild parent	*Blb*41	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.0 ± 0.00	-0.124_{13}; -0.405_{14}
Cultivar	Delikat	none	0(9)	0(9)	0(9)	0(9)	5.73 ± 1.17	0.311_{12}; 0.164_{15}; $0.477_{16;}$ 0.491_{18}
Cultivar	Baltica	*R3a, R3b*	nd	0(9)	HR(9)	HR(9)	3.92 ± 0.71	nd
Cultivar	Romanze	*R3b*	0(9)	0(9)	0(9)	HR(9)	5.20 ± 1.12	nd
Standard cultivar	Sarpo Mira	*R3a, R3b*	0(9)	0(9)	0(9)	HR(9)	2.11 ± 0.81	-0.573_{11}; -0.495_{12}; -0.332_{13}; $-0{,}567_{14}$; -0.286_{15}; -0.199_{16}; -0.128_{17}; -0.084_{18}
Standard cultivar	Adretta	nd	nd	nd	nd	nd	6.48 ± 1.57	0.246_{11}; 0.499_{12}; 0.255_{13}; 0.308_{14}; 0.416_{15}; 0.414_{16}; 0.364_{17}; 0.635_{18}
SH (Delikat + *blb*41)	95/1	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.40 ± 0.35	0.084_{13}
BC_1 (SH x Baltica)	95/1/3	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.69 ± 0.70	-0.308_{11}; 0.061_{12}
	95/1/4	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	2.94 ± 1.81	nd
BC_2 (BC_1 x Romanze)	95/1/4/11	*R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	1.19 ± 0.40	-0.138_{14}; -0.257_{15}; -0.051_{16}; -0.091_{17}; -0.096_{18}
BC_2 (BC_1 x Romanze)	95/1/4/59	*Rpi-blb1*, *R3a*, *R3b*	HR(9)	HR(9)	HR(9)	HR(9)	2.25 ± 1.65	0.147_{14}; 0.291_{15}
BC_1 (SH x Baltica)	95/1/7	*Rpi-blb1, -blb3, R3a, R3b*	HR(9)	HR(9)	HR(9)	HR(9)	3.13 ± 2.74	0.045_{12}

BC_1 clone with four stacked resistance genes, showing HR reaction and improved resistance to Pi in DLA and the field are highlighted in red; nd-not determined.

staining. For chromosome counts, at least five cells with well-spread metaphase chromosomes were used for each clone. Best spreads were photographed with a digital camera attached to an epifluorescent microscope (Olympus BX 60 with appropriate filter for DAPI). For all cytogenetic evaluations of the parental lines, the diploid *blb*41 and tetraploid cultivated potatoes were used as internal standards.

Determination of the Composition of the Genome Using mcGISH

A modified multicolour GISH protocol (Kruppa et al., 2013) was used to simultaneously visualize the chromosomes of *Solanum bulbocastanum* ($2n = 2x = 24$, A^bA^b) and *S. tuberosum* ($2n = 4x = 48$, AAAA). Therefore, total DNA was labeled with biotin-16-dUTP and digoxigenin-11-dUTP (Roche Diagnostics, Mannheim, Germany), respectively, using the random primed labeling protocol. The hybridization mixture contained 60 ng each of the labeled probes/slide, dissolved in a 15 μL mixture of 100% formamide, 20xSSC and 10% dextran-sulfate in the ratio of 5:1:4. In order to denature the probes DNA, the hybridization mixture was heated at 75°C for 8 min and then immediately put on ice. The chromosome spreads were denatured at 80°C in a mixture containing 100% formamide, 20xSSC and 10% dextran-sulfate in the ratio of 15:3:3. Hybridization was done at 42°C overnight. Streptavidin-FITC (Roche) and Anti-Digoxigenin-Rhodamine (Roche), respectively, were used in the detection phase. Counterstaining was performed using 2 μg/mL DAPI in Vectashield antifade mounting medium. The slides were screened using a Zeiss Axioskop-2 fluorescence microscope equipped with filters appropriate for DAPI, FITC, Rhodamine and the simultaneous detection of FITC and Rhodamine. Images were taken using a Spot CCD camera (Diagnostic Instruments) and processed using Image Pro Plus software (Media Cybernetics).

Resistance to Late Blight

Detached Leaf Assay (DLA)

This assay used five leaves per clone and was replicated at least twice. The description of the *Pi* strain is given in section "Field Test." One drop (1 μL) of a *Pi* suspension of 15×10^3 zoosporangia /ml was placed on the underside of the leaves, which were flipped after 1 day and incubated at 16°C and 95% relative humidity at a light intensity of 2.1 μmol m^2 s^{-1} and incubated for 5 days. After 5 days the size of the necrotic areas and mycelium development were estimated on a scale from 1 (no symptoms visible) to 9 (completely necrotic and covered with mycelium). The protocol for evaluating resistance using the DLA is described by Thieme et al. (2008). The *Pi* inoculum was maintained for 1 year on tuber slices of a highly susceptible variety without any R-genes (cv. 'Adretta') in order to maintain a high level of virulence.

Field Test

Field tests were carried out at the Julius Kühn-Institut (JKI) Federal Research Centre for Cultivated Plants Trial Station for Potato Research, Groß Lüsewitz 18190 Sanitz, Germany. Clones were cultivated in double-row plots with ten plants per plot. A strip of hemp 3 m in width was planted around the field to ensure wind protection and to maintain a humid environment. Irrigation was conducted in the early morning and in the evening if necessary. Control plots of varieties highly resistant (cv. 'Sarpo Mira') and susceptible (cv. 'Adretta') to late blight were also planted in the field. The inoculum of *Pi* consisted of a mixture of common races collected in the field in the year before and refreshed the following year with an inoculum collected from natural infections (Hammann et al., 2009). The virulence of the isolate was evaluated under field conditions using the SASA differential set carrying the *R* genes *R1* to *R11* (Black et al., 1953). It was found highly aggressive as it overcame the resistance genes *R1* to *R8* and *R10* and *R11*. Probably due to their later maturation, *R9* plants were less infected. A continuous monitoring of the composition of the clonal lineages of the pathogen was not possible throughout the test period, but analyses of samples from 2014 to 2017 identified EU_13 mating type A2 as the predominant genotype in local populations. The lowest leaves of each last plant in a row were inoculated with a 5 ml spore suspension (15×10^3 zoosporangia/ml) in the evening. The area of the potato tops attacked was scored as a percentage twice a week until the plants were mature. This was based on the Area Under Disease Progress Curve (AUDPC) and Relative Area Under Disease Progress Curve (rAUDPC) (Hammann et al., 2009). As rAUDPC is strongly associated with the maturity of the plants, it was transformed into delta (Δ) rAUDPC. RAUDPC represents the percentage of infection and Δ-rAUDPC the quantitative maturity-corrected transformation of rAUDPC (Shaner and Finney, 1977; Bormann, 2003; Truberg et al., 2009). The maturity of the parents, SHs and progenies was evaluated in a field trial that was treated with fungicides using a scale from 1 (very early) to 9 (very late). Fifteen plants per clone were cultivated and agronomic characters other than maturity were also assessed.

Evaluation of Flowering, Yield, Tuber Shape/Appearance and Tuber Processing Qualities

Assessment of Flowering, Yield, Tuber Shape and Appearance

The SHs and offspring were repeatedly cultivated in a greenhouse for 2–5 years depending on genotype and breeding step. In addition to the evaluation of their resistance to pathogens, morphological characteristics, such as, plant habitus, flowering and the fruit (berry) development were assessed. Crossbreeding experiments with pollen of different potato varieties were carried out. The number of flowers on a single greenhouse plant was assessed and expressed on a scale: + [low number ($n \leq 5$)], ++ [moderate number ($n = 6$–20)] and +++ [high number of flowers ($n > 21$)]. Fertile SHs and BC clones were selected for further use.

Tuber appearance in terms of shape, eye depth, skin texture were assessed and the results are summarized in **Table 4** [scale: + (poor), +++ (very good)].

Yield (kg of tubers per plot), total number of tubers and tuber shape were assessed of plants that developed from the 15 tubers planted in each plot in the field. Yield of SHs and

TABLE 4 | Assessment of the agronomic traits of BC progenies derived from *S. bulbocastanum,* GLKS 31741 (*blb*41) (+) cv. 'Delikat' hybrids grown in a field for 2 or 3 years, including flowering, tuber number, weight and starch content as well as processing qualities: Discolouration of raw tuber tissue, of cooked tuber flesh (average for 3 years); the BC clones with best yield, qualities for agronomic traits and processing are highlighted in green.

Genotype	Agronomic traits					Discolouration	
	Flower intensity	Appea-rance	Tubers (n/plant)	Tuber weight (kg/plant)	Starch content (%)	Raw tuber tissue	Cooking tuber flesh
Cv. Delikat	+	+++	11.4 ± 2.1	1.45 ± 0.5	17.5 ± 0.9	2.0 ± 0	4.7 ± 0.6
BC_1 82/4/8	+	+	13.1 ± 3.8	1.00 ± 0.2	16.3 ± 1.4	4.5 ± 0.7	7.7 ± 0.6
BC_1 82/4/27	+	+	14.3 ± 7.6	0.58 ± 0.3	12.5 ± 0.5	2.5 ± 0.7	6.0 ± 2.0
BC_1 82/4/30	(+)	+++	8.3 ± 3.8	0.52 ± 0.3	15.4 ± 2.2	2.5 ± 0.7	5.5 ± 2.1
BC_1 82/4/46	+	++	9.2 ± 1.7	1.23 ± 0.4	14.7 ± 2.3	2.0 ± 0	4.5 ± 2.1
BC_1 82/4/67	+	(+)	14.0 ± 0.9	1.38 ± 0.8	17.7 ± 1.4	5.0 ± 0	6.5 ± 2.1
BC_1 82/4/68	+	+	9.1 ± 1.9	0.87 ± 0.2	15.3 ± 2.0	2.0 ± 0	5.7 ± 1.2
BC_1 82/4/69	+	+	15.1 ± 6.6	0.52 ± 0.2	16.9 ± 0.3	3.5 ± 0.7	4.7 ± 0.6
BC_1 83/9/8	++	+	7.7 ± 1.3	0.94 ± 0.1	21.1 ± 0.6	3.5 ± 0.7	5.7 ± 1.5
BC_1 83/9/11	+++	+	6.8 ± 1.9	1.01 ± 0.2	17.6 ± 0.4	4.0 ± 0	6.0 ± 1.0
BC_1 83/9/35	+++	+	13.4 ± 5.5	0.53 ± 0.1	18.4 ± 1.1	3.5 ± 0.7	6.3 ± 1.2
BC_1 83/9/47	+	+	11.2 ± 2.0	0.95 ± 0.2	17.0 ± 1.0	4.5 ± 0.7	7.3 ± 0.6
BC_1 83/9/55	+	+++	8.8 ± 0.8	1.34 ± 0.6	17.6 ± 0.4	3.5 ± 0.7	5.3 ± 1.2
BC_1 83/9/63	(+)	++	10.1 ± 0.2	1.25 ± 0.2	15.8 ± 1.3	3.0 ± 0	4.7 ± 1.2
BC_2 82/4/68/11	+	+++	6.6 ± 1.2	2.67 ± 3.5	12.5 ± 1.1	3.5 ± 0.7	4.8 ± 1.3
BC_2 95/1/4/11	++	+	14.2 ± 1.6	0.75 ± 0.2	15.5 ± 0	4.5 ± 0	2.5 ± 0.7
BC_2 95/1/4/59	++	++	16.8 ± 0.6	1.10 ± 0.1	17.4 ± 2.6	4.0 ± 0	5.5 ± 0.7

Initial number, 15–20 plants per genotype and plot; tuber harvest per plot. Flower intensity on a scale: (+), +, low number (n ≤ 5); ++, moderate number (n = 6–20), and +++, high number of flowers (n > 21); Tuber appearance (including tuber shape, eye depth, skin texture) is summarized on a scale: (+), +, poor, +++, very good; Discolouration: 1 (best) – 9 (low quality).

derived progenies were compared to that of the parental potato cultivar 'Delikat'.

Processing quality was based on the discolouration of raw tuber tissue, flesh after cooking and the starch content.

Assessment of Discolouration of Raw Tuber Tissue

Four tubers per genotype were used 4–5 months after harvest. Using a cork-borer two tissue cylinders were cut out parallel to the long axis of each tuber. After 24 h, the discolouration was visually estimated in terms of intensity of gray to black on a scale from 1 (not discolored) to 9 (black). Cultivar 'Delikat' was used as a reference, as its reaction is well known.

Discolouration of Tuber Flesh After Cooking

Four medium-sized tubers per genotype were cooked, hot peeled, cut longitudinally and the cut surface placed facing down on a plate. Discolouration was estimated on a scale from 1 to 9 after 24 h and compared with the reference samples.

Starch Content

Starch content was determined, by calculating the difference between the weight of a tuber in air and in water using an under-water balance EORO-KUW-2000 (Fischer KG Bielefeld, Germany).

Statistical Analysis

The free statistical software R (R Development Core Team, 2010) was used for statistical calculations. The function mean() was used to calculate means and the function sd() was used to calculate standard deviations.

RESULTS

Identification of *Rpi-blb1*, *Rpi-blb3*, *R3a*, and *R3b* Genes in *S. bulbocastanum blb*41, SHs and Derived BC Progenies and Assessment of Resistance to Foliage Blight

Using specific markers it was possible to demonstrate the presence of the following resistance genes in *blb*41: *Rpi-blb1*, *Rpi-blb3, R3a*, and *R3b* and confirm their transfer into SHs and introgression into derived BC progenies (**Figure 1**, **Tables 1**–**3**, and **Supplementary Figure S1**). The fertile and best performing SH clones were further selected for ploidy (chromosome counts) and field assessments. All three SHs selected and the majority of their derived BCs carried both *Rpi* resistance genes (**Tables 1**–**3**). In one BC_1 (83/9/39) gene specific markers indicated the presence of only the *Rpi-blb3* gene (**Figure 1** and **Table 2**). This clone had a good score for resistance in the DLA (mean of 1.6), but a plus value for Δ-rAUDPC in the field (**Table 2**). The race-specific genes *R3a* and *R3b* were detected using gene specific markers in *blb*41 and the cultivars 'Baltica', 'Quarta' and 'Sarpo Mira', which were used as pollinators in back-crosses. Only *R3b* was detected in cv. 'Romanze' (**Tables 1**–**3** and **Supplementary Figure S1**). The selected SHs 82/4, 83/9 and 95/1 carried the

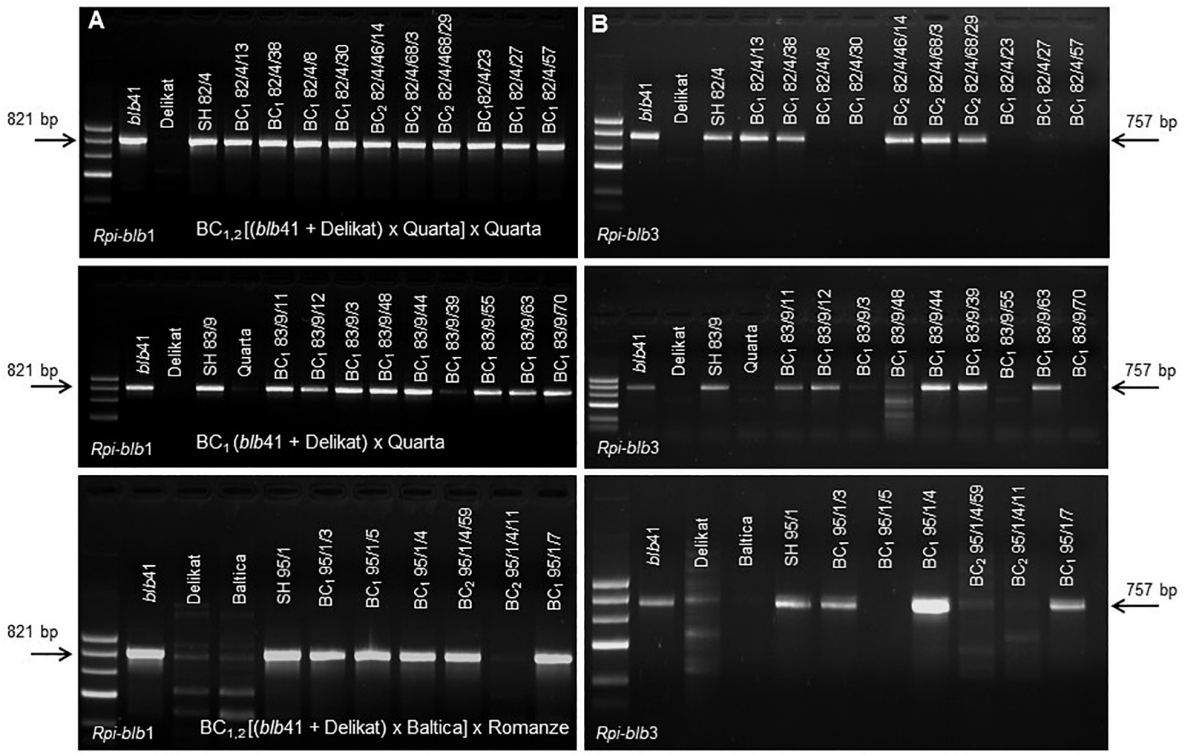

FIGURE 1 | Molecular analysis using gene specific markers to determine the presence of **(A)** *Rpi-blb1* and **(B)** *Rpi-blb3* genes in selected somatic hybrids, their parental lines *S. bulbocastanum* GLKS 33741 (*blb*41), cv. 'Delikat' and different segregating BC_1 progenies after back-crossing with cvs. 'Quarta' and 'Baltica' (from above to below, arrows indicate the target fragments).

two durable resistance genes *Rpi-blb1* and *Rpi-blb3* and race-specific *R3a* and *R3b* genes (**Tables 1–3**). Hence, eight clones of the BCs derived from the three selected SHs and had specific fragments of all the resistance genes (**Tables 1–3–highlighted** in red, **Figure 1**. The three selected SHs and 30 of the 36 derived BC clones (listed in **Tables 1–3**), had a significantly high degree of *Pi* resistance in field trials, during the 1–4 years in which the intensity of the foliage blight attack varied (**Tables 1–3** and **Figures 3**, **4A,B**). In the first years, a large number of SHs was evaluated in the field (**Figure 4A**) and the BC clones of the best three SHs were evaluated in the years 2014 to 2016 (**Figure 4B**). The level of resistance varied from year to year and the susceptible cultivar 'Delikat' had a rAUDPC of only 50% in 2013. The maturity-corrected values of Δ-rAUDPC for these SH clones are lower than that of the susceptible standard cv. 'Adretta' and cv. 'Delikat', but rather similar to the wild species *blb*41, used as a control, which did not show any symptoms of foliage blight attack (rAUDPC = 0%) or cv. Sarpo Mira, used as resistant standard cultivar. The area under the disease progress curve (rAUDPC in %) and maturity data are shown in **Figures 4A,B**, and Δ-rAUDPC in **Tables 1–3**, for each SH family. Minus values of Δ-rAUDPC indicate a low level of susceptibility. Three of the SHs selected and eleven BC_1–BC_2 clones had all the genes and were only slightly susceptible with minus values of Δ-rAUDPC in the field (see **Tables 1–3**). Other BCs (18) with only one, two or three resistance genes had minus values for Δ-rAUDPC in the field trials in at least 1 year in **Tables 1–3**. The maturity of SHs and derived BC_1 clones was intermediate between that of potato cultivar 'Delikat', with a maturity value of 5 and the very late maturing (value 9) wild species *blb*41 (**Figures 4A,B**). The majority of the BCs had values between 4.7 and 7.8, with some BC_1 clones maturing later with values 8 and 9 (82/4/8, /27, /67 and 83/9/11, /35). Many of the BC_1 clones had a low percentage rAUDPC with the last ones having low values in some years (**Figure 4B**). The SHs and their derived BC_1 had minus values for Δ-rAUDPC when tested in the field, which was confirmed by the DLA results with one exception, clone 82/4/68/29, which had a good value in DLA (2.36 ± 0.84), but was susceptible in the field with a plus value of Δ-rAUDPC in 1 year (**Table 1**). It needs to be verified by more assays in the future.

Agro-Infiltration With Effectors

The genes detected by gene specific markers after agro-infiltration with corresponding effectors were seen as elicitors of the hypersensitive response (HR) (**Figure 2**). Generally, after agro-infiltration these genes were associated with incompatible interactions and necrosis (**Tables 1–3**). There were also cases where genes were present but there was no hypersensitive reaction to the agro-infiltration of the *Avr* gene, as for example: BC_2 82/4/68/3 with the genes *Rpi-blb1* and *Rpi-blb3* but *Rpi-blb3* shows no HR (**Table 1**); or BC_1 83/9/8 with both *Rpi-blb3* and *R3a* genes showing no activity after agro-infiltration with *Avr2* and

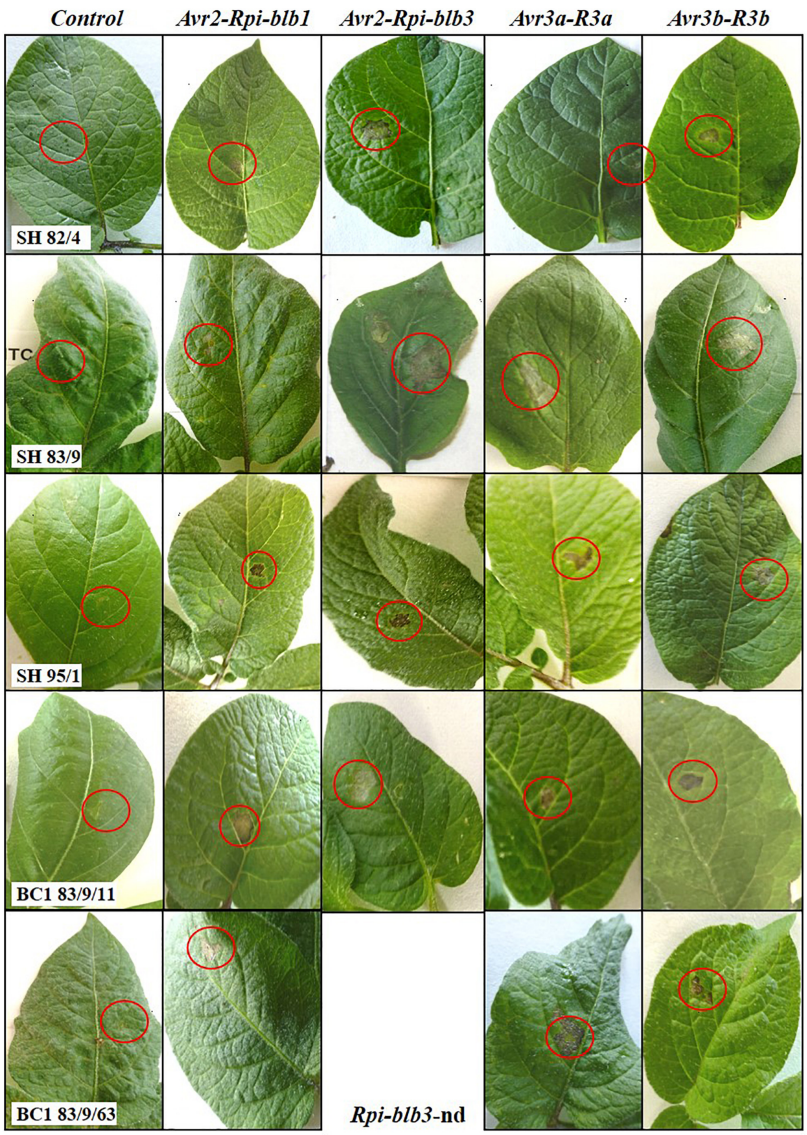

FIGURE 2 | Hypersensitive reaction of leaves of *Solanum bulbocastanum* (GLKS 33741, *blb*41) (+) cv. 'Delikat somatic hybrids (SHs) and selected back-cross progenies after agro-infiltration of *Avr*-effectors at 4*dpi*, which reveals the activity of corresponding resistance genes: *Rpi-blb1*, *Rpi-blb2*, *R3a*, or *R3b*. No visible reaction on leaves of SHs, BC_1 by infiltration of the vector without effectors (empty vector control), nd-not determined.

Avr3a; BC_1 83/9/55 and 83/9/70 in which the *R3a* gene, although present was inactive (**Table 2**). The eight BC clones with stacks of four genes that showed a HR are highlighted. Those are also best performing BCs in the DLA with constant minus values of Δ-rAUDPC in different years in the field after inoculation with *Pi* (**Tables 1–3**).

Ploidy and Genome Constitution of SHs and Derived BCs

As expected from previous investigations (Rakosy-Tican et al., 2015), the ploidy levels of the SH clones differed, one was hexaploid and two between pentaploid and hexaploid, i.e., aneuploids and asymmetric in their nuclear genetic constitutions (**Supplementary Table S1** and **Figures 5A–E**). Our results indicate that after back-crossing with the following tetraploid cultivars: 'Quarta', 'Baltica', 'Romanze', and 'Sarpo Mira', all BC_1 and BC_2 clones have lower chromosome numbers (**Supplementary Table S1**). Nevertheless, mcGISH indicates that between 4 and 24 of the chromosomes were from the wild parent *blb*41 and the introgression of the DNA of wild species into potato chromosomes (**Figure 5C**).

Yield of BC Clones in the Field and Processing Quality

The BC clones grown in the field varied in terms of yield and shape of the tubers at harvest (**Table 4** and **Figure 6**). The

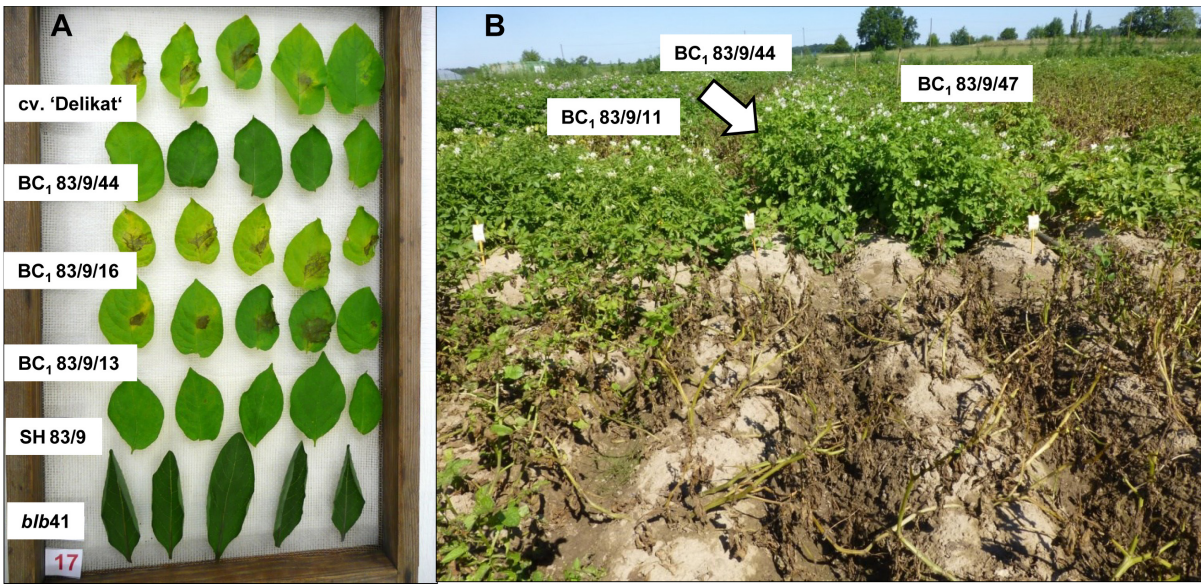

FIGURE 3 | Evaluation of resistance to late blight in a greenhouse and the field **(A)** Detached leaf assay (DLA) showing resistant *S. bulbocastanum* GLKS 33741 (*blb*41), a resistant somatic hybrid, two susceptible and one resistant BC_1 clones and the susceptible cv. 'Delikat' after infection with *Phytophthora infestans* (from below upward) **(B)** BC_1 clones 83/9/44 (arrow), 83/9/11 (left) and 83/9/47 (right) derived from *blb*41 (+) cv. 'Delikat' somatic hybrid 83/9 in a field trial in 2015 expressed improved horizontal resistance to foliage blight.

majority of the BC clones flowered and produced a satisfactory yield and number of tubers (**Table 4** and **Figure 6**). Moreover, three BC clones (BC_1 82/4/30; BC_1 83/9/55; BC_2 82/4/68/11) produced tubers of good appearance and similar to those of cultivar 'Delikat' (**Table 4** and **Figure 6**). The yield of two BC_1 clones (82/4/67, 83/9/63) and one BC_2 clone (95/1/4/59) was higher or similar to that of this cultivar, but the number of flowers and appearance of the tubers were low and acceptable, respectively. The starch content of the tubers of these clones varied from 12.5 to 21.1% and processing quality varied between scores 2.0 to 7.7 (**Table 4**). The BC_1 clones 82/4/46; 82/4/68; 83/9/35; 83/9/55 produced tubers with good processing qualities (**Figure 7**). Useful pre-breeding clones are also highlighted in green (**Table 4**).

DISCUSSION

Somatic hybridization and the identification of resistance genes using gene-specific markers were used to transfer a stack of genes determining durable resistance to late blight (*Rpi-blb1* and *Rpi-blb3*) and the race-specific genes (*R3a* and *R3b*) into the potato cv. 'Delikat'. This cultivar was shown to be a good candidate for somatic hybridization programs (Rakosy-Tican et al., 2015). By using *blb* accession GLKS 31741 (*blb*41), in which four R genes were identified and were introgressed into potato pre-breeding clones. Previously, it has not been possible to transfer this gene combination into potato either by classical breeding or by gene transfer.

The exploitation of the R genes for breeding research in potatoes has also been tested, in particular using genetic engineering approaches. Recently, more than 30 R genes found in species of wild potato resistant to late blight were mapped and 20 were cloned (Rodewald and Trognitz, 2013; Haverkort et al., 2016). These genes are available for the transfer into potato varieties and breeding lines. The company BASF developed the potato variety "Fortuna" derived from the cv. 'Fontane' by genetic transformation with two late blight R genes, *Rpi-blb1* and *Rpi-blb2* from *S. bulbocastanum* (Storck et al., 2012). This variety was not approved for the European market and it ceased to be cultivated. Zhu et al. (2012) have successfully stacked three other late blight resistance genes: *Rpi-sto1* (*S. stoloniferum*) homolog of *Rpi-blb1*, *Rpi-vnt1.1* (*S. venturii*) and *Rpi-blb3* (*S. bulbocastanum*). The susceptible cv. 'Désirée' became more resistant by introgressive hybridization and DLA indicated the presence of the stacked genes in the plants grown in a field over a period of 2 years (Zhu et al., 2012; Haesaert et al., 2015). In addition, two genes from *S. bulbocastanum* (*RB*, *Rpi-blb2*) and one R gene *Rpi-vnt1.1* from *S. venturii* were stacked using genetic transformation and transferred into the varieties preferred by farmers, 'Désirée' and 'Victoria', and transgenic clones of these plants were cultivated in Sub-Saharan Africa. These potato lines were completely resistant to late blight over three seasons in the field in Southwest Uganda (Gishlain et al., 2019). Currently, transgenic potato varieties are approved for cultivation in the United States, Canada and Bangladesh. In a 10 years research project on the durable resistance of potato to *Phytophthora* at Wageningen University and a Research Centre genetic engineering was used to produce potato plants resistant to late blight using R genes mainly from *Solanum bulbocastanum* (*blb*), *S. edinense, S. stoloniferum* (*sto*), *S. venturii* (*vnt*), and *S. chacoense* (*chc*). Four varieties were transformed with one to

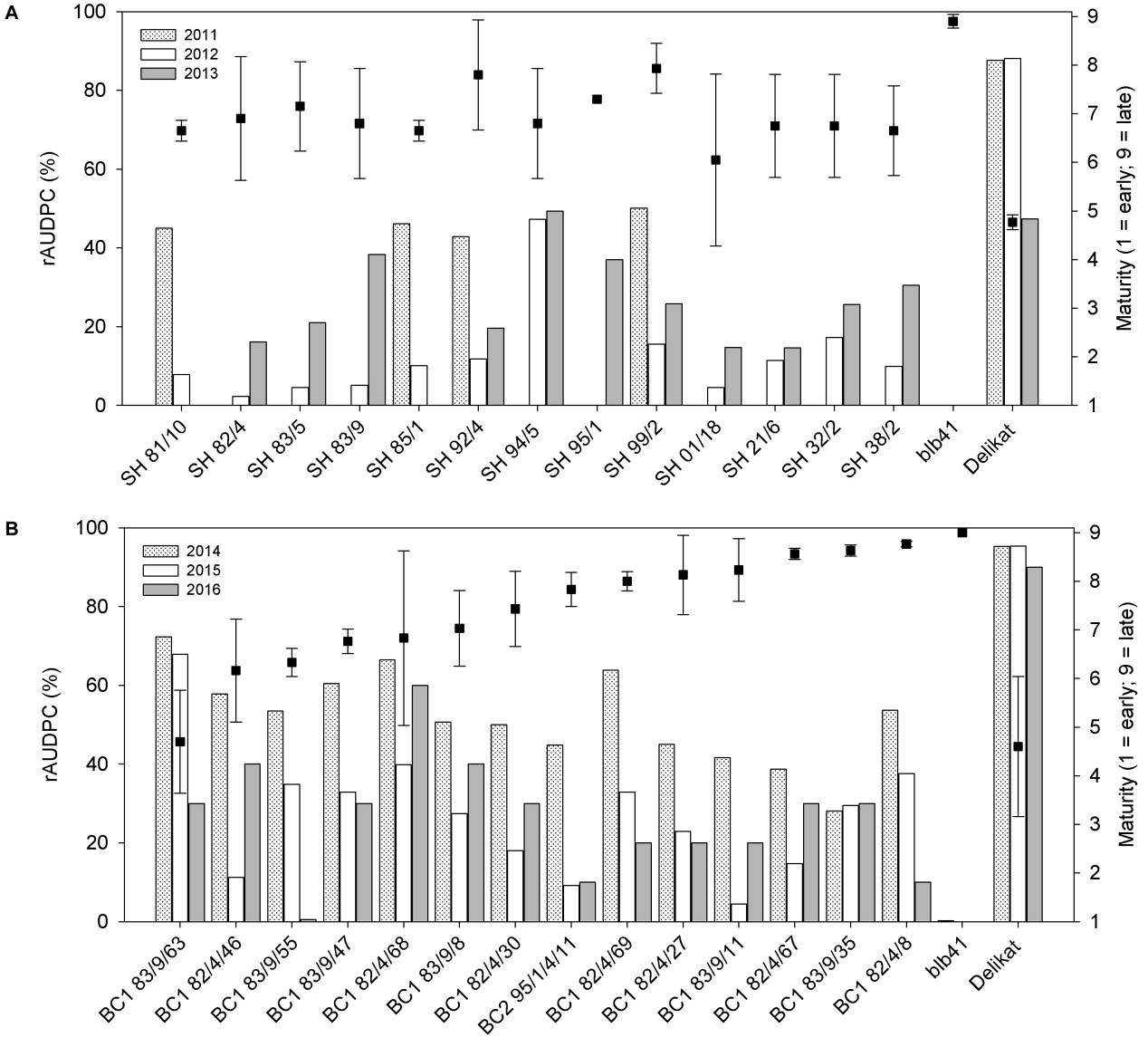

FIGURE 4 | Percentage of foliage area infected by late blight expressed as relative Area Under Disease Progress Curve (rAUDPC) and maturity (1 = very early; 9 = very late), evaluated in the field, after artificial inoculation of parents, **(A)** Somatic hybrids and **(B)** Their back-crossed progenies, potato (+) *Solanum bulbocastanum* GLKS 33741 (*blb*41); **(A)** years: 2011, 2012, 2013; **(B)** Values for years 2014–2016).

three R genes that originated from these wild potato species. The 'Désirée' clones containing stacks of either two R genes (*Rpi-blb3:sto1*; *Rpi-vnt1:chc1, Rpi-vnt1:sto1*) or three R genes (*Rpi-blb3:vnt1:sto1*) were not infected with *Pi* in field trials that lasted for 2 years (Haverkort et al., 2016). Recently, after using biotechnological improvement Simplot Plant Sciences has reported 'Innate' Gen 2 potatoes that are of high quality and show a strong resistance to late blight in the second generation[1]. Consequently, the use of biotechnology and MAS will increase the speed with which other genes conferring resistance can be incorporated into potato cultivars and their subsequent cultivation (Rietman, 2011; Chandel et al., 2015). Extensive investigations indicate that stacking of multiple R genes and monitoring how to deploy these stacks spatially and temporally could reduce fungicide use by over 80% (Haverkort et al., 2016).

Thus, genetic engineering could prove very useful if exempted from GMO rules in Europe. We used somatic hybridization, as the resulting plants are not genetically modified (directive 2001/18/EC–annex 1B)[2]. The use of somatic hybridization does enable to transfer new assortments of nuclear and cytoplasmic genes into regenerated plants and to overcome sexual incompatibility barriers (Rakosy-Tican et al., 2015; Thieme

[1]http://www.innatepotatoes.com/gen-two

[2]http://www.biosafety.be/PDF/2001_18.pdf

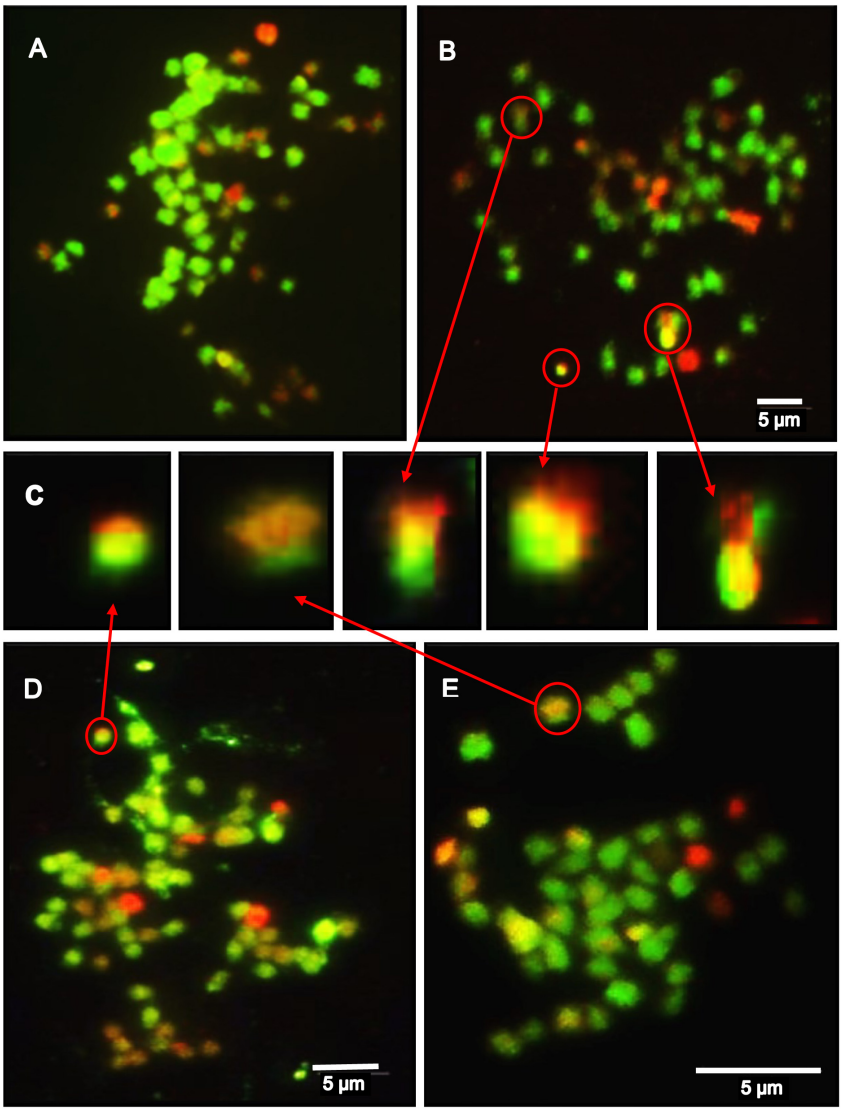

FIGURE 5 | The results of a mcGISH analysis of the genetic composition of somatic hybrids (SH), back-cross progenies (BC_1 and BC_2) – examples; green labels *S. tuberosum*, red labels *S. bulbocastanum* GLKS 33741 (*blb*41) chromosomes: **(A)** SH 83/9 - total 72 chromosomes, in red *S. bulbocastanum* = 24; **(B)** BC_1 83/9/3 - total chromosomes 54, 10 red from *blb*41 **(C)** Detailed magnifications of the introgressions from *blb*41 (red) into potato chromosomes (green) in BC_1 83/9/3, SH 95/1 and BC_2 95/1/4/59 (as indicated by red arrows); **(D)** SH 95/1 – total 64 chromosomes with 24 *blb*41 chromosomes in red; **(E)** BC_2 95/1/4/59 (total 48 chromosomes, 4 red from *blb*41 and recombinations – with red and green); Yellow regions denote the complementarity of the respective sequences; bar = 5 µm.

and Rakosy-Tican, 2017). Moreover, when asymmetric SHs are produced because of somatic incompatibilities (Rakosy-Tican et al., 2015), only a part of the nuclear DNA from the wild parent is present in the hybrids and derived BCs. In this study, mcGISH confirmed the introgression of *blb*41 DNA into BC clones, but more detailed analyses of chromosomes using FISH would identify the transferred resistance gene(s) and their location. Nevertheless, this study showed the introgression of DNA from wild species into potato cultivars using mcGISH.

The use of molecular markers and cytogenetic methods revealed that many of the SHs evaluated for resistance to *Pi* are asymmetric (Rakosy-Tican et al., 2015), which is confirmed in this study using indirect and direct cytogenetics and mcGISH. From the high number of the SHs between *blb*41 and five potato cultivars produced only those that had the basic potato characters in terms of morphology and fertility were selected (Rakosy-Tican et al., 2015) for assessing resistance to foliage blight and the important traits for a pre-breeding program. The DLA revealed that the three SH clones (*blb*41 + cv. 'Delikat') and BC_1 derived progenies with both *Rpi-blb1* and *Rpi-blb3* genes showed little symptoms of infection with *Pi*. The analysis of the specific alleles at the *Rpi-blb1, Rpi-blb3, R3a*, and *R3b* loci indicate that in the following generations

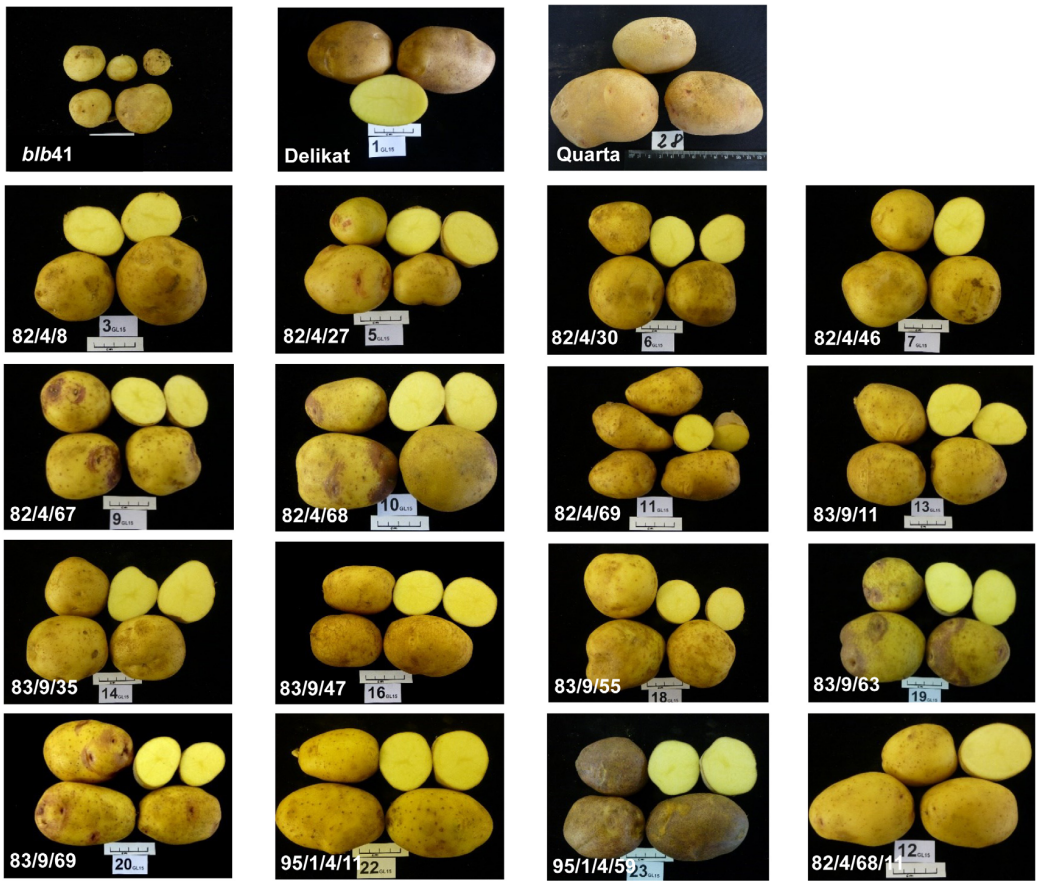

FIGURE 6 | Size and appearance of tubers of parental lines and selected BC_1 and BC_2 (the last three below) progenies of *S. bulbocastanum* GLKS 31741 (*blb*41) (+) cv. 'Delikat' somatic hybrids after cultivation in a field.

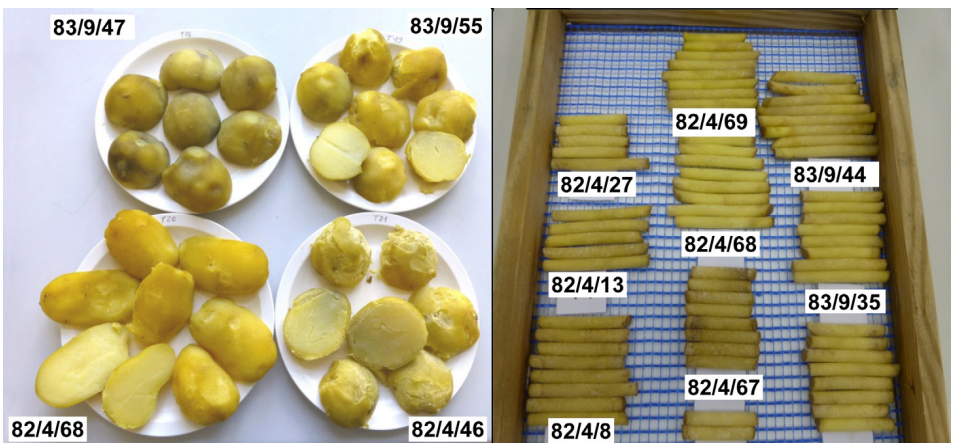

FIGURE 7 | Assessment of quality traits: Discolouration of cooked tuber flesh (left) and raw tuber tissue of different BC progenies derived from *S. bulbocastanum* GLKS 31741 (*blb*41) (+) cv. 'Delikat' somatic hybrids 82/4 and 83/9 (back-crossed with cv. 'Quarta').

the resistance genes segregated. Presence of race-specific genes in the following BC generations was also caused by their presence in some of the cultivars used for back-crossing, like 'Baltica', 'Sarpo Mira', and 'Quarta' with *R3a* and *R3b* genes, and 'Romanze' with only the *R3b* gene. In a few cases, a lack of a direct correlation between the presence of one of the

resistance genes and the level of resistance or functionality after agro-infiltration with the corresponding effector, indicate that the relevant genes were inactivated and the markers detected not only these resistance genes but also their analogs. Resistance is a complex trait with many effectors, epistatic or pleiotropic interactions (Lokossou et al., 2010; Aguilera-Galvez et al., 2018). It is known that in the susceptible haplotype (*rb*) there is a 18 bp deletion in the RB/*Rpi-blb1* gene (Bradeen et al., 2003). When the *Rpi-blb2* gene from *S. bulbocastanum* was transferred into potato cv. 'Désirée', there was not a positive correlation between gene expression in RT-qPCR assays and resistance to *Pi* in some cases (Orbegozo et al., 2016). Bearing in mind the genetic complexity of the SHs and BC clones (with many asymmetric nuclear constitutions in the SHs and a tendency for genetic stabilization in BC progenies), it is likely that the results of the RT-qPCR analysis will not always be correlated with the phenotypic resistance in DLA or field tests. Moreover, the evaluation of AUDPC in the field is also dependent on climate, which is the reason for presenting the data separately for each year in this study. The Δ-rAUDPC, however, confirmed the results of DLA in the field. It is likely that the tetrasomic inheritance in cultivated potato and the diverse nature of the genomic compositions of the SHs and derived progenies increased the complexity of the interactions, both genetically and between pathogen and host. Nevertheless, in the field trials SHs and four BC_1/BC_2 clones performed well in terms of tuber yield (**Table 4**) and resistance to foliage blight (**Tables 1–3**). For some of the hybrid clones the yield and number of tubers per plot was similar to those of the corresponding cultivar grown in the field trial, but tuber morphology in BC_2 progenies need further improvement by breeding. The negative values of the Δ-rAUDPC of SHs and BC_1 clones indicate that their resistance to foliage blight in the field assays was higher in at least 2 years. Compared to their parents more of the hybrids with both *Rpi-blb1* and *Rpi-blb3* genes survived an outbreak of late blight in the field in 2015 (data not shown). Moreover, the functionality of the genes identified by gene-specific markers was proven by agro-infiltration with corresponding *Avr*-effectors. The results demonstrate the functionality of stacked genes in the eight BCs, which were the most significant also in DLA and resistance in field assessments. The results demonstrate that the usage of somatic hybridization in combination with gene specific markers allow the transfer of multiple resistance traits (Thieme et al., 2008, 2010), which can also be tracked and stacked directly into potato cultivars. Incorporating more *R* genes into the potato might be the only way to increase the durability and level of resistance against late blight (Tan et al., 2010; Orbegozo et al., 2016). Other resistance mechanisms and traits can also be transferred along with those for which genes are already sequenced and, hence, can be easily tracked. Other authors (Davis et al., 2012) also report multiple resistances to late blight, viruses and aphids in the somatic fusion of potato and *S. bulbocastanum* and their crosses. In this study, tracking of these genes is not reported. Previous studies reported other resistant hybrids, which were fertile/infertile and either symmetric or asymmetric (Rakosy-Tican et al., 2015), which represent valuable material for further characterization and biotechnological improvement.

However, it is likely that *P. infestans* will eventually overcome the resistance to the newly released R gene variants. Therefore, various available strategies to control this insidious pathogen should be considered. It is necessary to deploy more cultivars that are both resistant and productive and our experiments showed that this is feasible. Moreover, some of our BC progenies of the three selected SHs are also good for producing chips and French Fries. Discolouration of the tissue of tubers, especially after cooking, and their starch content determine their table quality and value for the production of non-food products. Discolouration of the raw tuber flesh is caused by oxidation of phenolic compounds in the presence of polyphenol oxidase resulting in quinones, which is transformed into a dark pigment (Friedman, 1997). The level of phenolic compounds, activity of polyphenol oxidase and level of free amino acids also affect this process. Discolouration of tuber flesh after cooking is a key factor in determining their suitability for fresh consumption and processing (Wang-Pruski, 2007). The evaluation of the processing qualities of fusion hybrids and BC progenies will promote the exploitation of resistant clones as pre-breeding valuable genotypes. The results presented here underline the value of potato genotypes that are resistant and suitable for further processing.

In conclusion, this analysis of selected SHs and their BC_1 and BC_2 progenies has revealed eight progenies with a stack of four resistance genes, which are hypersensitive to corresponding *Avr*-effectors and resistant to late blight in DLA and the field (Δ-rAUDPC), the yield and processing qualities of some of which are good. These genotypes are valuable pre-breeding lines to produce potato cultivars with potentially durable resistance against foliage blight.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

ER-T and RT developed the idea, produced the plant material, wrote the manuscript and contributed to the DLA tests, agro-infiltration, and data analysis. MN carried out the marker analysis. T-ED, KK, and MM-L did the cytogenetics and mcGISH analysis. TH and JK performed the resistance in the field, yield evaluations and functional profiling of resistance genes by agro-infiltration. All authors read and approved the final manuscript.

FUNDING

ER-T acknowledges the financial support of a grant from the Romanian Authority for Scientific Research, CNCS-UEFISCDI

project number PNII-ID-PCE-2011-3-0586. The DAAD scholarship for 2018 awarded to ER-T is also acknowledged.

ACKNOWLEDGMENTS

Vivianne Vleeshouwers kindly provided *Agrobacterium tumefaciens* with effectors for checking the functionality of the stacked resistance genes. The authors wish to thank technical staff for excellent assistance, T. Dixon and N. Bernhardt for their helpful revision of the manuscript and Bernd Truberg for help with the statistical analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00699/full#supplementary-material

FIGURE S1 | An example of molecular analysis using gene specific markers to determine the presence of *R3a* and *R3b* genes in *S. bulbocastanum* GLKS 31741 (*blb*41), somatic hybrids *blb*41 (+) cv. 'Delikat' (SH) and their offspring (BC_1); M: DNA Ladder 100 bp.

TABLE S1 | Results of direct and indirect (flow cytometry) ploidy determination of selected somatic hybrids between potato cv. 'Delikat' and *Solanum bulbocastanum* (*blb*41) and their derived BC_1 and BC_2 progenies (nd-not determined).

REFERENCES

Aguilera-Galvez, C., Champouret, N., Rietman, H., Lin, X., Wouters, D., Chu, Z., et al. (2018). Two different R gene loci co-evolved with Avr2 of *Phytophthora infestans* and confer distinct resistance specificities in potato. *Stud. Mycol.* 89, 105–115. doi: 10.1016/j.simyco.2018.01.002

Black, W., Mastenbroek, C. W., Mills, R., and Peterson, L. C. (1953). A proposal for an international nomenclature of races of *Phytophthora infestans* and of genes controlling immunity in *Solanum demissum* derivatives. *Euphytica* 2, 173–179. doi: 10.1007/bf00053724

Bormann, C. A. (2003). *Genetic and Molecular Analysis Of Quantitative And Qualitative Late Blight Resistance In Tetraploid Potato*. Dissertation, Universität Hohenheim, Stuttgart.

Bradeen, J. M., Naess, S. K., Song, J., Haberlach, G. T., Wielgus, S. M., Buell, C. R., et al. (2003). Concomitant reiterative BAC walking and fine genetic mapping enable physical map development for the broad-spectrum late blight resistance region. *RB. Mol. Genet. Genomics* 269, 603–611. doi: 10.1007/s00438-003-0865-8

Bradshaw, J. E., Pande, B., Bryan, G. J., Hackett, C. A., McLean, K., Stewart, H. E., et al. (2004). Interval mapping of quantitative trait loci for resistance to late blight [*Phytophthora infestans* (Mont.) de Bary], height and maturity in a tetraploid population of potato (*Solanum tuberosum* subsp. *tuberosum*). *Genetics* 168, 983–995. doi: 10.1534/genetics.104.030056

Chakrabarti, S. K., Xie, C., and Tiwari, J. K. (eds) (2017). *The Potato Genome*. Berlin: Springer.

Chandel, P., Tiwari, J. K., Ali, N., Devi, S., Sharma, S., Sharma, S., et al. (2015). Interspecific potato somatic hybrids between *Solanum tuberosum* and *S. cardiophyllum*, potential sources of late blight resistance breeding. *Plant Cell Tiss. Organ. Cult.* 123, 579–589. doi: 10.1007/s11240-015-0862-8

Colton, L. M., Groza, H. I., Wielgus, S. M., and Jiang, J. M. (2006). Marker-assisted selection for the broad-spectrum potato late blight resistance conferred by gene RB derived from a wild potato species. *Crop Sci.* 46, 589–594. doi: 10.2135/cropsci2005.0112

Danan, S., Veyrieras, J.-B., and Lefebvre, V. (2011). Construction of a potato consensus map and QTL meta-analysis offer new insights into the genetic architecture of late blight resistance and plant maturity traits. *BMC Plant Biol.* 11:16. doi: 10.1186/1471-2229-11-16

Davis, J. A., Radcliffe, E. B., Thill, C. A., and Ragsdale, D. W. (2012). Resistance to aphids, late blight and viruses in somatic fusion and crosses of *Solanum tuberosum* L. *and Solanum bulbocastanum*. *Am. J. Pot. Res.* 89, 489–500. doi: 10.1007/s12230-012-9272-1

Doležel, J., Greilhuber, J., and Suda, J. (2007). Estimation of nuclear DNA content in plants using flowcytometry. *Nat. Protoc.* 2, 2233–2244. doi: 10.1038/nprot.2007.310

Du, J., Rietman, H., and Vleeshouwers, V. G. A. A. (2014). Agroinfiltration and PVX agroinfection in potato and *Nicotiana benthamiana*. *J. Visual Exp.* 2014:e50971. doi: 10.3791/50971

Friedman, M. (1997). Chemistry, biochemistry, and dietary role of potato polyphenoles. A review. *J. Agric. Food Chem.* 45, 1523–1540. doi: 10.1021/jf960900s

Gebhardt, C., and Valkonen, J. P. T. (2001). Organization of genes controlling disease resistance in the potato genome. *Annu. Rev. Phytopathol.* 39, 79–102.

Gishlain, M., Byarugaba, A. A., Magembe, E., Njoroge, A., Rivera, C., Román, M. L., et al. (2019). Stacking three late blight resistance genes from wild species directly into African highland potato varieties confers complete field resistance to local blight races. *Plant Biotechnol. J.* 17, 1119–1129. doi: 10.1111/pbi.13042

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., et al. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 46, 393–398.

Haesaert, G., Vossen, J. H., Custers, R., De Loose, M., Haverkort, A., Heremans, B., et al. (2015). Transformation of potato variety Désirée with single or multiple resistance genes increases resistance to late blight under field conditions. *Crop Prot.* 77, 163–175. doi: 10.1016/j.cropro.2015.07.018

Hammann, T., Truberg, B., and Thieme, R. (2009). "Improving resistance to late blight (*Phytophthora infestans*) by using interspecific crosses in potato (*Solanum tuberosum* spp.)," in *Crop Plant Resistance to Biotic and Abiotic Factors*, eds F. Feldmann, D. V. Alford, and C. Furk (Braunschweig: Gesellschaft), 407–414.

Hanson, P., Sitathani, K., Sadashiva, A., Yang, R.-Y., Graham, E., and Ledesma, D. (2007). Performance of Solanum habrochaites LA1777 introgression line hybrids for marketable tomato fruit yield in Asia. *Euphytica* 158, 167–178. doi: 10.1007/s10681-007-9440-4

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., et al. (2016). Durable late blight resistance in potato through dynamic varieties obtained by cisgenesis: scientific and societal advances in the DuRPh project. *Potato Res.* 59, 35–66. doi: 10.1007/s11540-015-9312-6

Haverkort, A. J., Struik, P. C., Visser, R. G. F., and Jacobsen, E. (2009). Applied biotechnology to combat late blight in potato caused by *Phytophthora infestans*. *Potato Res.* 52, 249–264. doi: 10.1007/s11540-009-9136-3

Hein, I., Birch, P. R. J., Lefebvre, S. D. V., Odeny, D. A., Gebhardt, C., Danan, S., et al. (2009). Progress in mapping and cloning qualitative and quantitative resistance against *Phytophthora infestans* in potato and its wild relatives. *Potato Res.* 52, 215–227. doi: 10.1007/s11540-009-9129-2

Hermsen, J. G. T., and Ramanna, M. S. (1973). Double-bridge hybrids of *Solanum bulbocastanum* and cultivars of *Solanum tuberosum*. *Euphytica* 22, 457–466. doi: 10.1007/bf00036641

Huang, S., van der Vossen, E. A. G., Kuang, H., Vleeshouwers, V. G. A. A., Zhang, N., Borm, T. J. A., et al. (2005). Comparative genomics enabled the isolation of the R3a late blight resistance gene in potato. *Plant J.* 42, 251–261. doi: 10.1111/j.1365-313x.2005.02365.x

Huang, S., Vleeshouwers, V. G., Werij, J. S., Hutten, R. C., van Eck, H. J., Visser, R. G. F., et al. (2004). The R3 resistance to *Phytophthora infestans* in potato is conferred by two closely linked R genes with distinct specificities. *Mol. Plant Microb. Interact* 17, 428–435. doi: 10.1094/mpmi.2004.17.4.428

Jo, K.-R., Zhu, S., Bai, Y., Hutten, R. C. B., Kessel, G. J. T., Vleeshouwers, V. G. A. A., et al. (2016). "Problematic crops: 1. potatoes: towards sustainable potato late blight resistance by cisgenic R gene pyramiding," in *Plant Pathogen Resistance Biotechnology*, ed. D. B. Collinge (Hoboken, NJ: John Wiley & Sons Inc), 171–191. doi: 10.1002/9781118867716.ch9

Johnston, S. A., den Nijs, T. P. M., Peloquin, S. J., and Hanneman, R. E. Jr. (1980). The significance of genic balance to endosperm development in interspecific crosses. *Theor. Appl. Genet.* 57, 5–9. doi: 10.1007/bf00276002

Kamoun, S., and Smart, C. D. (2005). Late blight of potato and tomato in the genomics era. *Plant Disease* 89, 692–699. doi: 10.1094/pd-89-0692

Kruppa, K., Türkösi, E., Szakács, É, Cseh, A., and Molnár-Láng, M. (2013). Development and identification of a 4HL.5DL wheat/barley centric fusion using GISH, FISH and SSR markers. *Cereal Res. Commun.* 41, 221–229. doi: 10.1556/crc.2012.0038

Li, G. C., Huang, S. W., Guo, X., Li, Y., Yang, Y., Guo, Z., et al. (2011). Cloning and characterization of R3b members of the R3 superfamily of late blight resistance genes show sequence and functional divergence. *Mol. Plant Microb. Interact* 24, 1132–1142. doi: 10.1094/mpmi-11-10-0276

Lokossou, A. A., Park, T.-H., van Arkel, G., Arens, M., Ruyter-Spira, C., Morales, J., et al. (2009). Exploiting knowledge of R/Avr genes to rapidly clone a new LZ-NBS-LRR family of late blight resistance genes from potato linkage group IV. *Mol. Plant Microb. Interact* 22, 630–641. doi: 10.1094/mpmi-22-6-0630

Lokossou, A. A., Rietman, H., Wang, M., Krenek, P., van der Schoot, H., Henken, B., et al. (2010). Diversity, distribution, and evolution of *Solanum bulbocastanum* late blight resistance genes. *Mol. Plant Microb. Interact* 23, 1206–1216.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassay with tobacco tissue culture. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Oosumi, T., Rockhold, D., Maccree, M., Deahl, K., McCue, K., and Belknap, W. (2009). Gene Rpi-bt1 from *Solanum bulbocastanum* confers resistance to late blight in transgenic potatoes. *Am. J. Potato Res.* 86, 456–465. doi: 10.1007/s12230-009-9100-4

Orbegozo, J., Roman, M. L., Rivera, C., Gamboa, S., Tovar, J. C., Forbes, G. A., et al. (2016). Rpi-blb2 gene from *Solanum bulbocastanum* confers extreme resistance to late blight disease in potato. *Plant Cell Tiss. Organ. Cult.* 125, 269–281. doi: 10.1007/s11240-016-0947-z

Paro, R., Triboni, G. M., Buccione, R., Rossi, G., Cellini, V., Canipari, R., et al. (2012). The fungicide mancozeb induces toxic effects on mammalian granulosa cells. *Toxicol. Appl. Pharmacol.* 260, 155–161. doi: 10.1016/j.taap.2012.02.005

R Development Core Team (2010). *Version 2.11.1*. Available online at: www.lsw.uni-heidelberg.de/users/christlieb/teaching/./R-refman.pdf (accessed September 2016).

Rakosy-Tican, E., Thieme, R., Nachtigall, M., Molnar, I., and Denes, T.-E. (2015). The recipient potato cultivar influences the genetic makeup of the somatic hybrids between potato cultivars and sexually incompatible species *Solanum bulbocastanum* Dun. *Plant Cell Tiss. Organ. Cult.* 122, 395–407. doi: 10.1007/s11240-015-0777-4

Rietman, H. (2011). *Putting the Phytophthora Infestans Genome Sequence At Work; Multiple Novel Avirulence And Potato Resistance Gene Candidates Revealed*. Ph. D. thesis, Wageningen University. Wageningen.

Rodewald, J., and Trognitz, B. (2013). *Solanum* resistance genes against *Phytophthora infestans* and their corresponding avirulence genes. *Mol. Plant Pathol.* 14, 740–757. doi: 10.1111/mpp.12036

Shaner, G., and Finney, R. E. (1977). The effect of nitrogen fertilization on the expression of slow-mildewing resistance in Knox wheat. *Phytopathology* 67, 1051–1056.

Shi, X., Tian, Z., Liu, J., van der Vossen, E. A. G., and Xie, C. (2012). A potato pathogenesis-related protein gene, StPRp27, contributes to race-nonspecific resistance against *Phytophthora infestans*. *Mol. Biol. Rep.* 39, 1909–1916. doi: 10.1007/s11033-011-0937-5

Simko, I. (2002). Comparative analysis of quantitative trait loci for foliage resistance to *Phytophthora infestans* in tuber-bearing *Solanum* species. *Am. J. Potato Res.* 79, 25–132.

Song, J., Bradeen, J. M., Naess, S. K., Raasch, J. A., Wielgus, S. M., Haberlach, G. T., et al. (2003). Gene RB cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9128–9133. doi: 10.1073/pnas.1533501100

Storck, T., Böhme, T., and Schultheiss, H. (2012). Status and perspectives of GM approaches to fight late blight. Thirteenth EuroBlight workshop, St. Petersburg (Russia). *PPO Spec. Rep.* 15, 45–48.

Tan, M. Y. A., Hutten, R. C. B., Visser, R. G. F., and van Eck, H. J. (2010). The effect of pyramiding *Phytophthora infestans* resistance genes RPi-mcd1 and RPi-ber in potato. *Theor. Appl. Genet.* 121, 117–125. doi: 10.1007/s00122-010-1295-8

Thieme, R., and Rakosy-Tican, E. (2017). "Somatic cell genetics and its application in potato breeding," in *The Potato Genome*, eds S. K. Chakrabarti, C. Xie, and J. K. Tiwari (Cham: Springer International Publishing), 217–269.

Thieme, R., Rakosy-Tican, E., Gavrilenko, T., Antonova, O., Schubert, J., Nachtigall, M., et al. (2008). Novel somatic hybrids and their fertile BC1 progenies of potato (*Solanum tuberosum* L.) + *S. tarnii*, extremely resistant to potato virus Y and resistant to late blight. *Theor. Appl. Genet.* 116, 691–700. doi: 10.1007/s00122-007-0702-2

Thieme, R., Rakosy-Tican, E., Nachtigall, M., Schubert, J., Hammann, T., Antonova, O., et al. (2010). Characterization of multiple resistance traits of somatic hybrids between *Solanum cardiophyllum* Lindl. and two commercial potato cultivars. *Plant Cell Rep.* 29, 1187–1201. doi: 10.1007/s00299-010-0905-x

Truberg, B., Hammann, T., Darsow, U., and Piepho, H.-P. (2009). Empirischer Vergleich verschiedener Methoden zur Reifekorrektur von Daten zum Befall mit Krautfäule (*Phytophthora infestans* (Mont.) de Bary) in Selektionsexperimenten bei der Kartoffel (*Solanum tuberosum* subsp. *tuberosum*). *J. Kulturpfl.* 61, 77–81.

Van der Vossen, E. A. G., Gros, J., Sikkema, A., Muskens, M., Wouters, D., Wolters, P., et al. (2005). The Rpi-blb2 gene from *Solanum bulbocastanum* is an Mi-1 gene homolog conferring broad spectrum late blight resistance in potato. *Plant J.* 44, 208–222. doi: 10.1111/j.1365-313x.2005.02527.x

Van der Vossen, E. A. G., Sikkema, A., Hekkert, B., Gros, J., Stevens, P., Muskens, M., et al. (2003). An ancient R gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. *Plant J.* 36, 867–882.

Vleeshouwers, V. G. A. A., Raffaele, S., Vossen, J. H., Champouret, N., Oliva, R., Segretin, M. E., et al. (2011). Understanding underexploiting late blight resistance in the age of effectors. *Annu. Rev. Phytopathol.* 49, 507–531. doi: 10.1146/annurev-phyto-072910-095326

Vleeshouwers, V. G. A. A., Rietman, H., Krenek, P., Champouret, N., Young, C., Oh, S.-K., et al. (2008). Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. *PLoS One* 3:e2875. doi: 10.1371/journal.pone.0002875

Vreugdenhil, D. (ed.) (2007). *Potato Biology and Biotechnology*. Amsterdam: Elsevier.

Wang, J., Jiang, J., and Wang, Y. (2013). Protoplast fusion for crop improvement and breeding in China. *Plant Cell Tissue Organ. Cult.* 112, 131–142. doi: 10.1007/s11240-012-0221-y

Wang, M., Allefs, S., van den Berg, R. G., Vleeshouwers, V. G. A. A., van der Vossen, E. A. G., and Vosman, B. (2008). Allele mining in solanum: conserved homologues of Rpi-blb1 are identified in *Solanum stoloniferum*. *Theor. Appl. Genet.* 116, 933–943. doi: 10.1007/s00122-008-0725-3

Wang-Pruski, G. (2007). The canon of potato science: 47. After-cooking darkening. *Potato Res.* 50, 403–406. doi: 10.1007/s11540-008-9061-x

Zhu, S., Li, Y., Vossen, J. H., Visser, R. G. F., and Jacobsen, E. (2012). Functional stacking of three resistance genes against *Phytophthora infestans* in potato. *Transgen. Res.* 21, 89–99. doi: 10.1007/s11248-011-9510-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 12 June 2020
doi: 10.3389/fpls.2020.00761

Introgression Breeding in Barley: Perspectives and Case Studies

Javier Hernandez, Brigid Meints and Patrick Hayes*

Department Crop and Soil Science, Oregon State University, Corvallis, OR, United States

Changing production scenarios resulting from unstable climatic conditions are challenging crop improvement efforts. A deeper and more practical understanding of plant genetic resources is necessary if these assets are to be used effectively in developing improved varieties. In general, current varieties and potential varieties have a narrow genetic base, making them prone to suffer the consequences of new and different abiotic and biotic stresses that can reduce crop yield and quality. The deployment of genomic technologies and sophisticated statistical analysis procedures has generated a dramatic change in the way we characterize and access genetic diversity in crop plants, including barley. Various mapping strategies can be used to identify the genetic variants that lead to target phenotypes and these variants can be assigned coordinates in reference genomes. In this way, new genes and/or new alleles at known loci present in wild ancestors, germplasm accessions, land races, and un-adapted introductions can be located and targeted for introgression. In principle, the introgression process can now be streamlined and linkage drag reduced. In this review, we present an overview of (1) past and current efforts to identify diversity that can be tapped to improve barley yield and quality, and (2) case studies of our efforts to introgress resistance to stripe and stem rust from un-adapted germplasm. We conclude with a description of a modified Nested Association Mapping (NAM) population strategy that we are implementing for the development of multi-use naked barley for organic systems and share perspectives on the use of genome editing in introgression breeding.

Keywords: genetic resources, multi-rust resistance, haplotype, high throughput genotyping, genetic diversity, genetic mapping

OPEN ACCESS

Edited by:
Mariola Plazas, Polytechnic University of Valencia, Spain

Reviewed by:
Michael Baum, International Center for Agricultural Research in the Dry Areas (ICARDA), Morocco
Shengguan Cai, Zhejiang University, China

****Correspondence:***
Javier Hernandez
hernanfr@oregonstate.edu

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *10 March 2020*
Accepted: *13 May 2020*
Published: *12 June 2020*

Citation:
Hernandez J, Meints B and Hayes P (2020) Introgression Breeding in Barley: Perspectives and Case Studies. Front. Plant Sci. 11:761. doi: 10.3389/fpls.2020.00761

INTRODUCTION

Introgression breeding has been an important method for improving barley since domestication, and it remains a key tool for expanding genetic diversity to meet current and future challenges to crop production. For the purposes of this chapter, we will define introgression as the transfer of one or several novel, favorable alleles from un-adapted germplasm to adapted germplasm. Barley serves as an excellent example for charting the history, current status, and future prospects for introgression breeding because it is a diploid genetic model for the Triticeae tribe and an important and versatile crop grown (nearly) from pole to pole. In our review, we will discuss the unique features of this crop, chart the evolution of tools for managing the introgression process, and look ahead to how introgression could be enhanced for both organic and conventional production systems.

DOMESTICATION, INTROGRESSION AND CURRENT STATUS

The domestication of cultivated barley (*Hordeum vulgare* L) from *Hordeum vulgare subsp. spontaneum* Koch) began ∼10,000 years before present (Zohary and Hopf, 2000). Key domestication traits included determinate growth habit, increased seed set, greater inflorescence number, non-shattering, larger seed size, and more rapid germination (Harlan et al., 1973). These early domestication efforts surely involved a first step towards managed introgression. As early domesticators identified plants with novel phenotypes, they exchanged seeds of these plants with neighbors. Blended with existing seed stocks, the novel seeds would have led to the formation of heterogeneous mixtures of nearly homozygous lines (land races). Selection for novel phenotypes would increase their frequency in these land races, and the naturally occurring outcrossing (∼2%) that occurs in barley (Abdel-Ghani et al., 2004) and/or the environmentally induced outcrossing the can occur in selfing species (van Ginkel and Flipphi, 2020) would increase the frequency of favorable alleles introgressed into locally adapted genomes. Interestingly, many of the initial domestication traits remain critical in modern barley breeding.

The initial domestication of barley, based on archeological evidence, most likely occurred in the Fertile Crescent (Harlan and Zohary, 1966) with other possible sites in Central Asia and Africa (Morrell and Clegg, 2007; Dai et al., 2012). From these initial domestication sites, barley moved quickly into Europe and Asia, with mechanical mixtures and outcrossing facilitating a march estimated to have occurred at a pace of nearly 10 km/year (Morrell and Clegg, 2007). Ten thousand years after domestication began, barley is the fourth most widely grown cereal after wheat, maize, and rice and is planted in a wide range of environments around the world (Hayes et al., 1993; FAOSTAT, 2017).

Two phenotypes that today define the principal germplasm groups of cultivated barley were selected ∼8–9,000 years ago: inflorescence type (two-row vs six-row) and hull adherence type (covered vs naked). The fertility of the triad of florets at each rachis node determines the head type barley: two-row (ancestral) or six-row (selected post-domestication). In the former, only the central florets are fertile, whereas in six-row types, all three florets are fertile. The *Vrs1* gene responsible for head type was cloned by Komatsuda et al. (2007) and these authors described a single dominant allele and a number of loss-of-function six-row alleles that may have been selected at different places and times. The predominance of a particular allele/head type in a geographical region can, in some cases, be traced to which type of inflorescence was first introduced (introgressed) and certainly to end use. Two-row types are used predominately in the brewing and distilling industries because of the higher likelihood of uniform, plump kernels (Schwarz and Li, 2011). Because barley was, and is, used primarily for malting and distilling in Europe, the wild type, two-row allele predominates in the region. The six-row phenotype, in contrast, was selected and maintained in North Africa, the Iberian Peninsula, and eastern Asia, where brewing was not as prevalent. Both head types are present in regions where barley is used for feed and food. An interesting case study in introgression and head type is the rapid switch from six-row malting types to two-row malting types in North America that occurred in response to a re-direction of the malting and brewing industries in the 1990's. Throughput the 20th century, the majority of North American malting barley was six-row due to perceived positive impact on beer flavor: today the American Malting Barley Association no longer supports research on six-row barley nor does it accept six-rows into its evaluation process (Craft Brewing Business, 2019). The rapid introgression of two-row spike morphology, in response to this shift, was accomplished thanks to targeted introgression of the *Vrs1* allele in two-row by six-row crosses and by the introduction of two-row varieties from Europe.

Today, most cultivated barleys are covered (hulled), meaning the lemma and palea adhere to the pericarp. Taketa et al. (2008) cloned the *Nud* gene, which is responsible for hull adherence. In *nud* genotypes, the seed threshes clean, as in wheat. The preferred botanical term for this phenotype is "naked," although the term "hull-less" is also common. Covered types are preferred by the malting and brewing industries, because the hulls are used as natural filters during the brewing process (Newman and Newman, 2008). Barley varieties selected for feed production are also often covered, as selection for this end-use is based primarily on grain yield. It is important to note, however, that the hull accounts for ∼10% of the yield and is composed of insoluble fiber (Rey et al., 2009). Due to the higher economic value of malting barley compared to feed and food barley, in Europe and North American the focus of breeding efforts has been on agronomic and quality performance in covered two-row types destined for the malting and brewing industry (Newman and Newman, 2008; Meints et al., 2016). Naked barley is preferred for human consumption as hull removal requires additional processing, e.g., mechanical removal of the hull by "pearling" (Meints and Hayes, 2020). Naked barley is a staple food crop in the Himalayan region, the Andes, and the Ethiopian highlands. In Morocco, average consumption was recently reported at 28 kg/year (Aldughpassi et al., 2016). Barley is currently gaining popularity in western diets due to its health and nutritional benefits (Meints et al., 2016). Naked barley currently represents a small percentage of world barley production, as most barley is grown for feed and malt (Newman and Newman, 2008). However, systematic introgression of the naked phenotype is a goal of an ongoing collaborative breeding effort in North America to develop naked multi-use barleys for organic systems, as described later in this chapter.

Introgression From Genetically Diverse Sources

With the rediscovery of Mendel's work and the application of it to plant breeding, introgression in barley was made more systematic via controlled crossing. Specific examples from the early days of barley breeding are not obvious, most likely due to a focus on quantitative traits such as yield and resistance to some diseases. With a focus on improvements in yield and grain quality, a common breeding strategy was based on crossing elite by elite

material. This led to an inevitable narrowing of the germplasm base (Bernardo, 2014) although, interestingly, selection responses were still achieved. Rasmusson and Phillips (1997) explored this question of continued response to selection in six-row malting barley adapted to the Upper Midwest of North America. Their insights and hypotheses were, alas, not followed up on in a systematic fashion due to the aforementioned curtailment of six-row malting barley production in favor of two-row types.

The recognition that genetic vulnerability and yield plateaus are an inevitable consequence of a narrow germplasm base (Gepts, 2006; McCouch et al., 2013) prompted a systematic search for usable genetic variation in the ancestor of wild barley (*H. vulgare subsp. spontaneum*), land races, and un-adapted germplasm. Recognizing that the distinction between land races and un-adapted germplasm is vague, much of the literature on expanding diversity in locally adapted, cultivated barley is focused on characterization with fewer concrete examples of introgression. While a comprehensive cataloging of germplasm characterization efforts and subsequent introgressions is not within the scope of the current review – the reader is referred to Von Bothmer et al. (2003) we will mention a few illustrative examples. Unique considerations and challenges apply to each of these classes of germplasm – in general it is more difficult to access useful alleles in *H. vulgare subsp. spontaneum*, and not as daunting for unadapted germplasm or land races. These considerations include cross incompatibility, infertility, reduced recombination, and introgression of undesirable alien genome segments resulting in linkage drag.

Starting with *H. vulgare subsp. spontaneum* (hereafter referred to as spontaneum), the potential value of the ancestral species has been well-documented via systematic characterization of phenotypic and genetic variation (Bedada et al., 2014; Sallam et al., 2017); Ongoing efforts to introgress low temperature tolerance alleles from spontaneum are promising (B. Steffenson, personal communication; Lei et al., 2019). Matus et al. (2003) developed a set of recombinant chromosome substitution lines (RCSLs) using a spontaneum donor and an elite cultivar recurrent parent. One of the RCSLs (RCSL-124) advanced to an on-farm trial for commercial assessment prior to release as a variety. Unfortunately, it did not have a yield advantage over the best available feed variety and therefore was as not released (unpublished data). Likewise, spontaneum is the source of novel lipoxygenase (LOX) alleles (Hirota et al., 2008) but the commercially deployed allele was identified in mutants generated in cultivated barley (Skadhauge et al., 2016).

Land races are often described as reservoirs of useful genetic variation for barley improvement and have been used for that purposed (Monteagudo et al., 2019). Historically, land races were a key resource for introgressing alleles into pure line varieties – an example is the *Rpg1* allele tracing to the landrace "Chevron" that subsequently protected North American barley from stem rust (incited by *Puccinia graminus* f.sp. *tritici*) for decades (Steffenson, 1992). In general, however, introgression of favorable alleles from land races into adapted germplasm involves choosing specific exemplars (accessions) for crossing – and this selection can obviate the stated advantages of the land race – which include heterogeneity and potentially heterozygosity (Poets et al., 2015). There is a rich literature on the improvement of land races, particularly in the context of farmer participatory plant breeding, and this was a key emphasis in the ICARDA barley improvement program, formerly based in Syria (Ceccarelli and Grando, 2000; Ceccarelli et al., 2000).

Germplasm collections such as the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK, Germany), the Okayama University Barley and Wild Plant Resource Center (Japan), the International Center for Agricultural research in the Dry Area (ICARDA) and the United States Department of Agriculture National Small Grains Collection (USDA-NSGC, United States) are excellent sources of genetic diversity. The latter contains 29,870 barley accessions including cultivars, breeding lines, land races, wild relatives and genetic stock from more than 100 countries (Bockelman and Valkoun, 2011) and has been extensively characterized for a range of economically important traits (Dahleen et al., 2012; Muñoz-Amatriaín et al., 2014a; Hemshrot et al., 2019). These collections, and others like them, are a rich source of germplasm for finding novel alleles for disease resistance (Czembor, 2000; Yun et al., 2006) drought tolerance (Talamé et al., 2004; Monteagudo et al., 2019) cold tolerance (Visioni et al., 2013; TCAP, 2014) yield (Nice et al., 2019) and other critical traits. One specific example of the effective use of the United States collection is introgressing resistance to the Russian Wheat Aphid (*Diuraphis noxia*). The cultivar "Burton" was developed with RWA resistance contributed by PI 366450 from Afghanistan (Bregitzer et al., 2005). In the case of the Okayama University collection, its accession have served as donors of alleles conferring resistance to Barley Yellow Mosaic Virus (Okada et al., 2004).

One consideration with documenting the effective use of germplasm collections is the time interval between introgression and variety release: pre-breeding can be a lengthy process that is not necessarily amenable to publication in peer reviewed journals. The availability of high-throughput tools, described in the next section, is setting the stage for effective introgression form germplasm collections.

Current Tools for Assessing and Exploiting Genetic Variation

Next-generation sequencing (NGS) technologies have provided cost-effective methods for surveying genome-wide variation and optimistically will facilitate not only germplasm characterization but also cost-effective and efficient introgression breeding. The use of high density single nucleotide polymorphism (SNP) genotyping platforms (Comadran et al., 2012; Bayer et al., 2017) has provided valuable insights into population structure in barley germplasm arrays that, in turn, generate clues regarding domestication, geographical origin, migration, recombination, and allelic diversity (Muñoz-Amatriaín et al., 2014b; Voss-Fels et al., 2015; Milner et al., 2019). Two of these platforms figure prominently in our own introgression efforts. Data from the re-sequencing of transcriptomes was used to develop the Illumina Infinium 9K assay, in which

7,842 SNPs can be tested simultaneously (Comadran et al., 2012). A more recent genotyping array was constructed based on DNA variant calling using exon capture (EC) in a range of European barley germplasm. This new Illumina Infinium 50K iSelect genotyping array integrates previous genotyping information from the 9K array to provide 43,461 SNPs (6,251 9K + 37,789 EC) that are available for genetic mapping and diversity analysis (Bayer et al., 2017). The sequenced barley genome (Mascher et al., 2017) along with bioinformatic tools, will facilitate the introgression of targeted genomic regions identified during the genetic characterization of diverse germplasm. Examples of this integration of SNP variation with barley genome sequence coordinates are provided in the haplotype visualizations we generated to describe outcomes of our introgression breeding efforts, as detailed in the following sections(**Figures 1a**, **1b**, **1c**).

A drawback to any array is ascertainment bias: the true variants affecting target traits, particularly INDELs, may not be represented in the germplasm used to develop the SNP array (Ganal et al., 2009; Davey et al., 2011). This drawback can be overcome, to some extent, by relying on the linkage disequilibrium (LD) of SNPs that are in LD with causal genes that are underlying the targeted phenotypic differences (Flint-Garcia et al., 2003; Myles et al., 2009; Lipka et al., 2015). Besides this potential downside, the use of NGS methods to rapidly discover thousands of genetic variants in coding or non-coding regions is becoming a standard tool for plant breeders to characterize existing germplasm, analyze genes/QTLs underlying traits of interest, estimate breeding values based on genotypic information, conduct marker assisted selection (MAS) and genomic selection (GS), and target specific alleles in the population (Muñoz-Amatriaín et al., 2014b; Varshney et al., 2014).

Of equal importance to high throughput genotyping are tools for identifying significant marker associations with traits of interest. Bi-parental populations and genome wide association studies (GWAS) are widely used by barley breeders and geneticists to reveal the genetic architecture of simple and complex traits. In barley, there are many examples of crossing two dissimilar parents to dissect simple and complex traits (Castro et al., 2003; Fisk et al., 2013; Esvelt et al., 2016). However, this approach has the constraints of testing just two alleles per locus at a time, low mapping resolution due to limited recombination events, and unrevealed polymorphisms between parents in some genomic regions linked with the trait (Bernardo, 2008; Würschum, 2012). GWAS takes advantage of historical recombination present in an uncontrolled population, which allows for higher resolution mapping (Rafalski, 2010; Bush and Moore, 2012). GWAS has been used in barley for characterizing the genetic basis of traits including growth habit, disease resistance, phenology, and end-use quality (Cuesta-Marcos et al., 2010; Visioni et al., 2013; Muñoz-Amatriaín et al., 2014a; Graebner et al., 2015; Gutiérrez et al., 2015), disease resistance (Sallam et al., 2017; Case et al., 2018), and drought and salt tolerance (Thabet et al., 2018; Xue et al., 2019). Quantitative trait loci (QTL) identified through either or both of these methods can then be targeted via (MAS) and/or used to monitor the effects of GS.

The identification of QTL through the integration of genotypic and phenotypic information sets the stage for introgression. In principle, this is as straightforward as using QTL as a platform for MAS (Babu et al., 2004; St Clair, 2010). Based on this approach, wild type barley accessions are a rich source of favorable alleles for yield, malting quality and disease resistance in barley. As an example, lines derived from *Hordeum bulbosum*, a secondary barley gene pool, have been used to characterize resistance and agronomic relevant traits (Pickering et al., 2004; Johnston et al., 2013; Czembor et al., 2019; Hoseinzadeh et al., 2020). Accessions derived from *spontaneum* are also a source of novel and potentially useful alleles (Matus et al., 2003; von Korff et al., 2008; Nice et al., 2019). The dilemma is that as the donors get more exotic, the more likely it is that there will be linkage drag. Matus et al. (2003) found that the *spontaneum* accession Caesarea 26–24 had a negative effect on the variety Harrington in terms of agronomic performance and malting quality. For this reason, during the introgression process, it is important to reduce the size of the chromosome section carrying the targeted genomic region. This is important because potentially undesirable genes with negative effects on important traits may be physically linked with the target donor allele(s) (Hospital, 2005). Without the use of markers defining QTL regions, these donor segments can be quite large, which increases the chance of undesirable genes ending up in the recurrent genetic background (Ribaut et al., 2002; Salina et al., 2003). One example of the application of markers to reduce linkage drag is the advanced back-cross method. This approach has been used in barley for several traits including disease resistance, malting quality, and yield (Matus et al., 2003; von Korff et al., 2008). High-resolution genotyping technologies can assist in overcoming the problem of linkage drag by providing better map/sequence resolution of the target allele(s) and, as a result, a reduction in the size of the introgressed DNA segments. However, even when the size of the introgression segments can be successfully reduced, the favorable alleles from the exotic germplasm may not have predictable phenotypic effects in new genetic backgrounds (Richardson et al., 2006). Therefore, a validation process of assessing novel qualitative, or quantitative, trait alleles is warranted (Bilgic et al., 2005; Richardson et al., 2006; Sharma et al., 2018; Hernandez et al., 2019).

A CASE STUDY IN CHARACTERIZATION AND INTROGRESSION: MULTI-RUST RESISTANCE

Stripe rust (incited by *Puccinia striiformis* f. sp. *hordei*) and stem rust (incited by *Puccinia graminis* f. sp. *tritici*) are barley diseases of worldwide importance. Stripe rust resistance has long been a focus of our program, due to its prevalence in the Pacific Northwest of the United States. Briefly, barley stripe rust (BSR) was first reported in the Americas in 1976, when it was discovered in Colombia (Dubin and Stubbs, 1986). The disease spread throughout the Americas, arriving in the

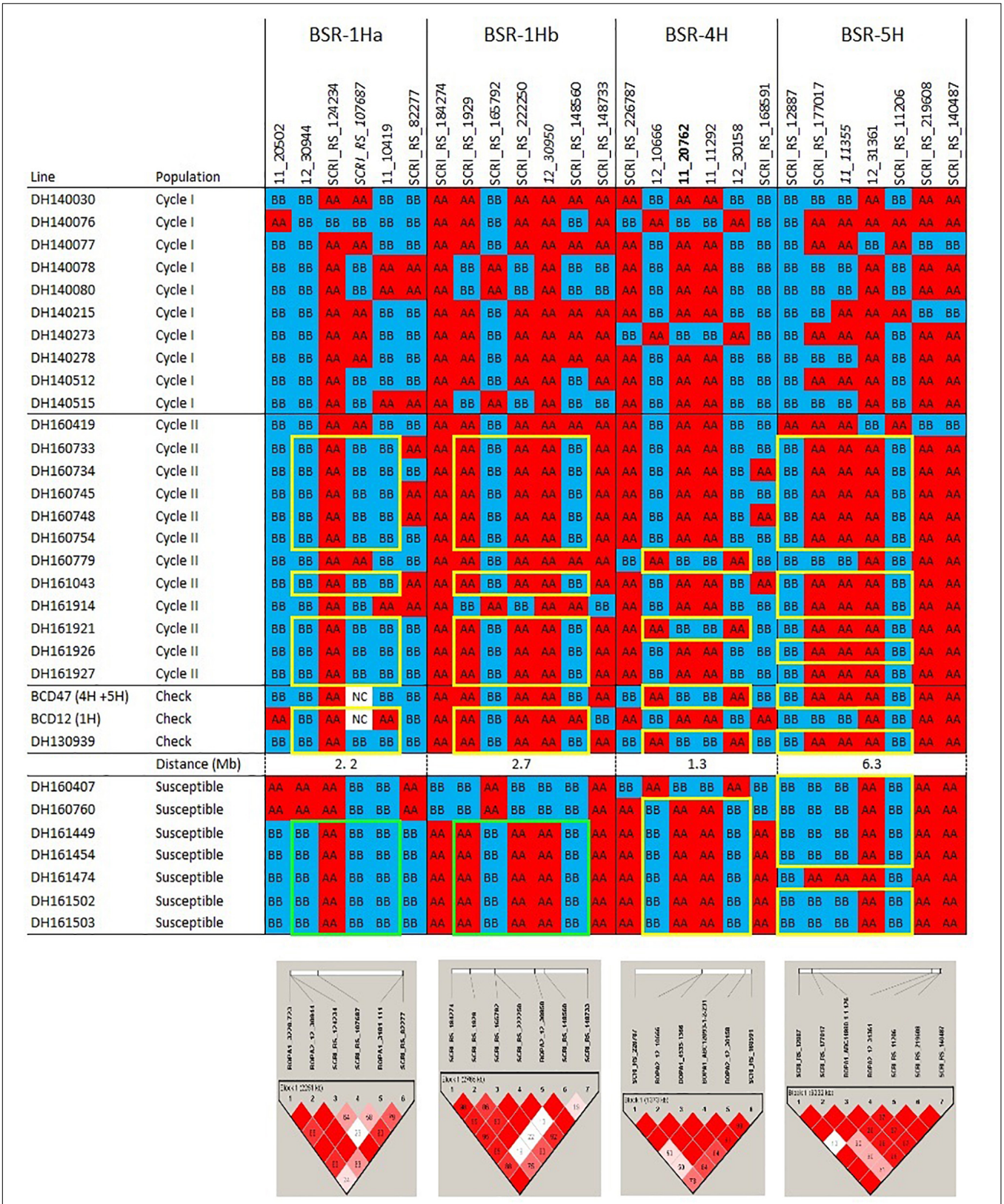

FIGURE 1a | Haplotypes and linkage disequilibrium heat maps based on high density SNP genotyping of Cycle I and Cycle II introgression lines that are resistant to barley stripe rust compared to resistant and susceptible checks. Details on the resistance QTLs are provided in the narrative. The size of each QTL interval (in Mb) is inferred from the barley consensus sequence. Most significant SNPs are in bold. Closer SNP from most significant marker in italic.

		BSR-5H							SR.5H-MN18							SR.5H-MN19							*rpg4/Rpg5*									
Line	Population	SCRI_RS_12887	SCRI_RS_177017	*11_11355*	12_31361	SCRI_RS_11206	SCRI_RS_219608	SCRI_RS_140487	JHI-Hv50k-2016-311731	JHI-Hv50k-2016-311734	JHI-Hv50k-2016-311764	**JHI-Hv50k-2016-311904**	JHI-Hv50k-2016-311906	JHI-Hv50k-2016-311908	JHI-Hv50k-2016-311927	JHI-Hv50k-2016-311989	JHI-Hv50k-2016-312001	JHI-Hv50k-2016-312044	**JHI-Hv50k-2016-312045**	JHI-Hv50k-2016-312058	JHI-Hv50k-2016-312060	JHI-Hv50k-2016-312082	SCRI_RS_194030	11_20334	11_20988	SCRI_RS_118326	**SCRI_RS_65371**	11_20829	SCRI_RS_10924	SCRI_RS_165835	SCRI_RS_159482	11_20826
DH140030	Cycle I	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140076	Cycle I	BB	AA	AA	AA	AA	AA	AA	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
DH140077	Cycle I	BB	AA	AA	BB	AA	BB	BB	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140078	Cycle I	BB	BB	BB	AA	BB	AA	AA	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140080	Cycle I	BB	BB	BB	AA	BB	AA	AA	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140215	Cycle I	BB	BB	AA	AA	AA	BB	BB	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140273	Cycle I	BB	AA	AA	AA	BB	AA	AA	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140278	Cycle I	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140512	Cycle I	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH140515	Cycle I	BB	BB	BB	AA	BB	AA	AA	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH160419	Cycle II	AA	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	BB	AA	BB	BB	AA	AA	AA	AA	BB	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH160733	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
DH160734	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
DH160745	Cycle II	BB	AA	AA	AA	BB	AA	AA	AA	BB	AA	AA	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH160748	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH160754	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
DH160779	Cycle II	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH161043	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH161914	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH161921	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
DH161926	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH161927	Cycle II	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
Q21861	Check	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	AA
UC1322	Check	BB	AA	AA	AA	BB	AA	AA	AA	BB	AA	AA	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
UC1266	Check	AA	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	BB	AA	BB	BB	AA	AA	AA	AA	BB	BB	AA	BB	BB	BB	BB	BB	AA	AA	BB
DH130939	Check	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	BB	AA	AA	AA	AA	AA	BB	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
	Distance (Mb)	6.3							0.69							0.49							2.7									
DH160793	Susceptible	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	BB	AA	BB	BB	BB	BB	BB	AA	AA	BB
DH160825	Susceptible	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	BB	AA	BB	BB	BB	BB	BB	AA	AA	BB
DH161498	Susceptible	BB	BB	BB	AA	BB	AA	AA	BB	AA	BB	BB	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	AA	AA	AA	BB	BB	AA	AA	BB	BB
DH161537	Susceptible	BB	AA	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	AA	AA	AA	AA	BB	BB	AA	AA
DH161930	Susceptible	AA	AA	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH161957	Susceptible	AA	AA	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	BB	BB	BB	BB	BB	BB	AA	BB	BB
DH162199	Susceptible	BB	AA	AA	AA	BB	AA	AA	BB	AA	BB	BB	AA	BB	AA	AA	AA	BB	BB	AA	BB	AA	AA	AA	AA	AA	BB	BB	AA	AA	BB	BB

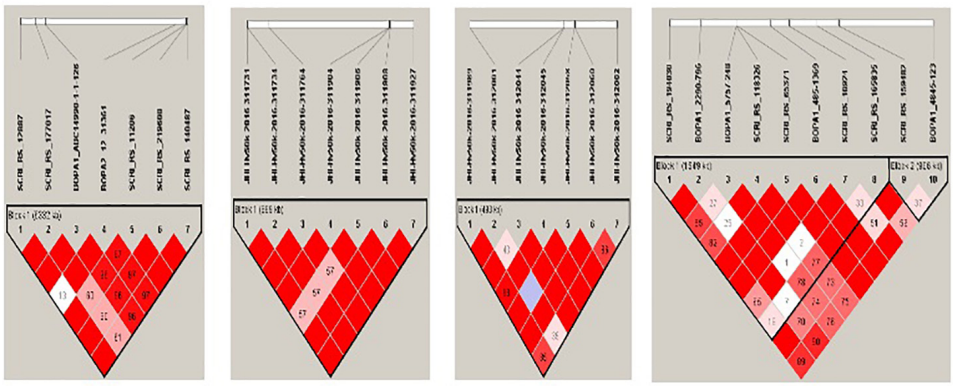

FIGURE 1b | Haplotypes and linkage disequilibrium heat maps based on high density SNP genotyping of Cycle I and Cycle II introgression lines resistant to barley stem rust compared to resistant and susceptible checks. Details on the resistance QTLs and *rpg4/Rpg5* are provided in the narrative. The size of each QTL interval (in Mb) is inferred from the barley consensus sequence. Most significant SNPs are in bold. Closer SNP from most significant marker in italic.

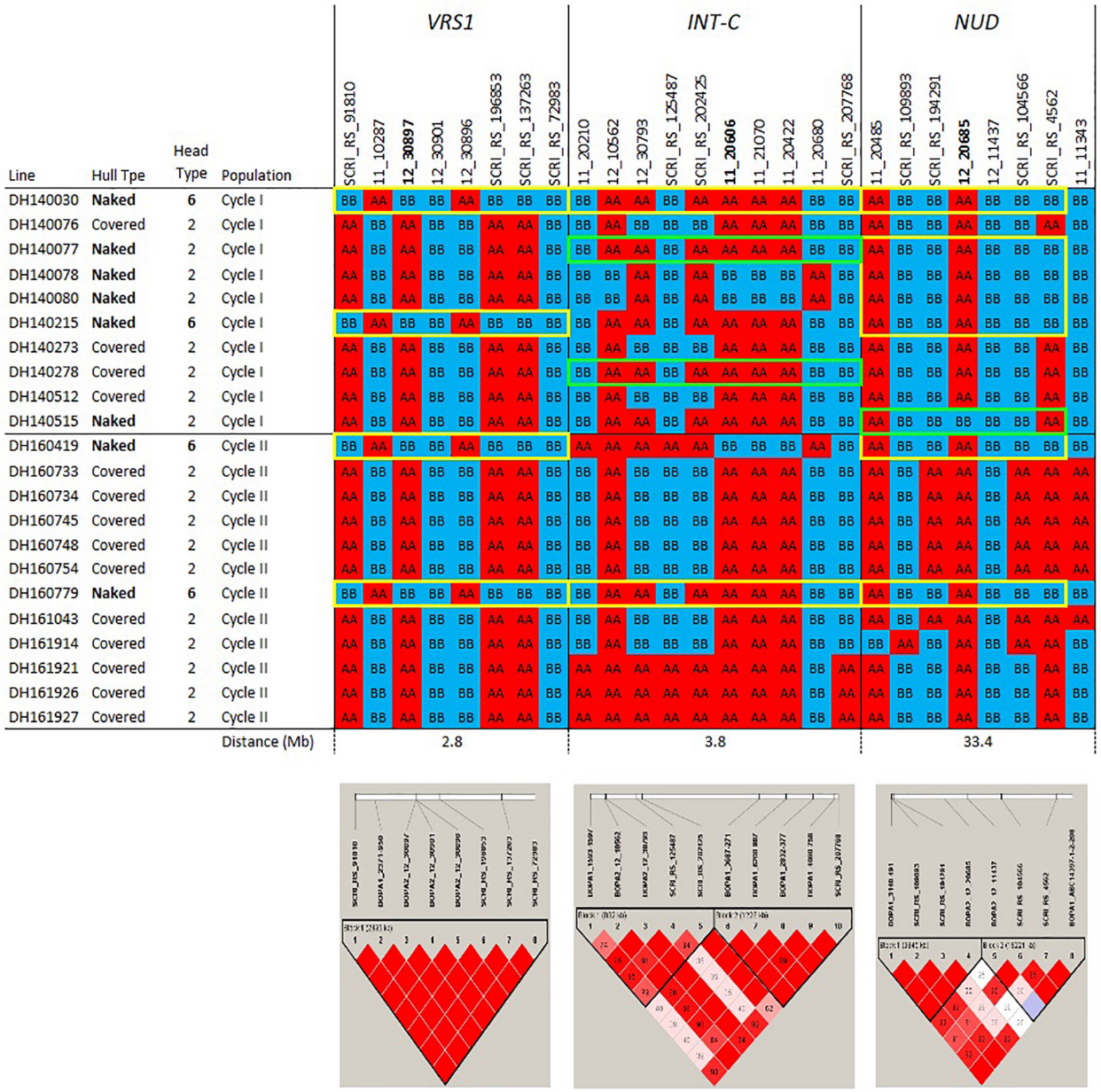

FIGURE 1c | Haplotypes and linkage disequilibrium heat maps based on high density SNP genotyping of selected Cycle I and Cycle II introgression lines at loci determining inflorescence type (*VRS1* and *INT-C*) and hull adherence (*NUD*). Details on the genes determining these morphological traits are provided in the narrative. The size of each introgression interval (in Mb) is inferred from the barley consensus sequence. Most significant SNPs are in bold. Closer SNP from most significant marker in italic.

United States in 1991. A long-term collaboration with the late Dr. Hugo Vivar, who led the former ICARDA barley program based at CIMMYT in Mexico, resulted in extensive literature on mapping resistance genes and QTLs (most recently reviewed by Belcher et al., 2018). Parallel to these efforts we conducted ongoing stripe resistance breeding efforts based primarily on phenotypic selection because (1) mapping efforts were conducted in spring growth habit barley and our breeding program is directed primarily at winter and facultative growth habit barley and (2) phenotypic selection is generally effective at our test sites due to high heritability and consistent natural disease epidemics. Periodically, however, QTL alleles were characterized in germplasm derived from phenotypic selection in our winter and facultative barley program (Belcher et al., 2018). We added stem rust to our resistance breeding efforts due to the threat posed by race TTKSK of this disease, which has yet to be reported in the Americas. Breeding for resistance to these two rusts has allowed us to integrate characterization, validation, and

introgression within a coordinated framework. A discussion of the framework of targeted introgression of resistance to race TTKSK, stripe rust, and, to a limited extent leaf rust now follows (Hernandez et al., 2020).

Cycle I Population

The resistance gene *Rpg1* has been used as the primary source of stem rust resistance since a major epidemic occurred in the United States in the 1940s (Steffenson, 1992). As noted earlier in this chapter, this gene was introgressed from a land race into the principal barley cultivars grown in the upper Midwest of the United States and the Canadian Prairies. During the late 1980s, *Rpg1* was defeated by race QCCJB, demonstrating the urgency of finding new sources of resistance to stem rust (Jin et al., 1994). The urgency became acute with race TTKSK, which is virulent to *Rpg1*: 95% of commercial barley cultivars are susceptible to this race (Steffenson et al., 2017). Fortunately, this race is yet to be reported in the Americas, which provides an opportunity to engage in defensive resistance introgression breeding. An accession – Q21861 – carries the resistance complex *rpg4/Rpg5* which confers resistance to several races of stem rust including QCCJB and TTKSK at all growth stages (Steffenson et al., 2009). This accession was developed by same Dr. Hugo Vivar with whom we had collaborated on stripe rust resistance. This unique and valuable accession can be considered an un-adapted elite line: it does not have the requisite productivity and end-use quality for direct release as a variety in North America. It did, however, provide a starting point for multiple programs to work defensively, via introgression, to prepare for the inevitable arrival of race TTKSK.

The process of introgression of the *rpg4/Rpg5* complex in our program is shown in **Figure 2**. As described by Hernandez et al. (2019), germplasm was obtained from Canadian programs that, in turn, was derived from introgression of *rpg4/Rpg5* from Q21861. This Canadian "pre-breeding" germplasm (five accessions) was used in crosses with varieties and elite germplasm from our own and other breeding programs (both covered and naked) that was resistant to stripe rust. In addition, we used as parents two accessions of Himalayan origin – one from the USDA-GRIN collection and one from the James Hutton Institute (Scotland) collection, and a land race from Washington, United States. One hundred and nineteen doubled haploid (DH) lines were generated from these F1s: this array is referred to as "Cycle I" – the first step in introgressing resistance to TTKSK into adapted germplasm. Using an allele-specific marker for *Rpg5* and SNP data from the Illumina 9K platform, we showed that in the Cycle I population the *rpg4/Rpg5* complex is required but not sufficient to confer stem rust resistance at the seedling stage in a diverse array of genetic backgrounds. Using GWAS, two other loci – one on chromosome 5H and one on chromosome 7H – were found to be associated with resistance and interacting with the *rpg4/Rpg5* complex. Sharma et al. (2018) also reported that an additional gene – in this case denominated *Rrr1* (required for *rpg4*-mediated resistance 1) – is required for the *rpg4/Rpg5* complex to confer resistance when introgressed into the variety "Pinnacle." The Cycle I germplasm was also phenotyped for adult plant resistance to stripe rust and leaf rust (incited by *Puccinia hordei*). Due to the limited number of environments where these diseases occurred, results were not included in the Hernandez et al. (2019) paper. In the context of this chapter, it is worth reporting the GWAS results and integrating them into the stem/stripe rust introgression story. Complete data are available at Barley World[1]. The combined GWAS of all available data (2 years, two locations) identified a significant QTL associated with stripe rust resistance on chromosome 5H at the same position as the adult plant resistance stripe rust QTL we found in Cycle II and reported in Hernandez et al. (2020). Based on one year/location of data, a QTL was identified on chromosome 7H, coincident with a leaf rust QTL reported by Gutiérrez et al. (2015) that traces back to ICARDA/CIMMYT/Mexico germplasm. Ten Cycle I doubled haploids were identified with resistance to all three rusts (**Table 1**). Of these ten, five (highlighted) were selected as parents for a second round of introgression and validation (Cycle II).

Cycle II Population

The Cycle II population is comprized of 358 doubled haploids derived from crosses of the five selected Cycle I lines with five elite lines from our program and three elite, un-adapted lines from the University of California – Davis barley breeding program (Hernandez et al., 2020). The goal of Cycle II was to continue introgression of multi-rust resistance alleles into elite germplasm. In Cycle II we assessed seedling resistance to TTKSK as we did with Cycle I, under tightly controlled environmental conditions, and added assessment of adult plant resistance under field conditions using race QCCJ as a surrogate for race TTKSK. We also assessed adult plant resistance to stripe rust, as we did with Cycle I. Leaf rust epidemics were not sufficiently severe in any of the field trials to generate data for QTL analysis. Reinforcing the importance of phenotyping introgression lines as extensively as possible, using GWAS we found different genes and/or QTL related to resistance to stem rust at the seedling and adult plant stages. While *rpg4/Rpg5* was a principal determinant of resistance at the seedling stage – it was not effective at the adult plant stage in one year. We hypothesized that the difference in resistance at the adult plant stage was due to temperature differences in the two years of testing: it is known that the *rpg4/Rpg5* complex does not confer resistance under high temperature conditions. A QTL on 5H, mapping to a different position than that identified in Cycle I, was associated with adult plant resistance under high temperature, and it is coincident with one of the three QTLs conferring resistance to stripe rust at the adult plant stage. Other significant stripe rust QTLs were identified on 1H and 4H. The 1H QTL is coincident with that reported by Toojinda et al. (2000), which traces to the variety Shyri, released in Ecuador by the ICARDA/CIMMYT program. Subsequently, Castro et al. (2002a) and Richardson et al. (2006) introgressed this allele into susceptible elite backgrounds and validated its effectiveness. The 4H QTL is also coincident with prior reports and traces to Calicuchima-sib, also from the ICARDA/CIMMYT

[1]https://barleyworld.org/barley-stripe-rust-bsr

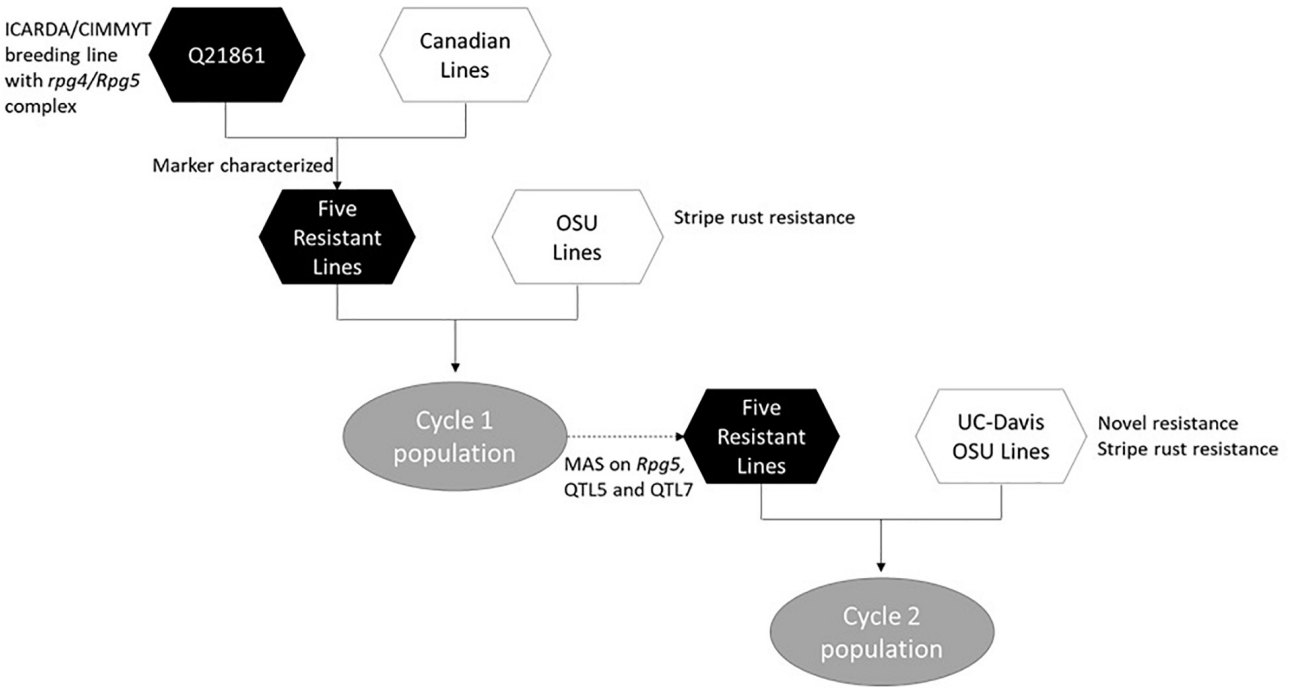

FIGURE 2 | Flow chart showing the introgression process involved in development of the Cycle I and Cycle II barley populations.

program (Chen et al., 1994). This 4H allele was subsequently validated in other elite genetic backgrounds after further cycles of introgression (Castro et al., 2002b) and was independently identified in unrelated germplasm by Esvelt et al. (2016). Calicuchima-sib was also the donor of the 5H QTL (Chen et al., 1994; Castro et al., 2002b). We have selected 12 doubled haploids from Cycle II with resistance to both stem and stripe rust. In summary, the introgression of multiple alleles from different regions of the genome was successful in conferring resistance to stripe, stem, and leaf rust. In the case of stripe rust and stem rust, the introgressed resistance alleles trace to the ICARDA/CIMMYT program based in Mexico – testimony to the effectiveness of this program in pyramiding resistance based on phenotype alone. Further research is needed to validate the leaf rust resistance QTL allele.

The pedigrees and phenotypes of the 22 doubled haploids selected from Cycles I and II are provided in **Table 1** and this germplasm is freely available for research purposes. Haplotype analysis provides insights into the genetic architecture of these introgression lines and addresses key issues in introgression breeding, such as the discriminatory power of marker haplotype information, extent of LD, and how these alleles interact with the genetic background they are introgressed into. The most phenotypically resistant lines from Cycle I and Cycle II were used to identify haplotypes associated with biotic and morphological traits based on high throughput genotyping arrays (**Figures 1a**, **1b**, **1c**). In the case of stripe rust resistance, a defined pattern is observed in lines carrying the resistance haplotype on 5H and 4H. These same haplotypes are observed in the resistant check BCD47 (Castro et al., 2003) and DH130939, a facultative breeding line with phenotypic resistance. For the QTLs on 1H, the haplotype is not as predictive. Q21861 is a well-known source of stem rust resistance carrying both *Rpg1* and the *rpg4/Rpg5* complex. This resistance was the foundation for mapping and introgression of stem rust resistance into the more adapted germplasm generated from the Cycle I and Cycle II populations. Based on the seedling stage resistance phenotype, there is a clear *rpg4/Rpg5* diagnostic haplotype for resistant lines in Cycle I and the Q21861 check. A similar pattern is observed in Cycle II, were half of the TTKSK-resistant lines have the same haplotype. Interestingly, the other half of the Cycle II resistant lines with seedling resistance to TTKSK and adult plant resistance to the surrogate race (QCCJ-B) have a distinct haplotype compared to Q21861. This haplotype is observed in the donors of a potentially new source of resistance (UC1322, UC1266 and, DH13939) and these donors share a haplotype in common at the adult plant resistance QTL on 5H. This 5H QTL for adult plant resistance to stem rust is coincident with the adult plant resistance QTL for stripe rust. For inflorescence type, at *VRS1*, all six-rows have a distinctive haplotype as compared to two-rows. At the *Int-C* locus (which determines the size of lateral florets), there is no defined haplotype. At the *Nud* locus, seven out eight naked lines share the same haplotype. The one exception merits further research. The LD among markers close to target loci was evaluated to identify if haplotype structures were constant across lines. In general, LD was high between the markers across all the loci evaluated for disease resistance and morphological traits. In a few cases (e.g., *Nud, INT-C* and *rpg4/Rpg5*), two blocks were identified among markers.

TABLE 1 | Selected doubled haploids from the Cycle I and Cycle II populations with resistance to stripe rust (incited by *Puccinia striiformis* f. sp. *hordei*), leaf rust (incited by *Puccinia hordei*) and stem rust (incited by *Puccinia graminis* f. sp. *tritici*).

Line	Pedigree	Population	GHF	FHD	IT-M (SR)[a]	Sev (SR)[b]	Sev (BSR)[c]	LR[d]	Head type	Hull type
DH140278[e]	SH98076/Full Pint	Cycle I	52	137.5	0;1-	10	10	13.7	Two	Covered
DH140078	SH98076/10.1151	Cycle I	75	139.5	1,0;	ND	10	3	Two	Naked
DH140512	SH98076/Full Pint	Cycle I	65	138.5	2,1	1	6	8.7	Two	Covered
DH140080	SH98076/10.1151	Cycle I	61	136	2,1,0;	10	1.7	7.3	Two	Naked
DH140515	SH98076/10.1151	Cycle I	68	138.5	1,0;	1	3.3	7.3	Two	Naked
DH140030	SH98076/10.1151	Cycle I	100	133	1,0;,2	10	11.7	10	Six	Naked
DH140077	Violetta/SH98076	Cycle I	104	137.5	2,1	1	3.3	10.3	Two	Naked
DH140076	MC0181-11/Full Pint	Cycle I	59	138.5	2,1	1	8.3	3.3	Two	Covered
DH140273	SH98076/Full Pint	Cycle I	50	133.5	1,2	10	3.3	10.3	Two	Covered
DH140215	SH98076/10.1151	Cycle I	42	130.5	1,2,0;	ND	3.3	10.7	Six	Naked
DH160419	UC1266/DH140213	Cycle II	63	121	0;1	7.25	6.4	ND	Six	Naked
DH160733	DH140512/UC1322	Cycle II	66	110	0;1-	5	2.5	ND	Two	Covered
DH160734	DH140512/UC1322	Cycle II	64	109.5	0;	10.75	2.5	ND	Two	Covered
DH160745	DH140512/UC1322	Cycle II	70	108.5	0;	8.25	2.6	ND	Two	Covered
DH160748	DH140512/UC1322	Cycle II	66	112	0;	2.5	2.6	ND	Two	Covered
DH160754	DH140512/UC1322	Cycle II	68	111	0;	4	2	ND	Two	Covered
DH161043	DH140512/UC1322	Cycle II	64	110.5	0;1	3	2.5	ND	Two	Covered
DH161921	DH140512/DH130004	Cycle II	55	122	0;1	11.75	3.5	ND	Two	Covered
DH161926	DH140512/DH130004	Cycle II	65	114	0;1	5	8.1	ND	Two	Covered
DH161927	DH140512/DH130004	Cycle II	65	117.5	0;1	13.75	7.5	ND	Two	Covered
DH161914	DH140512/10.0860	Cycle II	64	113.5	0;1-	6	3.1	ND	Two	Covered
DH160779	DH140030/UC1231L	Cycle II	62	115.5	10;	3.5	4.2	ND	Six	Naked

[a]Infection-type mode (IT-M) after infection by Puccinia graminis f. sp. tritici race TTKSK at seedling stage. Data was collected in the Biosafety Level 3 Containment Facility at University of Minnesota. [b]Disease severity of stem rust expressed as percentage at adult stage. For Cycle I, data collected under field evaluations in Kenia during 2017. For Cycle II, line means of data collected in field evaluations in Saint Paul, Minnesota during 2018 and 2019 seasons. [c]Disease severity of stripe rust expressed as percentage at adult stage. For Cycle I, data collected under field evaluations during 2017. For Cycle II, values represent the best linear unbiased predictor (BLUPs) for stripe rust severity after infection by Puccinia striiformis f. sp. hordei, across 2018-2019 seasons in Davis, CA and Corvallis, OR. [d]Disease severity of leaf rust expressed as percentage at adult stage. Data collected under field evaluations during 2017. [e]Lines in bold were used as parents for Cycle II population.

NEXT STEPS IN INTROGRESSION BREEDING: END-USE QUALITY

Currently, there is fragmentation in barley production for different end-uses due to the naked vs. covered grain phenotype, the genetic basis of which was described earlier in this chapter. We are initiating a collaborative effort to develop naked barley germplasm that will have sufficient quality and productivity to be used for food, malt, and/or feed. This effort is based on our recent review on the topic of breeding naked barley for multiple end-uses (Meints and Hayes, 2020) and will make use of NGS technologies. Faced with a lack of adapted naked barley germplasm resources, we have developed a plan for systematic introgression of target alleles, early generation marker-assisted selection, speed breeding, and SNP genotyping. The program is now at the implementation stage, and the general framework is a modified Nested Association Mapping (NAM) population that will allow for simultaneous introgression, development of enhanced germplasm and potential varieties, and gene discovery.

Nested association mapping (NAM) populations are multi-parent panels that are designed to combine the advantages of linkage analysis and association mapping in order to delve into genomic regions of interest (Maurer et al., 2015; Hemshrot et al., 2019). Additionally, these panels can be designed as breeding populations to select for new varieties with traits of interest. In order to breed for multi-use naked barley and explore regions of the genome associated with quality traits and the agronomic performance of naked barley, our modified NAM population will have three common parents each crossed to 25 regional parents selected by cooperating breeding programs, for a total of 75 crosses (see **Figure 3**). The three common parents are elite naked breeding lines and the 25 regional parents are a combination of un-adapted breeding lines and land races from USDA-GRIN chosen because they contain target alleles for traits of interest that will be introgressed into the elite parents. Thus, every cross will either be segregating for the naked grain trait or will be fixed for the trait. The overall breeding targets for the entire population are based on traits that are important for multi-use barley and include: naked caryopsis, facultative growth habit, two-row inflorescence, good threshability, and modest β-glucan.

In order to conserve space and select for highly heritable desired traits as soon as possible, a panel of high-throughput allele-specific markers will be used for MAS at the F_2 stage. The target loci are described in **Table 2**. All lines will be selected for the *nud* allele and a combination of the other alleles based on the traits targeted in that specific cross. By selecting for desired phenotypes at the F_2 stage, some genetic variation will be removed from the population, resulting in a modified NAM population rather than a true one. However,

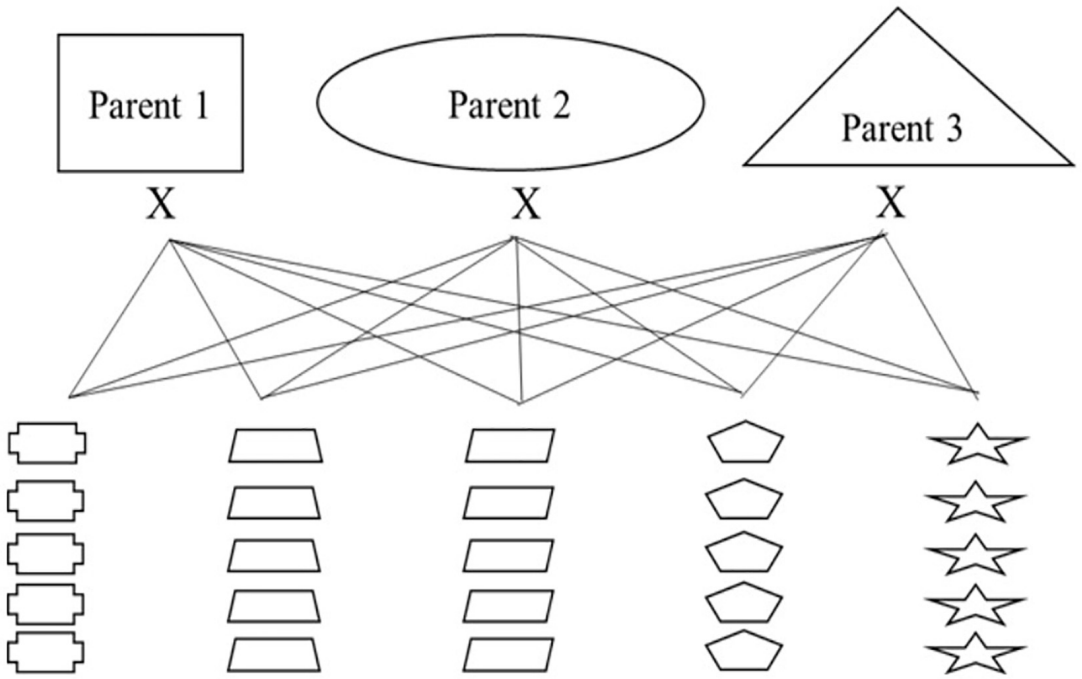

FIGURE 3 | An introgression breeding scheme for modified NAM population development using naked, multi-use barley for organic systems as a model.

TABLE 2 | Genes/QTLs targeted for marker assisted selection in the modified NAM population shown in **Figure 3**.

Targeted allele	Selection target	Citation
nud	Naked caryopsis	Taketa et al. (2008)
Vrs1	2-row spike	Komatsuda et al. (2007)
Wx	Normal starch	Patron et al. (2002)
Deletion at *Vrn-H2*	Facultative	Karsai et al. (2005)
Ppd-H2	Short photoperiod sensitivity	Laurie et al. (1995)
Rpg1	Stem rust resistance	Brueggeman et al. (2002)
Run8	Resistance to loose smut	Grewal et al. (2008)
Ruhq and Ruh1	Resistance to covered smut	Grewal et al. (2008)
Three locus haplotype	Resistance to spot blotch	Haas et al. (2016)

the population will still be useful for later GWAS. A potential drawback of using MAS at the F_2 stage to fix desirable alleles is that this will also reduce recombination in these regions and will create linkage blocks around the MAS targets. This will be problematic if undesirable alleles are linked to the favorable target alleles. To mitigate this potential impact, heterozygotes will be selected at the F_2 stage for target regions in a subset of crosses and advance them by single seed decent (SSD) imposing MAS for the heterozygote for several generations. After selfing the heterozygote, near-isogenic lines will be recovered for each target region in a selection of genetic backgrounds that can be used to (i) validate and quantify the value of the targeted loci; (ii) determine if there are undesirable traits linked to the targeted allele; (iii) develop a set of near isogenic parents that can be used to fine map the region and recover recombinants that resolve unfavorable linkages.

The modified NAM population will to be advanced through SSD using "speed breeding," a method in which increased daylight hours and temperatures decrease generation time (Heuschele et al., 2019). The population will be advanced in the greenhouse through at least the F_5 generation to increase homogeneity before field testing, where the lines will be assessed for agronomic performance, resistance to biotic and abiotic stresses, and end-use quality traits. This breeding scheme will allow for introgression of resistance to biotic and abiotic stresses and quality and agronomic traits into elite naked barley germplasm that will result in potential new multi-use cultivars and germplasm resources for other breeding programs.

CONCLUSION AND GENERAL PERSPECTIVES

Introgression breeding has been, remains, and will be a feature of barley improvement – providing an essential tool to meet the challenges of climate change, ensuring profitable and sustainable production, and enhancing both nutrition and flavor. The donors of alleles for introgression are, not surprisingly, more frequently reported in elite and/or "pre-bred" un-adapted germplasm, followed by land races and exotic accessions in germplasm collections, and finally by wild relatives. Breeders are, not surprisingly, loathe to range far afield in the gene pool, because this increases the risk of linkage drag and/or

disruption of the carefully constructed genome architectures that determine adaptation, meet productivity expectations, and ensure end-use quality. Someone, somewhere, however, needs to assume the risk of conducting the essential pre-breeding required to introgress alleles from the wild and exotic into more adapted backgrounds. An example of the effectiveness of such efforts was the ICARDA/CIMMYT barley program based in Mexico.

The availability of cost-effective, high throughput genotyping tools and analysis procedures has facilitated a plethora of germplasm characterization and allele-discovery studies. Effective mining of these alleles will optimistically follow, as the same tools can be used to track and validate the effects of introgression of these novel alleles into adapted germplasm. However, an interesting alternative emerges as QTLs are reduced to candidate genes and the costs of whole genome and targeted allele sequencing decline to the point of pricing them within reach of breeding programs.

The alternative is gene editing, as currently implemented by CRISPR-Cas9. If the target allele sequence is known - and its function understood - in un-adapted germplasm, a germplasm collection accession, or a wild relative, it is conceptually appealing to sidestep the Scylla and Charybdis of linkage drag and disruption of genome architecture by "simply" knocking-out/knocking-in the allele in adapted germplasm. "Simply" appears in quotes in the preceding sentence because although the gene editing process appears straightforward, it in fact requires a comprehensive understanding of gene function in order to know what to edit and how. Furthermore, germplasm specificity can limit what genotypes can be edited: this leads to the reliance on one or a few "workhorse" genotypes amenable to the transformation processes that can precede editing. In the cased of barley, this would mean that significant efforts would be required to introgress edited alleles from the highly transformable variety "Golden Promise" into target adapted backgrounds. At that point, breeders might question the merits of genome editing – which will likely involve regulatory hurdles and/or intellectual property costs – and instead choose to engage in the time-honored processes of crossing and selection.

AUTHOR CONTRIBUTIONS

JH and PH provided the outlines of the review and key concepts. JH wrote a draft layout of the manuscript. JH, BM, and PH contributed to the writing, editing and final draft of this review. All authors read and approved the manuscript.

FUNDING

Support was provided by Comision Nacional de Investigacion Cientifica y Tecnologica (CONICYT) Becas-Chile 72160554, the U.S. Department of Agriculture (Grants 58-2050-6-005 and 58-2072-5-037), and USDA-NIFA-OREI Grant 2017-51300-26809.

ACKNOWLEDGMENTS

We thank Shiaoman Chao (retired) USDA-ARS Fargo, North Dakota for Illumina 50K genotyping and allele calling, Rick Gadzinski and Sean Tracey at Functional Biosciences for genotyping related to Rpg5, Brian Steffenson and Austin Case for excellent technical assistance and phenotyping for TTKSK in the Biosafety Level-3 (BSL-3) Containment Facility at the University of Minnesota in St. Paul, Alicia del Blanco for excellent technical assistance and phenotyping for stripe rust at the UC-Davis.

REFERENCES

Abdel-Ghani, A. H., Parzies, H. K., Omary, A., and Geiger, H. H. (2004). Estimating the outcrossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. *Theor. Appl. Genet.* 109, 588–595. doi: 10.1007/s00122-004-1657-1

Aldughpassi, A., Wolever, T. M. S., and Abdel-Aal, E. S. M. (2016). "Barley," in *Encyclopedia of Food and Health*, eds B. Caballero, P. M. Finglas, and F. Toldrá (Burlington, VT: Elsevier Science), 328–331.

Babu, R., Nair, S. K., Prasanna, B., and Gupta, H. (2004). Integrating marker-assisted selection in crop breeding–prospects and challenges. *Curr. Sci.* 87, 607–619.

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., et al. (2017). Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* 8:1792. doi: 10.3389/fpls.2017.01792

Bedada, G., Westerbergh, A., Nevo, E., Korol, A., and Schmid, K. J. (2014). DNA sequence variation of wild barley *Hordeum spontaneum* (L.) across environmental gradients in Israel. *Heredity* 112, 646–655. doi: 10.1038/hdy.2014.2

Belcher, A. R., Cuesta-Marcos, A., Smith, K. P., Mundt, C. C., Chen, X., and Hayes, P. M. (2018). TCAP FAC-WIN6 elite barley GWAS panel QTL. I. barley stripe rust resistance QTL in facultative and winter six-rowed malt barley breeding programs identified via GWAS. *Crop Sci.* 58, 103–119. doi: 10.2135/cropsci2017.03.0206

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci.* 48, 1649–1664. doi: 10.2135/cropsci2008.03.0131

Bernardo, R. (2014). *Essentials of Plant Breeding*. Woodbury, MN: Stemma Press.

Bilgic, H., Steffenson, B., and Hayes, P. (2005). Comprehensive genetic analyses reveal differential expression of spot blotch resistance in four populations of barley. *Theor. Appl. Genet.* 111, 1238–1250. doi: 10.1007/s00122-005-0023-2

Bockelman, H. E., and Valkoun, J. (2011). "Barley germplasm conservation and resources," in *Barley: Improvement, Production, and Uses*, ed. S. E. Ullrich (Oxford: Wiley-Blackwell Press), 144–159. doi: 10.1002/9780470958636.ch7

Bregitzer, P., Mornhinweg, D., Hammon, R., Stack, M., Baltensperger, D., Hein, G., et al. (2005). Registration of 'Burton' barley. *Crop Sci.* 45, 1166–1168. doi: 10.2135/cropsci2004.0461CV

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., et al. (2002). The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9328–9333. doi: 10.1073/pnas.142284999

Bush, W. S., and Moore, J. H. (2012). Genome-wide association studies. *PLoS Comput. Biol.* 8:e1002822. doi: 10.1371/journal.pcbi.1002822

Case, A. J., Bhavani, S., Macharia, G., and Steffenson, B. J. (2018). Genome-wide association study of stem rust resistance in a world collection of cultivated barley. *Theor. Appl. Genet.* 131, 107–126. doi: 10.1007/s00122-017-2989-y

Castro, A., Capettini, F., Corey, A., Filichkina, T., Hayes, P., Kleinhofs, A., et al. (2003). Mapping and pyramiding of qualitative and quantitative resistance to

stripe rust in barley. *Theor. Appl. Genet.* 107, 922–930. doi: 10.1007/s00122-003-1329-6

Castro, A. J., Chen, X., Hayes, P. M., Knapp, S. J., Line, R. F., Toojinda, T., et al. (2002a). Coincident QTL which determine seedling and adult plant resistance to stripe rust in barley. *Crop Sci.* 42, 1701–1708. doi: 10.2135/cropsci2002.1701

Castro, A. J., Hayes, P. M., Fillichkin, T., and Rossi, C. (2002b). Update of barley stripe rust resistance QTL in the Calicuchima-sib x Bowman mapping population. *Barley Genet. Newsl.* 32, 1–12.

Ceccarelli, S., and Grando, S. (2000). "Barley landraces from the fertile crescent: a lesson for plant breeders," in *Genes in the Field: On-Farm Conservation of Crop Diversity*, ed. S. B. Brush (Boca Raton, FL: Lewis Publishers), 51–76.

Ceccarelli, S., Grando, S., Tutwiler, R., Baha, J., Martini, A. M., Salahieh, H., et al. (2000). A methodological study on participatory barley breeding I. Selection phase. *Euphytica* 111, 91–104. doi: 10.1023/A:1003717303869

Chen, F., Prehn, D., Hayes, P., Mulrooney, D., Corey, A., and Vivar, H. (1994). Mapping genes for resistance to barley stripe rust (*Puccinia striiformis* f. sp. *hordei*). *Theor. Appl. Genet.* 88, 215–219. doi: 10.1007/BF00225900

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., et al. (2012). Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44, 1388–1392. doi: 10.1038/ng.2447

Craft Brewing Business (2019). *The Malting Industry Continues to Move Away from Six-Row Barley (and other Updates from the Barley Improvement Conference) [Online]*. Available online at: https://www.craftbrewingbusiness.com/ (accessed March 6, 2020).

Cuesta-Marcos, A., Szûcs, P., Close, T. J., Filichkin, T., Muehlbauer, G. J., Smith, K. P., et al. (2010). Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in barley: implications for association mapping in germplasm arrays varying in size and structure. *BMC Genomics* 11:707. doi: 10.1186/1471-2164-11-707

Czembor, J. H. (2000). Resistance to powdery mildew in populations of barley landraces from Morocco. *Genet. Resour. Crop Evol.* 47, 439–449. doi: 10.1023/A:1008732919426

Czembor, J. H., Pietrusińska, A., Piechota, U., and Mańkowski, D. (2019). Resistance to powdery mildew in barley recombinant lines derived from crosses between *Hordeum vulgare* and *Hordeum bulbosum*. *Cereal Res. Commun.* 47, 463–472. doi: 10.1556/0806.47.2019.16

Dahleen, L. S., Bregitzer, P., Mornhinweg, D., and Jackson, E. W. (2012). Association mapping of Russian wheat aphid resistance in barley as a method to identify diversity in the National Small Grains Collection. *Crop Sci.* 52, 1651–1662. doi: 10.2135/cropsci2011.09.0525

Dai, F., Nevo, E., Wu, D., Comadran, J., Zhou, M., Qiu, L., et al. (2012). Tibet is one of the centers of domestication of cultivated barley. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16969–16973. doi: 10.1073/pnas.1215265109

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012

Dubin, H., and Stubbs, R. (1986). Epidemic spread of barley stripe rust in South America. *Plant Dis.* 70, 141–144.

Esvelt, K. K., Gordon, T., Bregitzer, P., Hayes, P., Chen, X., Del Blanco, I., et al. (2016). Barley stripe rust resistance QTL: development and validation of SNP markers for resistance to *Puccinia striiformis* f. sp. *hordei*. *Phytopathology* 106, 1344–1351. doi: 10.1094/PHYTO-09-15-0225-R

FAOSTAT (2017). *Production and Trade. Food and Agriculture Organization of the United Nations*. Rome: FAO.

Fisk, S. P., Cuesta-Marcos, A., Cistué, L., Russell, J., Smith, K. P., Baenziger, S., et al. (2013). FR-H3: a new QTL to assist in the development of fall-sown barley with superior low temperature tolerance. *Theor. Appl. Genet.* 126, 335–347. doi: 10.1007/s00122-012-1982-8

Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374. doi: 10.1146/annurev.arplant.54.031902.134907

Ganal, M. W., Altmann, T., and Röder, M. S. (2009). SNP identification in crop plants. *Curr. Opin. Plant Biol.* 12, 211–217. doi: 10.1016/j.pbi.2008.12.009

Gepts, P. (2006). Plant genetic resources conservation and utilization. *Crop Sci.* 46, 2278–2292. doi: 10.2135/cropsci2006.03.0169gas

Graebner, R. C., Wise, M., Cuesta-Marcos, A., Geniza, M., Blake, T., Blake, V. C., et al. (2015). Quantitative trait loci associated with the tocochromanol (vitamin e) pathway in barley. *PLoS One* 10:e0133767. doi: 10.1371/journal.pone.0133767

Grewal, T. S., Rossnagel, B. G., and Scoles, G. J. (2008). Validation of molecular markers for covered smut resistance and marker-assisted introgression of loose and covered smut resistance into hulless barley. *Mol. Breed.* 21, 37–48. doi: 10.1007/s11032-007-9107-9

Gutiérrez, L., Germán, S., Pereyra, S., Hayes, P. M., Pérez, C. A., Capettini, F., et al. (2015). Multi-environment multi-QTL association mapping identifies disease resistance QTL in barley germplasm from Latin America. *Theor. Appl. Genet.* 128, 501–516. doi: 10.1007/s00122-014-2448-y

Haas, M., Menke, J., Chao, S., and Steffenson, B. J. (2016). Mapping quantitative trait loci conferring resistance to a widely virulent isolate of *Cochliobolus sativus* in wild barley accession PI 466423. *Theor. Appl. Genet.* 129, 1831–1842. doi: 10.1007/s00122-016-2742-y

Harlan, J. R., De Wet, J., and Price, E. G. (1973). Comparative evolution of cereals. *Evolution* 27, 311–325. doi: 10.1111/j.1558-5646.1973.tb00676.x

Harlan, J. R., and Zohary, D. (1966). Distribution of wild wheats and barley. *Science* 153, 1074–1080. doi: 10.1126/science.153.3740.1074

Hayes, P. M., Blake, T., Chen, T. H., Tragoonrung, S., Chen, F., Pan, A., et al. (1993). Quantitative trait loci on barley (*Hordeum vulgare* L.) chromosome 7 associated with components of winterhardiness. *Genome* 36, 66–71. doi: 10.1139/g93-009

Hemshrot, A., Poets, A. M., Tyagi, P., Lei, L., Carter, C. K., Hirsch, C. N., et al. (2019). Development of a multiparent population for genetic mapping and allele discovery in six-row barley. *Genetics* 213, 595–613. doi: 10.1534/genetics.119.302046

Hernandez, J., del Blanco, I., Filichkin, T. P., Fisk, S. P., Gallagher, L. W., Helgerson, L. J., et al. (2020). A genome wide association study of resistance to *Puccinia striiformis* f. sp. *hordei* and *Puccinia graminis* f. sp. *tritici* in barley and development of resistant germplasm. *Phytopathology* 110, 1082–1092. doi: 10.1094/PHYTO-11-19-0415-R

Hernandez, J., Steffenson, B. J., Filichkin, T. P., Fisk, S. P., Helgerson, L. J., Meints, B., et al. (2019). Introgression of rpg4/Rpg5 into barley germplasm provides insights into the genetics of resistance to *Puccinia graminis* f. sp. *tritici* race TTKSK and resources for developing resistant cultivars. *Phytopathology* 109, 1018–1028. doi: 10.1094/PHYTO-09-18-0350-R

Heuschele, D., Case, A., and Smith, K. (2019). Evaluation of fast generation cycling in oat (*Avena sativa*). *Cereal Res. Commun.* 47, 626–635. doi: 10.1556/0806.47.2019.40

Hirota, N., Kaneko, T., Ito, K., and Takeda, K. (2008). Diversity and geographical distribution of seed lipoxygenase-1 thermostability types in barley. *Plant Breed.* 127, 465–469. doi: 10.1111/j.1439-0523.2008.01508.x

Hoseinzadeh, P., Ruge-Wehling, B., Schweizer, P., Stein, N., and Pidon, H. (2020). High resolution mapping of a *Hordeum bulbosum*-derived powdery mildew resistance locus in barley using distinct homologous introgression lines. *Front. Plant Sci.* 11:225. doi: 10.3389/fpls.2020.00225

Hospital, F. (2005). Selection in backcross programmes. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1503–1511. doi: 10.1098/rstb.2005.1670

Jin, Y., Steffenson, B. J., and Fetch, T. G. (1994). Sources of resistance to pathotype QCC of *Puccinia graminis* f. sp. *tritici* in barley. *Crop Sci.* 34, 285–288. doi: 10.2135/cropsci1994.0011183X003400010052x

Johnston, P. A., Niks, R. E., Meiyalaghan, V., Blanchet, E., and Pickering, R. (2013). Rph22: mapping of a novel leaf rust resistance gene introgressed from the non-host *Hordeum bulbosum* L. into cultivated barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 126, 1613–1625. doi: 10.1007/s00122-013-2078-9

Karsai, I., Szûcs, P., Mészáros, K., Filichkina, T., Hayes, P., Skinner, J., et al. (2005). The Vrn-H2 locus is a major determinant of flowering time in a facultative × winter growth habit barley (*Hordeum vulgare* L.) mapping population. *Theor. Appl. Genet.* 110, 1458–1466. doi: 10.1007/s00122-005-1979-7

Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., et al. (2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1424–1429. doi: 10.1073/pnas.0608580104

Laurie, D., Pratchett, N., Bezant, J., and Snape, J. (1995). RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter x spring barley (*Hordeum vulgare* L.) cross. *Genome* 38, 575–585. doi: 10.1139/g95-074

Lei, L., Poets, A. M., Liu, C., Wyant, S. R., Hoffman, P. J., Carter, C. K., et al. (2019). Environmental association identifies candidates for tolerance to low temperature and drought. *G3* 9, 3423–3438. doi: 10.1534/g3.119.400401

Lipka, A. E., Kandianis, C. B., Hudson, M. E., Yu, J., Drnevich, J., Bradbury, P. J., et al. (2015). From association to prediction: statistical methods for the dissection and selection of complex traits in plants. *Curr. Opin. Plant Biol.* 24, 110–118. doi: 10.1016/j.pbi.2015.02.010

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., et al. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544, 427–433. doi: 10.1038/nature22043

Matus, I., Corey, A., Filichkin, T., Hayes, P., Vales, M., Kling, J., et al. (2003). Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp. *vulgare* background. *Genome* 46, 1010–1023. doi: 10.1139/g03-080

Maurer, A., Draba, V., Jiang, Y., Schnaithmann, F., Sharma, R., Schumann, E., et al. (2015). Modelling the genetic architecture of flowering time control in barley through nested association mapping. *BMC Genomics* 16:290. doi: 10.1186/s12864-015-1459-7

McCouch, S., Baute, G. J., Bradeen, J., Bramel, P., Bretting, P. K., Buckler, E., et al. (2013). Agriculture: feeding the future. *Nature* 499, 23–24. doi: 10.1038/499023a

Meints, B., Cuesta-Marcos, A., Fisk, S., Ross, A., and Hayes, P. (2016). "Food barley quality improvement and germplasm utilization," in *Exploration, Identification and Utilization of Barley Germplasm*, eds G. Zhang and L. Chengdao (Cambridge, MA: Academic Press), 41–73. doi: 10.1016/b978-0-12-802922-0.00003-0

Meints, B., and Hayes, P. (2020). "Breeding naked barley for food, feed, and malt," in *Plant Breeding Reviews*, ed. I. Goldman (Hoboken, NJ: Wiley Publisher), 95–119. doi: 10.1002/9781119616801.ch4

Milner, S. G., Jost, M., Taketa, S., Mazón, E. R., Himmelbach, A., Oppermann, M., et al. (2019). Genebank genomics highlights the diversity of a global barley collection. *Nat. Genet.* 51, 319–326. doi: 10.1038/s41588-018-0266-x

Monteagudo, A., Casas, A. M., Cantalapiedra, C. P., Contreras-Moreira, B., Gracia, M. P., and Igartua, E. (2019). Harnessing novel diversity from landraces to improve an elite barley variety. *Front. Plant Sci.* 10:434. doi: 10.3389/fpls.2019.00434

Morrell, P. L., and Clegg, M. T. (2007). Genetic evidence for a second domestication of barley (*Hordeum vulgare* L) east of the Fertile Crescent. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3289–3294. doi: 10.1073/pnas.0611377104

Muñoz-Amatriaín, M., Cuesta-Marcos, A., Endelman, J. B., Comadran, J., Bonman, J. M., Bockelman, H. E., et al. (2014a). The USDA barley core collection: genetic diversity, population structure, and potential for genome-wide association studies. *PLoS One* 9:e94688. doi: 10.1371/journal.pone.0094688

Muñoz-Amatriaín, M., Cuesta-Marcos, A., Hayes, P. M., and Muehlbauer, G. J. (2014b). Barley genetic variation: implications for crop improvement. *Brief. Funct. Genomics* 13, 341–350. doi: 10.1093/bfgp/elu006

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., et al. (2009). Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21, 2194–2202. doi: 10.1105/tpc.109.068437

Newman, R. K., and Newman, C. W. (2008). *Barley for Food and Health: Science, Technology, and Products*. Hoboken, NJ: John Wiley & Sons, Inc.

Nice, L., Huang, Y., Steffenson, B. J., Gyenis, L., Schwarz, P., Smith, K. P., et al. (2019). Mapping malting quality and yield characteristics in a North American two-rowed malting barley × wild barley advanced backcross population. *Mol. Breed.* 39:121. doi: 10.1007/s11032-019-1030-3

Okada, Y., Kanatani, R., Arai, S., and Ito, K. (2004). Interaction between barley yellow mosaic disease-resistance genes rym1 and rym5, in the response to BaYMV strains. *Breed. Sci.* 54, 319–325. doi: 10.1270/jsbbs.54.319

Patron, N. J., Smith, A. M., Fahy, B. F., Hylton, C. M., Naldrett, M. J., Rossnagel, B. G., et al. (2002). The altered pattern of amylose accumulation in the endosperm of low-amylose barley cultivars is attributable to a single mutant allele of granule-bound starch synthase I with a deletion in the 5'-non-coding region. *Plant Physiol.* 130, 190–198. doi: 10.1104/pp.005454

Pickering, R., Niks, R., Johnston, P., and Butler, R. (2004). Importance of the secondary genepool in barley genetics and breeding. II. Disease resistance, agronomic performance and quality. *Czech J. Genet. Plant Breed.* 40, 79–85. doi: 10.17221/3703-CJGPB

Poets, A. M., Fang, Z., Clegg, M. T., and Morrell, P. L. (2015). Barley landraces are characterized by geographically heterogeneous genomic origins. *Genome Biol.* 16:173. doi: 10.1186/s13059-015-0712-3

Rafalski, J. A. (2010). Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13, 174–180. doi: 10.1016/j.pbi.2009.12.004

Rasmusson, D. C., and Phillips, R. L. (1997). Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. *Crop Sci.* 37, 303–310. doi: 10.2135/cropsci1997.0011183X003700020001x

Rey, J. I., Hayes, P. M., Petrie, S. E., Corey, A., Flowers, M., Ohm, J. B., et al. (2009). Production of dryland barley for human food: quality and agronomic performance. *Crop Sci.* 49, 347–355. doi: 10.2135/cropsci2008.03.0184

Ribaut, J.-M., Jiang, C., and Hoisington, D. (2002). Simulation experiments on efficiencies of gene introgression by backcrossing. *Crop Sci.* 42, 557–565. doi: 10.2135/cropsci2002.5570

Richardson, K. L., Vales, M. I., Kling, J. G., Mundt, C. C., and Hayes, P. M. (2006). Pyramiding and dissecting disease resistance QTL to barley stripe rust. *Theor. Appl. Genet.* 113, 485–495. doi: 10.1007/s00122-006-0314-2

Salina, E., Dobrovolskaya, O., Efremova, T., Leonova, I., and Röuder, M. (2003). Microsatellite monitoring of recombination around the Vrn-B1 locus of wheat during early backcross breeding. *Plant Breed.* 122, 116–119. doi: 10.1046/j.1439-0523.2003.00817.x

Sallam, A. H., Tyagi, P., Brown-Guedira, G., Muehlbauer, G. J., Hulse, A., and Steffenson, B. J. (2017). Genome-wide association mapping of stem rust resistance in *Hordeum vulgare* subsp. *spontaneum*. *G3* 7, 3491–3507. doi: 10.1534/g3.117.300222

Schwarz, P., and Li, Y. (2011). "Malting and brewing uses of barley," in *Barley: Production, Improvement, and Uses*, ed. S. E. Ullrich (Ames, IA: John Wiley & Sons), 478–521. doi: 10.1002/9780470958636.ch15

Sharma, P. R., Alhashel, A., Gross, T., Gross, P., and Brueggeman, R. S. (2018). Pyramiding rpg4 and Rpg1-mediated stem rust resistance in barley requires the Rrr1 gene for both to function. *Front. Plant Sci.* 9:1789. doi: 10.3389/fpls.2018.01789

Skadhauge, B., Lok, F., Breddam, K., Olsen, O., Bech, L. M., and Knudsen, S. (2016). *Barley with Reduced Lipoxygenase Activity and Beverage Prepared Therefrom*. U.S. Patent No 031,846,9A1. Washington, DC: U.S. Patent and Trademark Office.

St Clair, D. A. (2010). Quantitative disease resistance and quantitative resistance loci in breeding. *Annu. Rev. Phytopathol.* 48, 247–268. doi: 10.1146/annurev-phyto-080508-081904

Steffenson, B., Case, A., Pretorius, Z., Coetzee, V., Kloppers, F., Zhou, H., et al. (2017). Vulnerability of barley to African pathotypes of *Puccinia graminis* f. sp. *tritici* and sources of resistance. *Phytopathology* 107, 950–962. doi: 10.1094/PHYTO-11-16-0400-R

Steffenson, B., Jin, Y., Brueggeman, R., Kleinhofs, A., and Sun, Y. (2009). Resistance to stem rust race TTKSK maps to the rpg4/Rpg5 complex of chromosome 5H of barley. *Phytopathology* 99, 1135–1141. doi: 10.1094/PHYTO-99-10-1135

Steffenson, B. J. (1992). Analysis of durable resistance to stem rust in barley. *Euphytica* 63, 153–167. doi: 10.1007/BF00023920

Taketa, S., Amano, S., Tsujino, Y., Sato, T., Saisho, D., Kakeda, K., et al. (2008). Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4062–4067. doi: 10.1073/pnas.0711034105

Talamé, V., Sanguineti, M. C., Chiapparino, E., Bahri, H., Salem, M. B., Forster, B. P., et al. (2004). Identification of *Hordeum spontaneum* QTL alleles improving field performance of barley grown under rainfed conditions. *Ann. Appl. Biol.* 144, 309–319. doi: 10.1111/j.1744-7348.2004.tb00346.x

TCAP (2014). *Triticale Coordinated Agricultural Project [Online]*. Available online at: https://www.triticeaecap.org/ (accessed February 25, 2020).

Thabet, S. G., Moursi, Y. S., Karam, M. A., Graner, A., and Alqudah, A. M. (2018). Genetic basis of drought tolerance during seed germination in barley. *PLoS One* 13:e0206682. doi: 10.1371/journal.pone.0206682

Toojinda, T., Broers, L., Chen, X., Hayes, P., Kleinhofs, A., Korte, J., et al. (2000). Mapping quantitative and qualitative disease resistance genes in a doubled haploid population of barley (*Hordeum vulgare* L). *Theor. Appl. Genet.* 101, 580–589. doi: 10.1007/s001220051519

van Ginkel, M., and Flipphi, R. (2020). Why self-fertilizing plants still exist in wild populations: diversity assurance through stress-induced male sterility may

promote selective outcrossing and recombination. *Agronomy* 10:349. doi: 10.3390/agronomy10030349

Varshney, R. K., Terauchi, R., and McCouch, S. R. (2014). Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. *PLoS Biol.* 12:e1001883. doi: 10.1371/journal.pbio.1001883

Visioni, A., Tondelli, A., Francia, E., Pswarayi, A., Malosetti, M., Russell, J., et al. (2013). Genome-wide association mapping of frost tolerance in barley (*Hordeum vulgare* L.). *BMC Genomics* 14:424. doi: 10.1186/1471-2164-14-424

Von Bothmer, R., van Hintum, T., Knüpffer, H., and Sato, K. (2003). *Diversity in Barley (Hordeum vulgare)*. Amsterdam: Elsevier.

von Korff, M., Wang, H., Léon, J., and Pillen, K. (2008). AB-QTL analysis in spring barley: III. Identification of exotic alleles for the improvement of malting quality in spring barley (*H. vulgare* ssp. *spontaneum*). *Mol. Breed.* 21, 81–93. doi: 10.1007/s11032-007-9110-1

Voss-Fels, K., Frisch, M., Qian, L., Kontowski, S., Friedt, W., Gottwald, S., et al. (2015). Subgenomic diversity patterns caused by directional selection in bread wheat gene pools. *Plant Genome* 8, 1–13. doi: 10.3835/plantgenome2015.03.0013

Würschum, T. (2012). Mapping QTL for agronomic traits in breeding populations. *Theor. Appl. Genet.* 125, 201–210. doi: 10.1007/s00122-012-1887-6

Xue, W., Yan, J., Jiang, Y., Zhan, Z., Zhao, G., Tondelli, A., et al. (2019). Genetic dissection of winter barley seedling response to salt and osmotic stress. *Mol. Breed.* 39:137. doi: 10.1007/s11032-019-1042-z

Yun, S. J., Gyenis, L., Bossolini, E., Hayes, P. M., Matus, I., Smith, K. P., et al. (2006). Validation of quantitative trait loci for multiple disease resistance in barley using advanced backcross lines developed with a wild barley. *Crop Sci.* 46, 1179–1186. doi: 10.2135/cropsci2005.08-0293

Zohary, D., and Hopf, M. (2000). *Domestication of Plants in the Old World: The Origin and Spread of Cultivated Plants in West Asia, Europe and the Nile Valley*. Oxford: Oxford University Press.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 06 July 2020
doi: 10.3389/fpls.2020.00929

Introgression Leads to Genomic Divergence and Responsible for Important Traits in Upland Cotton

Shoupu He[1,2], Pengpeng Wang[1], Yuan-Ming Zhang[2], Panhong Dai[1], Mian Faisal Nazir[1], Yinhua Jia[1], Zhen Peng[1], Zhaoe Pan[1], Junling Sun[1], Liru Wang[1], Gaofei Sun[3] and Xiongming Du[1]**

[1] State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China, [2] College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [3] Department of Computer Science and Information Engineering, Data Mining Institute, Anyang Institute of Technology, Anyang, China

OPEN ACCESS

Edited by:
Pietro Gramazio, University of Tsukuba, Japan

Reviewed by:
Kai Wang, Fujian Agriculture and Forestry University, China
Xiaoming Wu, Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, China

***Correspondence:**
Gaofei Sun
sungaofei@sina.com
Xiongming Du
dujeffrey8848@hotmail.com

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *11 April 2020*
Accepted: *08 June 2020*
Published: *06 July 2020*

Citation:
He S, Wang P, Zhang Y-M, Dai P, Nazir MF, Jia Y, Peng Z, Pan Z, Sun J, Wang L, Sun G and Du X (2020) Introgression Leads to Genomic Divergence and Responsible for Important Traits in Upland Cotton. Front. Plant Sci. 11:929. doi: 10.3389/fpls.2020.00929

Understanding the genetic diversity and population structure of germplasms is essential when selecting parents for crop breeding. The genomic changes that occurred during the domestication and improvement of Upland cotton (*Gossypium hirsutum*) remains poorly understood. Besides, the available genetic resources from cotton cultivars are limited. By applying restriction site-associated DNA marker sequencing (RAD-seq) technology to 582 tetraploid cotton accessions, we confirmed distinct genomic regions on chromosomes A06 and A08 in Upland cotton cultivar subgroups. Based on the pedigree, reported QTLs, introgression analyses, and genome-wide association study (GWAS), we suggest that these divergent regions might have resulted from the introgression of exotic lineages of *G. hirsutum* landraces and their wild relatives. These regions were the typical genomic signatures that might be responsible for maturity and fiber quality on chromosome A06 and chromosome A08, respectively. Moreover, these genomic regions are located in the putative pericentromeric regions, implying that their application will be challenging. In the study, based on high-density SNP markers, we reported two genomic signatures on chromosomes A06 and A08, which might originate from the introgression events in the Upland cotton population. Our study provides new insights for understanding the impact of historic introgressions on population divergence and important agronomic traits of modern Upland cotton cultivars.

Keywords: Upland cotton, interspecific hybridization, introgression, genomic divergence, maturity, fiber quality

INTRODUCTION

The *Gossypium* genus (cotton) includes more than 50 species with wide distribution around the world (Wendel and Grover, 2015). Only four species from this genus have been domesticated during the history of cotton cultivation, including two diploid species (*G. herbaceum*, A_1 and *G. arboretum* A_2) and two tetraploid species (*G. hirsutum* $(AD)_1$ and *G. barbadense* $(AD)_2$), with *G. hirsutum* (Upland cotton) accounting for more than 95% of cotton fiber production in the modern world (Tyagi et al., 2014). The tetraploid cotton originated from a natural hybridization event

resulting in merging of the A and D genomes, approximately 1–2 million years ago (Wendel et al., 2010). In addition to the two domesticated tetraploid species (*G. hirsutum* and *G. barbadense*), five other wild species are distributed in the Hawaiian Islands (*G. tomentosum*, $(AD)_3$), Brazil (*G. mustelinum*, $(AD)_4$), Galapagos Islands (*G. darwinii*, $(AD)_5$), Dominican Republic (*G. ekmanianum*, $(AD)_6$) and Wake Atoll (*G. stephensii*, $(AD)_7$) (Percival et al., 1999; Grover et al., 2015; Gallagher et al., 2017). The suggested diversity-center of *G. hirsutum* is the Caribbean and Central America (southern Mexico and Guatemala), where seven geographical landraces have formed: *yucatanense*, *palmeri*, *morril*, *richmondi*, the extensively distributed races *punctatum*, *latifolium*, and *marie-galante*. The *yucatanense* race is considered as the most primitive form of *G. hirsutum*, and a subpopulation of *punctatum* was derived from this race (Wendel et al., 2010). *Palmeri*, *morrill*, and *richmondi* are three comparatively improved races distributed in several relatively small regions (Wendel et al., 1992; Percival et al., 1999). Modern elite cultivated *G. hirsutum* (Upland cotton) is reported to be derived from annual *latifolium* with better fiber quality due to lineage introgression from *G. barbadense* (Wendel et al., 2010).

DNA markers have been successfully applied in previous studies to exploit the cotton diversity and QTL mapping, majority of the molecular markers used in cotton depicted the low genetic diversity in modern cultivated Upland cotton germplasm (Brubaker and Wendel, 1994; Iqbal et al., 2001; Abdalla et al., 2001; Hinze et al., 2012; Fang et al., 2013; Tyagi et al., 2014; Fang et al., 2017; Wang et al., 2017; Ma et al., 2018). This genetic bottleneck impedes further gain in cotton improvement through conventional breeding techniques, especially considering future requirements for increased fiber quality and stress tolerance. Therefore, investigation and understanding of the genetic structure and elucidation of the genetic background of the existing Upland cotton germplasm are lager concern. Simple sequence repeats (SSRs) marker-based studies have been performed to investigate the genetic diversity of the Chinese Upland cotton germplasm (Chen and Du, 2006; Pang et al., 2006). However, these studies were limited by the sample size of the investigated population and molecular marker resolution, which resulted in an unclear depiction of the genetic background of Chinese Upland cotton germplasm. Advancement in the field of genomics, with the whole-genome sequence for cotton (Li et al., 2015; Zhang et al., 2015), lead to the rapid development of single nucleotide polymorphism (SNP) markers. Recently, most of the SNP-based studies focused on SNP marker development (Hulse-Kemp et al., 2015) and QTL detection (genetic map construction) primarily utilizing segregating populations (Wang S. et al., 2015; Zhang et al., 2016; Islam et al., 2016). Besides, genome-wide association studies were biased towards the identification of putative candidate genes for a subjective trait (Fang et al., 2017; Wang et al., 2017; Ma et al., 2018), whereas these studies lacked the incentive to provide basic information about genetic diversity of *G. hirsutum* population.

China is the world's largest cotton fiber producer and consumer (FAO, 2018). The early Chinese Upland cotton germplasm (introduced to Yangtze River and Yellow River regions) was mainly introduced from the United States and the former Soviet Union (Xinjiang province region). Cotton breeding programs in China resulted in the development of series of backbone parents with the integration of pedigree and hybrid breeding methods complemented by adaptation to the local environment (Liang et al., 2002). In addition, Chinese breeders utilized the wild relatives and landraces of *G. hirsutum* and generated abundant introgression lines to transfer favorable traits to commercial cultivars via interspecific hybridization (Liang et al., 2002). This endeavor extensively broadened the genetic pool of Chinese cotton germplasm and led to the further development of a series of elite lines with superior fiber quality and better stress response. However, the genetic basis of interspecific hybridization and its impact on the genomic structure and agronomic traits are still unknown.

A recent study, based on SNP microarray, demonstrated the presence of extensive genomic divergence in the Upland cotton source germplasm and suggested that divergent genomic regions might be related to maturity and heterosis (He et al., 2019). However, the origin of these regions is still unknown. Hereby, using RAD-seq technology, the genetic diversity of 582 tetraploid cotton accessions and their population structure was revealed. The origin of divergent genomic regions within cultivars was also confirmed. Through integration of the genome-wide association studies (GWAS) and previously reported QTLs, the agronomic contribution of variations within these regions was further discussed.

MATERIALS AND METHODS

Plant Materials, Sampling and DNA Extraction

A total of 582 tetraploid cotton accessions, including 470 accessions representing most of the genetic diversity of the Chinese *G. hirsutum* cultivars, 105 accessions belonging to eight geographic landraces, four *G. barbadense* accessions and three accessions from wild relatives, were examined in this study. detailed information for the accessions is provided in **Supplementary Table S1**. All cultivars were planted in the experimental field of the Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China. The landraces and wild relatives were sampled from the National Wild Cotton Nursery, Sanya, China. DNA from all samples was extracted from young fresh leaves following the CTAB method described by Paterson (Paterson et al., 1993).

Library Construction and Sequencing

Genomic DNA was first quantified on a Qubit 2.0 fluorometer (Invitrogen), after which the concentration was calculated. the DNA was diluted to 50 ng/μl; and 1 μg of each sample was transferred to a clean 200 μl PCR plate (Axygen). The genomic DNA in each well was digested with 1 μl of FastDigest TaqI (Fermentas) for 10 min at 65 °C in a volume of 30 μl. For the ligation reaction, 1 μl of barcoded adapters (10 μM) was added to individual wells, together with T4 DNA ligase (Enzymatics), in a total volume of 40 μl. The ligation reaction was incubated for 1 h at 22 °C and then heat-inactivated at 65 °C for 20 min. Twenty-four

ligation products from different samples were pooled into a single tube, and 2 μl of chloroform was added to inactivate the restriction enzyme. The mixtures were subsequently centrifuged at 12,000 rpm for 1 min, and the supernatant was transferred to a new tube. DNA fragments between 400 and 700 bp in length were screened in 2% agarose gels (Amresco) and purified using a QIA quick Gel Extraction Kit (QIAGEN). Next, the samples were resuspended in 50 μl of elution buffer and amplified via 10 cycles of PCR. Each amplification reaction included 8 μl of the library, 25 μl of Phusion Master Mix (Phusion high-fidelity, Finnzymes), 1 μl of the common primer (10 μM), 1 μl of the index primer (10 μM) and 15 μl of water. The amplified library was purified using a QIA quick PCR Purification Kit (QIAGEN), then quantified on an Agilent2100 Bioanalyzer (Agilent) and sequenced on an Illumina Hiseq2000 instrument (Illumina), as per the manufacturer's protocol.

Variant Calling

In this study, approximately 0.74 billion clean reads of 85 bp in length (597 GB) were generated from the Illumina platform. We used BWA (ver. 0.7.12) (Li and Durbin, 2009) to map all clean reads on the *G. hirsutum* genome (Zhang et al., 2015) with default parameters; only paired-end reads that were both mapped to the genome were retained for variant calling. SAMtools (ver. 1.1) (Li et al., 2009) was used for variant calling. First, reads with a quality of less than 20 were discarded, after which reads that passed quality filtering were assigned to call variants using 'bcftools'; (ver. 1.1); the ultimate SNP set (total 253,679 SNPs) was further filtered using the 'vcfutils.pl' script of 'bcftools';, with the following parameters: -Q 10 -d 2 -D 2000 -a 2 -w 3 -W 10 -1 0.0001 -2 1e-100 -3 0 -4 0.0001 -e 0.0001.

Phylogenetic Tree Construction and Population Genetic Analysis

A total of 9,868 SNPs screened from the set of 253,679 SNPs (MAF >0.05, Missing <0.1, and heterogeneity <0.3) were employed to construct the phylogenetic tree for all accessions using FastTree software (Price et al., 2009). We further screened out another SNP set of 68,118 SNPs from the 253,679 SNP set (MAF >0.05, Missing <0.2 and heterogeneity <0.3) to analyze population structure using ADMIXTURE software (Alexander et al., 2009). Principle component analysis (PCA) was performed using the 'smartpca' module of EIGENSOFT (https://www.hsph.harvard.edu/alkes-price/software/). The nucleotide diversity (π) was calculated by VCFTools (Danecek et al., 2011).

Physical Localization of SSR Markers and QTLs in the Genome

A total of 65,412 raw SSR marker clones were downloaded from COTTONGEN (www.cottongen.org). The sequences were aligned against the *G. hirsutum* genome (Zhang et al., 2015) using Blastn (ver. 2.2.30) (McGinnis and Madden, 2004). Only the two results (A and D subgenomes) with the smallest *p* values were retained as the possible physical positions of the SSR markers. For physical localization of the previously reported QTLs (references were listed in **Supplementary Tables S4** and **S7**), only the positions (genetic map chromosome assignment (Wang et al., 2006) of QTLs that agreed with the physical position of their corresponding SSR markers (genome chromosome assignment) were retained for further analysis. When two flanking markers were not located on the same chromosome, only markers that agreed with the genetic map were retained to represent that QTL.

Phenotyping and GWAS

All phenotypic data including development stage (maturity), boll weight, lint percentage, seed index, fiber length, fiber strength, micronaire and fiber elongation rate were investigated followed by "Descriptors and data standard for cotton" in three typical cotton-growing regions of China, including Anyang (Yellow River region), Nanjing (Yangtze River region) and Akesu (Xinjiang) for three years (2007–2009) with three replications for each environment. GWAS was performed by standard EMMAX procedure described by Kang et al. (2010) (http://genetics.cs.ucla.edu/emmax/).

Introgression Analysis

To identify introgression in Upland cotton cultivars, (1) the donor group was screened according to the result of population structure analysis (when $K = 4$) (selected from Group-0). Accessions in Group-0 containing a "deep blue" or "purple" lineage were selected as donor-group-blue (total of 50 accessions) or donor-group-purple (total of 73 accessions). (2) The receptor accessions were any cultivar that may have contained an introgressive fragment (selected from Group-1, Group-2 and Group-3). At each site for any given receptor accession, we first calculated the site identical sample count (site ISC) by comparing its genotype with the donor group (deep blue and purple were calculated separately). For instance, at one site, if the genotype of the receptor accession was "A" and donor group contained 20 "As", then the site ISC was recorded as 20 for that receptor accession. The window ISC was the sum of all site ISCs in the 1 Mb window region. Second, the window valid sample count (window VSC) was recorded as the total number of samples with known genotypes. Finally, the introgressive index was calculated using the following formula:

$$\text{Introgressive index} = \frac{\text{windowed ISC of donor group}}{\text{windowed VSC of donor group}} - \frac{\text{windowed ISC of acceptor group}}{\text{windowed VSC of acceptor group}}$$

All introgressive indexes larger than 0.15 were screened as introgression fragments and plotted in **Figure 3**.

RNA-SEQ DATA MANIPULATION

All RNA-seq data were downloaded from NCBI and were first mapped on the genome using TopHat (ver. 2.0.13) (Trapnell et al., 2009), after which the expression level was calculated

(fragments per kilobase of transcript per million mapped fragments, FPKM) using Cufflinks (Trapnell et al., 2010). The transcriptome data used in this study came from various tissues of TM-1 (PRJNA248163) and 10 DPA and 20 DPA fibers of *yucatanense* (SRP017061).

RESULTS

Genetic Relationships, Population Structure, and Genomic Divergence of Tetraploid Cotton

The whole sequencing panel contained 470 cultivars, 105 landraces, four *G. barbadense* accessions, and three wild relatives. Most of the cultivars were collected from China, and nearly all the landraces were from the United States, except *G. purpurascens* (**Supplementary Table S1**). After SNP filtering (MAF >0.05, missing <0.1, heterogeneity <0.3), we used 9,868 SNPs extracted from a raw SNP set (253,670 SNPs) to construct the phylogenetic tree. *G. mustelinum*, a wild and primitive tetraploid cotton species distributed in a small region of northeastern Brazil (Brubaker and Wendel, 1994), was used as the root for the phylogenetic tree (**Figure 1A**). Four *G. barbadense* cultivars were clustered together next to *G. mustelinum*, followed by two island species viz. *G. darwinii* and *G. tomentosum*. *G. hirsutum* clade was positioned next to *G. tomentosum* (**Figure 1B**). Almost all of the landraces (**Figure 1A**, orange branches) were separated from the clades of wild species earlier than the cultivars (**Figure 1A**, blue branches). We defined landraces together with the wild species and *G. barbadense* as "Group-0" (**Figures 1A, B**, red, purple, and orange branches). Remaining accessions belonged to *G. hirsutum* lines or cultivars (**Figure 1A**, blue branches), most of them (383 of 470, **Supplementary Table S1**) were elite genetic materials or commercial cultivars collected from various locations in China. In the phylogenetic tree, landraces could easily be distinguished from cultivars. However, random distribution of cultivars in this clade was in contrast with the geographic distribution of them (**Supplementary Table S1**) Overall, the phylogenetic tree showed that the cultivars branch was relatively compact than the landraces branch. Principal component analysis (PCA) complemented the compactness of the cultivars branch (**Figure 1C**). The whole-genome average nucleotide diversity (π) of landrace group (0.041×10^{-3}) was also greater than cultivars (0.022×10^{-3}). Therefore, we concluded that the genetic diversity of *G. hirsutum* landraces was higher than cultivars. The narrow genetic diversity of cultivars might have been caused by the limited number of early parentages.

Interestingly, in the model-based clustering analysis, Group-0 appeared to harbor two kinds of mixed lineages (**Figure 1D**, orange and blue, when $K = 2$), but the cultivar group was biased toward one of them (**Figure 1D**, blue, when $K = 2$). Among the cultivars, in particular, two distinct groups (**Figure 1D**, named as Group-1 and Group-2 harbored relatively more exotic

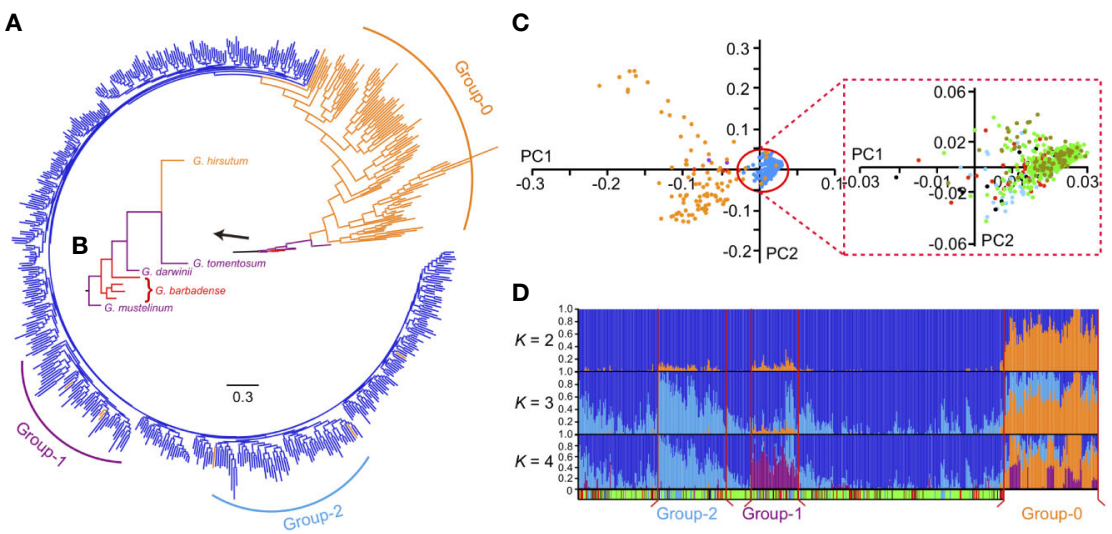

FIGURE 1 | The phylogenetic tree and genetic structure of tetraploid cotton. **(A)** The phylogenetic tree of all tetraploid cotton accessions. Three distinct groups are marked as purple (Group-1), blue (Group-2) and orange (Group-0) bars. **(B)** The zoomed view of wild tetraploid cottons and *G. barbadense* branches. Wilds (purple), *G. barbadense* (red) and *G. hirsutum* (orange) branches are marked with different colors, respectively. **(C)** The principal components analysis (PCA) of all accessions. First two principal components were used as *x* (PC1) and *y* (PC2) axis to plot 582 accessions in this study. In left panel, dots with different colors indicated wild species (purple), *G. barbadense* (red), *G. hirsutum* landraces (orange) and *G. hirsutum* cultivars (blue), respectively. In right panel, the red box indicated the zoomed view of PCA plots within *G. hirsutum* cultivars, different ecotypes including central Asia (deep blue), US (red), Yellow River region of China (green), Yangtze River region of China (asparagus) and other nations (black) were represented by dots with different colors. **(D)** The model-based clustering analysis with different clusters ($K = 2$ to 4). The y axis quantifies cluster membership, and the x axis represents the accessions; the order is identical with phylogenetic tree (the root was on the right). Group-1, Group-2, Group-0 and Group-3 (all remaining cultivars) are separated by vertical red lines. The colored bar at the bottom indicates the geographical origin of all accessions, light blue, Central Asia; red, United States; green, Yellow River region of China; asparagus, Yangtze River region of China; black, other nations; white, unknown.

introgression components (orange, when K = 2) than other accessions. Group-1 and Group-2 contained 54 accessions (including three landraces) and 76 accessions (including two landraces), respectively. The remaining 346 accessions (including one landrace) comprised Group-3 (**Supplementary Table S1**). Furthermore, Group-2 and Group-1 appeared to exhibit different lineage compositions at K = 4: more purple lineages were present in Group-1, whereas more blue lineages were observed in Group-2 (**Figure 1D**). According to the germplasm source information (**Supplementary Table S1**), we found that the purple lineage (Group-1) contained the accessions mostly collected from Yellow and Yangtze River regions. The blue lineage (Group-2) were mainly collected from high-latitude regions, including 33 accessions from the Chinese Yellow River region and 22 accessions from the former Soviet Union, Xinjiang Province and Liaoning Province of China. These accessions showed several typical characteristics, such as early maturity with small and compact plant architecture.

The Phenotypic Characteristic of Two Sub-Groups of *G. hirsutum* Cultivars

To investigate the phenotypical variation among three sub-groups, we analyzed eight major agronomic traits including maturity (development stage), yield (boll weight, seed index, and lint percentage), fiber quality (fiber length, fiber strength, micronaire, and fiber elongation rate) in three locations (Anyang, Nanjing, and Akesu) for three cropping seasons. Descriptive statistics for maturity significantly differentiated three groups, while Group-2 demonstrated early maturity than other groups. (**Figure 2**). Moreover, both Group-1 and Group-3 showed significantly better fiber quality (fiber length and fiber strength) than Group-2 (**Figure 2**). We also found that the fiber strength and micronaire of Group-1 were slightly better than Group-3. These results suggested that the accessions in Group-1 and Group-2 exhibited excellent fiber quality and significant early maturity, respectively.

Genomic Divergence and Introgression in the Upland Cotton Population

In the present study, two distinct sub-groups (Group-1 and Group-2) that contained relatively more exotic introgressions were identified through ancestry analysis (**Figure 1D**). To further investigate the specific introgressed genomic regions, we calculated the pairwise population differentiation statistic (Fst) for Group-1 vs. Group-3, Group-2 vs. Group-3, and Group-1 vs. Group-2 (Group-3 comprised remaining cultivars or lines except Group-1 and Group-2). The highly divergent regions (top 1%, Fst >0.364) identified among these groups were located in two regions, ranging from approximately 63.9 to 94.9 Mb on chromosome A06 and 21.8 to 71.7 Mb on chromosome A08 (**Figures 3A, D**, **Supplementary Table S2**). The genomic differentiation of these regions was dramatically higher than the average whole-genome level, indicating that they were the major genomic divergence regions within the *G. hirsutum* cultivar population. We identified 9, 3 and 3 highly differentiated regions on A06 in the comparison of Group-2 vs. Group-3 (average Fst = 0.437), on A08 for Group-1 vs. Group-3 (average Fst = 0.617) and for Group-1 vs. Group-2 (average Fst = 0.467), respectively (**Figures 3A, D**, **Supplementary Table S2**). A total of seven QTLs related to both yield and fiber quality were located in the divergent region (**Supplementary Table S3**). Interestingly, we found that most of these QTLs were derived from a parent with explicit wild introgression, such as *G. anomalum*, *G. barbadense*, or *G. arboretum* (**Supplementary Table S3**), which strongly suggested that these genomic regions could be related with genetic introgression during interspecific hybridization.

To confirm the genomic distribution of introgressive fragments in the cultivar population, by using Group-0 as a donor population (**Figure 1C**, when K = 4), we further calculated the distribution of the introgression index for the "purple" and "light blue" lineages of each accession on the At sub-genome (**Figures 3B, E**) and Dt sub-genome (**Figures 3C, F**), respectively. For all the cultivars, the introgressive fragments derived from Group-0 were mainly located on chromosomes A06, A08, D01, D08, and D09 (**Figures 3B, C, E, F**). In the two distinct groups, Group-1 carried "purple" lineage on A08 (**Figure 3B**, indicated by purple arrows), and Group-2 carried both "purple" and "light blue" lineages on A06 (**Figures 3B, C**, indicated by blue arrows). These unevenly distributed distinct introgressive fragments in Group-1 and Group-2 might be the cause of population differentiation in Upland cotton cultivars. We further found that the genotypes of Group-1 and Group-2 presented high heterozygosity and missing alleles in their specific introgression regions (A06 and A08) (**Supplementary Figure S4**). Besides, we found that these regions were mainly located at predicted pericentromeric regions on A06 and A08 (green bars at the bottom of **Figure 3A**) (Wang S. et al., 2015). In these regions, some QTLs related with yield and fiber quality were also detected in previous studies (**Figure 3A**, **Supplementary Table S3**). Therefore, these megabase-size regions with strong genetic linkage disequilibrium might have resulted from the low recombination frequency of pericentromeric regions.

The Potential Function of Divergence Regions on Chromosomes A06 and A08

Considering the specific pedigrees and phenotypes of the accessions in Group-1 and Group-2 with distinct genotypes showed exceptionally high divergence and harbored various QTLs (**Figures 2** and **3**), we speculated that the genomic variations in regions might be associated with important traits. According to the gene annotations, a total of 282 and 289 genes were annotated in the two regions on A06 and A08, respectively (**Supplementary Tables S4** and **S5**). Some genes in these regions showed different expression patterns in various tissues. We further screened some genes with tissue-specific expression that might regulate ovule or fiber development (being highly or expressed explicitly in ovules or fiber (**Supplementary Figures S5** and **S6**). For instance, several functionally characterized genes related to fiber, trichome, or root development were explicitly expressed in ovules or fibers. On chromosome 6, *meristem layer 1* (*ML1*, A06G1283) encoding a homeobox protein similar to GL2,

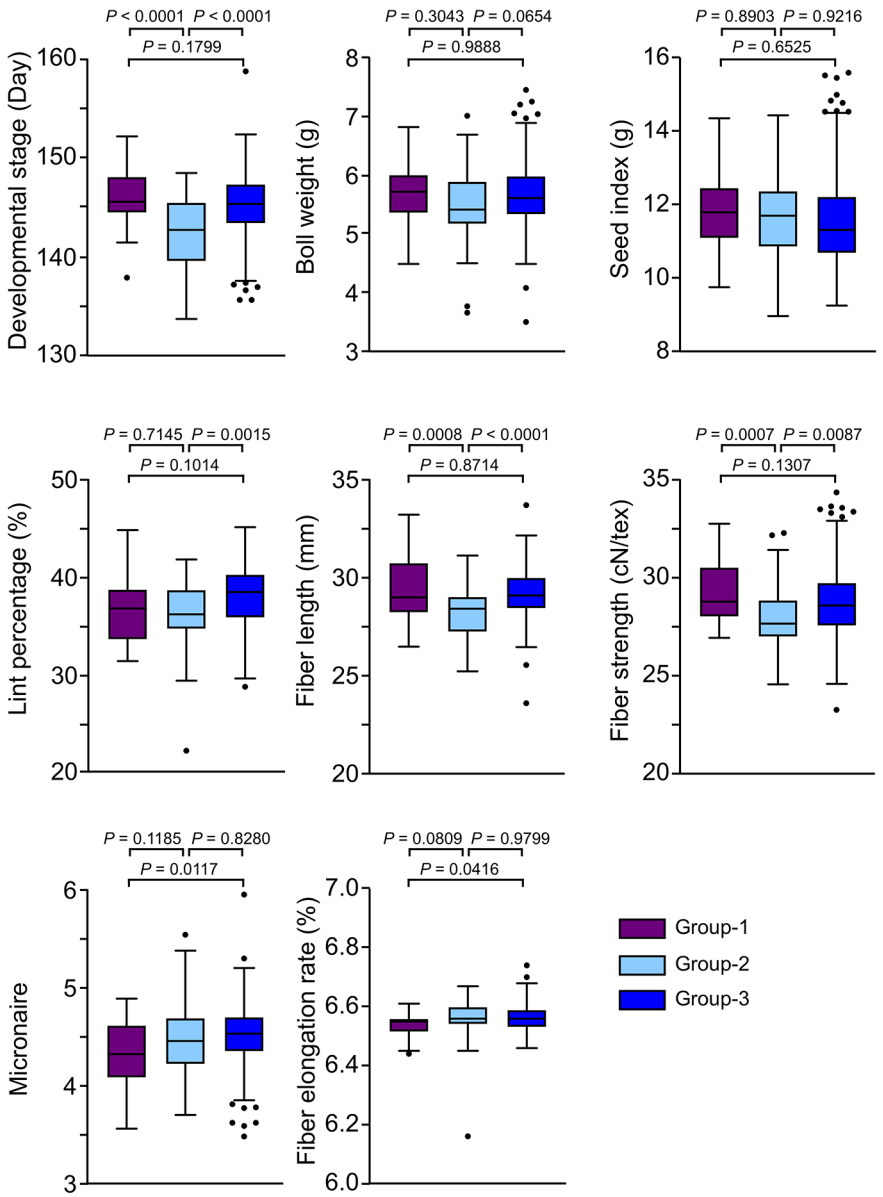

FIGURE 2 | Comparison of investigated traits among groups. In the box plots, the centerline, box limits, and whiskers indicate median, upper and lower quartiles, and 1.5× interquartile range, respectively. Points show outliers. The significances were tested by Tukey's multiple comparisons test.

is a transcription factor that interacts with MYB25 to further regulate trichome development in *Arabidopsis* (Zhang et al., 2010) and cotton (Ding M. et al., 2015). *β-ketoacyl-[acyl carrier protein] synthase I* (*KAS1*, A06G1195) is a crucial gene to regulate root development in rice (Ding, W. et al., 2015) and *1-Aminocyclopropane-1-Carboxylic acid Oxidase 4* (*ACO4*, A06G1341) is responsible for ethylene production and therefore influences cotton fiber growth (Qin et al., 2007) (**Supplementary Table S4**, **Supplementary Figure S5**). On chromosome A08, *glycosyl hydrolase 9C2* (*GH9C2*, A08G0869) is a gene that impacts cell wall development in plants (Glass et al., 2015). *MYB103* (A08G0993) is specifically expressed in 20 DPA and 25 DPA fibers and has been suggested to affect secondary cell wall biosynthesis and deposition (Sun et al., 2015) (**Supplementary Table S5**, **Supplementary Figure S6**). We also identified several putative genes that might be related to fiber development, i.e. the *cytochrome P450*, *family 77*, *subfamily B*, *polypeptide 1* (*CYP77B1*, A06G1290) and *RPM1 interacting protein 4* (*RIN4*, Gh_A06G1343) genes and an unannotated gene (GhA08G1000); these genes also showed specific expression patterns in ovules and fiber tissues (**Supplementary Table S5**, **Supplementary Figures S5** and **S6**). The causal nucleotide variations in these regions (or on genes) and their genetic functions in cotton should be further verified.

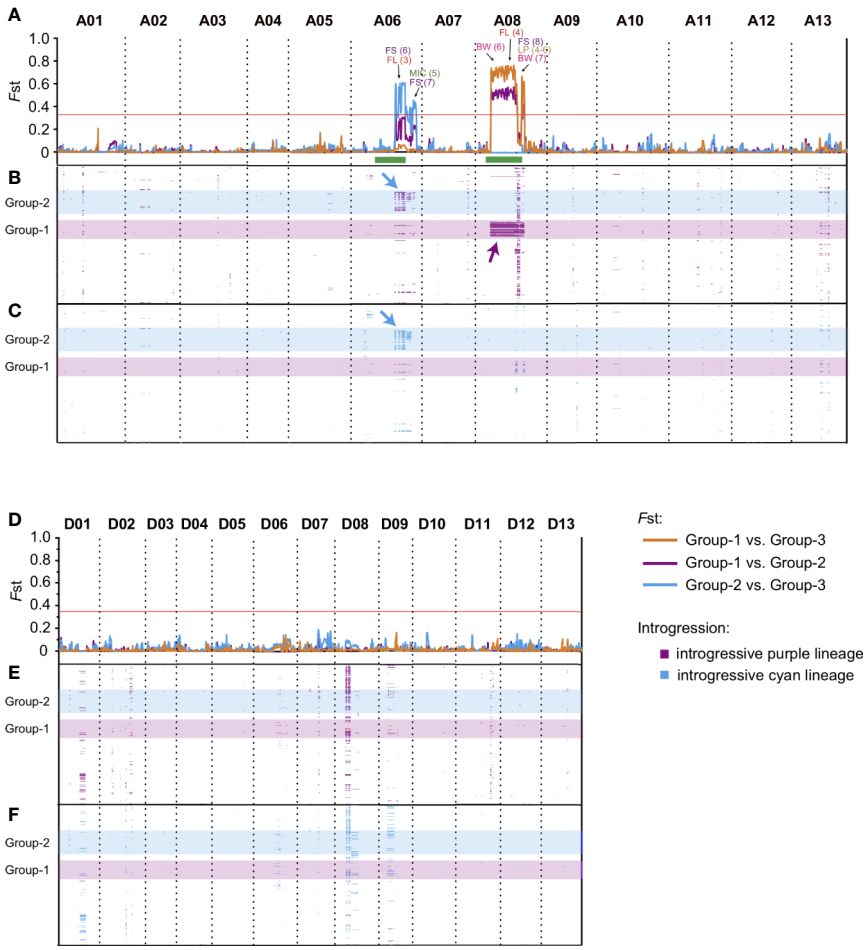

FIGURE 3 | The differentiation and introgression within cultivar population. The divergent genomic regions on At subgenome **(A)** and Dt subgenome **(D)** within Upland cotton cultivars. The *y* axis indicates the *F*st value, and three comparisons of Group-1 vs. Group-3 (orange), Group-2 vs. Group-3 (light blue) and Group-1 vs. Group-2 (purple) are represented by lines with different colors, respectively, the horizontal red lines represent the threshold value of *F*st (top 1%, *F*st >0.364), and the regions above lines indicates the divergent genomic regions within Upland cotton cultivars. QTLs corresponding to different traits were marked by different colors; the number in parentheses indicates the QTL IDs (detailed information of QTLs are listed in **Supplementary Table S3**). BW, boll weight; SI, seed index; LP, lint percentage; FL, fiber length; FS, fiber strength; MIC, micronaire. The green bars at the bottom of Chr. A06 and A08 indicate their putative pericentromeric regions (Wang S. et al., 2015). The introgression regions (introgression index >0.15) derived from "purple" **(B, E)** and "light blue" **(C, F)** lineages were presented by purple and light blue band, y-axis of **(B)**, **(C)**, **(E),** and **(F)** indicated the cultivars, the order of accessions were consistent with cluster result (**Figures 1A, B**). The position of Group-1 (light blue) and Group-2 (purple) were highlighted by a transparent band. The major introgression fragments on A06 (blue) and A08 (purple) were marked by arrows. All QTL references were listed in **Supplementary Table S3**.

Genome-Wide Association Study (GWAS) Further Confirmed the Function of Variations in Divergence Regions on Two Chromosomes

To further confirm the genetic function of divergence regions on chromosomes A06 and A08, we performed GWAS on 316 representative accessions selected from the whole panel of the population. We mainly focused on the significant GWAS signals (−logP >4) in the divergence regions on chromosomes A06 and A08. Interestingly, for chromosome A06, the majority of signals were associated with lint percentage (33/83) and development period (maturity) (17/83). However, for chromosome A08, the signals were associated with fiber quality traits, such as fiber strength (44/101) and fiber length (19/101) (**Supplementary Table S6**). Moreover, two adjacent LD blocks were identified in the overlapped regions between GWAS (development stage) and population divergence (*F*st) (**Figures 4A, B**). In these regions, a total of 12 genes were detected in block 1 (491 Kb) and block 2 (401 Kb) (**Figures 4C, D**), respectively. Furthermore, the allelic frequencies of signals A06_92148100 and A06_93448418 showed that Group-3 carried much more GG allele than Group-2 (**Figures 4E, F**). The development stage of accessions carrying genotype CC (92,148,100 and 93,448,418) showed significantly early maturity than other two genotypes (**Figures 4E, F**). In these regions, several genes such as *A06G1269*, *A06G1272*, *A06G1309*, and *A06G1314* were specifically expressed in floral organs

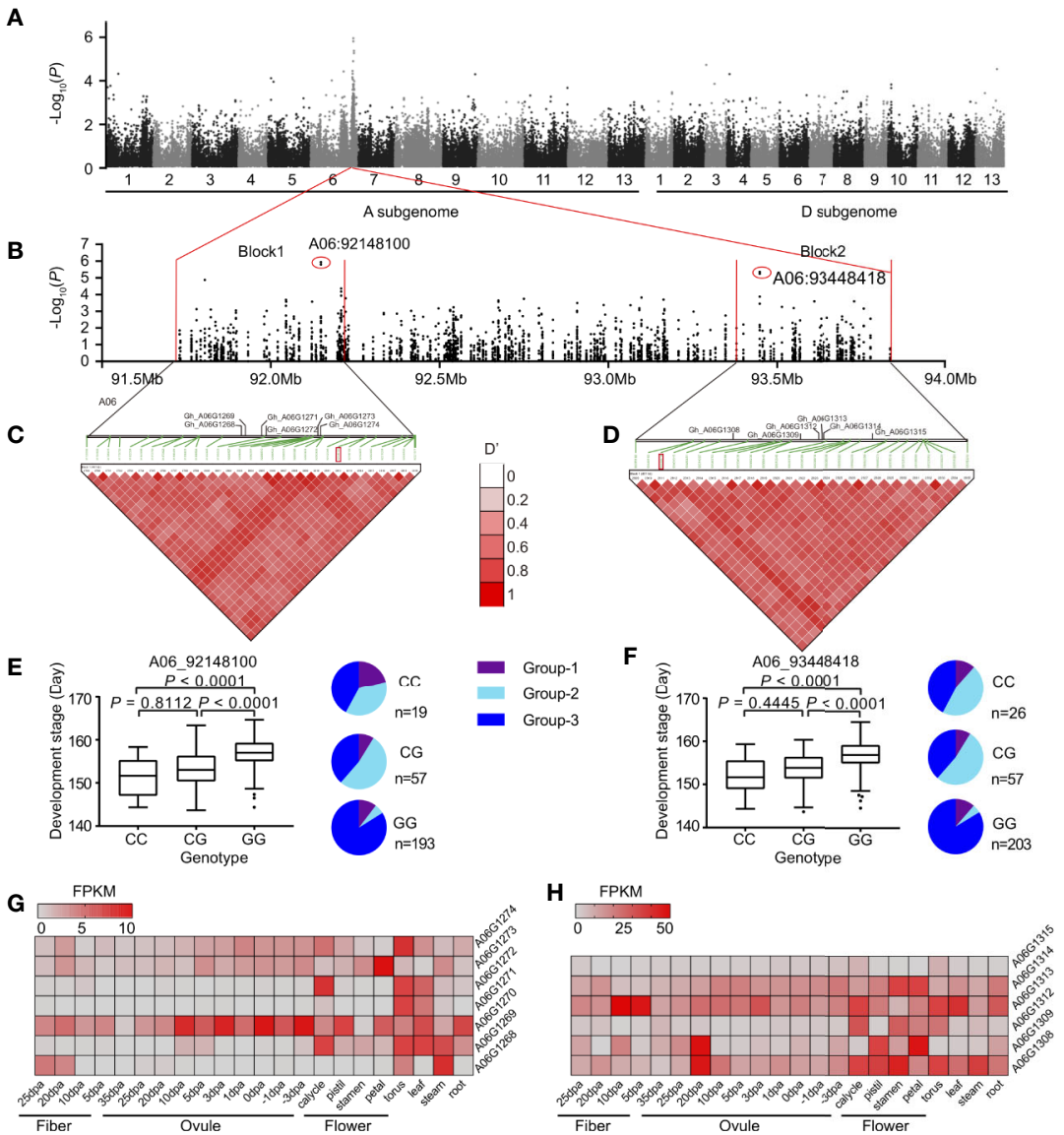

FIGURE 4 | The maturity trait-associated loci in the divergence region on chromosome A06. **(A)** Manhattan plots of GWAS for the development stage (Nanjing-2009). **(B)** the local Manhattan plot for the signals on chromosome A06, two strongest signals (A06_92148100 and A06_93448418) were marked by red circles. **(C, D)** The LD heatmap and annotated genes in two blocks. The location of the strongest signals was labeled by red rectangles. **(E, F)** Box plots for the development stage, according to the genotype of two strongest signals A06_92148100 (left) and A06_93448418 (right). In the box plots, the centerline, box limits, and whiskers indicates median, upper and lower quartiles, and 1.5× interquartile range, respectively. Points show outliers. Significances were tested by Dunn's multiple comparison test. The pie charts indicated the sub-group categorization of the strongest signals in GWAS population. The heatmaps indicated the level of genes in block 1 **(G)** and block 2 **(H)**, respectively. FPKM, fragments per kilobase per million.

(**Figures 4G, H**). These genes might control floral development and further regulating cotton maturity. In A08, a total of three similar blocks were detected associated with fiber strength (**Supplementary Figure S3**). We detected several genes in these blocks, possibly associated with fiber strength. For instance, *A08G0929* in block 2, located nearby the strongest signal of this block (A08_59869122), showed gradual increase in expression level from 10 DPA to 25 DPA in fiber. This gene involved in the flavonol and lignin biosynthetic process in *Arabidopsis* (Moinuddin et al., 2010) and *Brachypodium* (Ho-Yue-Kuang et al., 2016), which are two major biological processes to determine fiber quality in cotton. Taking into account the ecological distribution and morphological characteristics of three subgroups, our results emphasized the correlation of highly differentiated regions on chromosomes A06 and A08 with maturity and fiber strength, respectively.

DISCUSSION

Complex Genetic Background of the *G. hirsutum* Landrace Population and Narrow Diversity of the Cultivars Population

In this study, the relationship observed among wild tetraploid cottons, *G. barbadense*, and *G. hirsutum* cotton was consistent with a previous study based on SSR markers (Lacape et al., 2006). According to the phylogenetic tree and genetic structure analysis, within Group-0 (**Figures 1A, D**), the principal branch of the *marie-galante* race was the first to separate from wild species (except for one *morrill* accession). This landrace has widespread along the coast from southern North America to central South America and suggested to have been derived from introgression between *G. hirsutum* and *G. barbadense* (Wendel et al., 2010). Our results confirmed that *marie-galante* is the *G. hirsutum* landrace closest to *G. barbadense* and represents a potential source for increasing the diversity of improved lines in future cotton breeding. The *punctatum* and *latifolium* races include both perennial and annual forms and are widely distributed across Central America (Hutchinson et al., 1947). This extensive geographical distribution and frequent human interventions provided more opportunities for intercrossing with other indigenous races to increase the genetic variation of these races. Therefore, in the phylogenetic tree generated in the present study, these races were distributed in Group-0 (**Figures 1A, B**). Furthermore, these two races have been suggested to be the most probable original races of modern Upland cotton (Wendel et al., 1992). We also found that five *latifolium* accessions clustered in the cultivar clade (**Figure 1A**) (a total of six accessions clustered in the cultivar clade, including one *marie-galante* accession), implying that genetic background of *latifolium* accessions is very close to modern cultivars and most of Chinese Upland cotton germplasm was possibly originated from *latifolium*.

In the present study, both the phylogenetic tree and genetic structure analyses demonstrated the complex genetic background of *G. hirsutum* landraces. Although some of these landraces can be phenotypically differentiated, there are no apparent genetic characteristics for clearly distinguishing each race. This situation is very likely caused by the overlap in the habitats of these landraces and human activities, resulting in landraces that are genetically mixed (especially in three geographically widely distributed races: *marie-galante*, *punctatum*, and *latifolium*) containing true wild, feral and cultivated populations (Coppens d'Eeckenbrugge and Lacape, 2014). Based on SNP markers, we were able to classify the landraces into different subtypes for a better understanding of their genetic background and further utilization in breeding programs.

Initial reports concerning domestication of Upland cotton suggested that *G. hirsutum* cultivars should have an abundance of primitive gene-pool (Shands et al., 1991). However, narrow genetic diversity of cotton cultivars has been noted in several previous studies (Liang et al., 2002; Chen and Du, 2006; Lacape et al., 2006). This paradox might be the result of domestication bottleneck i.e., selection for early maturity which led to loss of several elite alleles and favorable genes during the expansion of Upland cotton into North America (Chen et al., 2006). Our results were consistent with lower diversity in the cultivar population than in landraces (**Figure 1A**), implying that there is great potential for improving modern cotton cultivars by utilizing landraces. Further classification and exploitation of these landraces, with the utilization of high throughput phenotyping and genotyping, can be useful to shed light on the domestication history of *G. hirsutum* leading towards improved future cotton breeding programs.

Genomic Differentiation Resulting From Landrace Introgression on Chromosomes A06 and A08 Was Responsible for Important Traits in Upland Cotton Cultivars

Nearly all modern cotton cultivars were primarily developed in the United States from four basic types (Acala, Plains, Delta, and Eastern type), which were originated from a diverse gene-pool (Petit Gulf) mixed with *G. hirsutum* and *G. barbadense* lineages (Shands et al., 1991). These four types of cotton were subsequently introduced to other major cotton production areas worldwide. American Upland cotton accessions were introduced in central Asian countries during the 1870–1880s. The initial germplasm was selected from a mixture of the American early-maturity varieties, including King, Triumph, and Russell's (Abdullaev and Abdullaev, 2013). Some elite varieties with excellent comprehensive characteristics (such as '108F'; and 'Tashkent series') were developed before regular breeding programs were established in the early 20th century (Abdullaev and Abdullaev, 2013). Two major ecotypes of Upland cotton were introduced into China in the early stage. The backbone parents of the central Asia type, exhibiting early maturity characteristics, were initially introduced into the northwestern and northeastern regions of China from the former Soviet Union ('King'; cultivar). For the other two traditional cotton production regions in China (the Yangtze River and Yellow River regions), early varieties were directly selected and developed from American commercial varieties such as Stoneville and Deltapine series exhibiting broad adaptation (Shands et al., 1991). Due to environmental differences among regions, local varieties subsequently developed the corresponding features for adaptation to the local environment. In brief, the introduction of central Asian-type cotton primarily contributed early maturity-related genetic resource, while the Stoneville/Deltapine germplasm contributed extensive adaptability to modern Chinese Upland cotton cultivars.

In this study, the genetic clustering showed no distinct geographic patterns (**Figure 1D**), which might be due to the extensive environmental adaptability and frequent germplasm exchanges among regions during the breeding process. Previously, we have identified several divergent genomic regions on chromosomes A06 and A08 related to maturity and heterosis, respectively (He et al., 2019). In this study, we further confirmed these divergent regions in a larger population. More importantly, through introgression analysis, these large-scale

variations were also detected in the landraces of *G. hirsutum*. Therefore, we suggested these variations in cultivars might have resulted from introgression of landraces and wild relatives. According to their geographic information, we found two distinct groups (Group-1 and Group-2) containing different introgressed genetic components that might be responsible for their different environmental adaptation. Our GWAS results further confirmed the specific haplotypes (or genes) on A06, which might regulate maturity in the cotton population (**Figure 4**). Although the origin of the accessions in Group-1 was mixed, according to the pedigree, wild or landrace lineage introgressed fragments could be identified in most of these accessions. In contrast to Group-2, there were more superior fiber accessions from the Yangtze River region (YZR) in Group-1. Therefore, we concluded that these two regions might be the significant genomic signatures for distinguishing Upland cotton germplasm adapted from different regions in China.

Interspecific hybridization breeding has played a critical role in the cotton breeding history worldwide; via hybridization among various *Gossypium* species, abundant introgression lines carrying excellent traits (i.e., high disease resistance, superior fiber quality) have been developed over the past decades (Liang et al., 2002). According to the GWAS results, we found that some regions (or genes) on A08 were strongly associated with fiber strength (**Supplementary Figure S3**).

Based on the results of genetic structure analysis (**Figure 1C**), population differentiation and introgression analyses (**Figure 3**), we clarified that the different introgressive components located on A06 and A08 might not only represent the major forces driving population differentiation in Upland cotton cultivars but also impact the major agronomic traits. In the breeding practices, introgression lines carrying excellent traits are often associated with certain disadvantages, such as superior fiber quality being associated with late maturity and a reduced yield; this phenomenon results from a linkage block region harboring antagonistic or pleiotropic genes. In tomato, co-localization of QTLs controlling multiple traits resulted from pericentromeric introgression from wild species (Budiman et al., 2004; Haggard et al., 2015). The repression of recombination in the pericentromeric region makes it difficult to precisely map the significant QTLs within this region (Haggard et al., 2015). Based on our results, the proportion of genotypes responding for early maturity and excellent fiber strength was declined in Group-3 (which was represented for most of the modern Chinese Upland cotton population) (**Figure 4F**). Therefore, considering the extensive range of introgressive regions on the chromosome, we conjecture that some major exotic QTLs controlling contrasted traits (i.e., early maturity and poor fiber quality) might be located at the same pericentromeric regions on A06 or A08 with strong linkage disequilibrium, resulting in the decrease proportion during the breeding process. According to genome annotations, although these regions span more than 100 Mb (**Supplementary Table S2**) in total, fewer than 300 genes are found in these regions on each chromosome. We also identified some genes that were specifically expressed in the ovule or fiber, which strongly suggested that they regulate fiber development (**Supplementary Tables S5** and **S6**, **Supplementary Figures S5** and **S6**). In the future, a comprehensive approach could potentially result in breaking the linkage utilizing functional genomics complemented by hybridization and make these genes into the application as a future breeding tool. Therefore, based on GWAS results and previously identified QTLs overlapping these regions, we strongly suggest that these regions are typical genomic signatures representing introgression lines in Upland cotton cultivars and genes in these regions are essential resources worthy of future study.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA353524.

AUTHOR CONTRIBUTIONS

XD, Y-MZ, and SH conceived and designed the experiments. PW, HX, ZPe, and ZPa performed library construction and sequencing. YJ, JS, and LW collected the field data. GS and PD performed the bioinformatics analysis. SH and GS analyzed the data. SH wrote the paper. MN edited the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Key Research and Development Program of China (2016YFD0100306, 2016YFD0100203) and the National Natural Science Foundation of China (31871677).

ACKNOWLEDGMENTS

We thank the National Wild Cotton Nursery (Sanya) and the National Mid-term Gene Bank for Cotton at the Institute of Cotton Research, Chinese Academy of Agricultural Sciences, for kindly providing the samples for DNA extraction.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00929/full#supplementary-material

REFERENCES

Abdalla, A. M., Reddy, O. U. K., El-Zik, K. M., and Pepper, A. E. (2001). Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor. Appl. Genet.* 102, 222–229. doi: 10.1007/s001220051639

Abdullaev, A., and Abdullaev, A. A. (2013). Cotton germplasm collection of Uzbekistan. *Asian Australas. J. Plant Sci. Biotechnol.* 7, 1–15.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109

Brubaker, C. L., and Wendel, J. F. (1994). Reevaluating the origin of domesticated cotton (*Gossypium hirsutum*; Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). *Am. J. Bot.* 81, 1309. doi: 10.2307/2445407

Budiman, M. A., Chang, S. B., Lee, S., Yang, T. J., Zhang, H.-B., de Jong, H., et al. (2004). Localization of *jointless-2* gene in the centromeric region of tomato chromosome 12 based on high resolution genetic and physical mapping. *Theor. Appl. Genet.* 108, 190–196. doi: 10.1007/s00122-003-1429-3

Chen, G., and Du, X.-M. (2006). Genetic diversity of source germplasm of Upland cotton in China as determined by SSR marker analysis. *Acta Genet. Sin.* 33, 733–745. doi: 10.1016/S0379-4172(06)60106-6

Coppens d'Eeckenbrugge, G., and Lacape, J.-M. (2014). Distribution and differentiation of wild, feral, and cultivated populations of perennial Upland cotton (*Gossypium hirsutum* L.) in Mesoamerica and the Caribbean. *PloS One* 9, e107458. doi: 10.1371/journal.pone.0107458

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

Ding, M., Ye, W., Lin, L., He, S., Du, X., Chen, A., et al. (2015). The hairless stem phenotype of cotton (*Gossypium barbadense*) is linked to a *copia*-like retrotransposon insertion in a Homeodomain-Leucine Zipper Gene (*HD1*). *Genetics* 201, 143–154. doi: 10.1534/genetics.115.178236

Ding, W., Lin, L., Zhang, B., Xiang, X., Wu, J., Pan, Z., et al. (2015). OsKASI, a β-ketoacyl-[acyl carrier protein] synthase I, is involved in root development in rice (*Oryza sativa* L.). *Planta* 242, 203–213. doi: 10.1007/s00425-015-2296-2

Fang, D. D., Hinze, L. L., Percy, R. G., Li, P., Deng, D., and Thyssen, G. (2013). A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (*Gossypium hirsutum* L.) cultivars from major cotton-growing countries. *Euphytica* 191, 391–401. doi: 10.1007/s10681-013-0886-2

Fang, L., Wang, Q., Hu, Y., Jia, Y., Chen, J., Liu, B., et al. (2017). Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits. *Nat. Genet.* 49, 1089–1098. doi: 10.1038/ng.3887

Food and Agriculture Organization (FAO) (2018). Available at: http://www.fao.org/faostat/en/#search/cotton.

Gallagher, J. P., Grover, C. E., Rex, K., Moran, M., and Wendel, J. F. (2017). A New Species of Cotton from Wake Atoll, *Gossypium stephensii* (Malvaceae) *Systematic Botany* 42, 115–123. doi: 10.1600/036364417X694593

Glass, M., Barkwill, S., Unda, F., and Mansfield, S. D. (2015). Endo-β-1,4-glucanases impact plant cell wall development by influencing cellulose crystallization. *J. Integr. Plant Biol.* 57, 396–410. doi: 10.1111/jipb.12353

Gover, C. E., Gallagher, J. P., Jareczek, J. J., Page, J. T., Udall, J. A., Gore, M. A., et al. (2015). Re-evaluating the phylogeny of allopolyploid Gossypium L. *Mol. Phylogenet. Evol.* 92, 45–52. doi: 10.1016/j.ympev.2015.05.023

Haggard, J. E., Johnson, E. B., and Clair, D. A. S. (2015). Multiple QTL for horticultural traits and quantitative resistance to Phytophthora infestans linked on *Solanum habrochaites* chromosome 11. *G3.* (*Bethesda*), 5 (2), 219–233. doi: 10.1534/g3.114.014654

He, S., Sun, G., Huang, L., Yang, D., Dai, P., Zhou, D., et al. (2019). Genomic divergence in cotton germplasm related to maturity and heterosis. *J. Integr. Plant Biol.* 61, 929–942. doi: 10.1111/jipb.12723

Hinze, L. L., Dever, J. K., and Percy, R. G. (2012). Molecular Variation Among and Within Improved Cultivars in the U.S. Cotton Germplasm Collection. *Crop Sci.* 52, 222. doi: 10.2135/cropsci2011.04.0202

Ho-Yue-Kuang, S., Alvarado, C., Antelme, S., Bouchet, B., Cézard, L., Le Bris, P., et al. (2016). Mutation in Brachypodium caffeic acid O-methyltransferase 6 alters stem and grain lignins and improves straw saccharification without deteriorating grain quality. *J. Exp. Bot.* 67, 227–237. doi: 10.1093/jxb/erv446

Hulse-Kemp, A. M., Lemm, J., Plieske, J., Ashrafi, H., Buyyarapu, R., Fang, D. D., et al. (2015). Development of a 63K SNP array for cotton and high-density mapping of intraspecific and interspecific populations of *Gossypium* spp. *G3* 5, 1187–1209. doi: 10.1534/g3.115.018416

Hutchinson, J. B., Silow, R. A., and Stephens, S. G. (1947). *The Evolution of Gossypium and the Differentiation of the Cultivated Cottons* (Trinidad: Empire Cotton Growing Corporation, Cotton Research Station).

Iqbal, M. J., Reddy, O. U. K., El-Zik, K. M., and Pepper, A. E. (2001). A genetic bottleneck in the 'evolution under domestication' of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. *Theor. Appl. Genet.* 103, 547–554. doi: 10.1007/PL00002908

Islam, M. S., Zeng, L., Thyssen, G. N., Delhom, C. D., Kim, H. J., Li, P., et al. (2016). Mapping by sequencing in cotton (*Gossypium hirsutum*) line MD52ne identified candidate genes for fiber strength and its related quality attributes. *Theor. Appl. Genet.* 129, 1071–1086. doi: 10.1007/s00122-016-2684-4

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548

Lacape, J. M., Dessauw, D., Rajab, M., Noyer, J. L., and Hau, B. (2006). Microsatellite diversity in tetraploid Gossypium germplasm: assembling a highly informative genotyping set of cotton SSRs. *Mol. Breed.* 19, 45–58. doi: 10.1007/s11032-006-9042-1

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R. J., et al. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* 33, 524–530. doi: 10.1038/nbt.3208

Liang, Z., Jiang, R., Zhong, W., He, J., Sun, C., Qou, Z., et al. (2002). Creation of the technique of interspecific hybridization for breeding in cotton. *Sci. China Ser. C-Life Sci.* 45, 331–336. doi: 10.1360/02yc9036

Ma, Z., He, S., Wang, X., Sun, J., Zhang, Y., Zhang, G., et al. (2018). Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nat. Genet.* 50, 803–813. doi: 10.1038/s41588-018-0119-7

McGinnis, S., and Madden, T. L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* 32, W20–W25. doi: 10.1093/nar/gkh435

Moinuddin, S. G. A., Jourdes, M., Laskar, D. D., Ki, C., Cardenas, C. L., Kim, K.-W., et al. (2010). Insights into lignin primary structure and deconstruction from *Arabidopsis thaliana* COMT (caffeic acid O-methyl transferase) mutant *Atomt1*. *Org. Biomol. Chem.* 8, 3928–3946. doi: 10.1039/c004817h

Pang, C., Du, X., and Ma, Z. (2006). Evaluation of the introgressed lines and screening for elite germplasm in *Gossypium*. *Chin. Sci. Bull.* 51, 304–312. doi: 10.1007/s11434-006-0304-4

Paterson, A. H., Brubaker, C. L., and Wendel, J. F. (1993). A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep.* 11, 122–127. doi: 10.1007/BF02670470

Percival, A. E., Wendel, J. F., and Stewart, J. M. (1999). "Cotton: origin, history, technology, and production," in *Taxonomy and Gemplasm Resources*. Eds. C. W. Smith and J. T. Cothren (New York: John Wiley & Sons, Inc.), 33–63.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

Qin, Y.-M., Hu, C.-Y., Pang, Y., Kastaniotis, A. J., Hiltunen, J. K., and Zhu, Y.-X. (2007). Saturated very-long-chain fatty acids promote cotton fiber and Arabidopsis cell elongation by activating ethylene biosynthesis. *Plant Cell* 19, 3692–3704. doi: 10.1105/tpc.107.054437

Shands, H. L., Wiesner, L. E., Meredith, W. R., and Wiesner, L. E. (1991). "Contributions of introductions to cotton improvement," in *Use of Plant Introductions in Cultivar Development: Proceedings of a Symposium Part 1*. Ed. H. L. Shands (Madison: Crop Science Society of America), 127–146.

Sun, X., Gong, S. Y., Nie, X. Y., Li, Y., and Li, W. (2015). A R2R3-MYB transcription factor that is specifically expressed in cotton (Gossypium hirsutum) fibers affects secondary cell wall biosynthesis and deposition in *Arabidopsis*. *Physiol. Plant.* 154, 420–432. doi: 10.1111/ppl.12317

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Tyagi, P., Gore, M. A., Bowman, D. T., Campbell, B. T., Udall, J. A., and Kuraparthy, V. (2014). Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). *Theor. Appl. Genet.* 127, 283–295. doi: 10.1007/s00122-013-2217-3

Wang, K., Song, X., Han, Z., Guo, W., Yu, J. Z., Sun, J., et al. (2006). Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence in situ hybridization mapping. *Theor. Appl. Genet.* 113, 73–80. doi: 10.1007/s00122-006-0273-7

Wang, S., Chen, J., Zhang, W., Hu, Y., Chang, L., Fang, L., et al. (2015). Sequence-based ultra-dense genetic and physical maps reveal structural variations of allopolyploid cotton genomes. *Genome Biol.*, 1–18. doi: 10.1186/s13059-015-0678-1

Wang, Y., Ning, Z., Hu, Y., Chen, J., Zhao, R., Chen, H., et al. (2015). Molecular mapping of restriction-site associated DNA markers in allotetraploid Upland cotton. *PloS One* 10, e0124781. doi: 10.1371/journal.pone.0124781

Wang, M., Tu, L., Lin, M., Lin, Z., Wang, P., Yang, Q., et al. (2017). Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nat. Genet.* 49, 579–587. doi: 10.1038/ng.3807

Wendel, J. F., and Grover, C. E. (2015). "Taxonomy and evolution of the cotton genus, *Gossypium*," in *Cotton 2nd*. Eds. D. D. Fang and R. G. Percy (Madison: American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc.), 1–2.

Wendel, J. F., Brubaker, C. L., and Percival, A. E. (1992). Genetic Diversity in Gossypium hirsutum and the Origin of Upland Cotton. *Am. J. Bot.* 79, 1291. doi: 10.2307/2445058

Wendel, J. F., Brubaker, C. L., and Seelanan, T. (2010). *Physiology of Cotton*. Eds. J. M. Stewart, D. M. Oosterhuis, J. J. Heitholt and J. R. Mauney (Netherlands: Springer), 1–18.

Zhang, F., Zuo, K., Zhang, J., Liu, X., Zhang, L., Sun, X., et al. (2010). An L1 box binding protein, GbML1, interacts with GbMYB25 to control cotton fibre development. *J. Exp. Bot.* 61, 3599–3613. doi: 10.1093/jxb/erq173

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207

Zhang, Z., Shang, H., Shi, Y., Huang, L., Li, J., Ge, Q., et al. (2016). Construction of a high-density genetic map by specific locus amplified fragment sequencing (SLAF-seq) and its application to Quantitative Trait Loci (QTL) analysis for boll weight in upland cotton (*Gossypium hirsutum.*). *BMC Plant Biol.* 16, 79. doi: 10.1186/s12870-016-0741-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 July 2020
doi: 10.3389/fpls.2020.01067

Diverse Wheat-Alien Introgression Lines as a Basis for Durable Resistance and Quality Characteristics in Bread Wheat

Eva Johansson[1], Tina Henriksson[2], Maria Luisa Prieto-Linde[1], Staffan Andersson[1], Rimsha Ashraf[1] and Mahbubjon Rahmatov[1]*

[1] Department of Plant Breeding, The Swedish University of Agricultural Sciences, Alnarp, Sweden, [2] Lantmännen Lantbruk, Svalöv, Sweden

OPEN ACCESS

Edited by:
Mariola Plazas, Polytechnic University of Valencia, Spain

Reviewed by:
Irina N. Leonova, Russian Academy of Sciences, Russia
Julio Huerta Espino, Agrícolas y Pecuarias (INIFAP), Mexico

***Correspondence:**
Eva Johansson
Eva.johansson@slu.se

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *09 April 2020*
Accepted: *29 June 2020*
Published: *17 July 2020*

Citation:
Johansson E, Henriksson T, Prieto-Linde ML, Andersson S, Ashraf R and Rahmatov M (2020) Diverse Wheat-Alien Introgression Lines as a Basis for Durable Resistance and Quality Characteristics in Bread Wheat. Front. Plant Sci. 11:1067. doi: 10.3389/fpls.2020.01067

Wheat productivity has been significantly improved worldwide through the incorporation of novel genes from various gene pools, not least from wild relatives of wheat, into the commonly cultivated bread and durum wheat. Here, we present and summarize results obtained from a diverse set of wheat-alien introgression lines with mainly introgressions of rye, but also of *Leymus* spp. and *Thinopyrum junceiforme* into bread-wheat (*Triticum aestivum* L.). From this material, lines carrying 2RL were found with good agronomic performance and multiple resistance not least towards several races of powdery mildew. A novel resistance gene, one of few showing resistance towards all today identified stem rust races, designated *Sr59*, was also found originating from 2RL. Lines with multiple introgressions from 4R, 5R, and 6R were found resistant towards the majority of the stripe rust races known today. Due to lack of agricultural adaptation in these lines, transfer of useful genes into more adapted wheat material is a necessity, work which is also in progress through crosses with the CS*ph1b* mutant, to be able to only transfer small chromosome segments that carry the target gene. Furthermore, resistance towards Russian wheat aphid was found in lines having a substitution of 1R (1D) and translocations of 3DL.3RS and 5AL.5RS. The rye chromosomes 1R, 2R, and 6R were found responsible for resistance towards the Syrian Hessian fly. High levels of especially zinc was found in several lines obtained from crosses with *Leymus racemosus* and *Leymus mollis*, while also some lines with 1R, 2R, or 5R showed increased levels of minerals and in particular of iron and zinc. Moreover, lines with 1R, 2R, 3R, and *Leymus* spp. introgressions were also found to have a combination of high iron and zinc and low cadmium concentrations. High variation was found both in grain protein concentration and gluten strength, measured as %UPP, within the lines, indicating large variation in bread-making quality. Thus, our study emphasizes the impact that wheat-alien introgression lines can contribute to current wheat lines and shows large opportunities both to improve production, resistance, and quality. To obtain such improvements, novel plant breeding tools, as discussed in this

paper, opens unique opportunities, to transfer suitable genes into the modern and adapted wheat cultivars.

Keywords: agronomic performance, baking quality, breeding, disease and pest resistance, *Leymus* spp., *Secale cereale* L., *Triticum aestivum* L.

INTRODUCTION

Wheat is one of the three major crops of importance for food security worldwide, the other two being rice and maize (FAO, 2016). Bread wheat (*Triticum aestivum* L.) is a hexaploid and the most commonly cultivated species of wheat (95%), belonging to the tribe Triticeae and the family Poaceae (McFadden and Sears, 1946; Dubcovsky and Dvorak, 2007). The second most commonly cultivated form of wheat is durum wheat (*Triticum durum* L.), contributing 5% to the total production (Dubcovsky and Dvorak, 2007). In total, wheat contributes 20% of the total calories and proteins consumed by the human population, thereby contributing to a higher total protein intake than the whole total meat consumption summed (Shewry and Hey, 2015).

Due to the high contribution of wheat to the daily human food intake, human food security is highly vulnerable to the increasing threats to wheat production from climate change, including global warming (Steenwerth et al., 2014). Wheat yield is also negatively affected by abiotic and biotic stresses resulting in economic losses to farmers (Husenov et al., 2020). The population growth predicted to be more than 9 billion people worldwide in 2050, result in additional demand on food production, simultaneously bringing an increasing competition for arable land for food production (FAO, 2016). To meet these challenges, novel wheat cultivars are urgently needed, adapted to contribute high yield under sustainable and demanding cultivation conditions (Shiferaw et al., 2013). For this purpose, novel plant breeding methodologies have to be developed in order to most beneficially use available genetic resources and smart and rapid plant development to produce the needed wheat materials in time to cope with needs and challenges.

Plant breeding to obtain sustainable, high resistance and high-quality crops are dependent on suitable genes for the wanted traits. For many traits, such genes are available within the breeding material in on-going breeding programs for the crop and will be easily transferred by breeders through ordinary crossing schemes. However, domestication and breeding practices have reduced the presence of rare and favorable allelic variation to biotic and abiotic stresses and environmental changes originally found in the wild relatives (Tanksley and McCouch, 1997; Singh et al., 2018). Therefore, wild relatives, landraces, and close relatives of wheat are a unique source of novel genetic variations for introgression into modern cultivars (Molnár-Láng et al., 2015). For wheat, several useful transfers of genes from landraces have been reported including e.g. the *Rht* dwarfing genes, the powdery mildew resistance gene *Pm24*, and several biotic and abiotic stress resistance genes (Kihara, 1983; Zeven, 1998; Huang and Röder, 2011; Cavanagh et al., 2013; Singh et al., 2018). Also, genes have been transferred to wheat from non-*Triticum* (alien) species, where transfers from e.g. rye (*Secale cereale*) have resulted in widely cultivated wheat cultivars (McIntosh et al., 1995; Friebe et al., 1996). The most successful alien transfer into the wheat genome is that of the 1RS chromosome segment, in the form of 1AL.1RS, 1BL.1RS, and 1DL.1RS translocations (Rabinovich, 1998; Mago et al., 2015), contributing several resistance genes for powdery mildew, leaf, stripe and stem rusts. Out of them, the 1BL.1RS wheat-rye translocation has contributed immensely to global wheat production as a source of resistance genes (*Sr31*/*Yr9*/*Lr26*/*Pm9*) to wheat fungal diseases (Schlegel, 2020), but it is also known to contribute weak and sticky dough (Dhaliwal et al., 1988). Rye is a unique source of many important traits for wheat improvement, e.g. the resistance genes *Sr27*, *Sr50*, *Sr1RS*^{Amigo}, *Lr25*, *Lr45*, *Pm7*, etc. have been identified from rye (The et al., 1991; Marais and Marais, 1994; McIntosh et al., 1995; Friebe et al., 1996), although these genes have contributed limitedly to agricultural production until now. More recently, some novel resistance genes from rye i.e. *Sr59*, *Yr83*, and *Pm56* have been introgressed into wheat (Rahmatov et al., 2016a; Hao et al., 2018; Li et al., 2020), which may be used as durable sources against fungal diseases. Herbicide-resistant evolution is challenging weed management; therefore, the allelopathic potential is a good solution to mitigate weed management in crop production. Bertholdsson et al. (2012), reported that rye is an excellent source of allelopathic potential that can be used for wheat breeding. In addition, Iron (Fe) and Zinc (Zn) deficiency are severely affecting human health, causing several physiological disorders, symptomatic anemia, stunting, etc., and therefore high content in staple crops such as wheat are of outmost importance (Johansson et al., 2014; Johansson et al., 2020). The recent great advancements in genomic and cytogenetic tools open opportunities to transfer alien resistance genes to wheat, simultaneously avoiding linkage drag issues.

The present paper is focusing on opportunities and challenges of the use of a diverse set of wheat-alien introgression lines with mainly introgressions of rye, but also of *Leymus* spp. and *Thinopyrum junceiforme* into bread-wheat (*T. aestivum* L.). This provides useful insight into the identification and characterization of wheat-alien introgression lines based on several studies through diseases and pests screening, agronomic performances and molecular markers. Resistances and quality characteristics of wheat within this material, connections to introgressed chromosomes, localization of genes, and status for transfer of these genes are described here. Finally, a short overview is given as to the impact of novel breeding strategies for the use of alien germplasm in modern breeding.

MATERIALS AND METHODS

Plant Materials

A set of winter and spring wheat-alien introgression lines maintained at the Plant Breeding Department at the Swedish

University of Agricultural Sciences were used in different part of the hereby presented studies. These lines were developed by crossing and backcrossing strategies during 1980 to 2000 by the late Professor Arnulf Merker at the Swedish University of Agricultural Sciences (**Table 1**). The wheat-alien introgression lines used for the present paper contained rye chromosomes with 1R, 2R, 3R, 4R, 5R and 6R in the form of a single disomic substitution wheat–rye translocations such as 1DL.1RS, 1BL.1RS, 2BS.2RL, 3DL.3RS and 5AL.5RS, lines with multiple combinations of rye chromosome substitutions such as 1R + 2R, 1R + 3R, 1R + 6R, 5R + 4R + 7R and 1R + 6R + 4R + 7R (Merker, 1979; Merker and Rogalska, 1984), and lines with introgressed chromatin from *Leymus mollis*, *Leymus racemosus*, and *T. junceiforme* (Ellneskog-Staam and Merker, 2001; Ellneskog-Staam and Merker, 2002). The full material used has previously been completely described in Rahmatov (2016) and Rahmatov et al. (2017).

Field Trials

A total of 180 of the winter wheat lines and 57 of the spring wheat lines were evaluated by field trials for multiple resistance and agronomic performance during two executive seasons, 2014 and 2015, in Svalöv, Sweden and in Harzhof and Laberweinting in Germany. During these seasons, the lines were continuously evaluated and scored (scale 1–9) for lodging (winter wheat) and presences of diseases (spring and winter wheat). Comparisons of presence of diseases and alien material were carried out (Andersson et al., 2016).

Diseases Screening

Stem rust seedling resistance assays with ten *Pgt* races and adult plant responses with three *Pgt* races (TTKSK + TTKST, TKTTF and MCCFC), were carried out on 185 and 94 of the winter and spring wheat-alien introgression lines under field conditions following the procedure described in Hysing et al. (2007) and Rahmatov et al. (2015); Rahmatov et al. (2016a, b). For the stripe rust evaluations, 189 of the winter and 73 of the spring wheat-alien introgression lines were tested in the seedling and adult plant stages. Twelve stripe rust races with different virulence/avirulence combinations and geographic origins were used for screening at the seedling stage along with adult plant evaluations in the field according to Rahmatov et al. (2017). Hysing et al. (2007), evaluated a set of 2BS.2RL wheat-rye translocation lines against stripe rust, leaf rust, and powdery mildew races.

Hessian Fly and Russian Wheat Aphid Screenings

A total of 57 spring and 185 winter wheat-alien introgression lines were evaluated in 2011 and 2012 at the seedling stage against Hessian fly (HF) and the Russian wheat aphid (RWA) in collaborations with ICARDA. In brief, the rearing rooms for HF experiments were kept at 20°C, Rh 70–80%, and the cycle of 16/8 h light/dark was used. Six or ten seeds per wheat accession were sown in hill plots in metal flats 55 × 45 × 10 cm, in total 48 accessions per box plus controls, in a mixture of soil:sand:peat (2:1:1). After 5–6 days, at the one-leaf stage, infestation by HF was done with about 30 females and 10 males under net for 3–4 days (El Bouhssini et al., 2013). The scoring took place 20 days after infestation, with the number of resistant and susceptible plants per accession. The first screening was conducted in the spring and winter materials in 2011, and a second screening was only conducted in the winter materials in 2012. Based on these two screenings, lines with 100% resistance reaction to HF were selected for further confirmations in four separate screenings.

The RWA biotype was collected from Tel Hadya, Syria, and thereafter reared on the susceptible wheat cultivar (Andersson et al., 2015). The experiments were carried out in a greenhouse at 19–20°C, with light/dark photoperiod 16/8 h and relative humidity of about 60%. The accessions were planted in a

TABLE 1 | Wheat-alien introgression lines and respective parents evaluated in this study.

Cross/Pedigree	Plant habit	No. of lines	Type	Reference
Triticale[a]	Spring and winter	5	×*Triticosecale*	Forsstrom and Merker, 2001
Wheat[a]	Spring and winter	8	*Triticum aestivum* and *Tr. carthlicum*	Forsstrom et al., 2002
Sv 876012 x H	Winter	37	Wheat–rye introgressions	Forsstrom and Merker, 2001
Sv 876032 x H x K	Winter	54	Wheat–rye introgressions	Forsstrom and Merker, 2001
Sv 856003 x H	Winter	6	Wheat–rye introgressions	Forsstrom and Merker, 2001
Sub 1R + 2R	Winter	42	Wheat–rye introgressions	Forsstrom and Merker, 2001
Malysh	Winter	6	Wheat–rye introgressions	Merker, 1984
Starke × Otello	Winter	7	Wheat–rye introgressions	Merker, 1984
Uno × Holme	Winter	8	Wheat–rye introgressions	Merker, 1984
Triticale VT828041	Spring	6	Wheat–rye introgressions	Merker, 1984
Triticale Drira	Spring	23	Wheat–rye introgressions	Merker, 1984
Triticale Beagle	Spring	12	Wheat–rye introgressions	Merker, 1984
Triticale VT83 615	Spring	2	Wheat–rye introgressions	Merker, 1984
Triticale VT83 591	Spring	4	Wheat–rye introgressions	Merker, 1984
Triticale VT 82 8039	Spring	5	Wheat–rye introgressions	Merker, 1984
3R BB14 (Cimmyt 1974)	Spring	4	Wheat–rye introgressions	Merker, 1984
Leymus mollis	Winter	42	Wheat–*L. mollis*introgressions	Merker and Lantai, 1997
Leymus racemosus	Spring	22	Wheat–*L. racemosus* introgressions	Merker and Lantai, 1997
Th. junceiforme	Spring	16	Wheat–*T. junceiforme* introgressions	Merker and Lantai, 1997
[a]Parental cultivars and Lines	309	TOTAL		

randomized (alpha design) order together with susceptible and resistant controls in each planting tray, in a mixture of soil, sand, and peat (2:1:1). An evaluation was done when symptoms were seen on susceptible checks, using the ICARDA RWA damage scale with a 1–3 scale for leaf rolling (LR) and 1–6 scale for leaf chlorosis (LC) (El Bouhssini et al., 2011). In the second advanced screening, selected accessions from the first screening results were repeated at four separate times (Andersson et al., 2015).

Allelopathic Potential of Wheat-Alien Introgression Lines

Allelopathic potential of the wheat–rye introgression lines were tested according to Bertholdsson et al. (2012). In this study, seeds of *Chenopodium alba*, *Lolium perenne*, *Brassica napus*, *Lactuca sativa*, *Eruca sativa*, *Sinapis indicum* and *Sinapis alba* were used to find high root growth inhibition when grown together with rye. In this investigation, four pregerminated cereal seedlings were planted along the wall of 400-ml Phytotech tissue culture vials (bottom diameter 75 mm) filled with 20 ml 0.35% water agar, and eight pregerminated mustard seedlings (*S. alba* cv. Medicus) were planted in a circle in the center of the vials. The experiment was tested in four replicates, and the dry weight of the shoot and root were measured (Bertholdsson et al., 2012).

Analysis of Grain Samples for Micronutrients Concentration and Protein Composition

A total of 40 of the lines were evaluated for micronutrients (e.g. Iron, Zinc, and Cadmium) content with Inductively Coupled Plasma Mass Spectrometry (ICPMS) at the University of Minnesota, similarly as described in Hussain et al. (2010) and Moreira-Ascarrunz et al. (2016). Briefly, all samples were ashed in a muffle furnace for 12 h at 485°C. Then, the ash was dissolved in 5 ml of 20% HCl followed by dilution with 5 ml of deionized water. The ICPMS provides concentration assays for several microelements, including zinc, iron, and cadmium in mg/Kg.

The complete set of winter wheat alien translocation lines were analysed with SE-HPLC according to Johansson et al. (2001) to evaluate the total amount of SDS-extractable proteins (TOTE) and percentage of unextractable polymeric protein in total polymeric protein (%UPP). A high correlation is known to exist between TOTE and grain protein concentration and between %UPP and gluten strength (Malik et al., 2011; Malik et al., 2013) and thereby this methodology can be used to understand relationships with bread-making quality (Hussain et al., 2012; Hussain et al., 2013; Vazquez et al., 2019).

Statistical Analyses

The statistical software SAS 9.3 (SAS, 2011) was used for principal component analyses (PCA) calculations to understand relationships between minerals and protein factors with evaluated wheat-alien introgression lines. In order to understand and visualize the distribution and relationship between variables and factors evaluated, principal component analysis (PCA) can be applied to orthogonally represent the variables in a data matrix vector. PCA is known to show the distribution of dependent variables and independent factors, in a loading and score plot, respectively (Wold et al., 1987). Values of content of Iron, Zinc and Cadmium were calculated by mini tab for wheat, Triticale, wheat–rye and wheat-*Leymus* lines and presented as boxplots with lowest and highest observations as well as lower and upper quartile and median.

RESULTS

Multiple Resistance and Agronomic Performance

The lines showed varying agronomic performance, with some lines being almost comparable to currently grown wheat in Sweden while others differed largely. Large variation was found in the material both for lodging and presence of diseases (**Figure 1**). However, the majority of the winter wheat lines had strong stem and with limited lodging, thus making them of interest as a source of lodging resistance (**Figure 1B**). Presence of 1R, 2R, 3R, 5R, 1R + 6R and *L. racemosus* correlated with decreased levels of infections with powdery mildew, *Zymoseptoria tritici (*causal agent of *Septoria triticae* blotch) and Fusarium head blight during field conditions. Lower levels of leaf, stem and stripe rusts infection responses were found in lines with 1R, 2R, 3R, 1R + 3R, 1R + 6R, and *L. racemosus*, respectively.

Rusts Screenings

From the stem rust seedling evaluation, eleven 2R (2B), three 2R (2D), and three 3R (3D) wheat–rye disomic substitution lines, and seven wheat-*T. junceiforme* were found to carry potentially new stem rust resistant gene/s (**Table 2**). Based on the ten *Pgt* races, known resistance genes could not be postulated because their reactions did not correspond to the avirulence/virulence profile of the races tested. All lines that were resistant at the seedling stage remained resistant at the adult plant stage against races TTKSK + TTKST in Kenya and TKTTF in Turkey. Trace resistance was found in several of the lines tested at St. Paul, Minnesota, against the race MCCFC (**Table 2**), although only a few number of lines were tested due to winter type of the material and limited seed available.

The wheat-alien introgression lines showed high variability in resistance/susceptibility reactions against the twelve stripe rust isolates applied to screen for resistance genes (**Table 3**). The screening resulted in 149 lines (57% of the lines), postulated to contain a combination of known *Yr* genes e.g. *Yr1*, *Yr2*, *Yr9*, and *Yr32*. However, six of the multiple wheat–rye introgression lines with 4R, 5R and 6R were identified as highly resistant against a total of 25 stripe rust races, including the twelve used for the full material (**Table 3**). Thus, these six lines might possess a new stripe rust resistance gene/s. Molecular cytogenetic analysis showed that the 4R, 5R and 6R rye chromosomes substituted 4D, 5D and 6D wheat chromosomes. Further studies are going on for determining the underlying genetic basis of these resistance gene/s.

Aphid and Hessian Fly Resistance

Among the total of 242 evaluated lines, 235 germinated and showed a high variation in resistance to RWA (**Figure 2**). A total

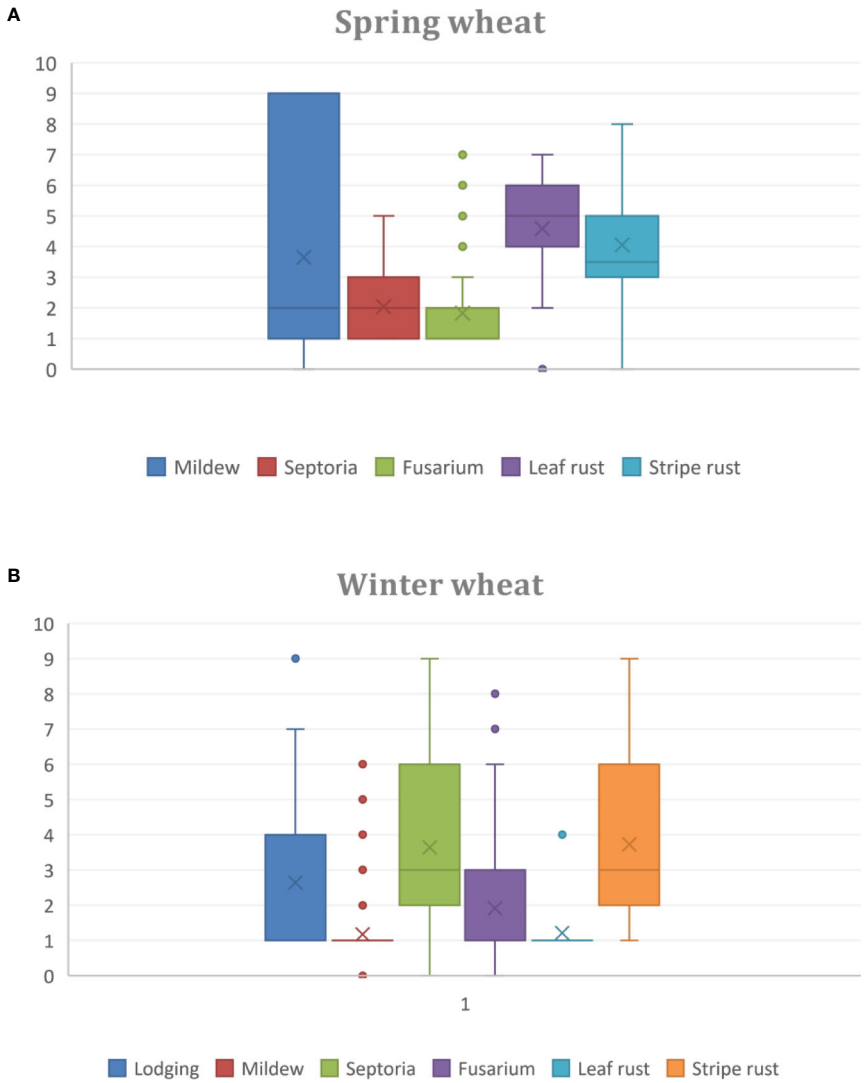

FIGURE 1 | Boxplots showing variation in lodging and various diseases based on scoring of the material from 0 to 9, in wheat alien introgression lines of **(A)** Spring wheat, and **(B)** Winter wheat, from field trials during two years in Sweden and Germany. In each boxplot, five bars are represented, indicating smallest observation, lower quartile, median, upper quartile, and largest observation, respectively. X marks the mean value.

of 23 accessions were identified as resistant against the RWA. Resistance was found to RWA, particularly in accessions having substitutions of 1R instead of 1D [1R (1D) or 1R (1D) + 6R (6D)], in translocations to 3D or 5A (3DL.3RS and 5AL.5RS) and accessions with introgressions of *L. mollis*.

The first screening (242 lines) for HF resistance showed 11 winter and two spring wheat accessions with 100% resistance, while in the second screening, nine of the 11 winter wheat accessions were proofed with 100% resistance, which also holds true for the additional four repeated screenings (**Table 4**). These fully resistant winter wheat accessions contained 1R, 1R + 6R, 1RS + 2RL, 1RL + 2RL, 2RL, and 2R translocations or substitutions. The presence of these genes in our alien wheat material might be one explanation for the HF resistance found although the presence of full resistance in accessions with the substitution 1R.1D in winter wheat and the translocation 1RS.1DL in spring wheat indicate the presence of additional unknown resistance genes in the present material. Besides, high and partial levels of resistance with the presence of 1R, 1RS, 2R, 3R, 3RS, 4R, 5R, 6RL, and *L. racemosus* and *L. mollis* substitutions and translocations were found promising sources against HF.

Nutritional Benefits

Principal component analyses indicated high levels of Cadmium (Cd) in the winter wheat lines as compared to the rest of the evaluated lines, while *Leymus* spp. was indicated as containing high levels of Iron (Fe) and Zinc (Zn; **Figure 3**). Mean values of

TABLE 2 | Stem rust seedling and adult plant resistance tests in the wheat–rye and wheat–*T. Junceiforme* introgression lines with potential sources of new stem rust resistance gene/s.

#	Chromosome	Seedling Resistance Test											Adult Plant Resistance		
		TTKSK, 1 Rep.	TTKSK, 2 Rep.	TPMKC	TTTTF	QTHJC	RKQQC	TTKST	TRTTF	TTTSK	TKTTF	MCCFC	TTKSK+ TTKST	TKTTF	MCCFC
SLU73	2R susbstituted 2B	2+	2	3+	;1	;01/3+	11+	2/2+	1+2/3+4	22+	1/2+	4	20MR	10MR	–
SLU74	2R susbstituted 2B	2+	2+	1 3+ Z	;1/1	;-1	;1	;11+	1+2/3+	22+	1/2+	4	30MR	10MR	–
SLU75	2R susbstituted 2B	2	2	3+	0;/0; 3+	;-1	;1	11+	33+	22+	11+	4	20MR	10MR	–
SLU76	2R susbstituted 2B	2	2	3+	;1	;1/2-	;11+	;11+	11+3	22+	3+	4	30MRMS	10MS	–
SLU77	2R susbstituted 2B	2	2	3+	2+3/4/;1	;1/1+3	;11+/2	;11+/2/3+	33+	22+	3+	4	30MRMS	10MRMS	–
SLU78	2R susbstituted 2B	2+	2+	3+	;/3-	1+3-	;11+3-	;11+2+	22+3+	22+	3+	4	30MRMS	10MS	–
SLU79	2R susbstituted 2B	2	2	3+	;/3-	;13-	;13-/3	;12+	11+/3+	22+	1	4	30MRMS	10MR	–
SLU80	2R susbstituted 2B	2+	2+	33+	3+	33+/;13-	33+/;13-	;11+/3+	1+2/3+	22+	3+	4	30MRMS	10MS	–
SLU81	2R susbstituted 2B	2+	2+	3+	3-/1;	11+	11+	1+3-	33+	22+	3+	4	40MRMS	10MS	–
SLU82	2R susbstituted 2B	2+	2+	3+	; 1 3-/4	;13-/3+	;13-/3+	;1/3	3+	22+	3+	4	40MRMS	10MS	–
SLU83	2R susbstituted 2B	2+	2+	4	; 1 3-	13-/3+	13-/3+	11+/3+	33+	22+/3+	3+	4	40MRMS	10MS	–
SLU210	2R susbstituted 2D	0;	0;	0;	1+	0;	;1	0;	;12	;1	1	0;	20RMR	5RMR	TR
SLU214	3R susbstituted 3D	0;	0;	11+	4	0;	0;	0;	;11+	3+	3+	0;	20R	10MRMS	TR
SLU219	3R susbstituted 3D	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	5TR	5R	TR
SLU222	3R susbstituted 3D	0;	0;	33+	4	1+2/3-	3	0;	22+	;	33+	11+	10RMR	70S	10R
SLU238	2R susbstituted 2D	1	1	2	22-	12-	;1	;11+	;01	;1	1-	;01-	10R	10RMR	TR
SLU239	2R susbstituted 2D	1	1-	2	22-	12-	;1	;01	;01	;1	1-	;01-	20RMR	5RMR	TR
SLU251	Th.-Wheat	0;	0;	3+	4	33+	3+	0;1	3+	0;	3+	4	–	–	–
SLU252	Th.-Wheat	0;	0;	33+	3+	33+	3+	0;	3+4	0;	3+	3+	–	–	–
SLU253	Th.-Wheat	1	0;	33+	3+	33+	3+	1+3-	2+3-	0;	3+	3+	–	–	–
SLU255	Th.-Wheat	0;	0;	3+	4	33+	3+	0,1	3+	0;	3+	3+	–	–	–
SLU256	Th.-Wheat	0;	0;	3+	3+	33+	3+	0;	3+4	0;	3+	3+	–	–	–
SLU274	Th.-Wheat	0;	0;	3+	4	33+	3+	0,1	3+	0;	3+	3+	–	–	5MS
SLU275	Th.-Wheat	0;	0;	33+	3+	33+	3+	0;	3+4	0;	3+	3+	–	–	30MS

Seedling infection types observed based on 0–4 scale (Stakman et al., 1962). The lines with;0–2+ types considered as resistant. The lines with 3–4 types considered as susceptible. Adult plant response was evaluated based on the Cobb Scale (Peterson et al., 1948) and host response to infection based on pustule type and size (Roelfs et al., 1992). TR, Trace Resistance; R, Resistance; MR, Moderately Resistance; MRMS, Moderately Resistance to Moderately Susceptible; and MS, Moderately Susceptible. A total of 94 lines of the total material were screened for adult plant resistance, explaining the lack of data for some of the lines presented here.

TABLE 3 | Resistance(R)/susceptibility(S) of wheat-alien introgression lines to isolates of *Puccinia striiformis tritici*.

Isolates	6 lines fromSv 876012 × H with 4R + 5R + 6R	256 lines			
		R	MR	MS	S
SE 205/12	R	12	23	102	119
UK 94/519	R	49	2	24	182
DK 66/02	R	84	4	29	140
Taj 01a/10	R	17	34	87	119
ER 02/03	R	174	15	54	14
DK 11/09	R	178	11	46	21
DK 71/93	R	66	34	99	56
AF 87/12	R	207	10	22	16
DK09/11	R	39	14	104	98
DK 122/09	R	6	15	126	109
SE 100/09	R	220	8	16	12
TR 34/11	R	170	41	8	32

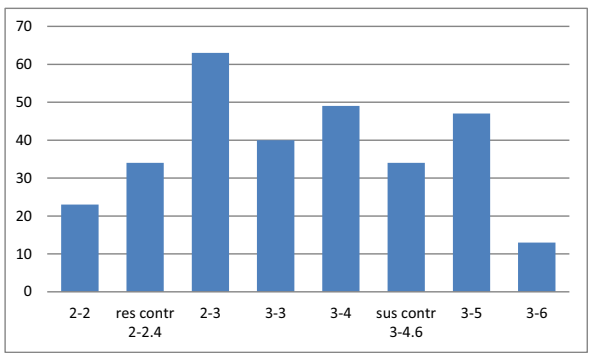

FIGURE 2 | Evaluation of 235 wheat alien introgression lines for resistance against Russian wheat aphid (RWA), resulting in number of lines with different scales of resistance. Scale: leaf rolling (LR)-leaf chlorosis (LC); 1-1 and 1-2 = highly resistant, 2-2 = resistant, 2-3 = moderately resistant, 3-3 moderately susceptible, 3-4, 3-5, 3-6 = susceptible. For the susceptible and resistant controls, mean values of 34 lines are used.

minerals content in the different types of material (Wheat–rye introgressions, *Leymus* spp. introgressions, wheat, and triticale) verified the high content of Zn in the *Leymus* spp. introgression lines and the high Cd content in the wheat lines (**Table 5**). A relatively high Fe content was found in two of the parental wheat lines used in the present study; Sonett (57.0 mg/kg) and Prins (60.6 mg/kg). Furthermore, the triticale parents, Drira (51.8 mg/kg) and Beagle (63.3 mg/kg), were observed to contain a high level of Zn (**Table 5**). Overall, the minimum 22.7 mg/kg and maximum 64.2 mg/kg for Fe concentrations were observed in the wheat–rye introgression lines with 1R, 2R, 3R, 5R, and 6R rye chromosomes (**Table 5**). The minimum and maximum Zn concentrations produced by these wheat–rye introgression lines were 32.9 mg/kg and 89.3 mg/kg, respectively. The overall grain Cd concentration ranged from 0.02 to 0.13 mg/kg, in which the lines with low Cd concentration were observed to be 0.015 to 0.017 mg/kg in the wheat–rye introgression 1R (1D), and the lines with *L. mollis* and *L. racemosus* chromosomes. Interestingly, nine of the lines with a high combination of Fe (ranged from 47.4 to 64.2 mg/kg) and Zn (ranged from 53.7 to 83.4 mg/kg) concentration and low Cadmium concentration (ranged from 0.02 to 0.07 mg/kg) were detected in the wheat-rye 1R (1D), 2R (2D), 2R (2B), 3R (3B), and *L. mollis* and *L. racemosus* intogression lines (**Table 5**).

Baking Quality

The evaluated alien introgression lines showed a high level of variability both in grain protein concentration and gluten strength (**Figure 4**). A total of 40% of the lines showed a higher grain protein concentration than the standard cultivar, Dragon, while 8% of the lines showed higher gluten strength than the standard. The 10% of the evaluated lines with the highest grain protein concentration (TOTE), were all found to have either addition of chromosome 1R, 2R, 4R, and 6R or a 1R/1D translocation (**Table 6**). Several of the high grain protein concentration lines also had additions of 1R and 6R. The lines with high gluten strength (% UPP) were found either to have introgressions from *Leymus* or additions of either 1R + 2R or 1R + 4R (**Table 6**).

TABLE 4 | Accessions of Swedish winter wheat with rye substitutions and translocations showing resistance for Hessian fly at separate screenings.

Acc.No.	Subs./transl.	1st screen		2nd screen		Mean 4 screens	
		Tot pl	% inf	Tot pl	% inf	Tot pl	% inf
Kr 08-59	1R.1D	5	0	10	0	9.25	0
Kr 08-60	1R.1D + 6R.6D	5	0	10	0	8.75	0
Kr 08-76	T1RS.1BL + T2BS.2RL	2	0	10	0	9.5	0
Kr 08-79	2R.2B	5	0	10	0	9.25	0
Kr 08-89	T1RL.1BS + T2BS.2RL	5	0	10	0	9.5	0
Kr 08-90	T1RL.1BS + T1BS.2RL	2	0	10	0	9.5	0
Kr 08-91	T2RL.2BS	5	0	10	0	9	0
Kr 08-94	T2RL.2BS	5	0	10	0	10	0
Kr 08-95	T2RL.2BS	5	0	10	0	9	0
Res Cont	6RL	5	0	10	0	10	0
Sus Cont	-	5	100	10	100	10	100

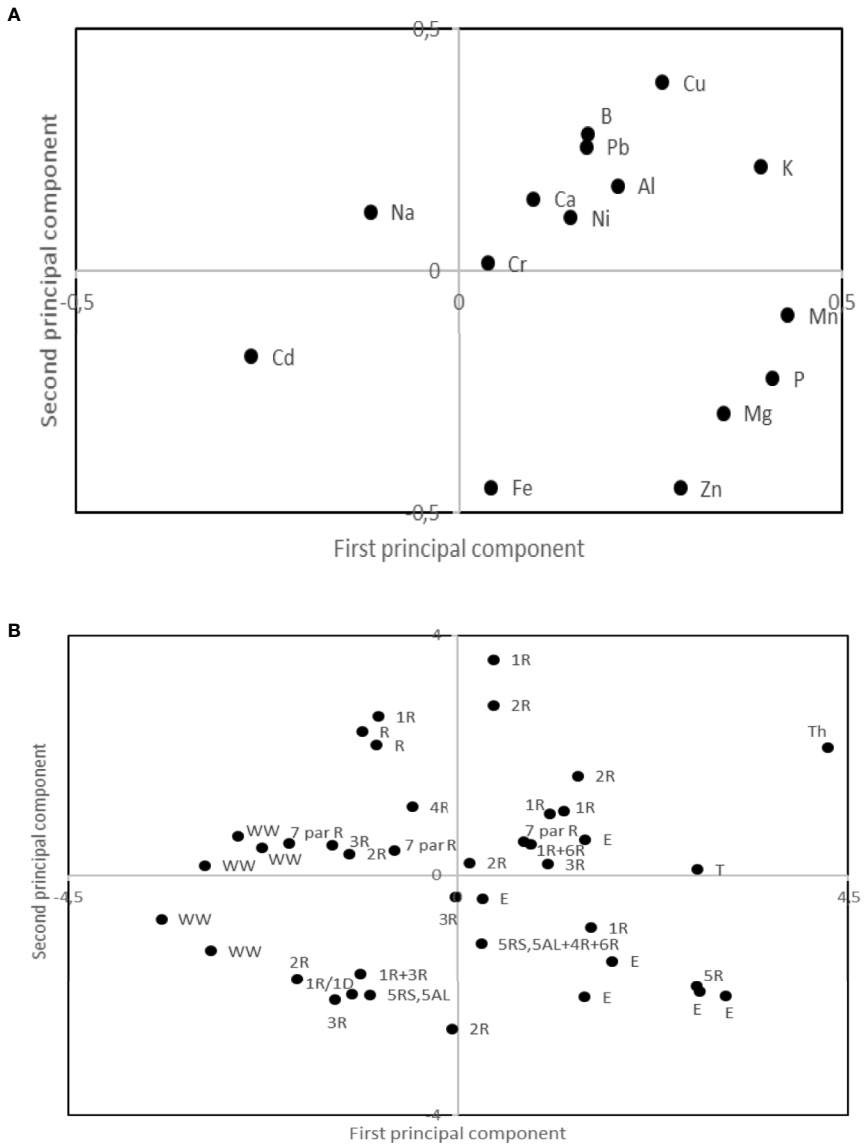

FIGURE 3 | Loading **(A)** and score **(B)** plot from principal component analyses of mineral composition in winter wheat (WW), Triticale (T), Rye (R), and alien substitution and translocation lines with rye introgressions (given as R and a what type of), Thinopyrum (Th) and Leymus (E) introgressions. The first and second principal component explained 23.0 and 17.0% of the variation, respectively.

DISCUSSION

New sources of genetic diversity are essential to improve yield, root growth, stand establishment, adaptation to climate change, nitrogen use efficiency, water use efficiency, resistance to abiotic and biotic stresses, biomass, photosynthetic potential, nutritional and end-use quality. In this paper, results from studies over a range of years are compiled to highlight the importance of wheat-alien introgression lines as a potential source of several important traits for wheat improvement. Our studies proved that these wheat-alien introgression lines carry various genetic variation e.g. resistance to diseases (rusts, powdery mildew, *S. triticae*, and Fusarium head blight), pests (Hessian fly and aphids), agronomic performance, weed competition, yield potential, microelements (Fe, Zn, Cd, etc.), fertility, alpha amylase activity, and positive end-use quality.

The evaluated 2R (2B) and 2R (2D) substitution lines showed resistance to all stem rust races at both the seedling and adult plant stages. Additionally, three of the 3R (3D) (SLU214, SLU219, and SLU222) substitution lines and seven of the wheat-*T. junceiforme* were found as potential sources of stem rust resistance genes. From the screening of a collection of wheat-alien introgression lines, the line SLU238 [2R (2D) wheat-rye disomic substitution] possessed resistance to many

TABLE 5 | Mean values of zinc, iron and cadmium concentrations in wheat, triticale, *Leymus* spp., wheat–rye introgression and wheat–*Leymus* spp. introgression lines.

Plant lines	Fe (mg/kg)		Zn (mg/kg)		Cd (mg/kg)	
	Mean	Range	Mean	Range	Mean	Range
Wheat (n = 5)	45.0	31.0–60.6	39.5	34.5–48.7	0.09	0.07–0.12
Rye (n = 2)	39.7	38.1–41.2	35.2	33.8–36.6	0.00	0.00–0.00
Triticale (n = 5)	37.9	29.5–45.0	48.9	38.7–63.3	0.09	0.07–0.15
Leymus **spp.** (n = 3)	49.4	41.4–59.1	75.8	62.4–83.4	0.02	0.00–0.02
Wheat–rye introgression (n = 22)	38.9	22.7–64.2	54.8	32.9–89.3	0.05	0.00–0.10
Wheat–*Leymus* **spp.** introgression (n = 3)	47.5	43.0–51.9	63.6	53.1–69.1	0.04	0.02–0.06

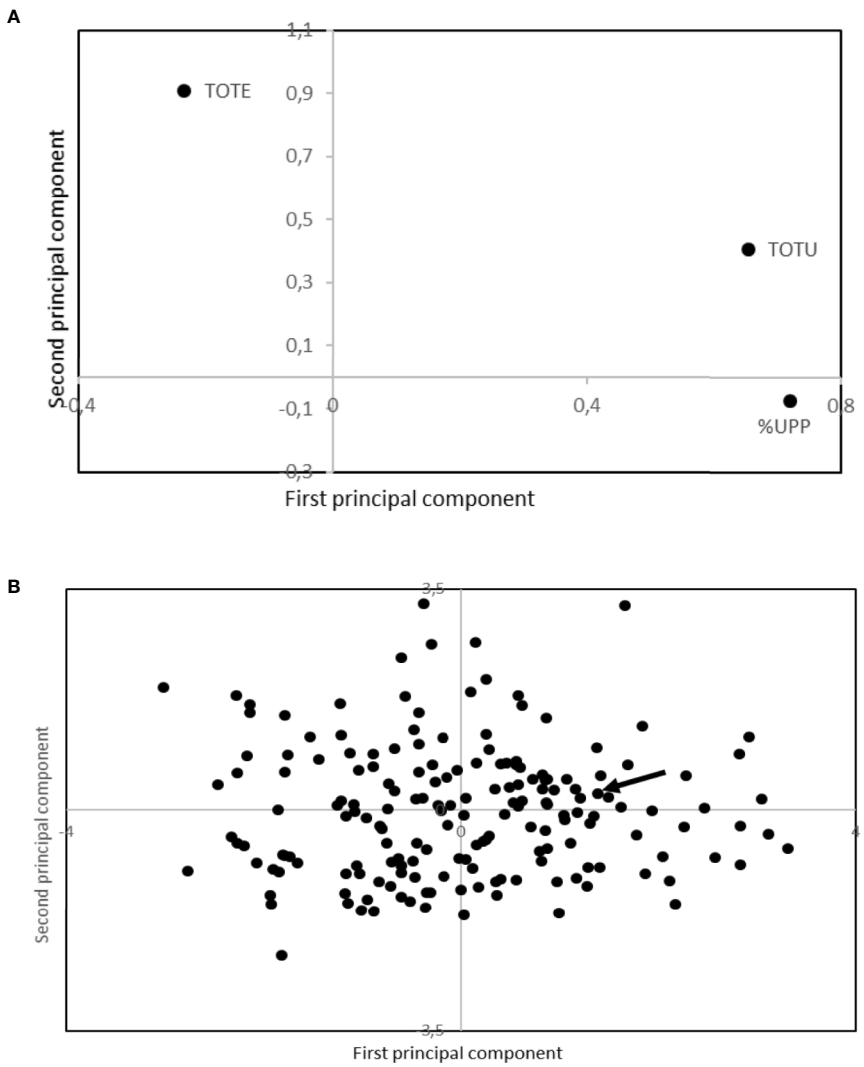

FIGURE 4 | Loading **(A)** and score **(B)** plot from principal component analyses of storage protein composition from SE-HPLC. The arrow is indicating the Swedish spring wheat line, used as a standard within the analyses. TOTE, total amount of SDS-extractable proteins; TOTU, total amount of SDS-unextractable proteins; and % UPP, percentage of unextractable polymeric protein in total polymeric protein. The first and second principal component explained 58.8 and 35.6% of the variation, respectively.

TABLE 6 | Accessions of Swedish winter wheat with substitutions and translocations (rye = R, Leymus) showing high (in descending order) total amount of SDS-extractable protein (TOTE—correlating to grain protein concentration) and percentage of unextractable polymeric protein in total polymeric protein (%UPP—correlating with gluten strength).

TOTE			%UPP		
Acc.No.	**Subs./transl.**	**Rel. values**	**Acc.No.**	**Subs./transl.**	**Values**
Kr 08-10	1R, 4R, 6R, 7R	1.74	Kr 08-109	*Leymus*	86.4
Kr 08-54	1R/1D	1.70	Kr 08-107	*Leymus*	85.7
Kr 08-16	1R, 4R, 6R, 7R	1.64	Kr 08-111	*Leymus*	85.1
Kr 08-57	1R/1D	1.63	Kr 08-104	*Leymus*	84.3
Kr 08-15	1R, 4R, 6R, 7R	1.60	Kr 08-100	2RL/2BS	81.9
Kr 08-08	1R, 4R, 6R, 7R	1.60	Kr 08-28	1R + 6R	81.6
Kr 08-55	1R/1D	1.59	Kr 08-79	1R + 2R	81.0
Kr 08-09	1R, 4R, 6R, 7R	1.59	Kr 08-80	1R + 2R	80.3
Kr 08-53	1R/1D	1.54	Kr 08-110	*Leymus*	79.3
Kr 08-30	IR + 6R	1.53	Kr 08-04	1R + 4R	79.2
Kr 08-143	5R/5A	1.53	Kr 08-108	*Leymus*	78.6
Kr 08-63	1R + 6R	1.52	Kr 08-77	1R + 2R	78.4
Kr 08-75	1RS + 2RL	1.50	Kr 08-106	*Leymus*	76.6
Kr 08-76	1RS + 2RL	1.50	Kr 08-01	1R + 4R	75.3
Kr 08-82	1R + 2R	1.47	Kr 08-95	1R + 2R	75.0
Kr 08-60	IR + 6R	1.47	Dragon		74.9
Kr 08-84	1R + 2R	1.45			
Kr 08-156	1BS/1RL	1.45			
Kr 08-52	1R/1D	1.45			
Kr08-20	IR+6R	1.44			
Dragon		1.22			

For TOTE the 20 accessions with highest value and their corresponding alien segments are shown while for %UPP, those higher than the standard (Dragon = Swedish winter wheat).

races of *Pgt*, including the widely virulent race TTKSK (Rahmatov et al., 2016a). In previous studies, Rahmatov et al. (2016b), reported that by the crossing of SLU238 and CS *ph1b* mutant, a new wheat-rye Robertsonian translocation 2DS·2RL was developed as the source of the gene *Sr59*. To this date, no stem rust resistance genes have been reported from the 2R chromosome, but chromosome 2R from different rye sources has been described as a source of resistance to various diseases and insects and also various agronomic traits. Previously, the resistance genes to leaf rust *Lr25* and *Lr45*, powdery mildew *Pm7* and Hessian fly resistance gene *H21* have been reported from the 2R chromosome (Friebe et al., 1996; Friebe et al., 1999). Furthermore, the resistance genes *Sr27*, *Sr31*/*Yr9*/*Lr26*, *Sr50*, *Sr1RS*^{Amigo} and, *SrSatu* have been described, originating from the rye chromosomes 1R and 3R, and these have been found to be effective against many of all the three rusts races (Marais and Marais, 1994; Mago et al., 2002; Singh et al., 2011; Olivera et al., 2013). Out of these resistance genes, *Sr31* has been deployed widely and provided durable resistance against stem rust races for over 30 years in agriculture (Singh et al., 2008).

Agronomic performances of some of the alien-wheat introgression lines were similar to wheat for grain yield, straw length, lodging, grain volume weight, 1000-kernel weight, fertility, grain a-amylase activity, and end-use quality (Hysing et al., 2007; Andersson et al., 2016) while some of the lines showed large variation in agronomic performance. Field studies indicated a correlation between the presence of rye (1R, 2R, 3R, 5R, 1R + 6R) and *L. racemosus* chromosomes, with low level of powdery mildew, *S. triticae* and Fusarium head blight infections (Andersson et al., 2016). Previous studies have reported an *Fhb3* resistance gene to Fusarium head blight derived from *L. racemosus* (Qi et al., 2008), which might also be present in our wheat—*L. racemosus* introgression lines. Therefore, future evaluation of these lines to other powdery mildew and *Z. triticae* isolates at seedling and adult plant stages are needed. Hysing et al. (2007), reported that red coleoptile color was correlated to the presence of the 2BS.2RL translocation allowing this character to be used as a morphological marker. Furthermore, lines with the 2BS.2RL translocation were demonstrated a high level of resistance against leaf rust and powdery mildew at the seedling stage (Merker and Forsström, 2000; Hysing et al., 2007) and adult plant resistance to TTKSK (Ug99; Rahmatov et al., 2015), thus indicating presence of uncharacterized resistance gene/s. Valuable rye chromosomes harboring beneficial genes from 4R, 5R, and 6R have also been identified (Rahmatov et al., 2017). These lines containing 4R, 5R, and 6R chromosomes are pointed out here as useful due to the fact that they are possessing novel stripe rust resistance genes. Further investigations are needed to understand the underlying genetic basis of this resistance. In various studies, stripe rust and powdery mildew resistance genes have reported on the 4R, 5R, and 6R chromosomes (An et al., 2015; Schneider et al., 2016; Xi et al., 2019), in which *Yr83* was mapped on the 6RL (Li et al., 2020). Besides this, chromosomes 4R and 6R have been demonstrated to contribute increased protein content and also to be associated with good pollinator traits (Nguyen et al., 2015; Schneider et al., 2016). Thus, there is a need to further exploit these wheat-alien introgression lines with various chromosome constitutions for wheat improvement.

High levels of resistance were identified in lines with the 1R, 3RS, 1R + 6R, 5R, and *L. mollis* chromosome introgressions against RWA. Resistances to RWA obtained from the wheat-alien introgression lines particularly lines with the 3R, 5R and *L. mollis* chromosomes have not previously been reported (Andersson et al., 2015). Previously, *Dn7*, *Gb2*, and *Gb6* resistance genes to cereal aphids have been reported on chromosome arm 1R (Friebe et al., 1996; Friebe et al., 1999; Anderson et al., 2003). Also, 1RSam.1AL and MA1S.1RLe(1B), 1Re(1D) wheat–rye translocation, and substitution lines were shown with a high level of resistance against HF and RWA, and these lines are now used in the international wheat breeding programs (Crespo-Herrera et al., 2019). The wheat-alien introgression lines with the presence of 1R, 1RS, 2R, 3R, 3RS, 4R, 5R, 6RL, and *L. racemosus* and *L. mollis* chromosomes provides resistance to the Syrian HF biotype. Previous studies have verified alien germplasm to contribute HF resistance in wheat through the *H21* and *H25* resistance genes from rye, located on 2R and 6R, respectively (Friebe et al., 1999). Hysing et al. (2007), reported that lines with the T2BS.2RL were susceptible to the HF biotypes thus this indicates different rye sources used for developing Swedish wheat-alien introgression lines. Host resistance to these insects is the most effective way of control, and various resistance genes have been derived from

alien species. The resistances to RWA and HF reported here originating from alien material have not previously been described and can, therefore, be useful to widen the pool of resistance genes in wheat breeding for resistance to RWA and HF.

The wheat-rye introgression lines displayed a good source of allelopathic potential, while lines with *L. mollis* chromosome showed a low level of allelopathic potential and the bread wheat genotypes showed no allelopathic activity. These wheat-alien lines can be used as a source of allelopathic potential and weed competitiveness in breeding programs to improve weed suppression ability for wheat improvement. Bertholdsson et al. (2012a, b), showed that the highest allelopathic potential was found in lines with 1R and 2R chromosomes. Moreover, some lines with multiple rye chromosomes (1R + 6R and 1R + 4R + 6R + 7R) were also showed high allelopathic activity (Bertholdsson et al., 2012). Previous studies have identified lines with 1R substitution showing early vigour, which can be positive for the root exudation of allelochemicals (Ehdaie et al., 2003). Breeding efforts for the allelopathic potential is considered as a complex trait (Bertholdsson, 2007), although successful examples are present on rice (Kong et al., 2006) and spring wheat (Bertholdsson, 2010). Quantitative trait loci (QTLs) linked to allelopathic traits have found on wheat chromosomes (Wu et al., 2003), thus, this indicates that allelopathic traits inherited quantitatively. The lines with high allelopathic potential identified in this study may be worthwhile for the breeding of allelopathic wheat, particularly for the purpose of organic wheat.

Various zinc, iron, and cadmium concentrations were identified in these lines. Wild relatives of wheat represent a reach source of micronutrient benefits because they have a huge and deep rooting system during its vegetation period that most efficiently uptake micronutrient if they are available in the soil (Borill et al., 2014). This has been proved by using natural genetic diversity for micronutrient uptake that can increase the nutrient content in wheat through genetic improvement (Velu et al., 2014). For instance, studies have indicated high levels of Fe and Zn to be encoded by a *Gpc-B1* locus, present in particular in wild emmer wheat (Uauy et al., 2006; Johansson et al., 2020). Thereby, genetic biofortification in wheat can be enhanced using these wheat-alien introgression lines as a source of natural genetic diversity.

Plant breeding is mostly targeting traits that improve yield potential, i.e. resistance to biotic and abiotic stresses, although for wheat improved baking and bread-making quality is also of outmost importance (Helguera et al., 2020). Wheat flour has, in particularly due to its unique protein properties, qualities which makes it outstanding for end-uses for daily food products such as bread, pastries, biscuits, porridge, cookies, etc. (Johansson et al., 2013). The gluten proteins, the gliadins, and the glutenins, encoded on group 1 and group 6 of the wheat chromosomes, are to a high extent responsible for the impact on the baking quality of wheat (Johansson et al., 2013). Alien introgressions into the wheat genome have often resulted in negative effects on the baking quality, e.g. the *Sec-1*, *Sec-2*, and *Sec-3* genes from rye instead of corresponding wheat genes at the group 1 chromosome of wheat (Kim et al., 2005). However, introgressions of rye from other parts of the genome than from the group 1 chromosomes might have less tremendous effects on the baking quality. Thus, previous results have indicated that 2BS.2RL wheat–rye translocations only had minor effects on baking quality (Hysing et al., 2007). These authors indicated that there were not any significant differences between the translocation and non-translocation groups like for grain a-amylase activity, grain starch, protein content, and other agronomic performances. Bread-making quality is known to be determined to a large extent by the gluten proteins, their amount and distribution (Johansson et al., 2013). Thus, the grain proteins concentration, the specific protein composition, the amount of specific proteins, and the amount and size distribution of polymeric protein are all factors of relevance for the bread-making quality (Finney and Barmore, 1948; Johansson et al., 2002; Johansson et al., 2003; Johansson et al., 2005; Johansson et al., 2008; Johansson et al., 2013). The evaluated alien introgression lines showed a high level of variability in both grain protein concentration and gluten strength. Thus, the alien material evaluated here, seems to have also interesting properties when it comes to specific quality breeding. Introgressions of *Leymus* seem to be able to contribute both high nutrition and high gluten strength to the material.

Alien Breeding Through Novel Tools

Introgression of desired genes from wild relatives into the bread wheat has become widely recognized as diversifying genetic diversity. However, wheat-alien chromosome additions often contribute negatively to the agricultural value of the line, therefore, desired gene/s has to be transferred into the wheat genome. Such transfers are normally blocked by the presence of a *Ph1* (*Pairing homoeologous*) allele, which strictly controls homologous chromosome pairing across the hexaploid genome to prevent hybridization between wheat and an alien species (Riley and Chapman, 1958). Anyhow, alien chromosome segments carrying gene/s of interest have been widely transferred into the wheat genome using the CS *ph1b* homoeologous recombination, radiation, and embryo culture techniques (Sears, 1977; Sears, 1993; Chen et al., 1994; Merker and Lantai, 1997). These approaches in a combination of molecular and cytogenetic manipulations were used to facilitate the introgression of *Sr26* and *Lr19* from *Thinopyrum ponticum*, *Sr39* from *Aegilops speltoides*, *Sr59* from *S. cereale*, etc. with small alien chromosome segments (Sharma and Knott, 1966; Merker and Lantai, 1997; Niu et al., 2011; Rahmatov et al., 2016a; Rahmatov et al., 2016b). More recently, reference genomes have been made available for wheat, (IWGSC, 2014), rye (Bauer et al., 2017), barley (IBGSC, 2012), rice (IRGSP, 2005), and Brachypodium (IBI, 2010), greatly facilitating the forward and reverse genetics in crops. Various high-throughput genotyping platforms such as the 9K and 90K Illumina Infinium SNP arrays and the 35K and 820K Affymetrix Axiom arrays have been developed for gene and QTL mapping (Wang et al., 2014; Winfield et al., 2016; Allen et al., 2017). In addition, genotyping-by-sequencing and exome capture sequencing opens

up more opportunities for markers development and gene isolation (Poland et al., 2012; Krasileva et al., 2017). All these genotyping platforms provide tremendous tools to assess the genetic diversity and allelic variation across plant genomes. However, a low level of SNP polymorphism between hexaploid wheat and wild relatives has been reported which negatively impact the use of the mentioned platforms (Winfield et al., 2016). Therefore, Tiwari et al. (2014) suggested the use of flow cytometric chromosome sorting to develop unique SNP markers for the mapping of alien genes to overcome these challenges. Whole-genome shotgun sequencing is becoming another valuable breeding tool in terms of time and cost, which are already used in major crops such as wheat (Brenchley et al., 2012), maize (Hufford et al., 2012), rice (Huang et al., 2012), and soybean (Fang et al., 2017). However, the transfer of desired alien gene/s remains a challenge, although some advances have been made in transferring resistance genes. Jupe et al. (2013) developed an exome capture and sequencing of nucleotide-binding leucine-rich repeat (NLR) genes in potato. Such resistance gene enrichment sequencing (RenSeq) allowed a rapid cloning of the *Sr22* and *Sr45* resistance genes through mutational genomics (Steuernagel et al., 2016). Another approach, MutChromSeq, has been applied through mutational genomics, chromosome flow sorting and sequencing that has resulted in the cloning of the *Pm2* resistance gene (Sánchez-Martín et al., 2016). Interestingly, another cloning approach suggested a combination of association genetics and R gene enrichment sequencing, which rapidly identified stem rust resistance genes for cloning (Arora et al., 2019). Besides, a combination of cisgenesis and genome editing tools may accelerate the plant breeding process (Cardi, 2016). Also, the use of speed breeding may significantly accelerate the generation times and breeding cycles (Watson et al., 2018). Therefore, integration of high-throughput genotyping and precise phenotyping tools may efficiently assist in transferring the introgression of small alien chromatin segments to develop new genetic diversity for wheat improvement. For example, the development of synthetic wheat and 1RS chromosome arm has made a great contribution to sustainable wheat production. Evidently, for the development of superior crop cultivars requires new genetic variation that meets sustainable agriculture and food security needs.

CONCLUSIONS—ALIEN GENES INTO MODERN WHEAT—FUTURE PERSPECTIVES

Every day, the human population is growing, and with that the demand of food from sustainable and healthy crop production. To adequately meet the global food demand required by 2050, there is a need to increase wheat yield annually. These can be achieved through the two unique opportunities; plant breeding and improved agronomic practices. Importantly, to meet projected food demand, the breeding programs need to broaden the existing genetic base, in particular by the use of alien species with the potential to improve yield, resistance to biotic and abiotic stresses and quality. Several of our studies have identified new sources of resistance to fungal diseases and insects in the wheat-alien introgression derivatives from *S. cereale*, *L. mollis*, *L. racemosus* and *T. junceiforme*. Also, these lines exhibiting good agronomic performances, high allelopathic potential, and superior end-use quality traits. Our results suggest that some of the lines could be used as a source of high Iron and Zinc and low Cadmium concentrations. These findings show that the wheat-alien introgressions are a potentially useful genetic resource for wheat improvement. The introgression of large alien chromosomes usually challenges researchers and breeders by causing linkage drag that can negatively effect on yield and quality properties. Fortunately, with the presence of high-throughput genotyping and phenotyping tools, opportunities increase to transfer desired gene/s with a small alien chromosome segment. Consequently, research is currently underway to transfer stem and stripe rust resistance genes into the elite wheat background to be used by breeders to develop superior wheat cultivars with new resistance genes. Further, additional research is also in progress for characterization and transferring of useful traits such as micronutrients (Zn, Fe, and Cd), allelopathic potential, diseases, and insect resistance as well as stable baking quality.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

EJ, TH, SA and MR planned various parts of the study, the hypothesis, and the objectives. TH, MP-L, SA, RA carried out various parts of the field and lab work. All authors contributed to compiling various parts of the results. EJ and MR planned the writing of this paper and did the first draft. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded by Partnership Alnarp grant nos. PA710 and PA1094, the Swedish Research Council VR grant no. 2016-05806 and Formas grant no. 2017-00514.

ACKNOWLEDGMENTS

The authors would like to thank our international research collaborators Matthew Rouse at the United States Department of Agriculture, ARS-CDL, Brian Steffenson at the University of Minnesota, Mogens Hovmøller at the GRRC, Aarhus University, Kumarse Nazari and Mustafa El-Bouhssini at the ICARDA. The authors also appreciate Kenyan Agricultural and Livestock Research Organization Food Crops Research Center and CIMMYT.

REFERENCES

Allen, A. M., Winfield, M. O., Burridge, A. J., Downie, R. C., Benbow, H. R., Barker, G. L. A., et al. (2017). Characterization of a wheat breeders' array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnol. J.* 15, 390–401. doi: 10.1111/pbi.12635

An, D., Zheng, Q., Luo, Q., Ma, P., Zhang, H., Li, L., et al. (2015). Molecular cytogenetic identification of a new wheat-rye 6R chromosome disomic addition line with powdery mildew resistance. *PloS One* 10 (8), e0134534. doi: 10.1371/journal.pone.0134534

Anderson, G. R., Papa, D., Peng, J., Tahir, M., and Lapitan, N. L. V. (2003). Genetic mapping of Dn7, a rye gene conferring resistance to the Russian wheat aphid in wheat. *Theor. Appl. Genet.* 107, 1297–1303.

Andersson, S. C., Johansson, E., Baum, M., Rihawi, F., and El Bouhssini, M. (2015). New resistance sources to Russian wheat aphid (*Diuraphis noxia*) in Swedish wheat substitution and translocation lines with rye (*Secale cereale*) and *Leymus mollis*. *Czech J. Genet. Plant Breed.* 51, 162–165. doi: 10.17221/72/2015-CJGPB

Andersson, S. C., Johansson, E., Henriksson, T., and Rahmatov, M. (2016). *Nya resistensgener för framtidens vete. LTV-fakultetens faktablad*. Available from: https://pub.epsilon.slu.se/13096/1/andersson_sc_etal_160225.pdf (Accessed March 09, 2020).

Arora, S., Steuernagel, B., Gaurav, K., and Chandramohan, S. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* 37, 139–143. doi: 10.1038/s41587-018-0007-9

Bauer, E., Schmutzer, T., Barilar, I., Mascher, M., Gundlach, H., Martis, M. M., et al. (2017). Towards a whole-genome sequence for rye (*Secale cereale* L.). *Plant J.* 89, 853–869. doi: 10.1111/tpj.13436

Bertholdsson, N.-O., Andersson, S. C., and Merker, A. (2012). Allelopathic potential of *Triticum* spp., *Secale* spp. and *Triticosecale* spp. and use of chromosome substitutions and translocations to improve weed suppression ability in winter wheat. *Plant Breed.* 131, 75–80. doi: 10.1111/j.1439-0523.2011.01895.x

Bertholdsson, N.-O. (2007). Varietal variation in allelopathic activity in wheat and barley and possibilities for use in plant breeding. *Allelopathy J.* 19, 193–201.

Bertholdsson, N.-O. (2010). Breeding spring wheat for improved allelopathic potential. *Weed Res.* 50, 49–57. doi: 10.1111/j.1365-3180.2009.00754.x

Borrill, P., Connorton, J. M., Balk, J., Miller, A. J., Sanders, D., and Uauy, C. (2014). Biofortification of wheat grain with iron and zinc: integrating novel genomic resources and knowledge from model crops. *Front. Plant Sci.* 5, 53. doi: 10.3389/fpls.2014.00053

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L. A., D'Amore, R., Allen, A. M., et al. (2012). Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491, 705–710. doi: 10.1038/nature11650

Cardi, T. (2016). Cisgenesis and genome editing. Combining concepts and efforts for a smarter use of genetic resources in crop breeding. *Plant Breed.* 135, 139–147. doi: 10.1111/pbr.12345

Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci.* 110, 8057–8062. doi: 10.1073/pnas.1217133110

Chen, P. D., Tsujimoto, H., and Gill, B. S. (1994). Transfer of *Ph*^{I} genes promoting homoeologous pairing from *Triticum speltoides* to common wheat. *Theor. Appl. Genet.* 88, 97–101. doi: 10.1007/BF00222400

Crespo-Herrera, L. A., Singh, R. P., Sabraoui, A., and El-Bouhssini, M. (2019). Resistance to insect pests in wheat-rye and *Aegilops speltoides* Tausch translocation and substitution lines. *Euphytica* 215, 123 . doi: 10.1007/s10681-019-2449-7

Dhaliwal, A. S., Mares, D. J., Marshall, D. R., and Skeritt, J. H. (1988). Protein composition and pentosan content in relation to dough stickiness of 1B/1R translocation wheats. *Cereal Chem.* 62, 143–149.

Dubcovsky, J., and Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316, 1862–1866. doi: 10.1126/science.1143986

Ehdaie, B., Whitkus, R. W., and Waines, J. G. (2003). Root biomass, water-use efficiency, and performance of wheat-rye translocations of chromosomes 1 and 2 in spring bread wheat "Pavon". *Crop Sci.* 43, 710–717. doi: 10.2135/cropsci2003.0710

El Bouhssini, M., Street, K., Amri, A., Mackay, M., Ogbonnaya, F. C., Omran, A., et al. (2011). Sources of resistance in berad wheat to Russian wheat aphid (*Diuraphis noxia*) in Syria identified using the Focused Identification of Germplasm Strategy (FIGS). *Plant Breed.* 130, 96–97. doi: 10.1111/j.1439-0523.2010.01814.x

El Bouhssini, M., Ogbonnaya, F. C., Chen, M., Lhaloui, S., Rihawi, F., and Dabbous, A. (2013). Sources of resistance in primary synthetic hexaploid wheat (*Triticum aestivum* L.) to insect pests: Hessian fly, Russian wheat aphid and Sunn pest in the fertile crescent. *Genet. Resour. Crop Evol.* 60, 621–627. doi: 10.1007/s10722-012-9861-3

Ellneskog-Staam, P., and Merker, A. (2001). Genome composition, stability and fertility of hexaploid alloploids between *Triticum turgidum* var. *carthlicum* and *Leymus racemosus*. *Hereditas* 13479– , 84. doi: 10.1111/j.1601-5223.2001.00079.x

Ellneskog-Staam, P., and Merker, A. (2002). Screening for Resistance to Powdery Mildew and Brown Rust in Wheat-*Leymus racemosus* and Wheat-*Thinopyrum junceiforme* Alloploids. *Acta Agric. Scand. Section B Soil Plant Sci.* 52, 158–161. doi: 10.1080/09064710300525

Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., et al. (2017). Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* 18, 161. doi: 10.1186/s13059-017-1289-9

FAO (2016). *Save and Grow in practice: maize, rice, wheat. A guide to sustainable cereal production*. Available from: http://www.fao.org/3/a-i4009e.pdf (Accessed March 09, 2020).

Finney, K. F., and Barmore, M. A. (1948). Loaf volume and protein content of hard winter and spring wheats. *Cereal. Chem.* 25, 291–312.

Forsström, P.-O., and Merker, A. (2001). Sources of wheat powdery mildew resistance from wheat-rye and wheat-*leymus* hybrids. *Hereditas* 134, 115–119. doi: 10.1111/j.1601-5223.2001.00115.x

Forsström, P., Merker, A., and Schwarzacher, T. (2002). Characterisation of mildew resistant wheat-rye substitution lines and identification of an inverted chromosome by fluorescent *in situ* hybridisation. *Heredity* 88, 349–355. doi: 10.1038/sj.hdy.6800051

Friebe, B., Jiang, J., Raupp, W. J., McIntosh, R. A., and Gill, B. S. (1996). Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91, 59–87. doi: 10.1007/BF00035277

Friebe, B., Kynast, R. G., Hatchett, J. H., Sears, R. G., Wilson, D. L., and Gill, B. S. (1999). Transfer of rye-derived Hessian fly resistance genes *H21* and *H25* from bread wheat into durum wheat. *Crop Sci.* 3, 1692–1696. doi: 10.2135/cropsci1999.3961692x

Hao, M., Liu, M., Luo, J., Fan, C., Yi, Y., Zhang, L., et al. (2018). Introgression of powdery mildew resistance gene *Pm56* on rye chromosome arm 6RS into wheat. *Front. Plant Sci.* 9, 1040. doi: 10.3389/fpls.2018.01040

Helguera, M., Abugalieva, A., Battenfield, S., Békés, F., Branlard, G., Cuniberti, M., et al. (2020). "Grain quality in breeding," in *Wheat quality for improving processing and human health*. Eds. T. M. Igrejas G. Ikeda and C. Guzmán (Switzerland: Springer), pp 273–pp 308.

Huang, X., and Röder, M. S. (2011). High-density genetic and physical bin mapping of wheat chromosome 1D reveals that the powdery mildew resistance gene *Pm24* is located in a highly recombinogenic region. *Genetica* 139, 1179–1187. doi: 10.1007/s10709-011-9620-y

Huang, X., Kurata, N., Wang, Z.-X., Wang, A., Zhao, Q., Zhao, Y., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501. doi: 10.1038/nature11532

Hufford, M. B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J.-M., Cartwright, R. A., et al. (2012). Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44, 808–811. doi: 10.1038/ng.2309

Husenov, B., Otambekova, M., Muminjanov, H., Morgounov, A., Asaad, S., Garkava-Gustavsson, L., et al. (2020). Constraints and perspectives for sustainable wheat production in Tajikistan. *Front. Sustain. Food Syst.* 4, 27. doi: 10.3389/fsufs.2020.00027

Hussain, A., Larsson, H., Kuktaite, R., and Johansson, E. (2010). Mineral composition of organically grown wheat genotypes: contribution to daily minerals intake. *Int. J. Environ. Res. Public Health* 7, 3442–3456. doi: 10.3390/ijerph7093442

Hussain, A., Larsson, H., Kuktaite, R., Prieto-Linde, M. L., and Johansson, E. (2012). Towards the understanding of bread-making quality in organically

grown wheat: Dough mixing behaviour, protein polymerisation and structural properties. *J. Cereal Sci.* 56, 659–666. doi: 10.1016/j.jcs.2012.07.016

Hussain, A., Larsson, H., Kuktaite, R., Prieto-Lindem, M. L., and Johansson, E. (2013). Amount and size distribution of monomeric and polymeric proteins in the grain of organically produced wheat. *Cereal Chem.* 90, 80–86. doi: 10.1094/CCHEM-05-12-0064-R

Hysing, S. C., Hsam, S. L. K., Singh, R. P., Huerta-Espino, J., Boyd, L. A., Koebner, R. M. D., et al. (2007). Agronomic performance and multiple disease resistance in T2BS.2RL wheat-rye translocation lines. *Crop Sci.* 47, 254–260. doi: 10.2135/cropsci2006.04.0269

IBGSC (The International Barley Genome Sequencing Consortium) (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543

IBI (The International Brachypodium Initiative) (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463, 763–768 . doi: 10.1038/nature08747

IRGSP (International Rice Genome Sequencing Project) (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800. doi: 10.1038/nature03895

IWGSC (The International Wheat Genome Sequencing Consortium) (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345, 1251788. doi: 10.1126/science.1251788

Johansson, E., Prieto-Linde, M. L., and Jönsson, J. (2001). Effects of wheat cultivar and nitrogen application on storage protein composition and bread-making quality. *Cereal Chem.* 78, 19–25. doi: 10.1094/CCHEM.2001.78.1.19

Johansson, E., Nilsson, H., Mazhar, H., Skerritt, J., MacRitchie, F., and Svensson, G. (2002). Seasonal effects on storage proteins and gluten strength in four Swedish wheat cultivars. *J. Sci. Food Agric.* 82, 1305–1311. doi: 10.1002/jsfa.1185

Johansson, E., Prieto-Linde, M., Svensson, G., and Jönsson, J. (2003). Influences of cultivar, cultivation year and fertilizer rate on amount of protein groups and amount and size distribution of mono- and polymeric proteins in wheat. *J. Agr. Sci.* 140, 275–284. doi: 10.1017/S0021859603003162

Johansson, E., Kuktaite, R., Andersson, A., and Prieto-Linde, M. L. (2005). Protein polymer built-up during wheat development: influences of temperature and nitrogen timing. *J. Sci. Food Agric.* 85, 473–479. doi: 10.1002/jsfa.2006

Johansson, E., Prieto-Linde, M., and Gissén, C. (2008). Influences of weather, cultivar and fertiliser rate on grain protein polymer accumulation in field-grown winter wheat, and relations to grain water content and falling number. *J. Sci. Food Agric.* 88:, 2011–2018. doi: 10.1002/jsfa.3312

Johansson, E., Malik, A. H., Hussain, A., Rasheed, F., Newson, W. R., Plivelic, T., et al. (2013). Wheat gluten polymer structures: the impact of genotype, environment, and processing on their functionality in various applications. *Cereal Chem.* 90, 367–376. doi: 10.1094/CCHEM-08-12-0105-FI

Johansson, E., Hussain, A., Kuktaite, R., Andersson, S. C., and Olsson, M. E. (2014). Contribution of organically grown crops to human health. *Int. J. Environ. Res. Public Health* 11, 3870–3893. doi: 10.3390/ijerph110403870

Johansson, E., Branlard, G., Cuniberti, M., Flagella, Z., Hüsken, A., Nurit, E., et al. (2020). "Genotypic and environmental effect on wheat technological and nutritional quality," in *Wheat quality for improving processing and human health*. Eds. T. M. Igrejas G. Ikeda and C. Guzmán (Switzerland: Springer), pp 171–204.

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G. J., et al. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *Plant J.* 76, 530–544. doi: 10.1111/tpj.12307

Kihara, H. (1983). "Origin and history of 'Daruma', a parental variety of Norin 10," in *Proceedings of the Sixth International Wheat Genetics Symposium*. Ed. S. Sakamoto (Kyoto, Japan, Beijing: Plant Germplasm Institute, University of Kyoto, China Agricultural Scientech Press).

Kim, W., Johnson, J. W., Baenziger, P. S., Lukaszewski, A. J., and Gaines, C. S. (2005). Quality effect of wheat-rye (1R) translocation in "Pavon 76". *Plant Breed.* 124, 334–337. doi: 10.1111/j.1439

Kong, C. H., Li, H. B., Hu, F., Xu, X. H., and Wang, P. (2006). Allelochemicals released by rice roots and residues in soil. *Plant Soil* 288, 47–56. doi: 10.1007/s11104-006-9033-3

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., et al. (2017). Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci.* 114, 913–921. doi: 10.1073/pnas.1619268114

Li, J., Dundas, I., Dong, C., Li, G., Trethowan, R., Yang, Z., et al. (2020). Identification and characterization of a new stripe rust resistance gene *Yr83* on rye chromosome 6R in wheat. *Theor. Appl. Genet.* 133, 1095–1107. doi: 10.1007/s00122-020-03534-y

Mago, R., Zhang, P., Vautrin, S., Šimková, H., Bansal, U., Luo, M.-C., et al. (2015). The wheat *Sr50* gene reveals rich diversity at a cereal disease resistance locus. *Nat. Plants.* 1, 15186. doi: 10.1038/NPLANTS.2015.186

Mago, R., Spielmeyer, W., Lawrence, G., Lagudah, E., Ellis, J., and Pryor, A. (2002). Identification and mapping of molecular markers linked to rust resistance genes located on chromosome 1RS of rye using wheat-rye translocation lines. *Theor. Appl. Genet.* 104, 1317–1324.

Malik, A. H., Prieto-Linde, M. L., Kuktaite, R., Andersson, A., and Johansson, E. (2011). Individual and interactive effects of cultivar maturation time, nitrogen regime and temperature level on accumulation of wheat grain proteins. *J. Sci. Food Agric.* 91, 2192–2200. doi: 10.1002/jsfa.4439

Malik, A. H., Kuktaite, R., and Johansson, E. (2013). Combined effect of genetic and environmental factors on the accumulation of proteins in the wheat grain and their relationship to bread-making quality. *J. Cereal Sci.* 57, 170–174. doi: 10.1016/j.jcs.2012.09.017

Marais, G. F., and Marais, A. S. (1994). The derivation of compensating translocations involving homoeologous group 3 chromosomes of wheat and rye. *Euphytica* 79, 75–80. doi: 10.1007/BF00023578

McFadden, E. S., and Sears, E. R. (1946). The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* 37, 107–116. doi: 10.1093/oxfordjournals.jhered.a105594

McIntosh, R. A., Wellings, C. R., and Park, R. F. (1995). *Wheat rusts: an atlas of resistance genes* (Dordrecht: Kluwer Academic Publishers).

Merker, A., and Forsström, P. (2000). Isolation of mildew resistant wheat-rye translocation lines from a double substitution line. *Euphytica* 115, 167–172. doi: 10.1023/A:1004018500970

Merker, A., and Lantai, K. (1997). Hybrids between wheats and perennial *Leymus* and *Thinopyrum* species. *Acta Agric. Scand. Section B Plant Soil Sci.* 47, 48–51. doi: 10.1080/09064719709362438

Merker, A., and Rogalska, S. (1984). The breeding behaviour of a double disomic wheat-rye substitution line. *Cereal. Res. Commun.* 12, 13–17.

Merker, A. (1979). The breeding behaviour of some rye wheat chromosome substitutions. *Hereditas* 91, 245–255. doi: 10.1111/j.1601-5223.1979.tb01667.x

Merker, A. (1984). The rye genome in wheat breeding. *Hereditas* 100, 183–191. doi: 10.1111/j.1601-5223.1984.tb00118.x

Molnár-Láng, M., Ceoloni, C., and Doležel, J. (2015). *Alien Introgression in Wheat. Cytogenetics, Molecular Biology, and Genomics* (Berlin/Heidelberg, Germany: Springer).

Moreira-Ascarrunz, S. D., Larsson, H., Prieto-Linde, M. L., and Johansson, E. (2016). Mineral nutritional yield and nutrient density of locally adapted wheat genotypes under organic production. *Foods* 5, 89. doi: 10.3390/foods5040089

Nguyen, V., Fleury, D., Timmins, A., Laga, H., Hayden, M., Mather, D., et al. (2015). Addition of rye chromosome 4R to wheat increases anther length and pollen grain number. *Theor. Appl. Genet.* 128, 953–964. doi: 10.1007/s00122-015-2482-4

Niu, Z., Klindworth, D. L., Friesen, T. L., Chao, S., Jin, Y., Cai, X., et al. (2011). Targeted introgression of a wheat stem rust resistance gene by DNA marker-assisted chromosome engineering. *Genetics* 187, 1011–1021. doi: 10.1534/genetics.110.123588

Olivera, P. D., Pretorius, Z. A., Badebo, A., and Jin, Y. (2013). Identification of resistance to races of Puccinia graminis f. sp. tritici with broad virulence in Triticale (×Triticosecale). *Plant Dis.* 97, 479–484.

Peterson, R. F., Campbell, A. B., and Hannah, A. E. (1948). A diagrammatic scale for estimating rust severity on leaves and stems of cereals. *Can. J. Res. Section C.* 26, 496–500. doi: 10.1139/cjr48c-033

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS One* 7, e32253. doi: 10.1371/journal.pone.0032253

Qi, L., Pumphrey, M., Friebe, B., Chen, P., and Gill, B. (2008). Molecular cytogenetic characterization of alien introgressions with gene *Fhb3* for

resistance to Fusarium head blight disease of wheat. *Theor. Appl. Genet.* 117, 1155–1166. doi: 10.1007/s00122-008-0853-9

Rabinovich, S. V. (1998). Importance of wheat-rye translocations for breeding modern cultivar of *Triticum aestivum* L. *Euphytica* 100, 323–340. doi: 10.1023/A:1018361819215

Rahmatov, M., Garkava-Gustavsson, L., Wanyera, R., Steffenson, B., Rouse, M., and Johansson, E. (2015). Stem rust resistance in 1BL.1RS and 2RL.2BS double wheat-rye translocation lines. *Czech J. Genet. Plant Breed.* 51, 148–154. doi: 10.17221/80/2015-CJGPB

Rahmatov, M., Rouse, M. N., Nirmala, J., Danilova, T., Friebe, B., Steffenson, B. J., et al. (2016a). A new 2DS·2RL Robertsonian translocation transfers stem rust resistance gene *Sr59* into wheat. *Theor. Appl. Genet.* 129, 1383–1392. doi: 10.1007/s00122-016-2710-6

Rahmatov, M., Rouse, M. N., Steffenson, B. J., Andersson, S. C., Wanyera, R., Pretorius, Z. A., et al. (2016b). Sources of stem rust resistance in wheat–alien introgression lines. *Plant Dis.* 100, 1101–1109. doi: 10.1094/PDIS-12-15-1448-RE

Rahmatov, M., Hovmøller, M. S., Nazari, K., Andersson, S. C., Steffenson, B. J., and Johansson, E. (2017). Seedling and adult plant stripe rust resistance in diverse wheat–alien introgression lines. *Crop Sci.* 57, 2032–2042. doi: 10.2135/cropsci2016.08.0664

Rahmatov, M. (2016). *Genetic characterisation of novel resistance alleles to stem rust and stripe rust in wheat-alien introgression lines. Doctoral thesis* (SLU Epsilon Open Archive). Available from: https://pub.epsilon.slu.se/13585/ (Accessed March 09, 2020).

Riley, R., and Chapman, V. (1958). Genetic control of the cytologically diploid behaviour of hexaploid wheat. *Nature* 182, 713–715. doi: 10.1038/182713a0

Roelfs, A. P., Singh, R. P., and Saari, E. E. (1992). *Rust diseases of wheat: Concepts and methods of disease management* (Mexico D.F., Mexico: CIMMYT).

Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., and Adamski, N. (2016). Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biol.* 17, 221. doi: 10.1186/s13059-016-1082-1

SAS (2011). *SAS Institute Inc 9.3* (NC, USA: SAS Institute Inc.).

Schlegel, R. (2020). *Current list of wheats with rye and alien introgression 1R.1B - 1R/1B - 1R-1B - 1RS.1BL - 1R(1B) - alien*. Available from: http://www.rye-gene-map.de/rye-introgression/index.html (Accessed March 09, 2020).

Schneider, A., Rakszegi, M., Molnár-Láng, M., Szakács, É., et al. (2016). Production and cytomolecular identification of new wheat-perennial rye (*Secale cereanum*) disomic addition lines with yellow rust resistance (6R) and increased arabinoxylan and protein content (1R, 4R, 6R). *Theor. Appl. Genet.* 129, 1045–1059. doi: 10.1007/s00122-016-2682-6

Sears, E. R. (1977). An induced mutant with homoeologous pairing in common wheat. *Can. J. Genet. Cytol.* 19, 585–593. doi: 10.1139/g77-063

Sears, E. R. (1993). Use of radiation to transfer alien chromosome segments to wheat. *Crop Sci.* 33, 897–901. doi: 10.2135/cropsci1993.0011183X003300050004x

Sharma, D., and Knott, D. R. (1966). The transfer of leaf rust resistance from *Agropyron to Triticum* by irradiation. *Can. J. Genet. Cytol.* 8, 137–143. doi: 10.1139/g66-018

Shewry, P. R., and Hey, S. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi: 10.1002/fes3.64

Shiferaw, B., Smale, M., Braun, H., Duveiller, E., Reynolds, M., and Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Seurc.* 5, 291–317. doi: 10.1007/s12571-013-0263-yn

Singh, R. P., Hodson, D. P., Huerta-espino, J., Jin, Y., Njau, P., Wanyera, R., et al. (2008). Will stem rust destroy the world's wheat crop? *Adv. Agron.* 98, 271–309. doi: 10.1016/S0065-2113(08)00205-8

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., et al. (2011). The emergence of Ug99 races of the stem rust fungus is a threat to the world wheat production. *Annu. Rev. Phytopathol.* 49, 465–481. doi: 10.1146/annurev-phyto-072910-095423

Singh, S., Vikram, P., Sehgal, D., Burgueño, J., Sharma, A., Singh, S. K., et al. (2018). Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Sci. Rep.* 8, 12527. doi: 10.1038/s41598-018-30667-4

Stakman, E. C., Stewart, D. M., and Loegering, W. Q. (1962). *Identification of physiologic races of Puccinia graminis var. tritici* (US Department of Agric., ARS E-617).

Steenwerth, K. L., Hodson, A. K., Bloom, A. J., Carter, M. R., Cattaneo, A., Chartres, C. J., et al. (2014). Climate-smart agriculture global research agenda: scientific basis for action. *Agric. Food Secur.* 3, 11. doi: 10.1186/2048-7010-3-11

Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., and Yu, G. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34, 652–655. doi: 10.1038/nbt.3543

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

The, T. T., Gupta, R. B., Dyck, P. L., Appels, R., Hohmann, U., and McIntosh, R. A. (1991). Characterization of stem rust resistant derivatives of wheat cultivar Amigo. *Euphytica* 58, 245–252. doi: 10.1007/BF00025256

Tiwari, V. K., Wang, S., Sehgal, S., Vrana, J., Friebe, B., Kubalakova, M., et al. (2014). SNP discovery for mapping alien introgressions in wheat. *BMC Genom.* 15, 273. doi: 10.1186/1471-2164-15-273

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC gene regulating senescence improves grain protein, zinc and irons content in wheat. *Science* 314, 1298–1300. doi: 10.1126/science.1133649

Vazquez, D., Berger, A., Prieto-Linde, M., and Johansson, E. (2019). Can nitrogen fertilization be used to modulate yield, protein content and bread-making quality in Uruguayan wheat? *J. Cereal Sci.* 85, 153–161. doi: 10.1016/j.jcs.2018.12.001

Velu, G., Ortiz-Monasterio, I., Cakmak, I., Hao, Y., and Singh, R. P. (2014). Biofortification strategies to increase grain zinc and iron concentrations in wheat. *J. Cereal Sci.* 59, 365–372. doi: 10.1016/j.jcs.2013.09.001

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Watson, A., Ghosh, S., Williams, M. J., Cuddy, S. W., Simmonds, J., Rey, M.-D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants.* 4, 23–29. doi: 10.1038/s41477-017-0083-8

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., et al. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnol. J.* 14, 1195–1206. doi: 10.1111/pbi.12485

Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemom. Intell. Lab. Syst.* 2 (1-3), 37–52. doi: 10.1016/0169-7439(87)80084-9

Wu, H., Pratley, J., Ma, W., and Haig, T. (2003). Quantitative trait loci and molecular markers associated with wheat allelopathy. *Theor. Appl. Genet.* 107, 1477–1481. doi: 10.1007/s00122-003-1394-x

Xi, W., Tang, Z., Luo, J., and Fu, S. (2019). Physical location of new stripe rust resistance gene (s) and PCR-based markers on rye (Secale cereal L.) chromosomes 5 using 5R dissection lines. *Agronomy* 9, 498.

Zeven, A. C. (1998). Landraces: a review of definitions and classifications. *Euphytica* 104, 127–139. doi: 10.1023/A:1018683119237

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 September 2020
doi: 10.3389/fpls.2020.567507

Potential Short-Term Memory Induction as a Promising Method for Increasing Drought Tolerance in Sweetpotato Crop Wild Relatives [*Ipomoea* series *Batatas* (Choisy) D. F. Austin]

Fernando Guerrero-Zurita[1], *David A. Ramírez*[2], *Javier Rinza*[3], *Johan Ninanya*[3], *Raúl Blas*[4] *and Bettina Heider*[1*]

[1] Genetics, Genomics and Crop Improvement Division, International Potato Center, Lima, Peru, [2] Latin American & Caribbean Regional Program, International Potato Center, Lima, Peru, [3] Crop and Systems Science Division, International Potato Center, Lima, Peru, [4] Crop Husbandry Department, Universidad Nacional Agraria La Molina, Lima, Peru

OPEN ACCESS

Edited by:
Pietro Gramazio,
University of Tsukuba, Japan

Reviewed by:
Robert Jarret,
United States Department of Agriculture, United States
Qingchang Liu,
China Agricultural University, China
Conxita Royo,
Institute of Agrifood Research and Technology (IRTA), Spain

***Correspondence:**
Bettina Heider
B.Heider@cgiar.org

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 29 May 2020
Accepted: 12 August 2020
Published: 03 September 2020

Citation:
Guerrero-Zurita F, Ramírez DA, Rinza J, Ninanya J, Blas R and Heider B (2020) Potential Short-Term Memory Induction as a Promising Method for Increasing Drought Tolerance in Sweetpotato Crop Wild Relatives [Ipomoea series Batatas (Choisy) D. F. Austin]. Front. Plant Sci. 11:567507. doi: 10.3389/fpls.2020.567507

Crop wild relatives of sweetpotato [*Ipomoea* series *Batatas* (Choisy) D. F. Austin] are a group of species with potential for use in crop improvement programs seeking to breed for drought tolerance. Stress memory in this group could enhance these species' physiological response to drought, though no studies have yet been conducted in this area. In this pot experiment, drought tolerance, determined using secondary traits, was tested in 59 sweetpotato crop wild relative accessions using potential short-term memory induction. For this purpose, accessions were subjected to two treatments, i) non-priming: full irrigation (up to field capacity, 0.32 w/w) from transplanting to harvest and ii) priming: full irrigation from transplanting to flowering onset (FO) followed by a priming process from FO to harvest. The priming process consisted of three water restriction periods of increasing length (8, 11, and 14 days) followed each by a recovery period of 14 days with full irrigation. Potential stress memory induction was calculated for each accession based on ecophysiological indicators such as senescence, foliar area, leaf-minus-air temperature, and leaf ^{13}C discrimination. Based on total biomass production, resilience and production capacity were calculated per accession to evaluate drought tolerance. Increase in foliar area, efficient leaf thermoregulation, improvement of leaf photosynthetic performance, and delayed senescence were identified in 23.7, 28.8, 50.8, and 81.4% of the total number of accessions, respectively. It was observed that under a severe drought scenario, a resilient response included more long-lived green leaf area while a productive response was related to optimized leaf thermoregulation and gas exchange. Our preliminary results suggest that *I. triloba* and *I. trifida* have the potential to improve sweetpotato resilience in dry environments and should be included in introgression breeding programs of this crop. Furthermore, *I. splendor-sylvae*, *I. ramosissima*, *I. tiliacea*, and wild *I. batatas* were the most productive species studied but given the

genetic barriers to interspecific hybridization between these species and sweetpotato, we suggest that further genetic and metabolic studies be conducted on them. Finally, this study proposes a promising method for improving drought tolerance based on potential stress-memory induction, which is applicable both for wild species and crops.

Keywords: senescence delay, foliar area, leaf temperature, ^{13}C discrimination, drought stress memory, Batatas complex

INTRODUCTION

In 2018, sweetpotato [*Ipomoea batatas* (L.) Lam.] was the root crop with the second highest global production after cassava [*Manihot esculenta* Crantz] (FAO, 2020). Sweetpotato is highly nutritious and outperforms most carbohydrate-based foods in terms of vitamins, minerals, dietary fiber, and total protein, making it a priority in crop-based strategies to enhance global food and nutrition security (Woolfe, 1992; Motsa et al., 2015). Sweetpotato is also known for its ability to grow in drought-prone soils with low external input (fertilizer and pesticide) requirements, while thanks to a short growing cycle, it is also recognized as an excellent crop for post-crisis (hurricane, flooding, refugee settlement, etc.) situations (Bradbury and Holloway, 1987; Mwanga and Ssemakula, 2011; Mekonnen et al., 2015). However, a range of biotic and abiotic stress factors such as viruses, weevils, and severe drought, are inhibiting farmers' ability to achieve sweetpotato's full yield potential (Valverde et al., 2007; Mwanga and Ssemakula, 2011; Agili and Nyende, 2012; Kivuva et al., 2015). Moreover, climate change's negative effect on sweetpotato productivity is of rising concern in regions affected by increasing global temperatures, such as tropical and subtropical areas of Sub-Saharan Africa (Low et al., 2009; Schulze, 2011; Knox et al., 2012). Resilient sweetpotato varieties could play an extremely important food and nutrition security role in the developing world, one that will increase in importance as the availability of good cropping areas becomes more limited.

Series *Batatas* [*Ipomoea* series *Batatas* (Choisy) D. F. Austin] is a subdivision within *Ipomoea*, the largest genus in the morning glory (Convolvulaceae) family. This group includes the cultivated hexaploid sweetpotato [*I. batatas* (L.) Lam.], wild tetraploid *I. batatas* (L.) Lam. (Ozias-Akins and Jarret, 1994), and 15 closely related wild species (Austin, 1978; McDonald and Austin, 1990; Austin et al., 1993; Wood et al., 2015; Wood et al., 2020). The wild species include *Ipomoea trifida* (H.B.K.) G. Don, *Ipomoea cordatotriloba* Dennstedt, *Ipomoea cynanchifolia* Meisn., *Ipomoea grandifolia* (Dammer) O'Donell, *Ipomoea lacunosa* L., *Ipomoea leucantha* Jacquin, *Ipomoea littoralis* Blume, *Ipomoea ramosissima* (Poir.) Choisy, *Ipomoea splendor-sylvae* House, *Ipomoea tabascana* McDonald and Austin, *Ipomoea tenuissima* Choisy, *Ipomoea tiliacea* (Willd.) Choisy in D. C., *Ipomoea triloba* L., *Ipomoea lactifera* J.R.I. Wood and Scotland, and *Ipomoea australis* (O'Donell) J.R.I. Wood & P. Muñoz. Since sweetpotato crop wild relatives (SP-CWR) are well adapted to diverse, even extreme environmental conditions (Iwanaga, 1988; Guarino and Lobell, 2011), they are considered a prominent genetic resource to improve both biotic and abiotic stress tolerance in cultivated crops (Iwanaga, 1988; Komaki, 2004; Nimmakayala et al., 2011). However, Khoury et al. (2015) pointed out that the full potential offered by SP-CWR, especially in terms of drought tolerance, is far from being fully exploited.

From a physiological point of view, drought tolerance encompasses all mechanisms that enable plants to avoid or tolerate dehydration, such as maintenance of cell turgor and water uptake, and reduction of water loss, among others (Fischer and Maurer, 1978; Turner, 1986; Turner, 1997). In sweetpotato, several physiological traits have been used for screening drought tolerant genotypes such as chlorophyll concentration (Mbinda et al., 2018; Mbinda et al., 2019), canopy cover (Laurie et al., 2014), leaf temperature (Laurie et al., 2014; Rukundo et al., 2017; Low et al., 2020), and ^{13}C discrimination (Low et al., 2020). The latter has been reported as one of the most accurate criteria for selecting drought tolerant genotypes (Tuberosa, 2012; Low et al., 2020) since it is a good indicator of stomatal conductance (Condon et al., 2004), water use efficiency (Turner, 1997), and photosynthetic performance (Jefferies and Mackerron, 1997; Dawson et al., 2002).

Previous studies in other plants such as *Arabidopsis thaliana* (Ling et al., 2018; Serrano et al., 2019), potato (Watkinson et al., 2006; Ramírez et al., 2015a), cassava (Cayón et al., 1997), wheat (Ahmed et al., 2016), and grasses (Walter et al., 2011) have shown that a previous exposure to stress in early development stages "prepares" the plant for a subsequent exposure to the same stress. This behavior, known as stress memory, is achieved through accumulation of transcription factor or signaling proteins and epigenetic changes, which are later translated into an enhanced physiological response (Bruce et al., 2007). In this study, we tested potential short-term memory occurrence, which involves tolerant responses within the plant life cycle to previous stress periods (Conrath et al., 2006; Bruce et al., 2007), as opposed to transgenerational memory, in which the tolerant response is inherited by the next plant generation (Han and Wagner, 2014). Wild species are considered to be possible sources of useful genes/alleles related to stress tolerance as they have evolved under natural selection to survive consecutives periods of climate extremes (Sharma et al., 2013). In the present study, we focused on the previously described physiological traits to determine potential stress memory induction in SP-CWR.

Farmers have traditionally selected drought tolerant genotypes by choosing those genotypes that maintained high yield under drought stress conditions. However, this selection strategy may penalize resilience or production stability, as shown

by Blum (1996), who found that some genotypes with high tolerance presented low yield performance in the absence of stress conditions. Fernandez (1992) classified plants into four groups based on yield performance in stress and non-stress conditions: Group A (accessions expressing uniform superiority in both stress and non-stress conditions), Group B (accessions presenting good performance only under potential conditions and not under stress conditions), Group C (accessions expressing a relatively higher performance only under stress), and Group D (accessions with poor performance in both environments). Recently, Thiry et al. (2016) proposed a yield-based definition of drought tolerance in terms of resilience and productivity of crops useful to facilitate the classification of genotypes for breeding programs. These authors stated that the resilience capacity index (RCI) expresses the yield decrease of genotypes under stress within a population, compared with yield potential conditions, whereas the production capacity index (PCI) expresses the mean production of genotypes under both stressed and non-stressed conditions within a population. Moreover, a genotype could be classified in any of Fernandez's (1992) Groups with those genotypes scoring the highest values being the most drought tolerant ones. Using combined indices such as resilience and productivity simultaneously in sweetpotato breeding programs would reduce costs and save time.

The aim of this study was i) to determine potential short-term memory induction in SP-CWR and its manifestation in ecophysiological traits like senescence, foliar area, leaf-minus-air temperature, and leaf ^{13}C discrimination and ii) to identify the memory-induced physiological mechanisms associated with the development of drought tolerance (in terms of resilience and productivity) in SP-CWR in order to identify new sources for breeding towards improved drought tolerance in sweetpotato.

MATERIAL AND METHODS

Experimental Conditions and Management

A pot experiment was conducted under greenhouse conditions from July 13th to December 18th, 2018 at the experimental station of the International Potato Center (CIP) in San Ramón, Junín, Peru (11° 7' 39.3" S, 75°21' 23.4" W, 850 m a.s.l.). The station is located in the mountainous Amazon area of central Peru. The region is characterized by a rainy, warm, and very humid climate (SENAMHI, 2020) with an average annual maximum temperature, average annual solar radiation and annual precipitation of 30.8 ± 0.46°C, 34.2 ± 1.51 MJ m^{-2} day^{-1}, and 1,294 mm, respectively (data from 2019, CIP–San Ramón weather station). Pots were distributed in two neighboring screenhouses covered with an anti-aphid mesh with an opening size of 0.26 × 0.82 mm. The screenhouses were also covered with a translucid plastic and a black mesh (hole size of 2 × 2 mm) to prevent high radiation stress and provide shelter from rain. During the study period, the average relative humidity, and maximum and minimum temperatures were 81.2 ± 0.96%, 32.7 ± 0.47°C, and 19.8 ± 0.21°C respectively. The average solar radiation and maximum vapor pressure deficit were 18.5 ± 0.94 MJ m^{-2} day^{-1} and 2.7 ± 0.13 kPa (**Supplementary Table S1**). These variables were measured every 5 min. Temperature, relative humidity (S-THB-M008 model), and solar radiation (S-LIB-M003 model) sensors were recorded with a HOBO U30 datalogger (Onset Computer Corporation, Bourne, MA, USA).

On July 13th eight seeds per accession were scarified individually and placed on moist filter paper in petri dishes. After germination seedlings were planted in peat pellets (Jiffy Products Ltd., Canada) for 15 days to promote root development. Subsequently, plantlets were transferred to pots (6.4 L) filled with 6.5 kg of a 2:1 mixture of sand and peat-based substrate (PRO-MIX, Premier Tech Horticulture, Canada). All pots were randomized regularly i.e. the positions of individual plants within the greenhouses were changed to avoid a significant effect of pot positions on the measured traits. In each screenhouse two HOBO U30 datalogger were placed at different positions (one near the central isle and another near the outside screen wall) to monitor temperature and humidity gradients.

Each pot received seven fertilizer applications scheduled every 2 weeks using 0.51 g N, 0.78 g P_2O_5, and 0.60 g, K_2O (Peters Professional ICL Ltd., Israel). To control thrips and whitefly, folding traps with pheromones were installed and insecticides (Ocaren, active ingredient: profenofos and fipronil, Interoc S.A., Peru and Vertimec, active ingredient: abamectina; Farmex S.A., Peru) were applied at a dose rate of 1 mL L^{-1} when necessary. Since the wild species included in this study are twiners, a wire-made spiral welded to three vertical rods was placed in each pot to support the plants, in order to facilitate vertical growth and the expansion of foliar area.

Plant Material

Fifty-nine accessions belonging to the *Ipomoea* series *Batatas* (Choisy) D. F. Austin, the closest wild relatives of sweetpotato, were selected from the SP-CWR collection of CIP's genebank (**Table 1**). The plant material included four accessions of cultivated hexaploid *I. batatas* (L.) Lam. (hereinafter "sweetpotato") namely "Beauregard," "Tanzania," and two accessions deriving from crosses between the two varieties: "B×T" (Wu et al., 2018). Also, two accessions of wild tetraploid (4x) *I. batatas* (L.) Lam. (Ozias-Akins and Jarret, 1994) and 53 accessions encompassing 10 species of the series *Batatas* were included. The wild relative species assessed in this study were: *I. australis* (O'Donell) J.R.I. Wood & P. Muñoz (5), *I. cordatotriloba* Dennstedt (1), *I. cynanchifolia* Meisn. (3), *I. grandifolia* (Dammer) O'Donell (5), *I. leucantha* Jacquin (2), *I. ramosissima* (Poir.) Choisy (8), *I. splendor-sylvae* House (3), *I. tiliacea* (Willd.) Choisy (2), *I. trifida* (H.B.K.) G. Don (16), and *I. triloba* L. (8) (Khoury et al., 2015; Wood et al., 2020). The basic chromosome number of all accessions in this study is x = 15. While most SP-CWRs of our taxon sample are diploid (2n = 2x = 30), four accessions are tetraploid (2n = 4x = 60) and the cultivated sweetpotato accessions are hexaploid (2n = 6x = 90) (G. Rossel, CIP, pers. comm.) (**Table 1**). The taxon sample of the wild species focused on Central and South American species and considered geographic distribution range and morphological variation between accessions of the same species.

TABLE 1 | Sweetpotato crop wild relatives [*Ipomoea* series *Batatas* (Choisy) D. F. Austin] accessions from CIP's genebank used in this study, with chromosome number and ploidy (basic chromosome number: x=15), country of origin and biological status for each of them.

Accession number	Species	Chromosome number and ploidy	DOI	Origin	Biological status	Groups
CIP 460360	*I. australis*	2n = 2x = 30†	10.18730/82SY	PRY	W	B
CIP 460296	*I. australis*	2n = 2x = 30†	10.18730/80T4	ARG	W	B
CIP 460345	*I. australis*	2n = 2x = 30†	10.18730/82AF	PRY	W	D
CIP 460585	*I. australis*	2n = 2x = 30†	10.18730/89GQ	ARG	W	D
CIP 460164	*I. australis*	2n = 2x = 30†	10.18730/7Y5Y	PRY	W	D
CIP 430434	*I. batatas*	2n = 4x = 60†	10.18730/852X	JAP	W	B
CIP 460577	*I. batatas*	2n = 4x = 60†	10.18730/898F	ECU	TC	B
CIP 440132	*I. batatas* "Beauregard"	2n = 6x = 90*	10.18730/65R1	USA	AC	B
CIP 440166	*I. batatas* "Tanzania"	2n = 6x = 90*	10.18730/66NY	UGA	TC	B
CIP 105269.232	*I. batatas* "B×T"	2n = 6x = 90*	10.18730/SK7J$	PER	Br	B
CIP 113641.086	*I. batatas* "B×T"	2n = 6x = 90*	10.18730/SK7K=	PER	Br	D
CIP 460077	*I. cordatotriloba*	2n = 2x = 30†	10.18730/7W75	MEX	W	B
CIP 460149	*I. cynanchifolia*	2n = 2x = 30†	10.18730/7XTK	BRA	W	B
CIP 460556	*I. cynanchifolia*	2n = 2x = 30†	10.18730/88M*	BRA	W	D
CIP 460555	*I. cynanchifolia*	2n = 2x = 30†	10.18730/88KZ	BRA	W	D
CIP 460610	*I. grandifolia*	2n = 2x = 30†	10.18730/89QY	BRA	W	B
CIP 460201	*I. grandifolia*	2n = 2x = 30†	10.18730/7YZK	ARG	W	D
CIP 460583	*I. grandifolia*	2n = 2x = 30†	10.18730/89EN	URU	W	D
CIP 460452	*I. grandifolia*	2n = 2x = 30†	10.18730/85MA	ARG	W	D
CIP 460337	*I. grandifolia*	2n = 2x = 30†	10.18730/8238	PRY	W	D
CIP 460619	*I. leucantha*	2n = 2x = 30†	10.18730/89T~	COL	W	B
CIP 460204	*I. leucantha*	2n = 2x = 30†	10.18730/7Z2P	ARG	W	D
CIP 460028	*I. ramosissima*	2n = 2x = 30†	10.18730/7V58	ECU	W	B
CIP 460032	*I. ramosissima*	2n = 2x = 30†	10.18730/7V9C	BOL	W	B
CIP 460566	*I. ramosissima*	2n = 2x = 30†	10.18730/88Y5	PER	W	B
CIP 460567	*I. ramosissima*	2n = 2x = 30†	10.18730/88Z6	PER	W	B
CIP 460005	*I. ramosissima*	2n = 2x = 30†	10.18730/7THS	PER	W	B
CIP 460047	*I. ramosissima*	2n = 2x = 30†	10.18730/7VNR	PER	W	B
CIP 460722	*I. ramosissima*	2n = 2x = 30†	10.18730/8B0$	ARG	W	B
CIP 460036	*I. ramosissima*	2n = 2x = 30†	10.18730/7VCF	BOL	W	D
CIP 460131	*I. splendor-sylvae*	2n = 2x = 30†	10.18730/7XE7	MEX	W	B
CIP 460373	*I. splendor-sylvae*	2n = 2x = 30†	10.18730/8355	NIC	W	B
CIP 460383	*I. splendor-sylvae*	2n = 2x = 30†	10.18730/83FF	NIC	W	B
CIP 460528	*I. tiliacea*	2n = 4x = 60†	10.18730/87V7	CUB	W	B
CIP 460531	*I. tiliacea*	2n = 4x = 60†	10.18730/87YA	CUB	W	B
CIP 460663	*I. trifida*	2n = 2x = 30†	10.18730/8A9B	MEX	W	C
CIP 113735.283	*I. trifida*	2n = 2x = 30†	10.18730/SK7P1	PER	Br	C
CIP 107665.9	*I. trifida*	2n = 2x = 30*	10.18730/SK7MU	PER	W	C
CIP 460026	*I. trifida*	2n = 2x = 30†	10.18730/7V36	COL	W	D
CIP 460430	*I. trifida*	2n = 2x = 30†	10.18730/84YS	CUB	W	D
CIP 460377	*I. trifida*	2n = 2x = 30†	10.18730/8399	NIC	W	D
CIP 460745	*I. trifida*	2n = 2x = 30†	10.18730/8BDA	GUA	W	D
CIP 460429	*I. trifida*	2n = 2x = 30†	10.18730/84XR	NIC	W	D
CIP 460096	*I. trifida*	2n = 2x = 30†	10.18730/7WKH	VEN	W	D
CIP 460195	*I. trifida*	2n = 2x = 30†	10.18730/7YSD	VEN	W	D
CIP 460022	*I. trifida*	2n = 2x = 30†	10.18730/7V03	COL	W	D
CIP 113735.258	*I. trifida*	2n = 2x = 30*	10.18730/SK7N0	PER	Br	D
CIP 113735.302	*I. trifida*	2n = 2x = 30*	10.18730/SK7Q2	PER	Br	D
CIP 113735.329	*I. trifida*	2n = 2x = 30*	10.18730/SK7R3	PER	Br	D
CIP 107665.19	*I. trifida*	2n = 2x = 30*	10.18730/SK7H~	PER	W	D
CIP 460021	*I. trifida*	2n = 2x = 30†	10.18730/7TZ2	VEN	W	D
CIP 460309	*I. triloba*	2n = 2x = 30†	10.18730/817H	PRY	W	B
CIP 460116	*I. triloba*	2n = 2x = 30†	10.18730/7X5=	COL	W	C
CIP 460560	*I. triloba*	2n = 2x = 30†	10.18730/88RU	PER	W	D
CIP 460052	*I. triloba*	2n = 2x = 30†	10.18730/7VSW	VEN	W	D
CIP 460093	*I. triloba*	2n = 2x = 30†	10.18730/7WJG	DOM	W	D
CIP 460517	*I. triloba*	2n = 2x = 30†	10.18730/87H$	ECU	W	D
CIP 460784	*I. triloba*	2n = 2x = 30†	10.18730/8C1Y	JAM	W	D
CIP 460078	*I. triloba*	2n = 2x = 30†	10.18730/7W86	MEX	W	D

*Groups were defined based on clustering analysis (see **Figure 4**). Groups B, C and D represent clusters I, II and III, respectively.*
*W, wild; TC, traditional cultivar/landrace; AC, advanced or improved cultivar; Br, breeding line. †(G. Rossel, CIP, pers. comm.). *Wu et al. (2018).*

Drought Priming Process

Eight individual plants per accession were distributed at random in both greenhouses. Four individual plants per accession (i.e. four replications) were randomly assigned to the two experimental treatments. Previous stress memory studies by Liu and Charng (2012); Ling et al. (2018) and Serrano et al. (2019) were used to guide the experimental design of this study. Treatments consisted of a series of water restriction periods of increasing length (priming), *versus* full irrigation (non-priming). After transplantation of seedlings into pots, all pots were watered until field capacity (0.32 w/w) three times per week, following the protocol of Rolando et al. (2015), until harvest. Non-primed plants received full irrigation during the entire experimental cycle while plants undergoing the water restriction treatment were watered until flowering onset (FO) (**Figure 1**). Priming was initiated when 50% of samples conforming an accession started to exhibit flower buds (defined as FO occurrence) (Thiry et al., 2016). Two different starting dates were established for early-flowering (20 accessions) and late-flowering (18 accessions) or non-flowering (21 accessions) accessions (**Supplementary Table S2**). Three priming events, or water restriction periods, were carried out with 8, 11, and 14 days of total water restriction. Each of these water restricted periods was followed by a recovery period of 14 days in which plants were irrigated until field capacity (**Figure 1**). Therefore, the priming process consisted of three water restriction periods: the first one started at FO whereas the second and third priming events were initiated at 68 and 94 days after transplanting (DAT) for early as well as 94 and 122 DAT for late and non-flowering accessions, respectively.

Drought stress responses were measured at the end of the first water restriction period to corroborate if 8 days was a sufficiently long period to induce a significant drought stress. For this purpose, maximum stomatal conductance at saturated light (g_{s_max}) was measured (*sensu* Medrano et al., 2002) in one young and sun-exposed leaf of eight plants of 19 randomly selected accessions using a portable photosynthesis system (LI-6400TX, LI-COR, Nebraska, USA). Measurements were carried out during early morning hours from 6:00 to 10:00 h local time. The following micro-environmental parameters were recorded: photosynthetic active radiation = 1,200 μmol m^{-2} s^{-1}, CO_2 concentration = 400 ppm, and atmospheric humidity = 50%.

Ecophysiological Measurements to Determine Potential Short-Term Memory Occurrence

Chlorophyll concentration (Chl_{SPAD}) of all accessions was measured using a portable chlorophyll meter (SPAD-502 Plus, Konica Minolta Inc., Osaka, Japan) on 17 occasions during the study period. An average of four readings from four young and sun-exposed leaves were taken per plant. Senescence (S) was estimated on each plant as the slope generated by fitting Chl_{SPAD} (from maximum leaf greenness to harvest) *vs.* time on a linear function (Li et al., 2019). Assuming that an extended senescence delay (high S) is associated with a higher probability of fixing more carbon during the lifespan, short-term memory proxy's (STM) effect on S (STM_S) was calculated as follows:

$$STM_S = S_{pr} - S_{npr} \quad (1)$$

Where S_{pr} and S_{npr} are S average values for primed and non-primed plants, respectively. Regular assessments of foliar area (FA) were carried out on both primed and non-primed plants before and after each water restriction and recovery period by taking visible images (using a Nikon model D7000 camera, Nikon Corp., Japan) of each plant. Images were acquired following CIP procedures (PSE-CIP, 2013) and processed with Image Canopy software (Barreda et al., 2017) that allows the calculation of foliar area through an image segmentation technique separating healthy green vegetation from other components within a picture. To test the effect of STM on FA (STM_{FA}) the following equation was calculated:

$$STM_{FA} = \frac{FA_{\max\_pr}}{FA_{\max\_npr}} \quad (2)$$

Where FA_{max_pr} and FA_{max_npr} are FA average maximum value of the temporal assessments corresponding to primed and non-primed plants, respectively. Leaf temperature was measured radiometrically from 14:00 to 15:00 h following the protocols of

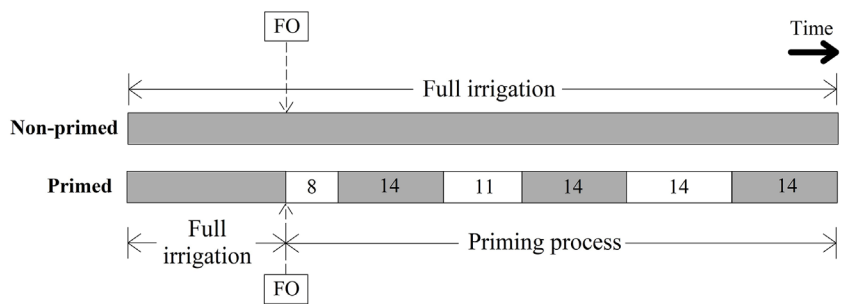

FIGURE 1 | Timeline representation for the watering treatments per accession: non-primed plants (no water restriction) and primed plants (water restriction periods). Gray and white blocks mean substrate watering with full irrigation or no irrigation, respectively. Duration of every period during the priming process is indicated in days within the blocks. The priming process started after flowering onset (FO).

Idso et al. (1981) and Rinza et al. (2019). Measurements were taken using an infrared thermometer (DT-882 model, CEM, China) during 14 measurement events throughout the study period. Air temperature was registered with four data loggers (HOBO U23 Pro v2 Temperature/Relative Humidity, Onset Computer Corporation, Bourne, MA, USA), two of which were located inside each greenhouse at canopy level. The difference between leaf and air temperature (dT) was calculated using average leaf temperature and average air temperature registered at the same time when leaf temperature was measured. STM effect on dT (STM_{dT}) was calculated as follows:

$$STM_{dT} = \frac{dT_{\min\_pr}}{dT_{\min\_npr}} \quad (3)$$

Where dT_{min_pr} and dT_{min_npr} are the average minimum dT over time reached by primed and non-primed plants of an accession. Finally, two composed leave samples per accession and treatment were collected at the end of each recovery period. Six assessments in total were carried out from FO until harvest at 66, 93, and 121 DAT for early flowering accessions and at 94, 121, and 148 DAT for late- or non-flowering accessions. Each composed sample consisted of 20 leaves—five young leaves per plant—which were oven dried at 60°C for 48 h (BLUE M Model POM-166EY, BLUE M Electronic Company, IL, USA). Dried leaves were milled with a ball miller (MBIX-100 model, MRC, Israel) and packed in tin capsules (Ramírez et al., 2015b). Capsules were sent to the Stable Isotope Facility at the University of Davis, USA, for carbon isotope composition ($\delta^{13}C$) analysis using a PDZ Europa ANCA-GSL elemental analyzer coupled to PDZ Europa 20-20 isotope ratio mass spectrometer (Sercon Ltd., Cheshire, UK). Leaf ^{13}C discrimination (Δ) was calculated as described by Farquhar et al. (1989):

$$\Delta(‰) = \left(\frac{\delta_a - \delta_p}{1 + \delta_p}\right) \times 1000 \quad (4)$$

Where δ_p is $\delta^{13}C$ of the sample and δ_a is the $\delta^{13}C$ of the atmospheric CO_2, -8‰. STM effect on Δ (STM_Δ) was estimated as follows:

$$STM_\Delta = \frac{\Delta_{\max\_pr}}{\Delta_{\max\_npr}} \quad (5)$$

Where Δ_{max_pr} and Δ_{max_npr} are the maximum Δ over time for primed and non-primed plants within an accession, respectively. The STM occurrence was defined when primed plants' response exceeded that of non-primed plants, i.e. when the STM value was higher than zero (STM>0) in the case of S in Equation 1, and higher than one (STM>1) in the case of FA, dT and Δ in Equation 2, 3, and 5, respectively. For all accessions, the priming process started at the same physiological stage (following recommendations from Thiry et al., 2016) so that all the ecophysiological indicators (S, A_{max}, dT_{min}, and Δ_{max}) used to determine potential STM occurrence were comparable between accessions regardless of the starting date of the priming process.

Drought Tolerance Indices and Statistical Analysis

At the end of the priming process, total biomass was harvested and subsequently oven dried at 60°C for 48 h to calculate the dry weight of total biomass of primed (Y_{pr}) and non-primed (Y_{npr}) plants. In this study, total biomass comprises above and below ground biomass and is generally expressed on a dry matter (DM) basis. The stress susceptibility index (SSI) (Fischer and Maurer, 1978) and the geometric mean productivity index (GMP) (Fernandez, 1992) were calculated as follows:

$$SSI = \frac{1 - (Y_{pr}/Y_{npr})}{1 - (\overline{Y_{pr}}/\overline{Y_{npr}})} \quad (6)$$

$$GMP = \sqrt{Y_{pr} \times Y_{npr}} \quad (7)$$

Where Y_{pr} and Y_{npr} are the average total biomass (DM) production of the primed and non-primed plants, respectively, and $\overline{Y_{pr}}$ and $\overline{Y_{npr}}$ are the overall average total biomass (DM) production of the primed and non-primed plants, respectively. SSI and GMP of a genotype were estimated as the resilience capacity index (RCI) and the production capacity index (PCI) following the score values method described by Thiry et al. (2016) (**Supplementary Appendix A**). For of both score indices (RCI and PCI), a score value of 1 means high susceptibility whereas a score of 10 represents high drought tolerance.

A repeated measures analysis of variance (rmANOVA) was run to evaluate the effects of the treatments (between subjects factor) and time (within subjects factor) on assessed ecophysiological variables (SPAD, FA and dT, except for Δ, due to its composed samples). A one-way ANOVA was used to assess differences (at $p < 0.05$) among accessions in total dry weight (Y). A t-student test was used to identify significant differences between treatments for every physiological evaluation of SPAD, FA, dT, Δ, and Y. A Pearson correlation analysis was performed to evaluate the relationship between STM traits and both RCI and PCI. A principal component analysis (PCA) was run to analyze the accessions ordination through the association of STM trait effects and RCI, as well as PCI. Finally, a cluster analysis following the Ward's method was computed using the R package "FactoMineR" (Lê et al., 2008) to classify all accessions into any of the four Groups of plants described by Fernandez (1992). All above described tests were performed with RStudio software (R Core Team, 2019).

RESULTS

Short-Term Memory Occurrence

Chl_{SPAD}, FA, dT, and Δ values ranged between 10.3–51.4 SPAD units, 0.0–1,537.3 cm^2, −12.2 – 4.1°C, and 20.4–25.3‰, respectively. Effects of the water treatment for Chl_{SPAD}, FA, and dT were significant ($p < 0.05$) for 33.9, 81.4, and 81.4% of the total accessions and the effect of time was significant in all accessions for Chl_{SPAD}, FA, and dT. The percentage of accessions with a significant ($p < 0.05$) interaction between watering

treatments and time for Chl_{SPAD}, FA, and dT were 93.2, 98.3, and 78.0%, respectively (**Supplementary Table S3**). At the end of the first water restriction period, at 53 DAT, primed plants had 78.6% less average g_{s_max} than non-primed plants (0.02 ± 0.03 and 0.10 ± 0.04 mol H_2O m^{-2} s^{-1} for primed and non-primed plants, respectively) of the 19 assessed accessions. Moreover, 16 accessions (81.3%) presented primed plants with average g_{s_max} below 0.05 mol H_2O m^{-2} s^{-1} (**Supplementary Table S4**). STM occurrence was detected in 23.7, 28.8, 50.8, and 81.4% of the total number of accessions for FA, dT, Δ, and S, respectively (**Figure 2**). For STM_S, *I. leucantha* presented the highest value (0.48 for CIP 460204) and *I. trifida* had the widest range (−0.09 – 0.35) (**Figure 3A**) while for STM_{FA}, *I. ramosissima* presented the highest value (1.5 for CIP 460032) as well as the widest range (0.76–1.5) (**Figure 3B**). *I. australis* showed both the highest STM_{dT} value (1.26 for CIP 460296) and the widest STM_{dT} range (0.66–1.26) (**Figure 3C**) whereas *I. batatas* had the highest STM_Δ value (1.04 for CIP 430434) and *I. trifida* the widest STM_Δ range (0.93–1.03) (**Figure 3D**). The raw data and RGB images of each plant are available online (Guerrero-Zurita et al., 2020).

Drought Tolerance, Resilience, and Productivity Under Drought and Their Relationship With Short-Term Memory Occurrence

SSI and GMP values ranged between −0.83 – 1.74 and 7.56–78.1, respectively (**Supplementary Table S5**). High coefficients of determination between the original values for SSI and GMP and their RCI and PCI score indices were obtained (R^2 = 0.970 for SSI *vs.* RCI and R^2 = 0.974 for GMP *vs.* PCI) indicating that score indices can be used as a surrogate of their original index

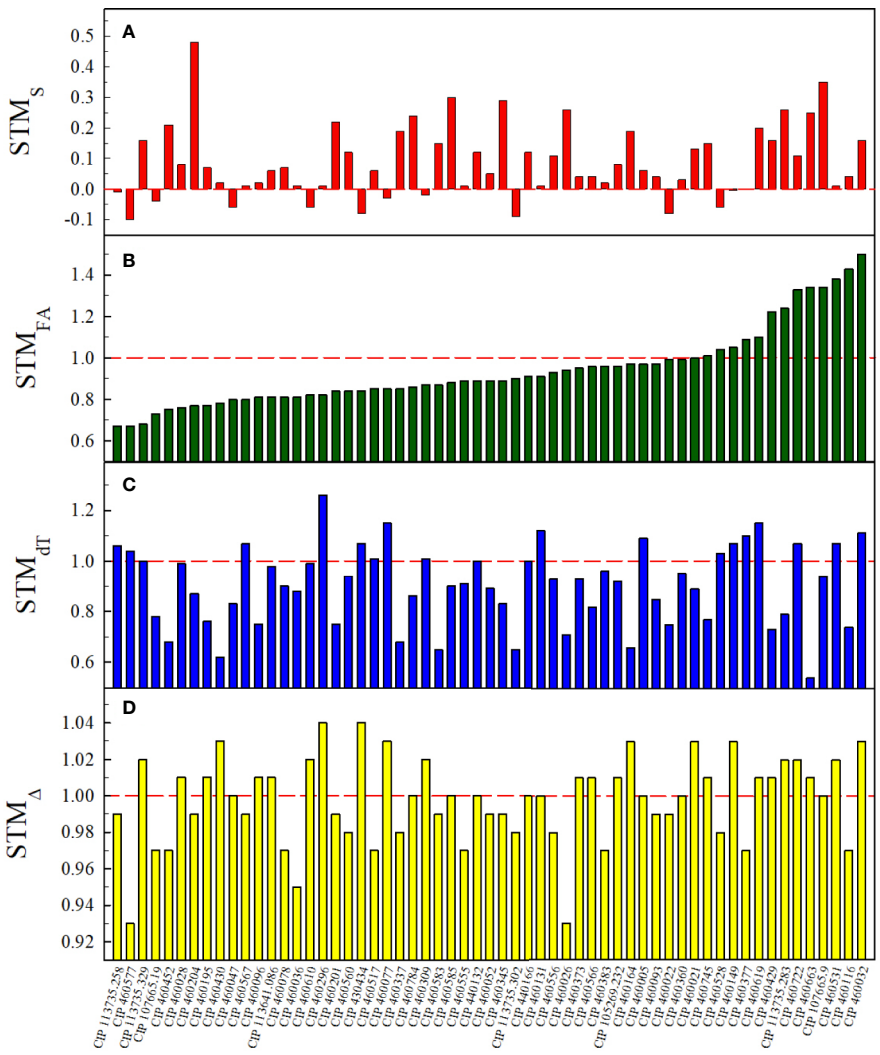

FIGURE 2 | Short-term memory (STM) values per accession for senescence (STM_S) **(A)**, foliar area (STM_{FA}) **(B)**, leaf-minus-air temperature (STM_{dT}) **(C)** and leaf ^{13}C discrimination (STM_Δ) **(D)**. Red dashed line indicates STM occurrence.

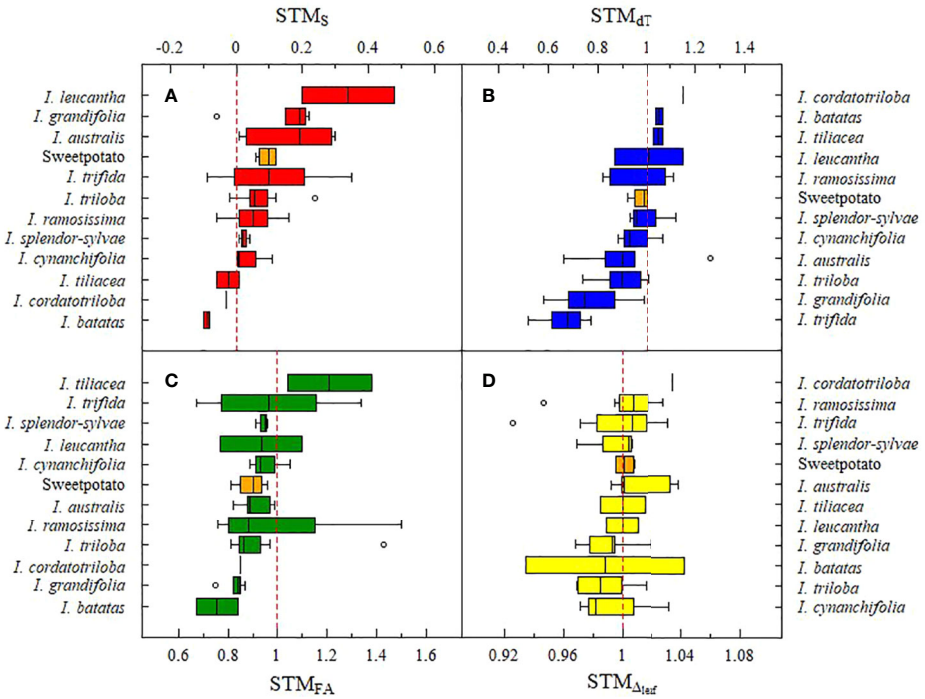

FIGURE 3 | Boxplot of short-term memory (STM) values per crop wild relatives (CWR) species and sweetpotato cultivars (orange boxplots) for senescence (STM_S) **(A)**, foliar area (STM_{FA}) **(B)**, leaf-minus-air temperature (STM_{dT}) **(C)** and leaf ^{13}C discrimination (STM_Δ) **(D)**. Red dashed line indicates the STM occurrence. In each specie's boxplot, black line represents the median value. Boxplot contains the variation between 25 and 75% and gray circles are the outlier values.

value. The highest RCI were found in CIP 460116 (8.0), CIP 107665.9 (7.5), and CIP 460663 (7.0) while *I. triloba* had the highest range of values (3.5–8.0) followed by *I. trifida* (2.0–7.5) and *I. grandifolia* (1.75–5.25). Regarding PCI, the highest values were found in CIP 440166 (9.75), CIP 460131 (6.0), and CIP 440132 (5.75) while sweetpotato cultivars presented the highest range of scores (2.75–9.75) followed by *I. splendor-sylvae* (4.25–6) and *I. ramosissima* (2.5–5.75). Moreover, *I. cynanchifolia*, *I. australis*, *I. grandifolia*, *I. trifida*, and *I. triloba* had higher RCI than PCI, whereas *I. leucantha*, *I. ramosissima*, *I. cordatotriloba*, *I. tiliacea*, wild *I. batatas*, *I. splendor-sylvae*, and the sweetpotato cultivars had higher PCI than RCI.

STM_{dT} showed a significant positive and negative correlation with PCI (r = 0.52) and RCI (r = −0.30), respectively, whereas STM_{FA} was the trait with the highest positive correlation with RCI (r = 0.38) (**Table 2**). The first three components of the PCA represented 72.7% of the total variance (**Table 3**). The first component was mainly explained by STM_S and RCI with a negative effect and STM_{dT} and PCI with a positive effect. STM_{FA} showed a higher weight in the second principal component. The third principal component was mainly explained by STM_{dT} with positive effect (**Table 3**).

Clustering analysis grouped the accessions into three clusters: I, II, and III (**Figure 4**). Cluster I, II, and III contained 24, 4, and 31 accessions, respectively (**Table 1**). When analyzing the average response per variable obtained in each group, cluster I

TABLE 2 | Pearson correlation coefficients matrix among short-term memory (STM) traits and drought tolerant indices.

	STM_S	STM_{FA}	STM_{dT}	STM_{dT}	RCI
STM_{FA}	0.22				
STM_{dT}	-0.26*	-0.01			
STM_Δ	0.01	0.20	0.20		
RCI	0.21	0.38**	-0.30*	-0.14	
PCI	-0.25	0.13	0.52**	0.19	-0.27*

*RCI, resilience capacity index; PCI, production capacity index. See STM traits abbreviations in **Figure 2**. $^{**}p < 0.01$, $^{*}p < 0.05$.*

TABLE 3 | Extracted components from Principal Component Analysis based on the ordination of short-term memory (STM) traits, drought tolerance resilience (RCI), and productivity (PCI).

Variable	PC 1	PC 2	PC 3
STM_S	-0.55	0.30	0.46
STM_{FA}	-0.21	0.85	-0.22
STM_{dT}	0.77	0.21	-0.14
STM_Δ	0.33	0.54	0.61
RCI	-0.66	0.37	-0.45
PCI	0.73	0.34	-0.25
Eigen-value	2.03	1.42	0.91
TCV (%)	31.86	57.47	72.72

*TCV, total cumulative variance; PC, principal component. See STM traits abbreviations in **Figure 2**.*

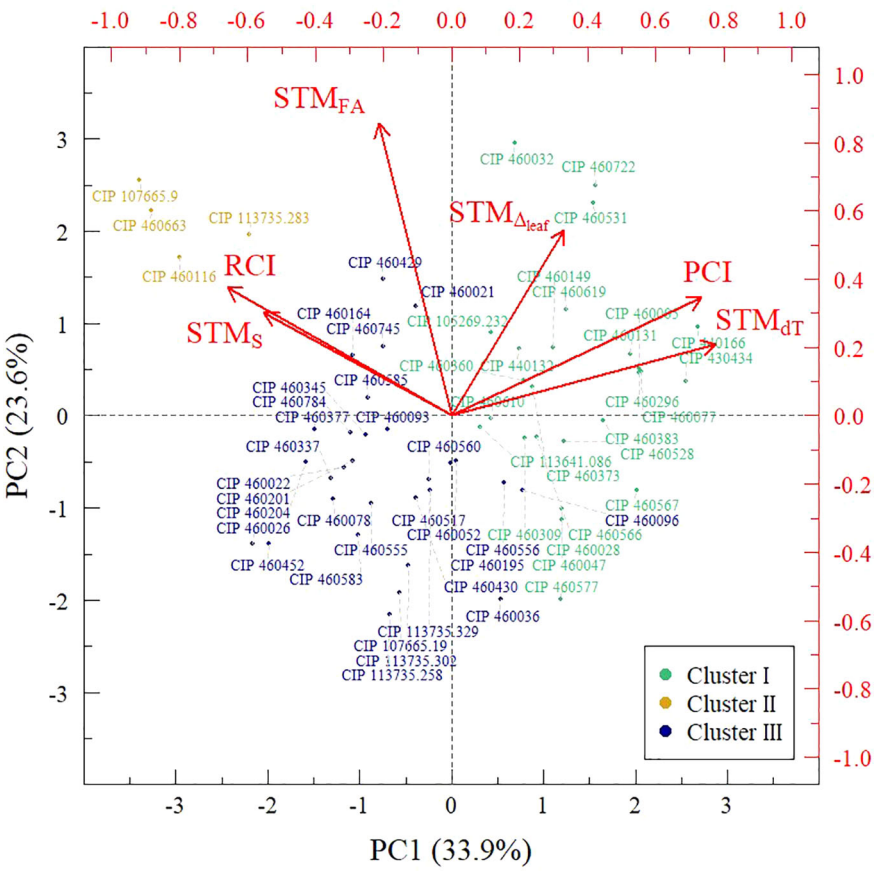

FIGURE 4 | Clustering analysis and 2D ordination of sweetpotato cultivars and its CWR species based on Principal Component Analysis for component loadings (STM traits, PCI and RCI). Clusters I, II and III correspond to Fernandez's (1992) Groups B, C, and D, respectively.

had the highest average for STM_{dT}, STM_{Δ}, PCI, and Y_{npr} (1.03 ± 0.1, 1.01 ± 0.0, 4.39 ± 1.7, and 49.1 ± 16.4 g, respectively) whereas cluster II was highest for STM_S, STM_{FA}, and RCI (0.23 ± 0.1, 1.34 ± 0.1, and 7.0 ± 1.1, respectively) (**Table 4**). Accessions belonging to cluster III displayed the lowest average response for all traits (**Table 4**). Cluster I contained the highest percentage of accessions with STM_{Δ} occurrence (70.8%) followed by cluster II (50%) and cluster III (35.5%) (**Figure 5**).

TABLE 4 | Average value ± standard error of STM traits, drought tolerance indices and total biomass production for each cluster (I, II and III) from Principal Component Analysis.

	Cluster		
Variable	**I**	**II**	**III**
STM_S	0.03 ± 0.01	0.23 ± 0.1	0.11 ± 0.1
STM_{FA}	0.96 ± 0.2	1.34 ± 0.1	0.87 ± 0.1
STM_{dT}	1.03 ± 0.1	0.63 ± 0.1	0.78 ± 0.1
STM_{Δ}	1.01 ± 0.0	1.00 ± 0.0	0.99 ± 0.0
RCI	3.17 ± 1.0	7.0 ± 1.1	3.77 ± 1.0
PCI	4.39 ± 1.7	1.81 ± 0.6	2.04 ± 0.9
Y_{npr} (g)	49.1 ± 16.4	17.6 ± 4.8	24.5 ± 9.3
Y_{pr} (g)	25.9 ± 9.4	16.6 ± 2.6	14.4 ± 5.0

*Y_{npr}, total biomass production of non-primed treatment; Y_{pr}, total biomass production of primed treatment. See STM traits, RCI, and PCI abbreviations in **Figure 2**.*

DISCUSSION

Priming Induction Was Expressed Mainly in Senescence Delay and Photosynthetic Performance Traits

The highest Chl_{SPAD} values (51.4, 50.5, and 49.4 SPAD units for accessions CIP 440166, CIP 430434, and CIP 460528, respectively) observed in this study were higher than previously recorded in root and tuber crops such as sweetpotato (~42.0 SPAD units; Mbinda et al., 2018; Mbinda et al., 2019), cassava (47 SPAD units; Ogaddee and Girdthai, 2019), and potato (49 SPAD units; Rolando et al., 2015 and ~45 SPAD units; Ramírez et al., 2014). Also, the lower dT_{min} average value obtained here (−12.2 ± 0.7°C, for primed CIP 460531), was much lower than minimum dT values reported in potato (about −6°C; Stark et al., 1991) or alfalfa (approximately −10°C; Idso et al., 1981) under no-stress

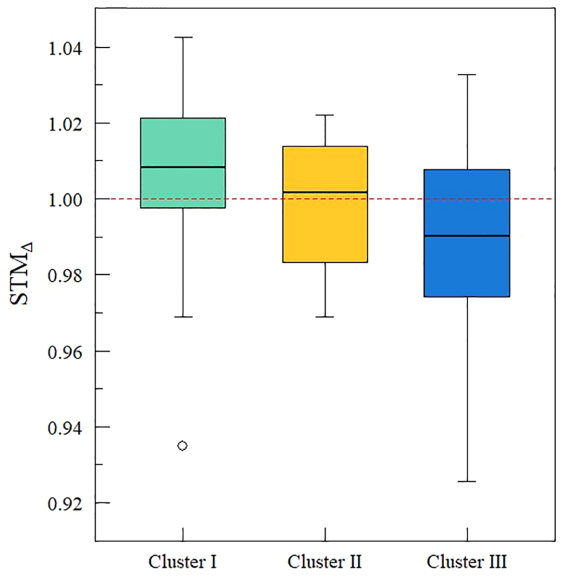

FIGURE 5 | Boxplot of short-term memory (STM) effect on leaf ^{13}C discrimination (STM_Δ) for each cluster. Red dashed line indicates the STM occurrence ($STM_\Delta > 1$). In each cluster's boxplot, black line represents the median value. Boxplot contains the variation between 25 and 75% and gray circles are the outlier values.

conditions. Given that the minimum value was achieved by primed plants in the recovery period of the third water restriction period, these results suggest an optimization in stomatal behavior as a result of the priming process. The highest Δ_{max} average value (25.3‰ for primed CIP 460583) obtained in this study in the third water restriction period, was higher than previously reported for sweetpotato (21.7 ‰, Zhang et al., 2015 and 23.6‰, Ramírez et al., 2017) or potato (23‰, Ramírez et al., 2015b).

Studies of stress memory response in *A. thaliana* (Ling et al., 2018; Serrano et al., 2019) found an improvement in heat-stress tolerance after a second priming event. The ecophysiological indicators (S, FA_{max}, dT_{min}, and Δ_{max}) assessed in our study indicated an improvement in primed plants after the first water restriction period. Delayed senescence was the most important trait increasing after water restriction events in 81.4% of the genotypes (**Figure 2A**). Delayed senescence is the result of a stress-response mechanism characterized by slower chlorophyll degradation over time, in comparison to unstressed genotypes (Thomas and Howarth, 2000; Rivero et al., 2007; Abdelrahman et al., 2017). In both potato (Rolando et al., 2015) and sweet potato (Smit, 1997; Bararyenya et al., 2020), delayed senescence has been attributed to outstanding yield, since it extends the period of time in which the plant can fix carbon. On the other hand, primed plants from 23.7% of the total number of accessions (**Figure 2B**) increased their foliar area (i.e. higher aerial biomass than non-primed plants) after the first water restriction period. These findings are corroborated by previous studies in which higher green leaf area was observed during recovery from early season drought in tolerant genotypes (Rivero et al., 2007; Puangbut et al., 2009; Lewthwaite and Triggs, 2012). Foliar area—a trait prioritized in breeding programs (Lenis et al., 2006; De Souza et al., 2017)—has been correlated to light interception and use efficiencies (Legg et al., 1979; Zhu et al., 2008; De Souza et al., 2017).

Another important trait to be considered in breeding programs that seek to improve drought tolerance is canopy temperature (Obidiegwu et al., 2015). The ability of plants to cool their leaves *via* stomata while at the same time saving water has been correlated with drought tolerance (Blum and Arkin, 1984; Blum, 2005; Hirayama et al., 2006; Ramírez et al., 2015b). The leaf-minus-air temperature (dT) has thus been used as an index of plant water status (Takai et al., 2010; Tuberosa, 2012; Iseki et al., 2018) or as an indirect method of indicating stomatal conductance (Jackson et al., 1977; Hatfield, 1983; Rinza et al., 2019). Previous studies have demonstrated that genotypes with the lowest dT values (here dT_{min}) had the highest transpiration and photosynthetic rates (Hirayama et al., 2006; Takai et al., 2010; Rukundo et al., 2017). These results are corroborated by our findings, in which primed plants from 17 accessions (**Figure 2C**) presented lower dT values in the second or third recovery period than those obtained by their respective non-primed plants throughout their lifespan. Our results suggest that a potential short-term memory induced cooling mechanism in primed plants enabled them to decrease leaf temperatures in response to hot and dry environments.

On the other hand, ^{13}C discrimination integrates the photosynthetic performance throughout the period of leaf tissue synthetization (Jefferies and Mackerron, 1997; Dawson et al., 2002). According to literature, this trait is expected to decrease in drought conditions, due to a reduced discrimination against ^{13}C by RuBisCO (Farquhar et al., 1989), and to partial recovery after re-watering (Xu et al., 2010; Ramírez et al., 2016; Silva-Díaz et al., 2020). In that respect, leaf ^{13}C discrimination from primed plants, measured in every recovery period, reflects the effects of previous water restriction periods. Memory effects were determined by calculating the maximum leaf ^{13}C discrimination (Δ_{max}) reached by primed plants and comparing it with the one achieved by non-primed plants. Primed plants from 30 accessions (**Figure 2D**) mitigated the negative effects of the first water restriction period, and even reached a higher Δ_{max}, at the end of the priming process, than non-primed plants grown under fully irrigated conditions suggesting a potential short-term memory occurrence. This improvement in photosynthetic performance has also been documented in potato (Ramírez et al., 2016; Silva-Díaz et al., 2020), wheat (Monneveux et al., 2005), and rice (Impa et al., 2005), but not on a level as high as observed in our study, induced by short-term memory in primed plants.

Of the four traits evaluated, the performance of senescence and photosynthesis were the most favorable following the priming process, which revealed their potential flexibility under stress memory induction. These results suggest that potential stress memory induction helps modify metabolic pathways related to chlorophyll degradation (Rivero et al.,

2007; Abdelrahman et al., 2017) and photosynthesis recovery (Ramírez et al., 2016). Also, despite their corroborated relation to drought tolerance, FA and dT presented a lower percentage of STM occurrence, suggesting a more complex response enhanced by stress memory. Species such as *I. leucantha*, *I. australis* and *I. grandifolia* presented even higher STM_S than sweetpotato cultivars (**Figure 3A**), leading us to recommend the consideration of these CWR species for use in breeding programs. Regarding photosynthetic performance, all SP-CWR species contained at least one accession in which STM occurrence (**Figure 3D**) suggests the potential of this group of species for easy recovery in environments with prolonged drought seasons.

A New Model for Drought Tolerance Improvement in Sweetpotato Based on Stress-Memory Induction and Wild Relatives

Resilience of a crop is the extent to which it is capable of surviving stress or other perturbations through physiological adaptations, with a minimum effect on yield (Holling, 1973; Trenbath, 1999). There is a differentiated capability of sweetpotato genotypes to allocate more carbon to above ground biomass than harvestable roots (e.g. Ejumula variety) under non-stress growth conditions (Coleman et al., 2006; Ramírez et al., 2017), however, the implications of this shifted resource allocation have not been explored yet. In our study, the capability to increase foliar area appears to be the main trait responsible for enhanced resilience (**Table 2**), an observation that is in agreement with previous studies on sorghum in which genotypes submitted to severe water restriction presented higher yields due to both increased leaf area and delayed senescence (Borrell et al., 2000). Moreover, in cassava, De Souza et al. (2017) highlight the importance of breeding for genotypes with increased leaf area to improve biomass production and drought response. Our results suggest that there is an association between delayed senescence and increased foliar area that elevated RCI response (**Figure 4**). In cereals, early-heading, which translates as a delayed senescence, is considered an escape strategy from adverse environmental conditions (Turner, 1979; Levitt, 1980; Dolferus, 2014) rather than representing a resilience response (Thiry et al., 2016). This study's results suggest that delayed senescence is the cause of increased leaf area. This resilience response includes more and longer-living foliage, which provides more time and space to fix additional CO_2 (Legg et al., 1979; Zhu et al., 2008; De Souza et al., 2017), and results in higher aerial biomass (Puangbut et al., 2009). However, the link between senescence delay and leaf biomass production is a relationship that should be explored more profoundly in the future.

Crop productivity has been associated with canopy temperature (e.g. dT) and the latter has been frequently used as a selection method for drought tolerance (Jackson et al., 1981; Jones, 2006; Takai et al., 2010; Zia et al., 2013). Some studies have shown that lower dT values are a consequence of higher stomatal conductance (Jackson et al., 1977; Hatfield, 1983; Rinza et al., 2019) and effective water uptake (Blum, 2009) representing the optimization of leaf thermoregulation and gas exchange. In our study, production capacity (PCI) was mainly related to STM_{dT} (**Table 2**) confirming the aforementioned results of other authors. The combination of dT and leaf ^{13}C discrimination were associated with elevated PCI values (**Figure 4**) suggesting that optimized leaf thermoregulation and gas exchange is a consequence of improved photosynthetic performance and recovery (Jefferies and Mackerron, 1997; Dawson et al., 2002), which ultimately leads to higher levels of productivity in SP-CWR.

Three groups of accessions performed differently in this study. Accessions from cluster I, presented higher biomass production in non-stress (Y_{npr}) than in water deficit conditions (Y_{pr}) (**Table 4**). Accessions in this cluster correspond to Group B of Fernandez's (1992) classification of plants based on biomass production (**Table 1**). Group B corresponds to Thiry et al. (2016) classification based on PCI and RCI, since accessions falling into group B also showed the highest average PCI value (4.39 ± 1.7). In cluster II, primed plants equaled the biomass production obtained from non-primed plants of the same accession (**Table 4**). This together with the highest average RCI value (7.0 ± 1.1), suggests that these accessions belong to Group C in Fernandez (1992) classification (**Table 1**). Cluster III of this study corresponds to Group D in Fernandez (1992) classification (**Table 1**) because the respective accessions had the lowest biomass production performance in both non-stress and stress conditions and the lowest PCI and RCI values (**Table 4**). Our results suggest that sweetpotato crop wild relatives possess outstanding physiological mechanisms to respond to both non-stress and water restriction scenarios. Primed SP-CWR species from cluster II (**Table 1**) such as *I. triloba* (CIP 460116) and *I. trifida* (CIP 107665.9, CIP 460663, and 113735.283) showed better biomass production performance under drought stress and hence could be considered for use in breeding programs (Thiry et al., 2016) to improve sweetpotato resilience, a trait much appreciated by breeders (Andrade et al., 2016). The aforementioned species coincide with previous studies (Iwanaga, 1988; Komaki, 2004; Zhang and Liu, 2005; Nimmakayala et al., 2011), which were based on biomass production under stress conditions but did not assess physiological performance. On the other hand, accessions from species such as *I. splendor-sylvae*, *I. ramosissima*, *I. tiliacea*, and wild *I. batatas* from cluster I could only be used as a potential genetic source of traits related to optimized leaf transpiration and photosynthetic performance.

The enhancement of drought tolerance in our study is also supported from a physiological point of view. Ramírez et al. (2015a) defined drought tolerance enhancement based on ^{13}C discrimination, when primed plants presented higher ^{13}C discrimination than non-primed plants. This physiological criterium matches our RCI-PCI drought tolerant genotypes from cluster I and II, as both of them contained the highest number of accessions with STM_{Δ} occurrence (**Figure 5**). Therefore, we have demonstrated, experimentally, the existence of drought tolerance and, physiologically, the drought tolerance enhancement induced by short-term memory.

Due to ploidy differences of the taxa and other barriers to hybridization, interspecific hybridization between sweetpotato

and related wild species is challenging and usually requires the employment of a pre-breeding approach. However, several authors have shown that the creation of interspecific hybrids is feasible using ovule culture (Kobayashi et al., 1994), somatic cell hybridization (Liu et al., 1994; Zhang et al., 2002; Yang et al., 2009), application of phytohormones in combination with controlled pollination (Cao et al., 2009), creation of synthetic hexaploids and triploids (Nishiyama et al., 1975; Shiotani and Kawase, 1987; Freyre et al., 1991), and interploidy hybridization (Orjeda et al., 1991). Successful production of hybrids between hexaploid sweetpotato and CWR species of the Batatas complex was reported for *I. trifida* (Orjeda et al., 1991; Kobayashi et al., 1994), *I. triloba* (Kobayashi et al., 1994; Liu et al., 1994; Yang et al., 2009), *I. lacunosa* (Kobayashi et al., 1994; Zhang et al., 2002), *I. grandifolia* (Cao et al., 2009), *I. littoralis* (Nishiyama et al., 1975), and *I. leucantha* (Nishiyama et al., 1975). Moreover, modern breeding techniques (e.g. CRISPR/Cas9-mediated genome editing) offer promising options for introgression of genes from wild relatives into the hexaploid cultivated genepool which need to be further explored (Wang et al., 2019).

CONCLUSION

Potential short-term memory induction constitutes a promising method to enhance physiological responses in SP-CWR. Primed accessions in this study showed physiological mechanisms (delayed senescence, increased foliar area, optimized leaf transpiration, or improved photosynthetic performance) that enable plants to cope with severe drought conditions. Because potential stress memory triggered the greatest increase in foliar area in *I. tiliacea* and some accessions from *I. ramosissima* (CIP 460032, CIP 460722) and *I. triloba* (CIP 460116), this trait might be more relevant in breeding for dual purpose sweetpotato cultivars, used to produce food as well as livestock feed. However, with a view toward long-term memory induction, more studies are required to elucidate the underlying molecular mechanisms (epigenetic processes, gene silencing, chromatin remodeling) responsible for drought tolerance improvement in sweetpotato as shown in this study. Furthermore, we showed that SP-CWR developed drought tolerance through two basic mechanisms: i) resilience, by developing more leaves with an increased time to fix carbon and ii) productivity, by optimizing leaf thermoregulation and gas exchange. The use of resilience capacity and productivity capacity, simultaneously, allowed us to easily identify genotypes from Group C, a group of plants that is highly appreciated by breeders. This study confirms the effectiveness of a potential short-term memory induction for enhancing plants' drought stress response, the potential applications of which include more efficient water use in irrigated crops and the production of more resilient sweetpotato planting material. It also sets a precedent in stress memory in SP-CWR and demonstrates that this group constitutes a potential and untapped source of valuable physiological traits for sweetpotato improvement programs. However, further field trials under different environmental conditions and in different years are necessary to confirm our preliminary results.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: [https://doi.org/10.21223/WNZR93].

AUTHOR CONTRIBUTIONS

BH and DR conceived and designed the study. FG-Z gathered the data, analyzed the results, and wrote the manuscript. FG-Z, BH, DR, JR, and JN performed the analysis. FG-Z, BH, DR, JN, JR, and RB edited the manuscript.

FUNDING

This investigation was led by North Carolina State University (NCSU), who managed the Sweetpotato CWR project in collaboration with the International Potato Center (CIP) with support from the CGIAR Research Program on Roots, Tubers, and Bananas (RTB). The study was part of the initiative "Adapting Agriculture to Climate Change: Collecting, Protecting and Preparing Crop Wild Relatives" which is supported by the Government of Norway. The initiative is managed by the Global Crop Diversity Trust with the Millennium Seed Bank of the Royal Botanic Gardens, Kew, UK, and implemented in partnership with national and international genebanks and plant breeding institutes around the world. For further information, see Dempewolf et al. (2014), and http://www.cwrdiversity.org/.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Craig Yencho, North Carolina State University (NCSU), who managed the Sweetpotato CWR project in collaboration with the International Potato Center (CIP). We also thank Carlos Flores, César Fernández, Genoveva Rossel, and Cecilia Silva for their technical support and assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.567507/full#supplementary-material

REFERENCES

Abdelrahman, M., El-Sayed, M., Jogaiah, S., Burritt, D. J., and Tran, L.-S. P. (2017). The "STAY-GREEN" trait and phytohormone signaling networks in plants under heat stress. *Plant Cell Rep.* 36, 1009–1025. doi: 10.1007/s00299-017-2119-y

Agili, S., and Nyende, B. (2012). Selection, yield evaluation, drought tolerance indices of orange-flesh Sweet potato (*Ipomoea batatas* (L.) Lam.) hybrid clone. *J. Nutr. Food Sci.* 02, 1–8. doi: 10.4172/2155-9600.1000138

Ahmed, M., Qadeer, U., Ahmed, Z.II, and Hassan, F. (2016). Improvement of wheat (*Triticum aestivum*) drought tolerance by seed priming with silicon. *Arch. Agron. Soil Sci.* 62, 299–315. doi: 10.1080/03650340.2015.1048235

Andrade, M.II, Naico, A., Ricardo, J., Eyzaguirre, R., Makunde, G. S., Ortiz, R., et al. (2016). Genotype × environment interaction and selection for drought adaptation in sweetpotato (*Ipomoea batatas* [L.] Lam.) in Mozambique. *Euphytica* 209, 261–280. doi: 10.1007/s10681-016-1684-4

Austin, D. F., Jarret, R. L., and Johnson, R. W. (1993). *Ipomoea gracilis* R. Brown (Convolvulaceae) and its allies. *Bull. Torrey Bot. Club.* 120, 49–59. doi: 10.2307/2996663

Austin, D. F. (1978). The *Ipomoea batatas* Complex-I. Taxonomy. *Bull. Torrey Bot. Club.* 105, 114–129. doi: 10.2307/2484429

Bararyenya, A., Olukolu, B. A., Tukamuhabwa, P., Grüneberg, W. J., Ekaya, W., Low, J., et al. (2020). Genome-wide association study identified candidate genes controlling continuous storage root formation and bulking in hexaploid sweetpotato. *BMC Plant Biol.* 20, 1–16. doi: 10.1186/s12870-019-2217-9

Barreda, C., Guerrero, J., and Cruz, M. (2017). *Image Canopy: Software to estimate canopy cover* (Lima, Peru: International Potato Center (CIP)). doi: 10.21223/P3/50TASS

Blum, A., and Arkin, G. F. (1984). Sorghum root growth and water-use as affected by water supply and growth duration. *Field Crops Res.* 9, 131–142. doi: 10.1016/0378-4290(84)90019-4

Blum, A. (1996). "Yield potential and drought resistance: are they mutually exclusive?," in *Increasing yield potential in wheat: breaking the barriers*. Eds. M. P. Reynolds, S. Rajaram and A. McNab (Mexico DF: CIMMYT), 90–100.

Blum, A. (2005). Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? *Aust. J. Agr. Res.* 56, 1159–1168. doi: 10.1071/AR05069

Blum, A. (2009). Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. *Field Crops Res.* 112, 119–123. doi: 10.1016/j.fcr.2009.03.009

Borrell, A. K., Hammer, G. L., and Douglas, A. C. L. (2000). Does Maintaining Green Leaf Area in Sorghum Improve Yield under Drought? I. Leaf Growth and Senescence. *Crop Sci.* 40, 1026–1037. doi: 10.2135/cropsci2000.4041026x

Bradbury, H., and Holloway, W. D. (1987). *Chemistry of tropical root crops: significance for nutrition and agriculture in the Pacific* (Canberra, Australia: ACIAR Monograph).

Bruce, T. J. A., Matthes, M. C., Napier, J. A., and Pickett, J. A. (2007). Stressful "memories" of plants: Evidence and possible mechanisms. *Plant Sci.* 173, 603–608. doi: 10.1016/j.plantsci.2007.09.002

Cao, Q., Zhang, A., Ma, D., Li, H., Li, Q., and Li, P. (2009). Novel interspecific hybridization between sweetpotato (*Ipomoea batatas* (L.) Lam.) and its two diploid wild relatives. *Euphytica* 169, 345–352. doi: 10.1007/s10681-009-9967-7

Cayón, M. G., El-Sharkawy, M. A., and Cadavid, L. F. (1997). Leaf gas exchange of cassava as affected by quality of planting material and water stress. *Phosynthetica* 34, 409–418. doi: 10.1023/A:1006868219162

Coleman, E., Maltby, J., Harper, S., O'Donnell, W., Hughes, M. M., Boreel, S., et al. (2006). *Improving sweetpotato agronomy to meet new market opportunities* (Sidney, Australia: Horticulture Australia).

Condon, A. G., Richards, R. A., Rebetzke, G. J., and Farquhar, G. D. (2004). Breeding for high water-use efficiency. *J. Exp. Bot.* 55, 2447–2460. doi: 10.1093/jxb/erh277

Conrath, U., Beckers, G. J. M., Flors, V., García-Agustín, P., Jakab, G., Mauch, F., et al. (2006). Priming: Getting Ready for Battle. *Mol. Plant Microbe Interact.* 19, 1062–1071. doi: 10.1094/MPMI-19-1062

Dawson, T. E., Mambelli, S., Plamboeck, A. H., Templer, P. H., and Tu, K. P. (2002). Stable Isotopes in Plant Ecology. *Ann. Rev. Ecol. Syst.* 33, 507–559. doi: 10.1146/annurev.ecolsys.33.020602.095451

De Souza, A. P., Massenburg, L. N., Jaiswal, D., Cheng, S., Shekar, R., and Long, S. P. (2017). Rooting for cassava: insights into photosynthesis and associated physiology as a route to improve yield potential. *New Phytol.* 213, 50–65. doi: 10.1111/nph.14250

Dempewolf, H., Eastwood, R. J., Guarino, L., Khoury, C. K., Müller, J. V., and Toll, J. (2014). Adapting Agriculture to Climate Change: A Global Initiative to Collect, Conserve, and Use Crop Wild Relatives. *Agroecol. Sust. Food* 38, 369–377. doi: 10.1080/21683565.2013.870629

Dolferus, R. (2014). To grow or not to grow: A stressful decision for plants. *Plant Sci.* 229, 247–261. doi: 10.1016/j.plantsci.2014.10.002

FAO (2020). *FAOSTAT Production Databases*. Available at: http://faostat.fao.org/ (Accessed May 14, 2020).

Farquhar, G. D., Ehleringer, J. R., and Hubick, K. T. (1989). Carbon isotope discrimination and photosynthesis. *Ann. Rev. Plant Physiol.* 40, 503–537. doi: 10.1146/annurev.pp.40.060189.002443

Fernandez, G. C. J. (1992). "Effective selection criteria for assessing plant stress tolerance," in *Proceeding of the International Symposium on Adaptation of Vegetables and other Food Crops in Temperature and Water Stress*. Ed. C. G. Kuo (Tainan, Taiwan: AVRDC), 257–270.

Fischer, R. A., and Maurer, R. (1978). Drought resistance in spring wheat cultivars. I. Grain yield responses. *Aust. J. Agric. Res.* 29, 897. doi: 10.1071/AR9780897

Freyre, R., Iwanaga, M., and Orjeda, G. (1991). Use of *Ipomoea trifida* (HBK.) G. Don germ plasm for sweet-potato improvement. 2. Fertility of synthetic hexaploids and triploids with 2 n gametes of *I. trifida*, and their interspecific crossability with sweet potato. *Genome* 34, 209–214. doi: 10.1139/g91-033

Guarino, L., and Lobell, D. B. (2011). A walk on the wild side. *Nat. Clim. Change* 1, 374–375. doi: 10.1038/nclimate1272

Guerrero-Zurita, F., Ramirez, D., Rinza, J., Ninanya, J., and Heider, B. (2020). "Replication data for: Potential Short-term Memory Induction as a Promising Method for Increasing Drought Tolerance Sweetpotato Crop Wild Relatives Ipomoea series Batatas (Choisy) Ed. D. F Austin. *Int. Potato Cent.* 1. doi: 10.21223/WNZR93

Han, S. K., and Wagner, D. (2014). Role of chromatin in water stress responses in plants. *J. Exp. Bot.* 65, 2785–2799. doi: 10.1093/jxb/ert403

Hatfield, J. L. (1983). The utilization of thermal infrared radiation measurements from grain sorghum crops as a method of assessing their irrigation requirements. *Irrig. Sci.* 3, 259–268. doi: 10.1007/BF00272841

Hirayama, M., Wada, Y., and Nemoto, H. (2006). Estimation of drought tolerance based on leaf temperature in upland rice breeding. *Breed. Sci.* 56, 47–54. doi: 10.1270/jsbbs.56.47

Holling, C. S. (1973). Resilience and Stability of Ecological Systems. *Ann. Rev. Ecol. Syst.* 4, 1–23. doi: 10.1146/annurev.es.04.110173.000245

Idso, S. B., Jackson, R. D., Pinter, P. J., Reginato, R. J., and Hatfield, J. L. (1981). Normalizing the stress-degree-day parameter for environmental variability. *Agr. Meteorol.* 24, 45–55. doi: 10.1016/0002-1571(81)90032-7

Impa, S. M., Nadaradjan, S., Boominathan, P., Shashidhar, G., Bindumadhava, H., and Sheshshayee, M. S. (2005). Carbon Isotope Discrimination Accurately Reflects Variability in WUE Measured at a Whole Plant Level in Rice. *Crop Sci.* 45, 2517–2522. doi: 10.2135/cropsci2005.0119

Iseki, K., Takahashi, Y., Muto, C., Naito, K., and Tomooka, N. (2018). Diversity of Drought Tolerance in the Genus Vigna. *Front. Plant Sci.* 9, 729. doi: 10.3389/fpls.2018.00729

Iwanaga, M. (1988). "Use of wild germplasm for sweetpotato breeding," in *Exploration, Maintaince and Utilization of Sweetpotato Genetic Resources* (Lima: International Potato Center (CIP), 199–210.

Jackson, R. D., Reginato, R. J., and Idso, S. B. (1977). Wheat canopy temperature: A practical tool for evaluating water requirements. *Water Resour. Res.* 13, 651–656. doi: 10.1029/WR013i003p00651

Jackson, R. D., Idso, S. B., Reginato, R. J., and Pinter, P. J. (1981). Canopy temperature as a crop water stress indicator. *Water Resour. Res.* 17, 1133–1138. doi: 10.1029/WR017i004p01133

Jefferies, R. A., and Mackerron, D. K. L. (1997). Carbon isotope discrimination in irrigated and droughted potato (*Solanum tuberosum* L.). *Plant Cell Env.* 20, 124–130. doi: 10.1046/j.1365-3040.1997.d01-5.x

Jones, H. G. (2006). Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance. *J. Exp. Bot.* 58, 119–130. doi: 10.1093/jxb/erl118

Khoury, C. K., Heider, B., Castañeda-Álvarez, N. P., Achicanoy, H. A., Sosa, C. C., Miller, R. E., et al. (2015). Distributions, ex situ conservation priorities, and genetic resource potential of crop wild relatives of sweetpotato [*Ipomoea batatas* (L.) Lam., *I.* series *Batatas*]. *Front. Plant Sci.* 6, 251. doi: 10.3389/fpls.2015.00251

Kivuva, B. M., Githiri, S. M., Yencho, G. C., and Sibiya, J. (2015). Screening sweetpotato genotypes for tolerance to drought stress. *Field Crops Res.* 171, 11–22. doi: 10.1016/j.fcr.2014.10.018

Knox, J., Hess, T., Daccache, A., and Wheeler, T. (2012). Climate change impacts on crop productivity in Africa and South Asia. *Environ. Res. Lett.* 7, 34032. doi: 10.1088/1748-9326/7/3/034032

Kobayashi, R. S., Bouwkamp, J. C., and Sinden, S. L. (1994). Interspecific hybrids from cross incompatible relatives of sweetpotato. *Euphytica* 80, 159–164. doi: 10.1007/BF00039646

Komaki, K. (2004). "Breeding value of wild species closely related to sweetpotato," in *Proceedings of international workshop on production, utilization and development of sweetpotato* (Korea: AIST), 164–172.

Laurie, R. N., Laurie, S. M., Du Plooy, C. P., Finnie, J. F., and Staden, J. V. (2014). Yield of Drought-tressed Sweet Potato in Relation to Canopy Cover, Stem Length and Stomatal Conductance. *J. Agric. Sci.* 7, 201–214. doi: 10.5539/jas.v7n1p201

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Software* 25 (1), 1–18. doi: 10.18637/jss.v025.i01

Legg, B. J., Day, W., Lawlor, D. W., and Parkinson, K. J. (1979). The effects of drought on barley growth: models and measurements showing the relative importance of leaf area and photosynthetic rate. *J. Agr. Sci.* 92, 703–716. doi: 10.1017/S0021859600053958

Lenis, J.II, Calle, F., Jaramillo, G., Perez, J. C., Ceballos, H., and Cock, J. H. (2006). Leaf retention and cassava productivity. *Field Crops Res.* 95, 126–134. doi: 10.1016/j.fcr.2005.02.007

Levitt, J. (1980). *Responses of plants to environmental stresses. 2d ed* (New York: Academic Press).

Lewthwaite, S. L., and Triggs, C. M. (2012). Sweetpotato cultivar response to prolonged drought. *Agr. New Zeal.* 42, 1–10.

Li, X., Ramírez, D. A., Qin, J., Dormatey, R., Bi, Z., Sun, C., et al. (2019). Water restriction scenarios and their effects on traits in potato with different degrees of drought tolerance. *Sci. Hortic.* 256, 108525. doi: 10.1016/j.scienta.2019.05.052

Ling, Y., Serrano, N., Gao, G., Atia, M., Mokhtar, M., Woo, Y. H., et al. (2018). Thermopriming triggers splicing memory in Arabidopsis. *J. Exp. Bot.* 69, 2659–2675. doi: 10.1093/jxb/ery062

Liu, H., and Charng, Y. (2012). Acquired thermotolerance independent of heat shock factor A1 (HsfA1), the master regulator of the heat stress response. *Plant Signal. Behav.* 7, 547–550. doi: 10.4161/psb.19803

Liu, Q. C., Wang, J. X., Li, W. J., and Zhou, H. Y. (1994). Protoplast fusion and regeneration of interspecific somatic hybrid plants between sweet potato (*Ipomoea batatas* (L.) Lam.) and its related species. *J. Agric. Biotechnol.* 2, 85–90.

Low, J., Lynam, J., Lemaga, B., Crissman, C., Barker, I., Thiele, G., et al. (2009). "Sweetpotato in Sub-Saharan Africa," in *The Sweetpotato*. Eds. G. Loebenstein and G. Thottappilly (Dordrecht: Springer Netherlands), 359–390. doi: 10.1007/978-1-4020-9475-0_16

Low, J. W., Ortiz, R., Vandamme, E., Andrade, M., Biazin, B., and Grüneberg, W. J. (2020). Nutrient-Dense Orange-Fleshed Sweetpotato: Advances in Drought-Tolerance Breeding and Understanding of Management Practices for Sustainable Next-Generation Cropping Systems in Sub-Saharan Africa. *Front. Sustain. Food Syst.* 4, 50. doi: 10.3389/fsufs.2020.00050

Mbinda, W., Ombori, O., Dixelius, C., and Oduor, R. (2018). *Xerophyta viscosa* Aldose Reductase, XvAld1, Enhances Drought Tolerance in Transgenic Sweetpotato. *Mol. Biotechnol.* 60, 203–214. doi: 10.1007/s12033-018-0063-x

Mbinda, W., Dixelius, C., and Oduor, R. (2019). Induced Expression of *Xerophyta viscosa* XvSap1 Gene Enhances Drought Tolerance in Transgenic Sweet Potato. *Front. Plant Sci.* 10, 1119. doi: 10.3389/fpls.2019.01119

McDonald, J. A., and Austin, D. F. (1990). Changes and Additions in *Ipomoea* Section *Batatas* (Convolvulaceae). *Brittonia* 42, 116. doi: 10.2307/2807625

Medrano, H., Escalona, J. M., Bota, J., Gulías, J., and Flexas, J. (2002). Regulation of photosynthesis of C3 plants in response to progressive drought: stomatal conductance as a reference parameter. *Ann. Bot.* 89, 895–905. doi: 10.1093/aob/mcf079

Mekonnen, B., Tulu, S., and Nego, J. (2015). Orange Fleshed Sweet Potato (*Ipomoea batatas* (L.) Lam.) Varieties Evaluated with Respect to Growth Parameters at Jimma in Southwestern Ethiopia. *J. Agron.* 14, 164–169. doi: 10.3923/ja.2015.164.169

Monneveux, P., Reynolds, M. P., Trethowan, R., González-Santoyo, H., Peña, R. J., and Zapata, F. (2005). Relationship between grain yield and carbon isotope discrimination in bread wheat under four water regimes. *Eur. J. Agron.* 22, 231–242. doi: 10.1016/j.eja.2004.03.001

Motsa, N. M., Modi, A. T., Mabhaudhi, T., and Mabhaudhi, T. (2015). Sweet potato (*Ipomoea batatas* (L.) Lam.) as a drought tolerant and food security crop. *South Afr. J. Sci.* 111, 1–8. doi: 10.17159/sajs.2015/20140252

Mwanga, R. O. M., and Ssemakula, G. (2011). Orange-fleshed sweetpotatoes for food, health and wealth in Uganda. *Int. J. Agr. Sustain.* 9, 42–49. doi: 10.3763/ijas.2010.0546

Nimmakayala, P., Vajja, G., and Reddy, U. K. (2011). "Ipomoea," in *Wild Crop Relatives: Genomic and Breeding Resources*. Ed. C. Kole (Berlin, Heidelberg: Springer), 123–132. doi: 10.1007/978-3-642-21102-7_7

Nishiyama, I., Miyazaki, T., and Sakamoto, S. (1975). Evolutionary autoploidy in the sweet potato (*Ipomoea batatas* (L.) Lam.) and its progenitors. *Euphytica* 24, 197–208. doi: 10.1007/BF00147186

Obidiegwu, J. E., Bryan, G. J., Jones, H. G., and Prashar, A. (2015). Coping with drought: stress and adaptive responses in potato and perspectives for improvement. *Front. Plant Sci.* 6, 542. doi: 10.3389/fpls.2015.00542

Ogaddee, P., and Girdthai, T. (2019). Physiological and ethylene accumulation responses of cassava under drought stress. *IOP Conf. Ser. Earth Environ. Sci.* 346, 12092. doi: 10.1088/1755-1315/346/1/012092

Orjeda, G., Freyre, R., and Iwanaga, M. (1991). Use of *Ipomoea trifida* germ plasm for sweet potato improvement. 3. Development of 4x interspecific hybrids between *Ipomoea batatas* (L.) Lam. (2n=6x=90) and *I. trifida* (H.B.K) G. Don. (2n=2x=30) as storage-root initiators for wild species. *Theoret. Appl. Genet.* 83, 159–163. doi: 10.1007/BF00226245

Ozias-Akins, P., and Jarret, R. L. (1994). Nuclear DNA Content and Ploidy Levels in the Genus Ipomoea. *J. Am. Soc Hortic. Sci.* 119, 110–115. doi: 10.21273/JASHS.119.1.110

Production Systems and Environment (PSE) and International Potato Center (CIP) (2013). Protocol for Designing and Conducting Potato Field Experiments for Modeling Purposes. (Lima, Peru: CIP), 1–16. doi: 10.4160/9789290604303

Puangbut, D., Jogloy, S., Vorasoot, N., Akkasaeng, C., Kesmalac, T., and Patanothai, A. (2009). Variability in yield responses of peanut (*Arachis hypogaea* L.) genotypes under early season drought. *Asian J. Plant Sci.* 8, 254–264. doi: 10.3923/ajps.2009.254.264

R Core Team (2019). *A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/

Ramírez, D. A., Yactayo, W., Gutiérrez, R., Mares, V., De Mendiburu, F., Posadas, A., et al. (2014). Chlorophyll concentration in leaves is an indicator of potato tuber yield in water-shortage conditions. *Sci. Hortic.* 168, 202–209. doi: 10.1016/j.scienta.2014.01.036

Ramírez, D. A., Rolando, J. L., Yactayo, W., Monneveux, P., Mares, V., and Quiroz, R. (2015a). Improving potato drought tolerance through the induction of long-term water stress memory. *Plant Sci.* 238, 26–32. doi: 10.1016/j.plantsci.2015.05.016

Ramírez, D. A., Rolando, J. L., Yactayo, W., Monneveux, P., and Quiroz, R. (2015b). Is Discrimination of ^{13}C in potato leaflets and tubers an appropriate trait to describe genotype responses to restrictive and well-watered conditions? *J. Agron. Crop Sci.* 201, 410–418. doi: 10.1111/jac.12119

Ramírez, D. A., Yactayo, W., Rens, L. R., Rolando, J. L., Palacios, S., De Mendiburu, F., et al. (2016). Defining biological thresholds associated to plant water status for monitoring water restriction effects: Stomatal conductance and photosynthesis recovery as key indicators in potato. *Agr. Water Manage.* 177, 369–378. doi: 10.1016/j.agwat.2016.08.028

Ramírez, D. A., Gavilán, C., Barreda, C., Condori, B., Rossel, G., Mwanga, R. O. M., et al. (2017). Characterizing the diversity of sweetpotato through growth parameters and leaf traits: Precocity and light use efficiency as important ordination factors. *S. Afr. J. Bot.* 113, 192–199. doi: 10.1016/j.sajb.2017.08.009

Rinza, J., Ramírez, D. A., García, J., de Mendiburu, F., Yactayo, W., Barreda, C., et al. (2019). Infrared Radiometry as a Tool for Early Water Deficit Detection:

Insights into Its Use for Establishing Irrigation Calendars for Potatoes Under Humid Conditions. *Potato Res.* 62, 109–122. doi: 10.1007/s11540-018-9400-5

Rivero, R. M., Kojima, M., Gepstein, A., Sakakibara, H., Mittler, R., Gepstein, S., et al. (2007). Delayed leaf senescence induces extreme drought tolerance in a flowering plant. *Proc. Natl. Acad. Sci.* 104, 19631–19636. doi: 10.1073/pnas.0709453104

Rolando, J. L., Ramírez, D. A., Yactayo, W., Monneveux, P., and Quiroz, R. (2015). Leaf greenness as a drought tolerance related trait in potato (*Solanum tuberosum* L.). *Environ. Exp. Bot.* 110, 27–35. doi: 10.1016/j.envexpbot.2014.09.006

Rukundo, P., Shimelis, H., Laing, M., and Gahakwa, D. (2017). Combining Ability, Maternal Effects, and Heritability of Drought Tolerance, Yield and Yield Components in Sweetpotato. *Front. Plant Sci.* 7, 1981. doi: 10.3389/fpls.2016.01981

Schulze, R. (2011). Approaches towards practical adaptive management options for selected water-related sectors in South Africa in a context of climate change. *Water SA* 37, 621–646. doi: 10.4314/wsa.v37i5.1

SENAMHI (2020). *Mapa Climático del Perú*. Available at: https://senamhi.gob.pe/?&p=mapa-climatico-del-peru (Accessed January 22, 2020).

Serrano, N., Ling, Y., Bahieldin, A., and Mahfouz, M. M. (2019). Thermopriming reprograms metabolic homeostasis to confer heat tolerance. *Sci. Rep.* 9, 1–14. doi: 10.1038/s41598-018-36484-z

Sharma, S., Upadhyaya, H. D., Varshney, R. K., and Gowda, C. L. L. (2013). Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Front. Plant Sci.* 4, 309. doi: 10.3389/fpls.2013.00309

Shiotani, I., and Kawase, T. (1987). Synthetic hexaploids derived from wild species related to sweet potato. *Jpn. J. Breed.* 37, 367–376. doi: 10.1270/jsbbs1951.37.367

Silva-Díaz, C., Ramírez, D. A., Rodríguez-Delfín, A., de Mendiburu, F., Rinza, J., Ninanya, J., et al. (2020). Unraveling Ecophysiological Mechanisms in Potatoes under Different Irrigation Methods: A Preliminary Field Evaluation. *Agron* 10, 827. doi: 10.3390/agronomy10060827

Smit, N. E. J. M. (1997). The effect of the indigenous cultural practices of in-ground storage and piecemeal harvesting of sweetpotato on yield and quality losses caused by sweetpotato weevil in Uganda. *Agr. Ecosyst. Environ.* 64, 191–200. doi: 10.1016/S0167-8809(97)00022-4

Stark, J. C., Pavek, J. J., and McCann, I. R. (1991). Using Canopy Temperature Measurements to Evaluate Drought Tolerance of Potato Genotypes. *J. Am. Soc Hortic. Sci.* 116, 412–415. doi: 10.21273/JASHS.116.3.412

Takai, T., Yano, M., and Yamamoto, T. (2010). Canopy temperature on clear and cloudy days can be used to estimate varietal differences in stomatal conductance in rice. *Field Crops Res.* 115, 165–170. doi: 10.1016/j.fcr.2009.10.019

Thiry, A. A., Chavez Dulanto, P. N., Reynolds, M. P., and Davies, W. J. (2016). How can we improve crop genotypes to increase stress resilience and productivity in a future climate? A new crop screening method based on productivity and resistance to abiotic stress. *J. Exp. Bot.* 67, 5593–5603. doi: 10.1093/jxb/erw330

Thomas, H., and Howarth, C. J. (2000). Five ways to stay green. *J. Exp. Bot.* 51, 329–337. doi: 10.1093/jexbot/51.suppl_1.329

Trenbath, B. R. (1999). Multispecies cropping systems in India: Predictions of their productivity, stability, resilience and ecological sustainability. *Agrofor. Syst.* 45, 81–107. doi: 10.1023/A:1006285319817

Tuberosa, R. (2012). Phenotyping for drought tolerance of crops in the genomics era. *Front. Physiol.* 3, 347. doi: 10.3389/fphys.2012.00347

Turner, N. (1979). "Drought resistance and adaptation to water deficits in crop plants," in *Stress physiology in crop plants*. Eds. H. Mussell and R. C. Staples (New York: John Wiley and Sons), 343–372.

Turner, N. (1986). Adaptation to Water Deficits: a Changing Perspective. *Funct. Plant Biol.* 13, 175. doi: 10.1071/PP9860175

Turner, N. (1997). "Further Progress in Crop Water Relations," in *Advances in Agronomy*. (Western Australia, Australia: Elsevier), 293–338. doi: 10.1016/S0065-2113(08)60258-8

Valverde, R. A., Clark, C. A., and Valkonen, J. P. T. (2007). Viruses and virus disease complexes of sweet potato. *Plant Viruses* 1, 116–126.

Walter, J., Nagy, L., Hein, R., Rascher, U., Beierkuhnlein, C., Willner, E., et al. (2011). Do plants remember drought? Hints towards a drought-memory in grasses. *Environ. Exp. Bot.* 71, 34–40. doi: 10.1016/j.envexpbot.2010.10.020

Wang, H., Wu, Y., Zhang, Y., Yang, J., Fan, W., Zhang, H., et al. (2019). CRISPR/Cas9-Based Mutagenesis of Starch Biosynthetic Genes in Sweet Potato (*Ipomoea Batatas*) for the Improvement of Starch Quality. *Int. J. Mol. Sci.* 20, 4702. doi: 10.3390/ijms20194702

Watkinson, J. I., Hendricks, L., Sioson, A. A., Vasquez-Robinet, C., Stromberg, V., Heath, L. S., et al. (2006). Accessions of *Solanum tuberosum* ssp. andigena show differences in photosynthetic recovery after drought stress as reflected in gene expression profiles. *Plant Sci.* 171, 745–758. doi: 10.1016/j.plantsci.2006.07.010

Wood, J. R. I., Carine, M. A., Harris, D., Wilkin, P., Williams, B., and Scotland, R. W. (2015). *Ipomoea* (Convolvulaceae) in Bolivia. *Kew Bull.* 70, 1–124. doi: 10.1007/s12225-015-9592-7

Wood, J. R. I., Muñoz-Rodríguez, P., Williams, B. R. M., and Scotland, R. W. (2020). A foundation monograph of *Ipomoea* (Convolvulaceae) in the New World. *PhytoKeys* 143, 1–823. doi: 10.3897/phytokeys.143.32821

Woolfe, J. A. (1992). *Sweet potato: an untapped food resource* (Cambridge, England: Cambridge University Press).

Wu, S., Lau, K. H., Cao, Q., Hamilton, J. P., Sun, H., Zhou, C., et al. (2018). Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement. *Nat. Commun.* 9, 1–12. doi: 10.1038/s41467-018-06983-8

Xu, Z., Zhou, G., and Shimizu, H. (2010). Plant responses to drought and rewatering. *Plant Signal. Behav.* 5, 649–654. doi: 10.4161/psb.5.6.11398

Yang, Y. F., Guan, S. K., Zhai, H., He, S. Z., and Liu, Q. C. (2009). Development and evaluation of a storage root-bearing sweetpotato somatic hybrid between *Ipomoea batatas* (L.) Lam. and *I. triloba* L. *Plant Cell Tiss Organ Cult.* 99, 83–89. doi: 10.1007/s11240-009-9578-y

Zhang, S. S., and Liu, L. F. (2005). "Utilization of the wild relatives in sweetpotato breeding," in *Sweetpotato Breeding and Industrialization in China*. Ed. Q. C. Liu (Beijing: China Agricultural University Press), 29–32.

Zhang, B. Y., Liu, Q. C., Zhai, H., and Zhou, H. Y. (2002). Production of fertile interspecific somatic hybrid plants between sweetpotato and it wild relative, *Ipomoea lacunosa*. *Acta Hortic.* 583, 81–85. doi: 10.17660/ActaHortic.2002.583.8

Zhang, H., Zhu, L., Zhang, C., Ning, Y., and Zhang, Y. (2015). Effect of water stress on dry mass accumulation and allocation in sweet potato based on stable isotope analysis. *Can. J. Plant Sci.* 95, 263–271. doi: 10.4141/cjps-2014-170

Zhu, X. G., Long, S. P., and Ort, D. R. (2008). What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Curr. Opin. Biotech.* 19, 153–159. doi: 10.1016/j.copbio.2008.02.004

Zia, S., Romano, G., Spreer, W., Sanchez, C., Cairns, J., Araus, J. L., et al. (2013). Infrared Thermal Imaging as a Rapid Tool for Identifying Water-Stress Tolerant Maize Genotypes of Different Phenology. *J. Agron. Crop Sci.* 199, 75–84. doi: 10.1111/j.1439-037X.2012.00537.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 16 September 2020
doi: 10.3389/fpls.2020.567425

Introgression Breeding in Cowpea [*Vigna unguiculata* (L.) Walp.]

*Ousmane Boukar[1], Michael Abberton[2], Olaniyi Oyatomi[2], Abou Togola[1], Leena Tripathi[3] and Christian Fatokun[2]**

[1] Cowpea Breeding Unit, International Institute of Tropical Agriculture, Kano, Nigeria, [2] Genetic Resources Center, International Institute of Tropical Agriculture, Ibadan, Nigeria, [3] Biosciences, International Institute of Tropical Agriculture, Nairobi, Kenya

The narrow base of genetic diversity characteristic of cowpea can be attributed to it being self-pollinating, evolving from narrow wild germplasm and exhibiting very limited gene flow between wild and cultivated types. Backcrossing to introduce simply inherited desirable traits and utilization of improved breeding lines and varieties as parents in crossing programs further narrowed the genetic base of cowpea varieties. In most cowpea breeding programs, genes for resistance and market traits were pyramided into lines characterized by high levels of acceptance to farmers and consumers. Besides predisposing widely distributed improved varieties to genetic vulnerability, a narrow base of genetic variation may be contributing to the plateauing in cowpea grain yield, which compromises genetic gains. Cross compatible wild relatives have not been used in variety development because breeders shy away from them due to their tiny seed size, unattractive seed coat color and texture, pod shattering, and susceptibility to viruses. A number of wild cowpea relatives, both within and outside section *Catiang* of *Vigna* species, have been evaluated for their reaction to cowpea insect pests and diseases. *Vigna vexillata* lines were resistant to the legume pod borer (*Maruca vitrata*), the cowpea weevil (*Callosobruchus maculatus*), and *Striga gesnerioides* but are cross incompatible with cultivated cowpea. Some lines among the cross compatible wild relative *V. unguiculata* ssp. *dekindtiana* were found to be resistant to aphid in the seedling stage, while others showed good levels of drought and heat tolerance. Molecular markers are being generated to identify quantitative trait loci (QTL) with effects on some desirable attributes in cowpea. Modern breeding tools, including transgenics, can be applied for the improvement of cowpea, bypassing the natural barriers of traditional breeding. Transgenic cowpea with *Bt* gene *cry1Ab* showing resistance to *M. vitrata* has been released in Nigeria. Genome editing, a powerful emerging tool, can also be used for developing improved cowpea varieties with durable resistance to pests and diseases.

Keywords: cowpea, *Vigna unguiculata*, crop wild relatives, introgression, genetic diversity, genomics, new plant breeding techniques

OPEN ACCESS

Edited by:
Mariola Plazas, Polytechnic University of Valencia, Spain

Reviewed by:
Ira Ashri Herniter, University of California, Riverside, United States
Steven B. Cannon, United States Department of Agriculture, United States

***Correspondence:**
Christian Fatokun
c.fatokun@cgiar.org

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *29 May 2020*
Accepted: *31 August 2020*
Published: *16 September 2020*

Citation:
Boukar O, Abberton M, Oyatomi O, Togola A, Tripathi L and Fatokun C (2020) Introgression Breeding in Cowpea [Vigna unguiculata (L.) Walp.]. Front. Plant Sci. 11:567425. doi: 10.3389/fpls.2020.567425

INTRODUCTION

Cowpea, also known as black-eyed pea, belongs to section *Catiang* (DC) Verdc. genus *Vigna*, tribe *Phaseoleae* in Family *Fabaceae* (Maréchal et al., 1978). It is a self-pollinating diploid with chromosome number 2n = 22 and a genome size of about 613 (Arumuganathan and Earle, 1991) to 640.6 Mb (Lonardi et al., 2019). It is cultivated worldwide especially in Africa, Central and South America, Asia, Oceania, Southern Europe, and USA (Quin, 1997) while most is produced in the dry savannah regions of sub-Saharan Africa (SSA) in companion with mainly sorghum and millet in the same fields (Steele, 1972). Only a small proportion of SSA farmers grow cowpea as a sole crop (Steiner, 1982). In comparison to many other crops, cowpea is more adapted to drought stress and even performs relatively better in depauperized soils that are characteristic of the agro-ecologies where the crop is most extensively grown in SSA (Mortimore et al., 1997). In addition, being a legume, it has the capacity of fixing atmospheric nitrogen, some of which it utilizes for growth and development, while some do remain in the soil for the benefit of following crops (Quin, 1997). The protein rich grains are commonly eaten across the regions in different food dishes while the leaves are consumed as a pot herb, especially in East Africa. Farmers in the dry savannah areas feed their livestock with cowpea haulm, which has high nutritional value. Many efforts were devoted previously to developing cowpea varieties with high grain yield, while in recent times attention is being focused on developing dual purpose varieties with both high grain and fodder yields. It was reported that in four years, 75% of farmers in Kano State, Northern Nigeria adopted a dual-purpose cowpea variety (IT89KD-288) due to its additional fodder yield (Inaizumi et al., 1999). Many of the farmers' traditional varieties (local varieties) show attributes similar to dual purpose types, but the former are generally more adapted to intercropping. Progress can be made through breeding to increase both grain and fodder yields simultaneously because grain and fodder yields in cowpea tend to be positively correlated (Samireddypalle et al., 2017). Highly significant variations were also detected among cowpea lines for measured livestock nutrition traits such as nitrogen (N), fibre fractions, *invitro* digestibility, and metabolizable energy content. The authors also reported the absence of any trade-offs between grain yield and haulm fodder quality traits. Interestingly, haulm acid detergent lignin, and grain yield were observed to be inversely correlated, suggesting that high grain yielding varieties had lower haulm lignin content.

The center of diversity for cultivated cowpea is reported to be in West and Central Africa while that for wild relatives is in southern parts of Africa. Wild cowpea relatives are mostly distributed from Namibia through Botswana, Zambia, Zimbabwe, Mozambique, South Africa, and Swaziland (Padulosi and Ng, 1997). These authors based their suggestion on primitive characteristics such as perenniality, small seed size, pod shattering, hairiness of plant parts, distinct exine on pollen surface and outcrossing among other traits that are associated with wild relatives. They further suggested that the Transvaal Region of South Africa represents the center of speciation of *Vigna unguiculata* since the most primitive forms of the wild relatives, especially varieties *rhomboidea*, *protracta*, *tenuis*, and *stenophylla*, all in section *Catiang* are mostly found there. The variety *rhomboidea* has, the narrowest range of geographical distribution, mainly in the area from 20°S to 27°S and 26°E to 32°E and an isolated presence around Cape Town in South Africa. The wild relatives of cowpea include *Vigna unguiculata* subsp. *dekindtiana*, *V. unguiculata* subsp. *stenophylla*, *V. unguiculata* subsp. *tenuis*, *V. nervosa*, *V. vexillata*, *V. oblongifolia*, *V. frutescens*, *V. reticulata*, *V. luteola*, *V. pygmaea*, *V. gazensis*, and *V. nuda* (Padulosi and Ng, 1990). Samples of each of these wild cowpea relatives were collected from different parts of Zimbabwe. In addition, samples of *V. kirkii*, *V. platyloba*, and *V. wittei* were collected from Tanzania, further showing that the center of diversity of wild cowpea is in the southern parts of Africa. Therefore, collections made from this sub-region should, to a large extent, represent the most significant diversity among wild cowpea relatives.

The center of domestication for cultivated cowpea is still to be agreed among taxonomists. Some suggested that the area of cultivated cowpea domestication is located from Senegal in West Africa to Eritrea in the East (Kouam et al., 2012). Following a single-nucleotide polymorphism (SNP) marker analysis carried out on 1,200 cowpea lines, two gene pools were identified, one each in West Africa and in East Africa (Huynh et al., 2013). The authors concluded that cowpea has dual domestication. Another molecular marker diversity study also confirmed West and Central Africa as the region of cowpea domestication (Xiong et al., 2016). The authors suggested that the yard-long-bean, *V. unguiculata* ssp. *sesquipedalis* and *V. unguiculata* ssp. *biflora* evolved in India and south East Asia after cowpea was introduced into the region from Africa. The domestication of cowpea experienced a double bottleneck (Pasquet, 1996). The first was from its wild progenitor, which resulted in primitive cultivar group cv.-gr. Biflora and cv.-gr. Textilis and the second from the primitive cultivar-group to the evolved cultivar-groups cv.-gr. Melanophthalmus in West Africa and cv.-gr. Sesquipedalis in Asia. However, the generally accepted immediate progenitor of cowpea is *Vigna unguiculata* ssp. *dekindtiana*, which is widely distributed across Africa (Padulosi and Ng, 1997). This wide distribution of the immediate progenitor across the region may contribute to the lack of clarity as to the exact location of the domestication of cultivated cowpea. The oldest evidence so far is from archaeological excavations made in Ghana, in the West Africa sub-region, which suggest that cowpea domestication took place before 1500 BC (D'Andrea et al., 2007).

Despite the availability of a fairly large number of cowpea germplasm being conserved in some gene banks, the genetic base of the crop remains narrow. An evaluation of genetic diversity among improved cowpea varieties and breeding lines obtained from IITA breeding nursery using simple sequence repeat (SSR) markers revealed that improved cowpea varieties, in general, have a narrow genetic base (Li et al., 2001). This can be attributed to breeders' consistently using improved elite lines as parents in

crosses to generate segregating populations in their programs. In addition, the backcross method of breeding is often used to introduce simply inherited traits to varieties as a means of correcting some of the weaknesses that may be present in existing varieties. Being a highly self-pollinated crop evolved from a single wild progenitor also contributed to its narrow genetic base. In order to broaden cowpea's genetic base, it is necessary to utilize alien germplasm, especially from among cross compatible wild relatives.

Genetically modified (GM) cowpea is being developed at some research stations in SSA (ACB, 2015). The currently available GM cowpea variety carries an insecticidal *Cry1Ab* gene encoding a *Bacillus thuringiensis* (*Bt*) toxin. Ghana, Malawi, Burkina Faso, and Nigeria are the countries where confined field evaluations of transgenic *Bt* cowpea lines have been performed (Togola et al., 2017). Genetic engineering was undertaken for the development of cowpea with resistance to the legume pod borer *Maruca vitrata* Fabricius (Lepidoptera: Crambidae), the most damaging and economically important post-flowering insect pest of cowpea in SSA, following the inability to successfully cross cowpea with those of its wild relatives that exhibit resistance to this pest (Togola et al., 2017). The first successful report of an insertion of a transgene into cowpea through transgenesis involved the *Cry1Ab Bt* gene (Popelka et al., 2006). The *Bt* gene has now been transferred through backcrossing to some improved and released cowpea varieties in four SSA countries. The transgenic *Bt* cowpea lines, when evaluated in the field, showed high levels of resistance to *Maruca* as the larvae which are responsible for causing damages on cowpea flowers, pods, seeds, and young shoots were killed following their feeding on the plant (Mohammed et al., 2014). One major limitation of the *Bt* genes is their poor expression in higher eukaryotes (Bett et al., 2017), and a second limitation is their selective properties that target mostly Lepidopteran species. This implies that cowpea farmers who adopt these transgenic varieties will still need to protect their crops against other insect pests that are not controlled by *Bt* gene.

The first GM insect resistant cowpea variety [*SAMPEA 20-T*, Pod Borer Resistant (PBR) Cowpea] has recently been approved for commercialization in Nigeria (Crop Biotech, 2019). This commercialized variety has protection against *Maruca vitrata*, a pest that can cause grain yield losses of up to 60% (Singh et al., 1990). The new variety was developed by scientists at the Institute for Agricultural Research (IAR), Ahmadu Bello University, Zaria Nigeria, in collaboration with several partners under the coordination of the African Agricultural Technology Foundation (AATF) (Mohammed et al., 2014). Although the PBR cowpea confers resistance to *M. vitrata*, it is still susceptible to other insect pests, which are capable of causing major problems to cowpea cultivation. The crop is attacked at every stage in its life cycle by different species of insects, each capable of causing significant grain yield losses. The most damaging insect pests of cowpea, in addition to the pod borer are the cowpea aphid (*Aphis craccivora* Koch), flower bud thrips (*Megalurothrips sjostedti* Trybom), and pod sucking bugs (*Clavigralla tomentosicollis* Stål, *Riptortus dentipes* Fabricius) and storage pests such as bruchid (*Callosobruchus maculatus* Fabricius and *Bruchidius atrolineatus* Pic) (Jackai and Singh, 1988). These pests have devastating effects on cowpea production in the field and seeds in storage. Insect pest control in cowpea in SSA farmers' fields remains abysmally low due mainly to high costs of synthetic insecticides that are imported. This has continued to depress the productivity of improved cowpea varieties being released in several countries. Therefore, there is a need to generate broad-spectrum resistance to these pests by stacking pest resistance genes through application of the transgenic approach or editing of the host plant genome. In this paper, we review studies carried out to identify wild cowpea relatives with traits such as resistance to insect pests, *Striga*, drought tolerance and high nutritional quality that could be exploited in the development of new improved cowpea varieties and some other potential traits that could be introgressed into cowpea as well as efforts made to introgress desirable traits from wild to improved cultivars.

EXPLORING THE DIVERSITY AND POTENTIAL OF CWRS FOR INTROGRESSION BREEDING

The value of conserved genetic resources depends largely on the uniqueness of the samples in the collection and the extent of the diversity captured (Upadhyaya et al., 2006). The wild relatives of cowpea represent valuable genepools yet to be tapped for cowpea variety development. Wild relatives of cowpea contain genotypes that independently evolved within specific environments and are potentially crucial in cowpea genetic improvement in the context of climate change. The success of a plant-breeding program depends largely on the availability of a large germplasm collection, representing the genetic diversity of the crop species and the knowledge of important morphological and agronomic features of the accessions in the collection. For a long-term crop improvement program, a large and diverse germplasm collection is an invaluable source of supply of parental strains for hybridization and subsequent development of improved varieties (Chheda and Fatokun, 1982). Since its establishment in 1967, the International Institute of Tropical Agriculture (IITA) has accumulated a large collection of cowpea germplasm exceeding 15,000 accessions of cultivated varieties drawn from over 100 countries and over 1,500 accessions of wild *Vigna* species. This collection contains cowpea genotypes that display variations in many traits such as plant pigmentation, plant type, plant height, leaf type, growth habit, photo-sensitivity or insensitivity, maturity, nitrogen fixation, fodder quality, heat and drought tolerance, root architecture, pod traits, seed traits, grain quality and resistance to major diseases, root-knot nematodes, insect pests (aphids, bruchid, thrips), and parasitic weeds (*Striga* and *Alectra*). Wild forms and closely related species of cowpea, therefore, could have potential as additional sources of useful genes for cowpea improvement (Baudoin and Marechal, 1985; Padulosi and Ng, 1990).

Characterization of germplasm accessions was largely concerned with the description of accession composition and

morphological characteristics or phenotypic expressions (Longwe, 1996). This characterization process involves recording those characters, which are heritable and visible by observation as expressed in any environment. This type of characterization is an account of the plant morphology, either throughout its life cycle or only at maturity. In more recent times and with the advent of molecular markers, it has become a common practice to use markers for characterization of crops' germplasms. These markers are available in abundant numbers, neutral to environmental influence and, therefore, more robust in discriminating among germplasm lines. The molecular markers that have been used in cowpea characterization include restriction fragment length polymorphism (Fatokun et al., 1993), amplified fragment length polymorphism (Coulibaly et al., 2002), random amplified polymorphic DNA (Ba et al., 2004), SSRs (Li et al., 2001), inter SSR analysis (Ajibade et al., 2000), gene derived markers (Wang et al., 2008), and SNPs (Xiong et al., 2016).

Studies reporting the use of wild cowpea relatives in the development of improved varieties are rare to come by, if any. Breeders have shied away from utilizing wild relatives in cowpea variety development because of undesirable attributes. Generally, wild cowpea relatives have small seed size, unattractive seed coat color and texture, tendency to be susceptible to several virus diseases, pod shattering, weedy, and indeterminate growth habit (Rawal et al., 1976) among the cross compatible ones. With the genomic tools being developed for cowpea, breeders should be more inclined to embark on pre-breeding activities to make good progress by exploiting the potential benefits of wild cowpea relatives. Small seed size, a characteristic of wild cowpea relatives, appears to be dominant to large size (Rawal et al., 1976). In SSA, the larger the size of the cowpea grain the higher is the consumer preference; hence, breeders place some emphasis on seed size while making selections. The gene action on seed size in cowpea is mostly additive but with significant additive x additive epistatic effects and conditioned by at least eight effective factors (Drabo et al., 1984). The possibility of linkage drag of some undesirable traits that are characteristic of the wild relatives can now be readily eliminated by applying molecular tools.

It is interesting to note that two QTLs for the number of seeds per pod were found in cowpea on chromosomes Vu05 and Vu09 (Lo et al., 2018). The allele *CSp09* on chromosome Vu09, which accounted for 21.09% of the variation for number of seeds per pod, was donated by the wild parent. A transcription factor identified in the region of the QTL *CSp09* was found in Arabidopsis to be associated with ovule development (Wei et al., 2015). A higher number of seeds per pod should result in higher grain yield as have been reported in soybean (Van roekel and Purcell, 2016), peanut (Songsri et al., 2009), and rapeseed (Yang et al., 2016). Hence, the generally small seed size of the wild cowpea relatives notwithstanding, they are capable of contributing to higher grain yield in cultivated cowpea through an increase in the number of seeds in each pod, i.e., increasing the number of ovaries per pod. The QTLs associated with this trait suggest that the number of seeds per pod is heritable and can, therefore, be selected for in breeding of cowpea for higher grain yield. Long peduncles enable plants to carry their pods above the canopy thereby helping to reduce the amount of damage caused by *Maruca vitrata* (Oghiakhe et al., 2011), and allowing easier harvesting of pods. they are also a source of fibre (Pasquet, 1996). This trait is present in the wild cowpea line used in generating the linkage map. A single QTL on chromosome VU05 that explains 71.83% of phenotypic variation was detected (Lo et al., 2018). Understanding the genetic basis of perenniality should facilitate success in perennializing cultivated cowpea, which could result in the development of new potentially higher-yielding varieties.

USEFUL TRAITS PRESENT IN SOME WILD COWPEA RELATIVES

Efforts have been made to identify wild cowpea relatives, which exhibit traits that are desirable and could be transferred to improved cowpea varieties. Some of the traits detected following these efforts are discussed below:

Insect Resistance

Aphid (*Aphis craccivora* Koch) is the first major insect pest that affects cowpea at an early developmental stage. It attacks seedlings by sucking sap from the plant, especially when drought occurs. Seedlings can be killed if the infestation is high and accompanied by delayed rainfall (Singh and Jackai, 1985). Until recently, a dominant gene (*Rac*) detected in a germplasm line (TVu-3000) conferred resistance to this pest and the gene was incorporated in breeding lines developed at IITA (Bata et al., 1987). However, this gene has become ineffective and all plants having the gene now succumb to the insect, thus calling for the detection of other sources of aphid resistance genes. A cowpea wild relative, TVNu-1158, showed resistance to aphid in the seedling stage (Souleymane et al., 2013). This wild relative has been successfully crossed to cowpea, and a set of RILs developed, which have been used to generate a linkage map of cowpea. In addition, QTLs with effects on domestication related traits have been detected using this RIL population (Lo et al., 2018). The set of RILs have been evaluated for resistance to aphid, and some of them were found to be resistant. These are now being used as parents in a crossing program to some elite lines with the aim of transferring the resistance to the latter. In a recent study, three cultivated cowpea accessions TVu-6464, TVu-1583, and TVu-15445 with good levels of resistance to *A. craccivora* comparable to the level found in an existing resistant TVu-801 were reported (Togola et al., 2020). These new sources of aphid resistance in both wild and cultivated cowpea lines need to be tested for allelism. Should all or some of them turn out to be non-allelic, pyramiding them in good genetic backgrounds would lead to the development of new resistant cowpea varieties that will be resilient to aphid. In addition, the resistance will remain present in such varieties over a longer period. The resistance mechanism in the three accessions listed above was established to be linked to their low sucrose levels and high levels of kaempferol and quercetin (aglycones of phenolic compounds) (Togola et al., 2020). A flavonoid HPLC fingerprint carried out on some wild

and cultivated cowpea also showed a positive relationship between aphid resistance and high levels of flavonoid glycosides (Lattanzio et al., 1997).

Some lines of *Vigna vexillata* have been reported to exhibit resistance to the legume pod borer (*Maruca vitrata*), the most cosmopolitan of cowpea insect pests capable of reducing grain yield by up to 60% (Singh et al., 1990). It is the cause of severe damages to pods, seeds, and young tender plant tissues. All of five accessions of *V. vexillata*, 10 of *V. oblongifolia*, two of *V. macrosperma*, and one *V. reticulata*, all wild cowpea relatives, were completely resistant to Maruca (Singh et al., 1990). Trichomes present on the pods of two *V. vexillata* accessions (TVNu-72 and TVNu-73) were partly responsible for their resistance to *Maruca*, although when trichomes were removed the insect's larvae developed but not optimally (Jackai and Oghiakhe, 2009). On the other hand, the two lines were resistant to the pod sucking bug, *Clavigralla tomentosicollis* Stål (Hemiptera, Coreidae) as the insect's nymphs could not survive whether trichomes were present or removed. However, adult pod sucking bugs' feeding damage score was slightly higher when trichomes were removed from the pods. They concluded that the mechanisms of insect resistance present in the two *V. vexillata* lines were antibiosis and antixenosis. All accessions of *V. luteola* (3), *V. vexillata* (17), *V. macrosperma* (2), and *V. angustifolia* (3), were resistant to *Clavigralla tomentosicollis* Stål (Singh et al., 1990). All six and 27 accessions of *V. luteola* and *V. vexillata*, respectively, showed resistance to the cowpea seed weevil, *C. maculatus* Fabricius (Singh et al., 1990).

Striga and Alectra Resistance

Striga gesnerioides [Wild.] Vatke and *Alectra vogelii* [Benth], members of the family Scrophulariaceae (Kuijt, 1969), are parasitic weeds that attack cowpea plants in the field. They are root pathogens with *Striga* alone capable of causing up to 50% (Parker, 1991) or even 83% to 100% (Emechebe et al., 1991; Cardwell and Lane, 1995) yield loss in cowpea. A single *Striga* plant can produce a large quantity of seeds (up to 90,000), many of which may remain viable in the soil for up to 15 to 20 years (Parker, 1991). *Striga*, which is generally more virulent, is predominantly present in the dry savannah (Sudan and Sahel) agroecologies where most cowpea is produced, while *Alectra* is found in the moist savanna areas such as the guinea savannah. The development of varieties that are resistant to these parasites is the most efficient way of controlling them in SSA farmers' fields. Genes that confer resistance to both parasites have been identified and incorporated into improved varieties that farmers are growing in different countries. A major constraint to disseminating resistant cowpea varieties across the region is the existence of many races of *Striga* present in the different countries. Five races were initially identified (Lane et al., 1997). Race 1 is found in Burkina Faso, race 2 in Mali, race 3 in Nigeria and Niger, race 4 in the Republic of Benin, and race 5 in Cameroon. Three of these races (1, 3, and 5) were later reported to be present in Nigeria (Singh and Emechebe, 1997). Two previously unknown races parasitizing cowpea have since been reported one each in Senegal—race 6 and Zakpota in the Republic of Benin—SG4z (Botanga and Timko, 2006) making the number of identified races seven. The *Striga* resistance genes *Rsg-1*, *Rsg-2*, and *Rsg-3* were obtained from two different cowpea lines (Fery and Singh, 1997). The cowpea line B301 is the source of two duplicate genes *Rsg-1* and *Rsg-2*. Some wild cowpea relatives were evaluated for their reactions to the parasite in highly *Striga* infested fields at Minjibir near Kano in northern Nigeria located in the Sudan savannah agroecology and some of them showed high levels of resistance. Among the wild cowpea relatives that showed immunity to *Striga* are TVNu-1070, TVNu-1083, TVNu-585, TVNu-1535, TVNu-1537, TVNu-1647, and TVNu-491 belonging to the following *Vigna* species: *ambacensis*, *parkeri*, *oblongifolia*, and *reticulata* (Oyatomi et al., 2016). Only one *Striga* resistant accession, TVNu-1589, among the tested wild relatives belong to section *Catiang*. *Striga* seeds cannot be moved across national boundaries due to plant quarantine regulations but seeds of resistant cultivars and cross compatible wild relatives can be shared with cowpea breeders across the West African sub-region for evaluation against the different races present in their respective countries. Two duplicate genes with symbols *Rav-1* and *Rav-2* control resistance of cowpea to *Alectra* (Singh et al., 1993). Genes conferring resistance to *Striga* and *Alectra* are neither allelic nor linked (Atokple et al., 1995). However, additional sources of resistance genes could be present in the resistant wild cowpea relatives and these should boost resistance levels if non-allelic to any of the reported dominant resistance genes that have been incorporated into improved varieties. The resistance genes present in these wild cowpea relatives could serve as additional sources of resistance to potentially new *Striga* races that may develop as a result of climate change effects.

Nutritional Qualities

Eight wild *Vigna* s pp. (*V. vexillata*, *V. vexillata macrosperma*, *V. luteola*, *V. oblongifolia*, *V. unguiculata dekindtiana*, *V. racemosa*, *V. reticulata*, and *V. ambacensis*) were evaluated for chemical characteristics such as protein, amino acid profiles, starch digestibility as well as for anti-metabolic compounds, such as trypsin inhibitors, cysteine proteinase inhibitors, lectins, phytic acid, and tannins (Marconi et al., 1997). The aim was to identify potentially useful materials for improving the nutritional and insect resistance aspects of cowpea. *V. vexillata* showed a high protein content of up to 29.3% in the grains, while all the accessions contained high sulphur amino acids as a result of which they all showed high chemical scores. Starch content in their grains ranged from 64% to 75%. There was wide variability found in the levels of trypsin inhibitors, tannins, and lectins in the grains. Also, *V. luteola* contained high levels of these compounds, while *V. reticulata* and *V. ambacensis* as well as the immediate progenitor of cultivated cowpea, *V. unguiculata dekindtiana*, had very low levels. Significant positive correlations were found between bruchid resistance and trypsin inhibitor, tannin, and starch content. Despite the high protein content in grains of wild cowpea relatives, it was observed that its digestibility is low compared with cultivated lines (Marconi et al., 1990). These authors also reported that protein availability was slightly higher in wild relatives than in tested cultivated lines.

Some of the wild *Vigna* species can be utilized by humans in different ways besides the protein rich grains. For example, the tuberous roots of *V. vexillata* and *V. lobatofolia* which contain higher protein content than potato by up to 15% and six times the level found in cassava are eaten in several communities (Wehmeyer et al., 1969; Chandel et al., 1972). Some wild cowpea relatives exhibit traits that could be useful in enhancing the food value of cowpea, such as higher protein levels in grains.

Drought Tolerance

When compared with many other crops, cowpea performs relatively better under drought conditions. However, the occurrence of drought, especially from seedling to flowering stage can still have adverse effects on its productivity. The existing level of drought tolerance in improved varieties can be enhanced through breeding. The habitats from where some of the accessions of the wild relatives were collected, and their growth habits, suggest that some of them could be potential sources of genes for drought tolerance. For example, accessions of the subspecies *tenuis* and *stenophylla* were collected mostly in the dry savannah agro-ecologies of Mozambique and Zimbabwe where the soils are generally sandy (Padulosi and Ng, 1997). It is conceivable that some of the wild cowpea relatives collected from these dry environments should be potential sources of drought tolerance genes. This is because they have become adapted to such environments. A recent study on Mesoamerican common bean showed that wild relatives of the crop collected from dry areas of South America are found to be good sources of drought tolerance (Teran et al., 2020). Some wild cowpea lines are characterized by perennial growth habit and this attribute could also contribute to drought tolerance as they retain their greenness, and are able to survive from one cropping season to another through the intervening dry season.

Longevity

Cultivated cowpea is an annual crop with improved varieties being either extra early, i.e., maturing in 60 days, early (65–75 days), medium (75–100 days), or late (more than 100 days). Most of the traditional farmers' varieties are, on the other hand, late maturing types. These latter types also tend to spread, cover the ground quickly (Rawal, 1975), and are day length sensitive (Craufurd et al., 1997). Farmers who prefer to grow cowpea as a sole crop usually choose day neutral, extra early or early maturing varieties while those who grow cowpea in intercrop prefer the dual purpose and late maturing types. In the dry savannah regions of SSA, cowpea fodder is appreciated by livestock farmers because of its relatively high nutritional quality. It has been observed that in the dry savannah agro-ecologies, cowpea farmers make a reasonable income from sales of cowpea fodder (Inaizumi et al., 1999; Samireddypalle et al., 2017). Even after insect pests such as pod borer and pod sucking bugs have damaged pods and seeds on the crop while growing in the field, farmers are still able to harvest the cowpea fodder for sale to livestock owners. Cut and carry systems of cowpea fodder are well developed in Asia and Australia where yields can reach up to 4 tons per hectare (Heuze et al., 2015). Some cross compatible cowpea relatives are perennial and can therefore be grown by farmers and herders as long-term sources of fodder for animals. A set of RILs generated from a bi-parental cross between cultivated and a wild cowpea relative with the latter exhibiting perennial growth habit was evaluated, and some of the RILs demonstrated longevity by surviving and staying green for more than 700 days from planting in pots (Lo et al., 2020). An understanding of the genetic basis of perenniality should facilitate efforts in perennializing cultivated cowpea which could result in the development of new potentially higher yielding varieties. Additionally, cowpea farmers in the Sahelian region who keep livestock may have a preference for such long surviving plants as they will be good sources of fodder over an extended period. Varieties with stay green characteristics over a long time could be of immense benefit to the itinerant herdsmen, who are based mainly in the dry savannah areas but migrate to the more humid coastal areas of West Africa for greener fodder during the dry season when the savannah vegetation has dried up. The availability of such varieties will also help stem the farmers/herdsmen clashes that are commonly reported from different parts of Nigeria and elsewhere in the West African sub-region.

INTERSPECIFIC HYBRIDIZATION AND BACKCROSS BREEDING: BARRIERS AND OVERCOMING THEM

Some traits desired in improved cowpea varieties are present in several of the crop's wild relatives. For example, genes for resistance to the legume pod borer (*Maruca vitrata*) and pod sucking bugs were found in accessions of *V. oblongifolia* and *V. vexillata* (IITA, 1988; Singh et al., 1990). Among the wild cowpea relatives outside of section *Catiang*, *V. vexillata*, which belongs to section *Plectrotropis*, was reported to be the most phylogenetically close to cowpea based on RFLP analysis (Fatokun et al., 1993). This notwithstanding all attempts made to cross cowpea with *V. vexillata* failed to produce any interspecific hybrid (Fatokun, 2002). Embryological studies were carried out to determine the causes of interspecies incompatibility between the two (Barone and Ng, 1990). According to the authors, incompatibility results from lack of fertilization in most instances and the collapse of ovules following pollination. Several crosses were made between cowpea and *V. vexillata* with the aim of transferring the desirable genes for resistance to insect pests such as pod borer, flower bud thrips, and seed storage weevil from the wild to cultivated genotype. Some of the efforts made to overcome the cross incompatibility between cowpea and *V. vexillata* (Fatokun, 2002) were as follows:

a. Making crosses among several accessions of both species. This was attempted from the recognition that certain cross combinations are more productive than others. In some crop species such as *Nicotiana* s pp., it was observed that crosses between some accessions were more successful than others (Pittarelli and Stavely, 1975). However, in the case of cowpea

and *V. vexillata*, none of the cross combinations resulted in an interspecies hybrid.

b. Application of growth hormones such as 2,4-D and NAA on the pistil before or after pollination was also attempted to promote a successful interspecies cross between cowpea and *V. vexillata*. Growth hormones have been successfully used to prolong pod retention in interspecific crosses in *Phaseolus vulgaris* from 15 to 30 days (Al-Yasiri and Coyne, 1964). In the case of cowpea, spraying low concentrations of 2,4-D on flowers of *V. vexillata* before and after pollination with cowpea resulted in pod retention until pod maturity. However, no viable hybrid seed was obtained from the retained pods (Fatokun, 2002).
c. Rescue of embryos extracted from ovules retained for up to 4 days following interspecific cross pollination—*V. vexillata* x *V. unguiculata* ssp. *unguiculata* was attempted. The embryo and endosperm obtained following crosses between *V. vexillata* and cowpea collapsed between five to eight days after pollination (Barone et al., 1992). Embryos were found to form and developed up to the globular stage, after which further development stopped when cowpea was crossed with *V. vexillata* (Fatokun, 1991). Placing these interspecific embryos in MS culture media containing growth hormones did not result in further development (Fatokun, 2002).
d. Polyploidization of both species prior to being used for making interspecific hybridizations was carried out. Both cowpea and *vexillata* are diploids with a chromosome number of 2n = 22. Only accessions of cultivated cowpea responded positively to treatments with low concentrations of colchicine. However, the polyploid cowpea lines generated were fertile and set seeds, still could not produce any hybrid when crosses were made between them and *V. vexillata* in both directions, i.e., using the polyploid as male and female (Fatokun, 2002).
e. The use of a parthenocarpic cowpea line (R1 36) obtained from the University of California Riverside as a female parent to crosses with accessions of *V. vexillata* was attempted. However, no successful hybrid resulted (Fatokun, 2002).

From the foregoing listed unsuccessful attempts to cross cowpea *V. vexillata*, there is a very strong cross incompatibility between cowpea and its wild relatives outside the section *Catiang*. This strong cross incompatibility barrier has so far prevented the introgression of useful genes in *V. vexillata* to cowpea for variety development.

The wide crosses that have so far succeeded with cowpea are those involving members of section *Catiang* (DC) Verdc., which seem to contain the primary and secondary gene pools for cowpea. Unlike what has been observed among African *Vigna* species, successful interspecific hybridizations have been reported among Asiatic *Vigna* species that belong to section *Ceratotropis* (Piper) Verdc. Successful crosses were made between *V. radiata* x *V. dalzelliana*, *V. radiata* x *V. radiata* var. *sublobata*, *V. radiata* x *V. mungo* var. *silvestris*, and *V. umbellata* x *V. radiata* all of which are Asiatic *Vigna* species (Pandiyan et al., 2010). The successful interspecific crosses among the Asiatic *Vigna* species have enhanced genetic diversity and made it possible to take advantage of some attributes in the wild relatives for the development of high performing improved mungbean (*V. radiata*) varieties.

DEVELOPMENT OF POPULATIONS WITH INTROGRESSIONS FROM CWRS: INTROGRESSION LINES, CHROMOSOME SUBSTITUTION LINES, ADVANCED BACKCROSSES, AND OTHERS

Successful crosses between cowpea and its wild relatives have only been reported when the crosses involve those belonging to section *Catiang*. Even with some accessions from this section, it was necessary to carry out embryo rescue to be able to successfully cross some cowpea *Vigna unguiculata* ssp. *unguiculata* lines with *Vigna unguiculata* ssp. *pubescens* (Fatokun and Singh, 1987). The F_1 plants derived from this cross exhibited partial sterility due to a low number of fertile pollen grains. In common bean (*Phaseolus vulgaris* L.), a grain legume with 11 pairs of chromosomes and a member of *Phaseolinaea* as cowpea, inbred populations were developed from three wild x domesticated backcrosses (Teran et al., 2020). The BC_1S_4 populations were evaluated in the fields located at three environments comprising two fully irrigated trials during two cropping seasons and an imposed terminal drought in the second season. The study revealed that the two populations derived from wild parents obtained from lower rainfall regions produced lines that gave higher yield compared to the domesticated parent in the three environments. They further reported that 20 QTLs for yield were detected in 13 regions on eight of common bean's 11 chromosomes. Five of the QTLs showed at least one wild allele that contributed to increased yield compared to the domesticated parent. In cowpea, an advanced backcross has been generated from the cross between an improved variety (IT99K-573-1-1) and a wild relative (TVNu-1158). The wild relative was identified as an aphid resistant line (Souleymane et al., 2013). The F_1 was backcrossed to the improved variety, and the BC_1F_1 further advanced to F3 generation. This population has been further advanced but have not been phenotyped.

CHARACTERIZATION AND EVALUATION OF POPULATIONS WITH INTROGRESSIONS FROM CWRS FOR SIMPLE AND COMPLEX TRAITS

A set of recombinant inbred lines (RILs) was generated from a cross between a cultivated and a wild cowpea relative. The set of RILs has been evaluated for nine different traits, including those related to cowpea domestication such as pod shattering, 100-seed weight, pod size, and flower characteristics. In addition, the RILs were genotyped using the Cowpea iSelect Consortium Array, which assays 51,128 SNPs. The 17,739 of these SNPs that passed

quality controls were used to develop a high-density linkage map of cowpea (Lo et al., 2018). Sixteen quantitative trait loci (QTL) were detected across seven chromosomes for the traits measured. One major finding reported by these authors is the co-location of QTL for four traits controlling increased organ size, which are very important during the crop's domestication process, namely, primary leaf length and width, 100-seed weight, and pod length in the same region of chromosome *Vu08*. Increased leaf size should result in a higher amount of photosynthate, which would support an increase in biomass and subsequently higher grain yields. In *Vigna umbellata* (rice bean), QTL for leaf size was closely located to those controlling seed and pod size related traits (Isemura et al., 2010). A study carried out in cowpea showed that two major QTLs are associated with flowering in the wide cross with the wild parent conferring late flowering on some RILs (Lo et al., 2018). The main QTL for the number of days from sowing to flowering in their study is located on chromosome Vu09. Using the same set of RILs, the authors reported the detection of three important loci for plant longevity and flower scent in cowpea (Lo et al., 2020). The QTL for plant longevity is located on chromosome number Vu08, the same chromosome on which some organ size related QTL were reported in their earlier study. Within this perenniality QTL region, they observed the presence of genes encoding an F-box protein (*Vigun08g215300*) and two kinases (*Vigun08g217000*, *Vigun08g217800*), both involved in physiological processes that include flowering time regulation and plant longevity. QTL for flower scent, an important trait in insect induced cross pollination, was detected on two chromosomes, Vu01 and Vu11.

"OMICS" APPLICATIONS TO INTROGRESSION BREEDING

Compared with many other crops, the development and application of genomic tools for cowpea improvement has lagged behind and only a few relevant studies have been reported. Some progress has been recorded in recent times following identification of molecular markers associated with some important traits in the crop, but marker application in variety development is still very limited. In the Republic of Niger, marker assisted backcross was used to transfer *Striga* resistance gene from the breeding line IT93K-693-2 into three farmers' preferred varieties, *viz*., IT90K-372-1-2, KVx30-309-6G, and TN5-78 (Salifou et al., 2016). The microsatellite marker SSR1 was used to track and introgress the resistance gene in the segregating populations. Marker-assisted breeding based on SNP genotyping was used to stack large seed haplotypes into a CB27 background with 22 g/100 seeds using a rare haplotype associated with large seeds at the *Css-1* locus from cowpea variety IT82E-18 (18.5 g/100 seeds) (Lucas et al., 2015). These authors used foreground and background selections during two cycles of backcrossing based on genome-wide SNP markers and obtained families with very large seeds (28–35 g/100 seeds). Three major QTLs associated with bacterial blight, one on Vu09 (*qtlblb-1*), and two on Vu04 (*qtlblb-2* and *qtlblb-3*), accounting for 30.58%, 10.77%, and 10.63% phenotypic variations, respectively, have been identified (Dinesh et al., 2016). They successfully introgressed the QTL on Vu09 from cultivar V-16 into the bacterial leaf blight susceptible variety C-152 through marker-assisted backcrossing (MABC). These are the first reports on the application of marker assisted breeding in cowpea improvement. With these demonstrated successful introgression of desired genes into cowpea from cultivated lines, breeders can now apply marker assisted technology in exploiting the hidden potentials present in wild cross compatible cowpea wild relatives. Carrying out genome-wide association studies (GWAS) using wild crop relatives can be a good and successful starting point for identifying homologous genes in other species belonging or not to the same botanical family (Huard-chauveau et al., 2013). Cowpea can also benefit from utilizing its cross compatible wild relatives as exemplified in the example of common bean mentioned above (Teran et al., 2020).

GENE EDITING OF COWPEA TO FACILITATE THEIR USE IN BREEDING

Genome-editing is a new breeding tool that enables the efficient and precise targeted modification of plant genome and therefore has a lot of potential for crop improvement. It has been applied in a wide variety of plant species for functional gene analysis and the improvement of economically important traits. Several tools like Zinc Finger Nucleases (ZFNs), TAL effector proteins (TALENs), RNA-guided nucleases (RGENs), and CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 (CRISPR associated protein 9) have been applied for improvements of crops through targeted genome editing. All these methods are based on the formation of double-strand break (DSB) at specific loci and triggering DNA repair mechanisms (Weithal and Gurel, 2016). CRISPR/Cas9 has rapidly become the most popular of the genome engineering approaches as it is comparatively simple and easy to adapt. CRISPR/Cas9 technology has been successfully applied in several plant species such as Arabidopsis, *Nicotiana banthemiana*, rice, wheat, maize, sorghum, tomato, soybean, apple, citrus, poplar, and coffee (Song et al., 2016; Scheben et al., 2017; Breitler et al., 2018). Recently, the CRISPR/Cas9-based genome editing system was developed for cowpea, demonstrating the disruption of the endogenous representative symbiotic nitrogen fixation (SNF) gene (Ji et al., 2019).

Genome editing has enormous potential for improving agronomic traits of crops such as resistance to diseases and insect-pests. Resistance to diseases and insect-pests in cowpea crop can be induced using genome editing technique following several approaches such as knocking down of the susceptibility genes, manipulating the effector-target interaction, modifying the receptor gene to boost the host immune system and/or altering the plant hormones responsible for antagonistic action of defence leading to enhanced broad-spectrum resistance (Bisht et al., 2019). Genome editing can be applied to both cultivated and wild cowpea relatives following which any useful traits

exhibited by the modified lines can be introgressed in the breeding programs.

CONCLUSIONS

The wild relatives of cowpea abound in the southern parts of Africa, where available evidence suggest their center of diversity exists. Seeds of a number of these wild relatives have been collected and conserved in the genetic resources center at IITA Ibadan, Nigeria, and elsewhere. Breeders, plant physiologists, and entomologists, in particular, have evaluated some of the wild cowpea relatives for the purpose of identifying those with the potential to contribute genes for resistance to insect pests and drought tolerance in cultivated cowpea. These efforts have led to the discovery of accessions of wild relatives that show resistance to insect pests of cowpea. For example, accessions of *V. oblongifolia* and *V. vexillata* that were evaluated showed resistance to insect pests such as aphid, flower bud thrips, pod borers, and bruchids. However, very strong cross incompatibility exists between cultivated cowpea and *vexillata* and this has made it impossible for breeders to introgress the resistance genes into the former. However, few cross compatible wild cowpea relatives show resistance to some constraints, and these wild relatives are being used as sources of resistance genes. The development of molecular markers in cowpea will help breeders utilize these cross compatible wild relatives in cowpea variety development. Undesirable traits such as small seed size, unattractive seed coat color, etc. present in the wild relatives can be more readily selected using molecular marker technologies. The new breeding technologies will be of immense benefit while introgressing additional traits such as perennial growth habit and stay green characteristics existing in wild relatives into cultivated cowpea.

AUTHOR CONTRIBUTIONS

OB led several of the studies reported here and wrote most of the sections of the paper with strong contributions of CF who also designed the paper and is the corresponding author. MA and OO shared the wild accessions from the IITA Genetic Resources Centre and conducted the screening of wild relatives for *Striga r*esistance. LT wrote the section on the genetic engineering and gene editing. AT wrote the sections on insect pests. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors wish to thank the Bill and Melinda Gates Foundation for funds provided under the TLIII project used to cover the costs of some of the studies reported here. Similarly, some of the studies reviewed were carried out as part of the initiative "Adapting Agriculture to Climate Change: Collecting, Protecting and Preparing Crop Wild Relatives", which is supported by the Government of Norway and managed by the Global Crop Diversity Trust with the Millennium Seed Bank of the Royal Botanic Gardens, Kew. We also thank IITA for providing facilities used for several of the studies reviewed in this paper as well as IITA's cowpea breeding staff in Ibadan and Kano, Nigeria for their support.

REFERENCES

African Centre for Biodiversity (ACB) (2015). *GM and seed industry eye Africa's lucrative cowpea seed markets: The political economy of cowpea in Nigeria* (Ghana and Malawi: Burkina Faso). Available at: https://acbio.org.za/cowpea/.

Ajibade, S., Weeden, N., and Chite, S. (2000). Inter simple sequence repeat analysis of genetic relationships in the genus *Vigna*. *Euphytica* 111, 47–55. doi: 10.1023/A:1003763328768

Al-Yasiri, A., and Coyne, D. P. (1964). Effects of growth regulators in delaying pod abscission and embryo abortion in interspecific cross *Phaseolus vulgaris* x *P. accutifolius*. *Crop Sci.* 4, 433–435. doi: 10.2135/cropsci1964.0011183X000400040032x

Arumuganathan, K., and Earle, E. (1991). Nuclear DNA content of some important plant species. *Plant Mol. Bio. Rep.* 9, 208–218. doi: 10.1007/BF02672069

Atokple, I. D. K., Singh, B. B., and Emechebe, A. M. (1995). Genetics of resistance to *Striga* and *Alectra* in cowpea. *J. Hered.* 86, 45–49. doi: 10.1093/oxfordjournals.jhered.a111524

Ba, F. S., Pasquet, R. S., and Gepts, P. (2004). Genetic diversity in cowpea [*Vigna unguiculata* (L.) Walp.] as revealed by RAPD markers. *Genet. Res. Crop Evol.* 51, 539–550. doi: 10.1023/B:GRES.0000024158.83190.4e

Barone, A., and Ng, Q. (1990). "Embryological study of crosses between Vigna unguiculata and V. vexillata," in *Cowpea Genetic Resources*. Eds. N. Q. Ng and L. M. Monti (Nigeria: International Institute of Tropical Agriculture (IITA)), pp. 151– 160.

Barone, A., Del Guidice, A., and Ng, Q. N. (1992). Barriers to interspecific hybridization between *Vigna unguiculata* and *V. vexillata*. *Sex Reprod.* 5, 195–200. doi: 10.1007/BF00189811

Bata, H. D., Singh, B. B., Singh, S. R., and Ladeinde, T. A. O. (1987). Inheritance of resistance to aphid in cowpea. *Crop Sci.* 27, 892–894. doi: 10.2135/cropsci1987.0011183X002700050011x

Baudoin, J. P., and Marechal, R. (1985). "Genetic diversity in Vigna," in *Cowpea Research, Production and Utilization*. Eds. S. R. Singh and K. O. Rachie (Chichester: John Wiley and Sons), 3–11.

Bett, B., Gollasch, S., Moore, A., James, W., Armstrong, J., Walsh, T., et al. (2017). Transgenic cowpeas (*Vigna unguiculata* L. Walp.) expressing *Bacillus thuringiensis* Vip3Ba protein are protected against the Maruca pod borer (*Maruca vitrata*). *Plant Cell Org. Tiss. Cult. (PCTOC)* 131, 335–345. doi: 10.1007/s11240-017-1287-3

Bisht, D. S., Bhatia, V., and Bhattacharya, R. (2019). Improving plant-resistance to insect-pests and pathogens: The new opportunities through targeted genome editing. *Semin. Cell Dev. Biol.* 96, 65–76. doi: 10.1016/j.semcdb.2019.04.008

Botanga, C. J., and Timko, M. P. (2006). Phenetic relationships among different races of *Striga gesnerioides* (Willd.) Vatke from West Africa. *Genome* 49 (11), 1351–1365. doi: 10.1139/g06-086

Breitler, J. C., Dchamp, E., Campa, C., Rodrigues, L. A. Z., Guyot, R., Marraccini, P., et al. (2018). CRISPR/Cas9-mediated efficient target mutagenesis has the potential to accelerate the domestication of *Coffea canephora*. *Plant Cell Org. Tiss. Cult.* 134, 383–394. doi: 10.1007/s11240-018-1429-2

Cardwell, K. F., and Lane, J. A. (1995). Effect of soils, cropping system and host phenotype on incidence and severity of *Striga gesnerioides* on cowpea in West Africa. *Agric. Econ. Environ.* 53 (3), 253–262. doi: 10.1016/0167-8809(94)00575-y

Chandel, K. P. S., Arora, R. K., and Joshi, B. S. (1972). *Vigna capensis* Walp. (*V. vexillata*) an edible root legume. *Curr. Sci.* 41, 537.

Chheda, H. R., and Fatokun, C. A. (1982). Numerical analysis of variation patterns in okra (*Abelmoschus esculentus* [L.] Moench). *Bot. Gaz.* 143, 253–261. doi: 10.1086/337298

Coulibaly, S., Pasquet, R. S., Papa, R., and Gepts, P. (2002). AFLP analysis of the phenetic organization and genetic diversity of *Vigna unguiculata* L. Walp. reveals extensive gene flow between wild and domesticated types. *Theor. Appl. Genet.* 104, 358–366. doi: 10.1007/s001220100740

Craufurd, P. Q., Summerfield, R. J., Ellis, R. H., and Roberts, E. H. (1997). "Photoperiod, temperature and the growth and development of cowpea," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. 75–86.

Crop Biotech Update (2019). *International Services for the Acquisition of Agric-Biotechnology Applications (ISAAA) Brief. "2019 Dec 18th. Ithaca, New York, USA: ISAAA · c/o Global Development-CALS B75 Mann Library · Cornell University.*

Dinesh, H. B., Lohithaswa, H. C., Viswanatha, K. P., Singh, P., and Rao, A. M. (2016). Identification and marker-assisted introgression of QTL conferring resistance to bacterial leaf blight in cowpea (*Vigna unguiculata* (L.) Walp.). *Plant Breed.* 135 (4), 506–512. doi: 10.1111/pbr.12386

Drabo, I., Redden, R., Smithson, J. B., and Aggarawal, V. D. (1984). Inheritance of seed size in cowpea (*Vigna unguiculata* (L.) Walp.). *Euphytica* 33, 929–934. doi: 10.1007/bf00021923

D'Andrea, A. C., Kahlheber, S., Logan, A. L., and Watson, D. J. (2007). Early domesticated cowpea (*Vigna unguiculata*) from Central Ghana. *Antiquity* 81, 686–698. doi: 10.1017/S0003598X00095661

Emechebe, A. M., Singh, B. B., Leleji, O.II, Atokple, I. D. K., and Adu, J. K. (1991). *Combating striga in Africa: proceedings of the international workshop held in Ibadan, Nigeria, 22-24 August 1988*. Ibadan, Nigeria: International Institute of Tropical Agriculture. pp. 18–28, 1991ref.13.

Fatokun, C. A., and Singh, B. B. (1987). Interspecific hybridization between *V. pubescens* and *V. unguiculata* (L.) Walp. through embryo rescue. *Plant Cell Org. Tiss. Cult.* 9, 229–233. doi: 10.1007/BF00040808

Fatokun, C. A., Danesh, D., Young, N. D., and Stewart, E. L. (1993). Molecular taxonomic relationships in the genus *Vigna* based on RFLP analysis. *Theor. Appl. Genet.* 86, 97–104. doi: 10.1007/BF00223813

Fatokun, C. A. (1991). Wide hybridization in cowpea: problems and prospects. *Euphytica* 54, 137–140. doi: 10.1007/BF00039599

Fatokun, C. A. (2002). "Breeding cowpea for resistance to insect pests: attempted crosses between cowpea and Vigna vexillata," in *Challenges and Opportunities for Enhancing Sustainable Cowpea Production. Proceedings of the World Cowpea Conference III held at the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. 4-8 September 2000*. Eds. C. A. Fatokun, S. A. Tarawali, B. B. Singh, P. M. Kormawa and M. Tamo. Ibadan, Nigeria: International Institute of Tropical Agriculture. pp. 52–61.

Fery, R. L., and Singh, B. B. (1997). "Cowpea genetics: a review of the recent literature," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. 13–29.

Heuze, V., Tran, G., Bastianelli, D., and Lebas, F. (2015). *Cowpea (Vigna unguiculata) forage. Feedipedia. A programme by INRA, CIRAD, AFZ and FAO*. Rome: Food and Agriculture Organization. Available at: https://www.feedipedia.org/node/233.

Huard-chauveau, C., Perchepied, L., Debieu, M., Rivas, S., Kroj, T., Kars, I., et al. (2013). An atypical kinase under balancing selection confers broad-spectrum disease resistance in Arabidopsis. *PloS Genet.* 9, e1003766. doi: 10.1371/journal.pgen.1003766

Huynh, B. -L., Close, T. J., Roberts, P. A., Hu, Z., Wanamaker, S., Lucas, M. R., et al. (2013). Gene pool and genetic architecture of domesticated cowpea. *Plant Genome* 6 (3), 8.

Inaizumi, H., Singh, B. B., Sanginga, C. P., Manyong, V. M., Adesina, A. A., and Tarawali, S. A. (1999). *Impact: Adoption and impact of dry-season dual-purpose cowpea in the semiarid zone of Nigeria* (Ibadan, Nigeria: International Institute of Tropical Agriculture), 25pp.

International Institute of Tropical Agriculture (IITA) (1988). *Annual Report for 1987 (GLIP)* (Ibadan, Nigeria: International Institute of Tropical Agriculture (IITA)).

Isemura, T., Kaga, A., Tomooka, N., Shimizu, T., and Vaughan, D. A. (2010). The genetics of domestication of rice bean, *Vigna umbellata*. *Ann. Bot.* 106, 927–944. doi: 10.1093/aob/mcq188

Jackai, L. E. N., and Oghiakhe, S. (2009). Pod wall trichomes and resistance of two wild cowpea *Vigna vexillata*, accessions to *Maruca testulalis* (Geyer) (Lepidoptera: Pyralidae) and *Clavigralla tomentosicollis* Stål (Hemiptera: Coreidae). *Bull. Entomol. Res.* 79, 595–605. doi: 10.1017/S0007485300018745

Jackai, L. E. N., and Singh, S. R. (1988). Screening techniques for host plant resistance to insect pests of cowpea. *Trop. Grain Legumes Bull.* 35, 2–18.

Ji, J., Zhang, C., Sun, Z., Wang, L., Duanmu, D., and Fan, Q. (2019). Genome editing in cowpea *Vigna unguiculata* using CRISPR-Cas9. *Int. J. Mol. Sci.* 20:2471. doi: 10.3390/ijms20102471

Kouam, E. B., Pasquet, R. S., Campagne, P., Tignegre, J.-B., Thoen, K., Gaudin, R., et al. (2012). Genetic structure and mating system of wild cowpea populations in West Africa. *BMC Plant Biol.* 12:113. doi: 10.1186/1471-2229-12-113

Kuijt, J. (1969). *The Biology of Parasitic Flowering Plants* (Berkeley, CA, USA: University of California Press), 346 pp.

Lane, J. A., Moore, T. H. M., Child, D. V., and Bailey, J. A. (1997). "Variation in virulence of Striga gesnerioides on cowpea: new sources of crop resistance," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. 225–230.

Lattanzio, V., Cardinali, A., Linsalata, V., Perrino, P., and Ng, N. Q. (1997). "Flavonoid HPLC fingerprints of wild Vigna species," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. 66–74.

Li, C. D., Fatokun, C. A., Ubi, B., Singh, B. B., and Scoles, G. J. (2001). Determining genetic similarities and relationships among cowpea breeding lines and cultivars by microsatellite markers. *Crop Sci.* 41, 189–197. doi: 10.2135/cropsci2001.411189x

Lo, S., Muñoz-Amatriaín, M., Boukar, O., Herniter, I., Cisse, N., Guo, Y.-N., et al. (2018). Identification of QTL controlling domestication-related traits in cowpea (*Vigna unguiculata* L. Walp.). *Sci. Rept.* 8, 6261. doi: 10.1038/s41598-018-24349-4

Lo, S., Fatokun, C., Boukar, O., Gepts, P., Close, T., and Muñoz-Amatriaín, M. (2020). Identification of QTL for perenniality and floral scent in cowpea (*Vigna unguiculata* [L.] Walp.). *Plos One* 15 (4). doi: 10.1371/journal.pone.0229167

Lonardi, S., Muñoz-Amatriain, M., Liang, Q., Shu, S., Wanamaker, S.II, Lo, S., et al. (2019). The genome of cowpea (*Vigna unguiculata* [L.] Walp.). *Plant J.* 98, 767–782. doi: 10.1111/tpj.14349

Longwe, L. (1996). "Malawi: Country Report" in *FAO International Technical Conference on Plant Genetic Resources, Leipzig, Germany, 17 – 23, June, 1996*. Rome: Food and Agriculture Organisation. pp 1–pp47.

Lucas, M. R., Huynh, B.-L., Roberts, P. A., and Close, T. J. (2015). Introgression of a rare haplotype from Southeastern Africa to breed California blackeyes with larger seeds. *Front. Plant Sci.* 6:126:126. doi: 10.3389/fpls.2015.00126

Marconi, E., Lombardi-Boccia, G., Carnovale, E., and Ng, N. Q. (1990). "Nutritional evaluation of wild and cultivated species of cowpea," in *Cowpea Genetic Resources*. Eds. N. Q. Ng and L. M. Monti (Ibadan, Nigeria: International Institute of Tropical Agriculture (IITA), Pp. 101–118.

Marconi, E., Ruggeri, S., and Carnovale, E. (1997). Chemical evaluation of wild under-exploited *Vigna* s pp. Seeds. *Food Chem.* 59, 203–212. doi: 10.1016/S0308-8146(96)00172-0

Maréchal, R., Mascherpa, J. M., and Stainier, F. (1978). Etude taxonomique d'un groupe d'espèces des genres *Phaseolus* et *Vigna* (*V. unguiculata*) sur la base de données morphologiques et polliniques, traitées pour l'analyse informatique. *Boissiera* 28, 1–273.

Mohammed, B. S., Ishiyaku, M. F., Abdullahi, U. S., and Katung, M. D. (2014). Response of transgenic *Bt* cowpea and their hybrids under field conditions. *J. Plt. Breed. Crop Sci.* 6, 91–96. doi: 10.5897/JPBCS2013.0401

Mortimore, M. J., Singh, B. B., Harris, F., and Blade, S. F. (1997). "Cowpea in traditional cropping systems," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. 99–Pp:113.

Oghiakhe, S., Jackai, L. E. N., and Makanjuola, W. A. (2011). Cowpea plant architecture in relation to infestation and damage by the legume pod borer, *Maruca testulalis* Geyer (Lepidoptera: Pyralidae)—I. Effect of canopy structure and pod position. *Int. J. Insect Sci.* 12, 193–199. doi: 10.1017/S1742758400020695

Oyatomi, O., Fatokun, C., Boukar, O., Abberton, M., and Ilori, C. (2016). "Screening wild Vigna species and cowpea (Vigna unguiculata) landraces for sources of resistance to Striga gesnerioides)," in *Enhancing crop genepool use: capturing wild relatives and landrace diversity for crop improvement*. Eds. N.

Maxted, E. M. Dulloo and B. V. Ford-Lloyd (Wallingford, United Kingdom: CAB Int), pp 27–pp 31. doi: 10.1079/9781780646138.0027

Padulosi, S., and Ng, N. Q. (1990). "Wild Vigna species in Africa. Their collection and potential utilization," in *Cowpea Genetic Resources*. Eds. N. Q. Ng and L. M. Monti (Ibadan, Nigeria: International Institute of Tropical Agriculture), 58–77.

Padulosi, S., and Ng, Q. (1997). "Origin, taxonomy and morphology of Vigna unguiculata (L.) Walp," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. 1–Pp:12.

Pandiyan, M., Natesan, S., Ramamoorthi, N., and Muthiah, A. (2010). Interspecific hybridization of Vigna radiata x 13 wild Vigna species for developing MYMV donor. *Electronic J. Plant Breed* 1 (4), 600–610.

Parker, C. (1991). Protection of crops against parasitic weeds. *Crop Prot.* 10, 6–22. doi: 10.1016/0261-2194(91)90019-N

Pasquet, R. S. (1996). "Wild cowpea (Vigna unguiculata) evolution," in *Advances in Legume Scum. Jour. System. 8: Legumes of Economic Importance*. Eds. B. Pickersgill and J. M. Lock (United Kingdom: Royal Botanic Garden Kew), 95–100.

Pittarelli, G. W., and Stavely, J. R. (1975). Direct hybridization of *Nicotiana rependa* x *N. tabacum*. *J. Hered.* 66, 281–284. doi: 10.1093/oxfordjournals.jhered.a108630

Popelka, J. C., Gollasch, S., Moore, A., Molvig, L., and Higgins, T. J. V. (2006). Genetic transformation of cowpea (*Vigna unguiculata* L.) and stable transmission of the transgenes to progeny. *Plant Cell Rep.* 25, 304–312. doi: 10.1007/s00299-005-0053-x

Quin, F. M. (1997). *Introduction. in Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), pp. ix–xv.

Rawal, K. M., Rachie, K. O., and Franckowiak, J. D. (1976). Reduction in seed size in crosses between wild and cultivated cowpea. *J. Hered.* 67, 253–254. doi: 10.1093/oxfordjournals.jhered.a108725

Rawal, K. M. (1975). Natural hybridization among wild, weedy and cultivated *Vigna unguiculata* (L.) Walp. *Euphytica* 24 (3), 699–707. doi: 10.1007/BF00132908

Salifou, M., Tignegre, J. B. L. S., Tongoona, P., Offei, S., Ofori, K., and Danquah, E. (2016). Introgression of Striga resistance gene into farmers' preferred cowpea varieties in Niger. *Int. J. Plant Breed. Genet.* 3 (6), 233–240.

Samireddypalle, A., Boukar, O., Grings, E., Fatokun, C., Kodukula, P., Devulapalli, R., et al. (2017). Cowpea and groundnut haulms fodder trading and its lessons for multidimensional cowpea improvement for mixed crop livestock systems in West Africa. *Front. Plant Sci.* 8, 9. doi: 10.3389/fpls.2017.00030

Scheben, A., Wolter, F., Batley, J., Puchta, H., and Edwards, D. (2017). Towards CRISPR/Cas crops- bringing together genomics and genome editing. *New Phytol.* 216, 682–698. doi: 10.1111/nph.14702

Singh, B. B., and Emechebe, A. M. (1997). "Advances in research on cowpea Striga and Alectra," in *Advances in Cowpea Research. Co-publication of International Institute of Tropical Agriculture (IITA) and Japan International Research Center for Agricultural Sciences (JIRCAS)*. Eds. B. B. Singh, D. R. Mohan Raj, K. E. Dashiell and L. E. N. Jackai (Ibadan, Nigeria: IITA), Pp. 215–224.

Singh, S. R., and Jackai, L. E. N. (1985). *Insect pests of cowpeas in West Africa: their life cycle, economic importance and potential for control in Cowpea Research, Production and Utilization*. Eds. S. R. Singh and K. O. Rachie (Chichester, England: John Wiley and Sons), 217–232.

Singh, S. R., Jackai, L. E. N., Dos Santos, J. H. R., and Adalla, C. B. (1990). "Insect pests of cowpea," in *The Insect Pests of Tropical Food Legumes*. Ed. S. R. Singh (Chichester, England: John Wiley & Sons Ltd, Baffins Lane), Pp 43–Pp 89.

Singh, B. B., Emechebe, A. M., and Atokple, I. D. K. (1993). Inheritance of Alectra resistance in cowpea genotype B301. *Crop Sci.* 33, 70–72. doi: 10.2135/cropsci1993.0011183X003300010011x

Song, L., Huang, S. C., Wise, A., Castanon, R., Nery, J. R., and Chen, H. (2016). A transcription factor hierarchy defines an environmental stress response network. *Science* 354, 6312. doi: 10.1126/science.aag1550

Songsri, P., Vorasoot, N., Jogloy, S., Kesmala, T., Akkasaeng, C., Patanothai, A., et al. (2009). Evaluation of yield and reproductive efficiency in peanut (*Arachis hypogea* L.) under different available soil water. *Asian J. Plant Sci.* 8, 465–473. doi: 10.3923/ajps.2009.465-473

Souleymane, A., Aken'Ova, M. E., Fatokun, C. A., and Alabi, O. Y. (2013). Screening for resistance to cowpea aphid (*Aphis craccivora* Koch) in wild and cultivated cowpea (*Vigna unguiculata* L. Walp.) accessions. *Int. J. Sci. Environ. Tech.* 2, 611–621.

Steele, W. (1972). *Cowpea in Africa. PhD Thesis* (United Kingdom: University of Reading, Reading U.K).

Steiner, K. G. (1982). *Intercropping in tropical smallholder agriculture with special reference to West Africa. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH, No. 137* (Eschborn, Germany), 303 pp.

Teran, J. C. B. M., Konzen, E. R., Palkovic, A., Tsai, S. M., and Gepts, P. (2020). Exploration of the yield potential of Mesoamerican wild common beans from contrasting eco-geographic regions by nested recombinant inbred populations. *Front. Plant Sci.* 11:346:346. doi: 10.3389/fpls.2020.00346

Togola, A., Boukar, O., Belko, N., Chamarthi, S. K., Fatokun, C., Tamo, M., et al. (2017). Host plant resistance to insect pests of cowpea (*Vigna unguiculata* L. Walp.): achievements and future prospects. *Euphytica* 213, 239. doi: 10.1007/s10681-017-2030-1

Togola, A., Boukar, O., Servent, A., Chamarthi, S., Tamo, M., and Fatokun, C. (2020). Identification of sources of resistance in cowpea mini core accessions to *Aphis craccivora* Koch (Homoptera: Aphididae) and their biochemical characterization. *Euphytica* 216, 1–15. doi: 10.1007/s10681-020-02619-5

Upadhyaya, H. D., Gowda, C. L. L., Buhariwalla, H. K., and Crouch, J. H. (2006). Efficient use of crop germplasm resources: Identifying useful germplasm for crop improvement through core and mini-core collections and molecular marker approaches. *Plant Genet. Res.* 4 (1), 25–35. doi: 10.1079/PGR2006107

Van roekel, R. J., and Purcell, L. C. (2016). *Understanding and increasing soybean yields. Proceedings of the Integrated Crop Management Conference 4*. Ames, IA, USA: Iowa State University. pp 31–35. Available at: https://lib.dr.iastate.edu/icm/2016/proceedings/4.

Wang, M. L., Barkley, N. A., Gillaspie, G. A., and Pederson, G. A. (2008). Phylogenetic relationships and genetic diversity of the USDA *Vigna* germplasm collection revealed by gene-derived markers and sequencing. *Genet. Res. Camb.* 90, 467–480. doi: 10.1017/S0016672308009889

Wehmeyer, A. S., Lee, R. B., and Whiting, M. (1969). The nutrient composition and dietary importance of some vegetable foods eaten by the Kung Bushmen. *South Afr. J. Nutr.* 5, 1529–1530.

Wei, B., Zhang, J., Pang, C., Yu, H., Guo, D., Jiang, H., et al. (2015). The molecular mechanism of sporocyteless/nozzle in controlling Arabidopsis ovule development. *Cell Res.* 25, 121. doi: 10.1038/cr.2014.145

Weithal, D. M., and Gurel, F. (2016). "Plant genome editing and its application in cereals," in *Genetic Engineering - An Insight into the Strategies and Applications*. Ed. F. Jamal (London, UK: Intech Open). Retrieved from: https://www.intechopen.com/books/genet-engineering-an insight into strategies and applications/plant-genome editing and its application in cereals.

Xiong, H., Shi, A., Mou, B., Qin, J., Motes, D., Lu, W., et al. (2016). Genetic diversity and population structure of cowpea (*Vigna unguiculata* L. Walp.). *Plos One* 11 (8), e0160941. doi: 10.1371/journal.pone.0160941

Yang, Y., Shi, J., Wang, X., Liu, G., and Wang, H. (2016). Genetic architecture and mechanism of seed number per pod in rapeseed: elucidated through linkage and near-isogenic line analysis. *Sci. Rep.* 6, 24124. doi: 10.1038/srep24124

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 September 2020
doi: 10.3389/fpls.2020.575373

Spike-Stalk Injection Method Causes Extensive Phenotypic and Genotypic Variations for Rice Germplasm

Yuanyi Hu[1,2†], Bigang Mao[1,3†], Yumei Xia[1,3†], Yan Peng[1,4], Dan Zhang[1], Li Tang[1], Ye Shao[1,4], Yaokui Li[1,4] and Bingran Zhao[1,3,4]*

[1] State Key Laboratory of Hybrid Rice, Hunan Hybrid Rice Research Center, Changsha, China, [2] Molecular Breeding Laboratory, National Innovation Center of Saline-Alkali Tolerant Rice in Sanya, Sanya, China, [3] Long Ping Branch, Graduate School of Hunan University, Changsha, China, [4] College of Agricultural, Hunan Agricultural University, Changsha, China

Genetic diversities or favorable genes within distantly related species are the important resources for crop genetic improvement and germplasm innovation. Spike-Stalk injection method (SSI) has long been applied in rice genetic improvement by directly introducing genetic materials from non-mating donor species, while its inheritance patterns and the underlying mechanisms are poorly elucidated. In this study, a rice variant *ERV1* with improved yield-related traits was screened out in the way of introducing genomic DNA of *Oryza eichingeri* (2n=24, CC genome) into *RH78* (*Oryza sativa L*. 2n=24, AA genome) using SSI method. Genome-wide comparison revealed that the genomic heterozygosity of *ERV1* was approximately 8-fold higher than *RH78*. Restriction-site associated DNA sequencing technology (RAD-seq) and association analysis of the *ERV1* inbred F2 population identified 5 quantitative trait loci (QTLs) regions responsible for these yield-related traits, and found that genomic heterozygosity of *ERV1* inbred lines was significantly lower than *ERV1*, while spontaneous mutation rate of the *ERV1* inbred lines was significantly higher than *ERV1*. Our results preliminarily uncovered the inheritance patterns of SSI variant rice, and the potential genomic regions for traits changes, which yielded novel insights into the mechanisms of SSI method, and may accelerate our understanding of plant genome evolution, domestication, and speciation in nature.

Keywords: rice, genome sequencing, genetic diversity, exogenous DNA transfer, quantitative trait locus mapping

OPEN ACCESS

Edited by:
Jaime Prohens, Universitat Politècnica de València, Spain

Reviewed by:
Kejian Wang, Chinese Academy of Agricultural Sciences, China
Dejun Li, Chinese Academy of Tropical Agricultural Sciences, China

***Correspondence:**
Bingran Zhao
brzhaohhrrcrice@163.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *23 June 2020*
Accepted: *08 September 2020*
Published: *25 September 2020*

Citation:
Hu Y, Mao B, Xia Y, Peng Y, Zhang D, Tang L, Shao Y, Li Y and Zhao B (2020) Spike-Stalk Injection Method Causes Extensive Phenotypic and Genotypic Variations for Rice Germplasm. Front. Plant Sci. 11:575373. doi: 10.3389/fpls.2020.575373

INTRODUCTION

Rice (*Oryza sativa L.*) is one of the most important food crops in the world, providing 21% of the total calorie intake of the global population and up to 76% of calorie intake in Southeast Asia(Fitzgerald et al., 2009; Chen, 2010). Increasing rice yield is therefore of great importance to global food security. Two breakthroughs in rice yield that led to major increases in rice productivity over the last century (Yuan and Virnani, 1989) include the application of semi-dwarf rice varieties in the 1950s–1960s (Peng, 1994; Ashikari et al., 2005), and the development of Three-line and Two-line hybrid rice (Zhou et al., 2012; Chen and Liu, 2014), which were commercially developed in the 1970s and 1990s, respectively. These breakthroughs were mainly due to the discovery and application of new genes or novel alleles from wild rice species (Brar and Khush, 1997; Zhao et al., 2007). Genetic diversities or favorable genes from

distantly related species are therefore crucial for genetic improvement and germplasm innovation in rice, and can help guarantee the sustained growth of rice yield in the future (Gill et al., 2011).

Transferring genetic materials from wild relatives of rice into modern cultivated rice is a promising way to use genetic information and favorable genes (Zamir, 2001; Wang et al., 2017). In the genus *Oryza*, the wild species can be used to genetically improve rice because they have abundant genetic variation with many agronomical useful genes (Chen, 2010; Jena, 2010). Generally, researchers have relied on distant hybridization to introduce "exotic" allelic diversity, or on genetic engineering to introduce "foreign" genes (Zhou et al., 1983; Zhao et al., 2005) from wild rice. Over the last few decades, several genes, or QTLs controlling traits of economic importance, such as cytoplasmic male sterility (CMS), yield-related QTLs, resistance of bacterial leaf blight, blast and brown plant-hopper, have been tagged or cloned (Wang et al., 2005), and transferred from wild rice into elite rice cultivars by means of distant hybridization and genetic engineering. However, several incompatibility barriers such as low crossability, linkage drag, and limited recombination between chromosomes of wild and cultivated species limit the application of distant hybridization. The limited number of cloned genes or QTLs controlling traits of economic importance, and the security of genetically modified food also hampered the development and large-scale application of genetic engineering.

In the past three decades, a large number of new crop germplasms with improved characteristics in yield, plant type, grain quality, and stress tolerance have been created in China (Hu et al., 2014; Peng et al., 2016), benefiting from a widely used series of methods called Exogenous DNA Transfer (EDT) methods, mainly including pollen tube pathway method (Zhou et al., 1983; Wang et al., 2005), seed immersion method (Luo and Wu, 1989), and the Spike-Stalk Injection method (de la Peña et al., 1987; Zhao et al., 2007). Researchers previously proved that genomic DNA of wild species could be introduced into the rice panicle at the differentiation stage, the ovary at the meiosis stage, or into dry seeds at the water absorbency stage (Zhao et al., 2007). Zhao et al. (2005) analyzed a genetically stable variant, *YVB*, that was developed from a mutant of *V20B* (a rice maintainer line) through transformation of genomic DNA of wild rice (*Oryza minuta*). They found high DNA polymorphisms between *YVB* and *V20B*, and preliminary confirmed the integration of special DNA fragment from *O. minuta* in *YVB* genome. Meanwhile, they identified the genes responsible for the trait changes in *YVB* relative to *V20B*, and revealed that allelic variation might be the main factor contributing to the abundant phenotypic changes of this variant (Peng et al., 2016). However, the possible molecular mechanisms underlying trans-species transfer of genetic information, especially genome-wide genetic variation and inheritance of variants in the early generations compared to the recipients, are poorly elucidated.

Comparative genomic analyses among closely related species have allowed for insight into plant gene and genome evolution (Bennetzen, 2007; Paterson et al., 2010). For instance, studies of yeast, *Drosophila*, and human genomes have revealed mechanisms of gene and genome evolution in fungi and animals (Scannell et al., 2006; Clark et al., 2007). Meanwhile, comparison of *O. nivara*, *O. glaberrima*, *O. barthii*, *O. glumaepatula*, and *O. meridionalis* genomes unveiled how morphological and reproductive diversity have arisen in specific lineages as a result of expansions or contractions of gene families, providing more evidence for the processes of adaptation and speciation (Zhang et al., 2014). However, current comparative genomic analyses mainly focus on the comparison between different varieties, or between closely related species of flowering plants; few are carried out to query the fundamental rules of inheritance from the parent to the offspring by means of horizontal and/or vertical transmission, especially in the early generations.

In this study, we conducted whole-genome sequencing and comparative genomic analysis of *ERV1*, an original variant with greatly improved yield traits, and an inbred *ERV1* population. *ERV1* was derived from an *indica* variety *RH78* by transferring genomic DNA of *O. eichingeri* to the recipient rice *RH78* through the SSI method. Our results would accelerate the process for unveiling the possible mechanisms underlying SSI method, and may enhance our understanding of plant genome evolution, domestication, and speciation in nature.

MATERIALS AND METHODS

Spike-Stalk Injection Method

SSI method is a modified technique based on the procedure invented by *Pena* (de la Peña et al., 1987). To perform the technique, we waited until recipient rice had just undergone meiosis (up to 2–4 days after meiosis) and selected spikes that had a distance of 1–5 cm from the petiole of the flag leaf to the top of the 2nd leaf. Part of the leaf sheath of the top 3rd leaf was removed, and the uppermost internode was exposed. Donor DNA (50–100 μl) at a concentration of about 500 ng/μl was mixed with 0.2×SSC solution, and 0.02 M $CaCl_2$ solution was then injected into the internode with a microsyringe at an inclined angle of about 30–45°. The injected spike was immediately bagged with parchment (to avoid the pollution of exogenous pollen) until harvest. The harvested seeds were planted uniparted. From initial heading stage to the maturity stage, rice plants with significant trait differences from recipient rice were considered variant rice.

Plant Materials and Traits Measurement

The variant *ERV1* was discovered in the experimental field of the Hunan Hybrid Rice Research Center, Changsha, in September 2013. The recipient rice *RH78* was planted in the same field at the same time, and with the same management to compare agronomic traits. *ERV1* was selfed to generate the F2 selfing population. *ERV1*, five mature plants of *RH78*, and 216 *ERV1* inbred F2 lines were sampled to determine mean values of yield characters, including heading stage (HS), flag leaf length (FLL), flag leaf width (FLW), plant height (PH), effective tiller number (TN), main panicle length (PL), spikelet per panicle (SP), seed setting rate (SSR), and 100-grain weight (GW). HS refers to the time from seed sowing to first-earing phase. FLL and FLW represent the maximum length and width of the flag leaf measured by an electronic digital caliper. PH

refers to the distance from the ground to the highest point of one plant. TN is the number of tillers that can develop mature panicles. SP is the spikelet number of the main panicle. SSR is the number of filled grains divided by the spikelet number of the main panicle. GW refers to the weight of 100 filled grains of the main panicle. Student's *t*-tests were used for all trait data of *ERV1* and *RH78* to identify significant differences.

DNA Extraction and Genome Sequencing

Genomic DNA of *ERV1*, *RH78*, and 216 *ERV1* inbred F2 lines were extracted following the standard procedure of a PureLink™ Genomic Plant DNA Purification Kit (Invitrogen). For genome re-sequencing, the genomic DNA of *RH78* and *ERV1* were extracted, and then fragmented randomly. The DNA fragments of desired length (500 base pairs) were visualized with agarose gel electrophoresis, adapter primers were added to generate DNA clusters, and DNA clusters were subjected to HiSeq 2000 paired-end sequencing. For RAD-seq of *ERV1* inbred F2 lines, the genomic DNA of each of the 216 inbred lines was digested with the TaqI restriction endonuclease, and DNA fragments between the sizes of 300 and 700 bp were isolated with agarose gel electrophoresis. Adapters were ligated to the isolated DNA fragments for cluster preparation, and then subjected to the HiSeq 2000 platform for 100 base paired-end sequencing (Davey et al., 2011).

Bioinformatics Analysis Process for Re-Sequencing

The raw data generated by the HiSeq 2000 sequencer was subjected to a series of filters to generate a high-quality data set (Huang et al., 2009; Gu et al., 2013). The high-quality data was then aligned to the *Nipponbare* genome (allowing three mismatches) *in silico* using the SOAPaligner software (Yu et al., 2001; Ouyang et al., 2005). Sequencing coverage and depth of each sample were calculated based on the alignment result. Based on the consensus sequence between the reference and each of the sequenced samples, single nucleotide polymorphisms (SNPs) in *ERV1*, *RH78*, and each of the *ERV1* F2 inbred lines relative to the *Nipponbare* genome were identified using the SOAPsnp software with the following filtering steps: mapping quality > 20; coverage depth between 1 and 300; distance of adjacent variation > 5 bp; and copy number < 2. According to the filtered SNPs set, SNPs between *ERV1* and *RH78* were further identified with the following filter steps: target mapping quality > 20; the sequencing depth of homologous site should be > 2; and the sequencing depth of heterozygous sites should be > 6. The final SNP set was then annotated according to the reference genome sequence of the *japonica* cultivar *Nipponbare*.

Bioinformatics Analysis Process for RAD-Seq

The sequencing reads from each line were obtained according to the sequence tag, and then mapped to the *Nipponbare* genome sequence for alignment using the SOAPaligner software (Gu et al., 2013). The results of SOAP alignment were then transformed and subjected to SAMtools (Li et al., 2009) to generate the input files suitable for realSFS software (version 0.983). Nucleotide sites with the likelihood of a genotype for each line with a probability higher than 0.95 and a total population depth higher than 40 were extracted as candidate SNPs. The recombination breakpoints of the F2 population were then identified according to the method developed by the research group of Huang et al. (2009) with some modification (Duan et al., 2013). The Bin map of each line was constructed using a PERL script, and QTLs were identified by composite interval mapping using MapQTL5 software.

Estimate of Genomic Heterozygosity and Spontaneous Mutation Rate

Mutation sites in *ERV1* refer to the genotype in *ERV1* that was different from the genotype of *RH78*. For example, if the genotype of *RH78* was A, and *ERV1* was homologous G, or C, or T, this genotype was defined as a homologous SNP; if the *ERV1* was heterozygous A/G, A/C, A/T, this genotype was defined as a heterozygous SNP; if the genotype of *ERV1* was T/C, or T/G, or C/G, we defined it as a mutation SNP. The genomic heterozygosity was calculated as the number of heterozygous SNPs divided by the length of covered bases for the sequencing sample. One site where the genotype of >20 F2 line was not identical to the genotype of *RH78* and/or *ERV1* was defined as a spontaneous mutation. The spontaneous mutation rate of the *ERV1* inbred line was calculated as the number of spontaneous mutations (>20 lines) in the F2 population divided by the length of covered bases for the sequencing sample.

RESULTS

Trait Comparison Between *ERV1* and *RH78*

A large proportion of traits in *ERV1* were the intermediate forms between the recipient *RH78*, and the donor *O. eichingeri*. This made ERV1 seems to be a hybrid F1 cross between *RH78* and *O. eichingeri* (**Figure 1A**). We surveyed the main agronomic characters of *RH78* and *ERV1* in the mature stage and found that *ERV1* had higher PH (**Figure 1B**), longer inflorescence length (**Figure 1C**), longer and wider flag leaves (**Figures 1D, E**), larger panicles (**Figure 1F**), and a higher SSR (**Figure 1G**) compared to *RH78*. The spikelet number and SSR of *ERV1* were significantly higher than *RH78*, while the TN and GW of *ERV1* were almost equal to *RH78* (**Table S1**), meaning that *ERV1* has a higher yield potential in breeding practice.

Genome Sequencing of *ERV1* and *RH78*

Using Illumina sequencing technology with a whole-genome shotgun sequencing approach, a total of 11.4 and 10.2 G bases (Gb) of raw sequencing data were generated for *ERV1* and *RH78*, respectively (Raw sequence data of *ERV1* and *RH78* have been deposited in the NCBI Short Read Archive with access number SRA7923627). After a series of corrections and filters, 10.81 Gb and 10.10 Gb clean data for *ERV1* and *RH78*, respectively, were retained for downstream analysis. The clean data were then aligned to the *Nipponbare* genome, a high-quality rice reference genome with an effective genome size of 372,317,567 base pairs (bp), for whole-genome comparison using the SOAPaligner software algorithm

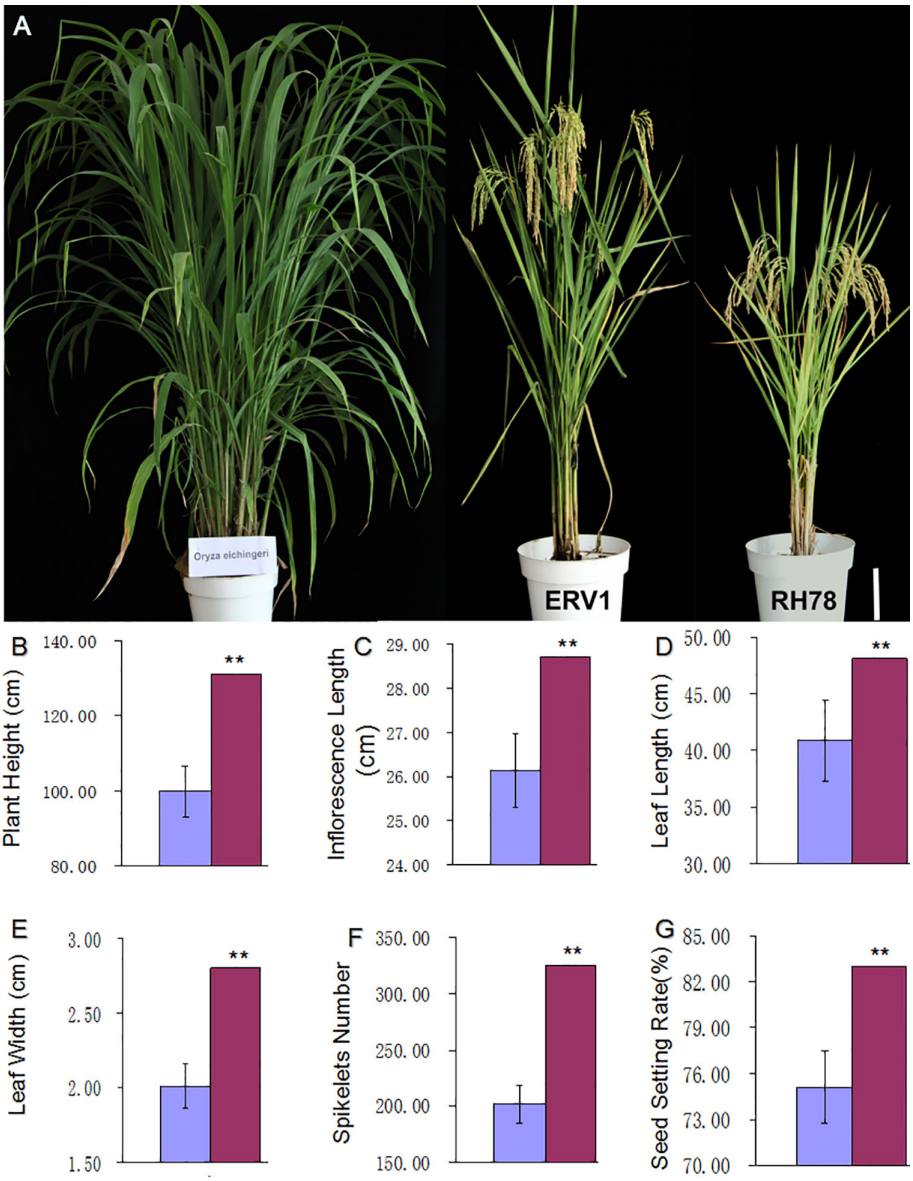

FIGURE 1 | Trait comparison of *ERV1* and *RH78*. **(A)** Plant phenotype of the donor rice *O. eichingeri* (left), the recipient rice *RH78* (right), and the variant rice *ERV1* (middle), the white bar = 20 cm. **(B–G)** is the comparison of main agronomic characters of *RH78* and *ERV1* in the mature stage; **(B)** plant height, **(C)** inflorescence length, **(D)** flag leaf length, **(E)** flag leave width, **(F)** spikelet number, and **(G)** seed setting rate. The main agronomic characters of RH78 and ERV1 were compared and analyzed by one-way ANOVA method, and the symbol "**" indicate significant difference detected by LSD tests at $P < 0.01$.

(SOAP2, version 2.20). The *ERV1* and *RH78* clean data covered 312,397,330 bp (83.66%) and 303,082,602 bp (81.06%) of the *Nipponbare* reference genome, providing an effective sequencing depth of 29.03-fold for *ERV1* and 27.14-fold for *RH78* (**Table 1**).

Identification and Annotation of SNPs

The wide genomic coverage and high sequencing depth of *ERV1* and *RH78* sequencing data were considered suitable for use in the identification of single nucleotide polymorphisms (SNPs). After a series of corrections and filtering steps, a total of 1,397,635 SNPs, including 1,257,597 homozygous SNPs (homo-SNPs), and 101,822 heterozygous SNPs (hetero-SNPs) were detected in *RH78* compared to *Nipponbare* genome sequence. Hetero-SNPs accounted for 7.29% of the total SNPs, and the genomic heterozygosity of *RH78* genome was 3.36×10^{-4} (**Table S2**). For *ERV1*, 1,898,506 SNPs, including 1,087,532 homo-SNPs and 810,974 hetero-SNPs were identified relative to *Nipponbare* genome. The proportion of hetero-SNPs was 42.72% of the total SNPs, and the genomic heterozygosity of *ERV1* was as high as 25.96×10^{-4} (**Table S3**), approximately 8-fold higher than that of *RH78*.

TABLE 1 | Statistics of the sequencing data of *ERV1* and *RH78*.

Sample	Reference length (bp)	Raw data (Gb)	Clean data (Gb)	Covered base (bp)	Sequencing depth (fold)	Coverage (%)
RH78	372,317,567	10.31	10.10	303,082,602	27.14	81.06
ERV1	372,317,567	11.41	10.81	312,397,330	29.03	83.66

Reference length refers to the assembled chromosome length of the Nipponbare genome (IRGSPv6).

The distribution patterns of SNPs in *ERV1* and *RH78* genomes were presented using the SNP density, which was calculated as the total number of SNPs divided by the total covered bases in each chromosome. The average total SNP density across the whole genome was 4.65 SNPs/kb in the *RH78* genome (**Table S2**), slightly lower than 6.15 SNPs/kb in the *ERV1* genome (**Table S3**). While, the density of hetero-SNPs in *ERV1* (2.64 hetero-SNPs/kb) was much higher than that in *RH78* (0.34 hetero-SNPs/kb). At the chromosome level, the highest density of total SNPs and homo-SNPs was identified in chromosome 11 for both *RH78* and *ERV1*, and the lowest in chromosome 5 for *RH78* and chromosome 4 for *ERV1*. The highest density of hetero-SNPs was discovered in chromosome 11 for *RH78*, but in chromosome 12 for *ERV1*, and the lowest in chromosome 1 for *RH78* and chromosome 11 for *ERV1* (**Figure 2**). To answer whether these SNPs in *ERV1* came from O. *eichingeri*, 2 DNA fragments (less than 1000 bp) in each chromosome of *Nipponbare* genome that were simultaneous covered by the sequencing reads of *RH78*, ERV1, and *O. eichingeri* were randomly selected, but only 6 DNA fragments were simultaneous amplified in *RH78*, *ERV1*, and *O. eichingeri* using PCR amplification method (**Figure 3A**), of which only one DNA fragment was sequenced with good quality. Comparing the sequencing data of this DNA fragment among *RH78*, *ERV1*, and *O. eichingeri* found that 3 heterozygous SNPs in *ERV1* relative to *RH78* were totally identical to *O. eichingeri* (**Figure 3B**), which partially proved that these SNPs in *ERV1* came from *O. eichingeri*. However, more experiments are needed.

The genome-wide SNPs in *ERV1* compared to *RH78* were filtered out from the SNP data mentioned above. A total of 455,244 SNPs were identified in *ERV1*, of which 98.79% (449,726) SNPs were heterozygous (**Figure S1**); only 5,518 (1.21%) SNPs were homozygous (**Table 2**). The hetero-SNPs in *ERV1* were annotated according to the reference genome sequence of *japonica* cultivar *Nipponbare* because they were a very high proportion (98.79%) out of the total SNPs. After filtering, a total of 14,955 Non-synonymous hetero-SNPs (Ns-hetero-SNPs) that might affect the expression and function of the allele were identified in 7,325 genes (**Table S4**), of which 94.67% (6,934) genes had less than 5 Ns-hetero-SNPs. These genes were then functionally annotated using gene ontology (GO) classification (http://www.geneontology.org) and were categorized into 39 functional groups based on sequence homology. In each of the three main categories (biological process, molecular function, and cellular component) of the GO classification, there were 19, 11, and 9 functional groups, respectively (**Figure 4**).

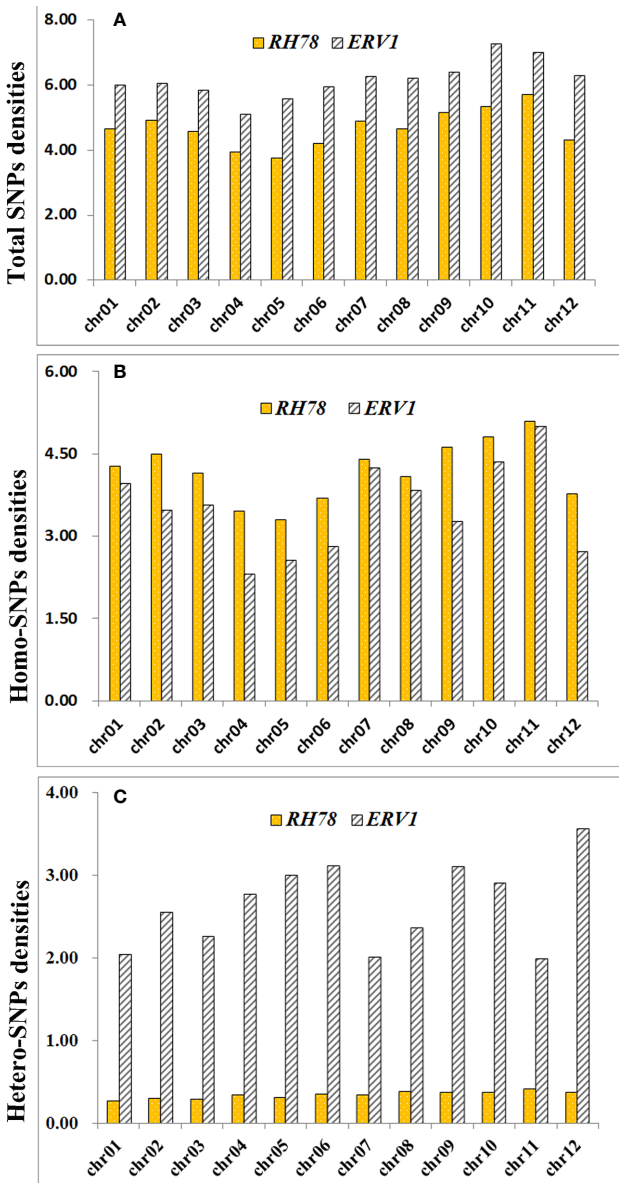

FIGURE 2 | Distribution of single nucleotide polymorphism (SNP) densities of *ERV1* and *RH78* compared to the *Nipponbare* genome in all chromosomes. **(A)** Total SNPs densities of *RH78* (yellow bars) and *ERV1* (gray bars) in 12 chromosomes; **(B)** homo-SNPs densities of *RH78* (yellow bars) and *ERV1* (gray bars) in 12 chromosomes; **(C)** hetero-SNPs densities of *RH78* (yellow bars) and *ERV1* (gray bars) in 12 chromosomes. All horizontal axes of **(A–C)** represent the chromosome number, and the vertical axes of **(A–C)** represent the SNP densities on each chromosome.

Sequencing and Association Analysis of the *ERV1* Inbred F2 Population

Restriction-site associated DNA sequencing technology (RAD-seq) was carried out on the *ERV1* inbred F2 population. The

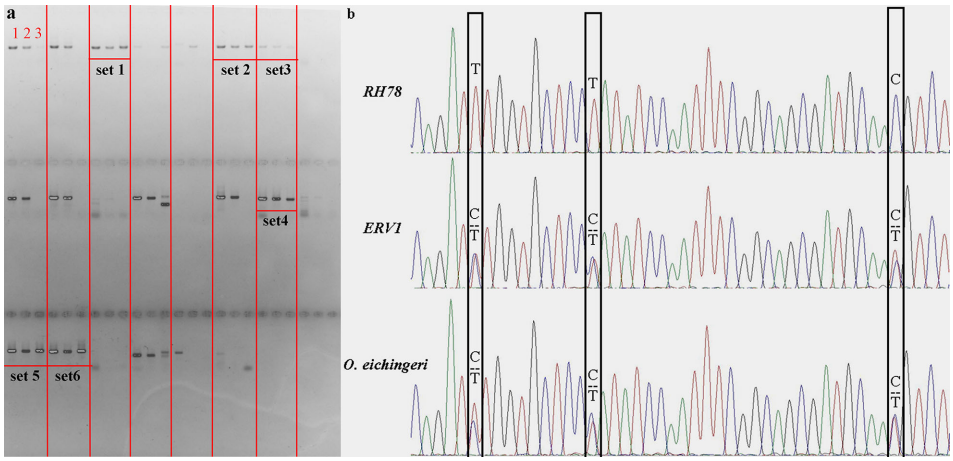

FIGURE 3 | PCR amplification, Sanger sequencing and comparison of the homologous sequences among *RH78*, *ERV1*, and *O. eichingeri*. **(A)** PCR amplification of 24 randomly selected DNA fragments in *RH78*, *ERV1*, and *O. eichingeri*. The PCR bands in each set were presented with the same order as number 1 for RH78, number 2 for ERV1 and number 3 for *O. eichingeri*. The 6 DNA fragments that were simultaneous amplified in *RH78*, *ERV1* and *O. eichingeri* were noted as set1, set2, set3, set4, set5, and set6, and only PCR products of set4 were sequenced with good quality. **(B)** Result of Sanger sequencing and comparison of the set4 sequences. The single nucleotide polymorphisms (SNPs) or mutations among *RH78*, *ERV1*, and *O. eichingeri* were labeled with black boxes.

TABLE 2 | Statistics of single nucleotide polymorphisms (SNPs) in *ERV1* relative to *RH78*.

Chromosome	SNP number	Homo-SNPs	Hete-SNPs
chr01	43,225	343	42,882
chr02	50,240	738	49,502
chr03	45,020	495	44,525
chr04	42,737	656	42,081
chr05	48,879	391	48,488
chr06	51,818	564	51,254
chr07	20,626	236	20,390
chr08	24,943	303	24,640
chr09	34,147	518	33,629
chr10	29,288	464	28,824
chr11	16,024	225	15,799
chr12	48,297	585	47,712
Total	455,244	5,518	449,726

RAD-seq data were then used to construct a genetic linkage map with high-density SNPs for use in association analysis exploring genotype-phenotype relationships at the whole-genome level.

RAD-seq libraries of 216 *ERV1* F2 inbred lines were constructed, yielding approximately 30.85 Gb raw sequencing data (**Table S5**, Raw sequence data of *ERV1* F2 inbred lines have been deposited in the NCBI Short Read Archive with access number SRA7923627), and providing an average sequencing coverage of 5.97%, and a mean sequencing depth of 3.23-fold for each F2 line. After manual filtration, a total of 32,644 population SNPs were selected for the F2 inbred population (**Table 3**). The distribution of the population SNPs along the chromosomes was uneven, with high-density and low-density regions (**Figure S2**). The genotype and bin information of each line of the F2 population was identified according to the recombination breakpoints. Using Joinmap 4.1 software, a total of 1,465 bins were used to construct the genetic map (**Table 3**). The size of the genetic map was 1489.7 centimorgans (**Figure S3**).

According to the genetic map and the phenotypic data of the F2 population (**Table S6**), gene mapping of the 9 yield-related traits was carried out in QTLcartographer software using the composite interval mapping method. A total of 5 QTLs were detected (**Table 4**). The numbers of QTLs, as well as their contributions to phenotypic variance are summarized in **Table 4**. One QTL region, located at bin22, had an LOD value of 4.1, explained 8.4% of the PH variation, and was detected on chromosome 1. One QTL for PL was discovered on chromosome 2; it had a LOD score of 3.55 and explained 7.3% of the phenotypic variation. One QTL for GW was identified on chromosome 4, located at bin 525, had a LOD score of 3.55, and explained 7.3% of the phenotypic variation. Two QTLs for TN were discovered on chromosome 2, and 11, separately. The first QTL was located at bin260, had an LOD score of 3.9, and explained 8% of the phenotypic variation; the other QTL for TN was located at bin1207, had an LOD score of 3.53, and explained 6.8% of the phenotypic variation.

Genomic Heterozygosity and Spontaneous M in the *ERV1* Inbred F2 Lines

RAD-seq data of the *ERV1* inbred F2 population provided us the opportunity for detailed analysis of the genomic condition of each F2 line. We first analyzed the ratio of homo-SNPs out of total population SNPs for each *ERV1* inbred F2 line. After manual filtering, the genotypes of 135 F2 lines were estimated (**Table S7**). The average proportion of homo-SNPs was 83.49% of the total SNPs, much higher than the theoretical value of 75%; 3.7% (n=5) F2 lines ranged from 63.99%–70%, 26.67% (n=36) ranged from 70%–80%, 49.63% (n=67) ranged from 80%–90%, and 20% (n=27) ranged from 90%–96.02%. The lowest ratio (63.99%) of

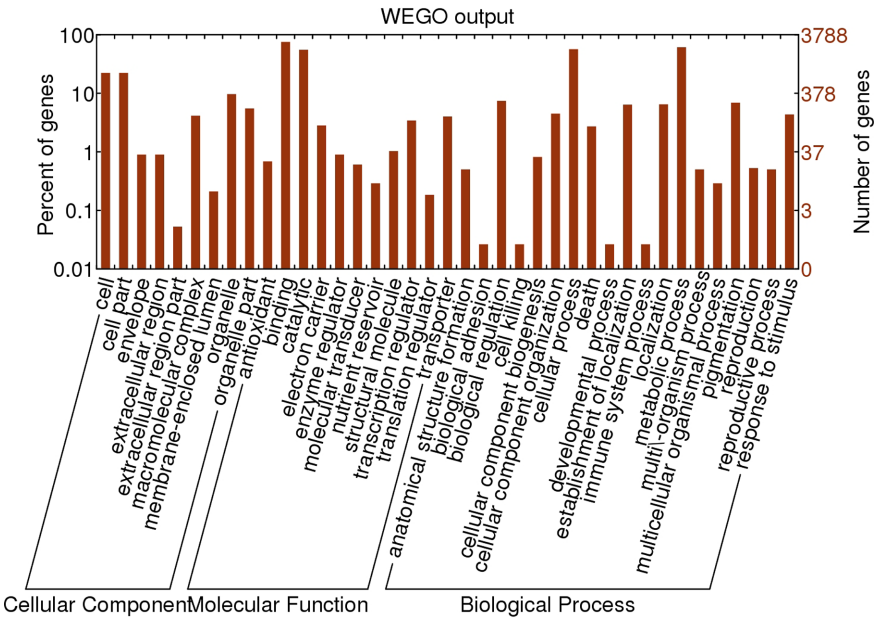

FIGURE 4 | Gene Ontology (GO) classification of the genes affected by Non-synonymous hetero-SNPs in *ERV1* compared to *RH78*. The horizontal axis represents the GO items, including cellular component, molecular function, and biological process. The left vertical axis represents the percent of genes affected by Non-synonymous hetero-SNPs of the corresponding GO item, and the right vertical axis represents the number of genes affected by Non-synonymous hetero-SNPs of the corresponding GO item.

TABLE 3 | Single nucleotide polymorphisms (SNPs) used for genotyping and bin number and genetic distance on 12 rice chromosomes for the F2 population.

Chromosome	Heterozygous SNPs in *ERV1*	Number of SNPs (population)		Number of bins	Genetic distance (cM)
		Before filtration	After filtration		
Chr01	42,881	7,523	2,816	169	234.8
Chr02	49,500	9,189	3,440	174	150.9
chr03	44,525	8,205	2,568	167	187.2
chr04	42,078	8,113	2,204	123	95.8
chr05	48,486	8,774	6,184	140	91.7
chr06	51,254	9,763	3,151	149	98.7
chr07	20,390	4,424	977	75	126.5
chr08	24,640	4,953	1,078	81	126.8
chr09	33,629	6,721	2,724	93	92.2
chr10	28,824	6,091	2,329	86	78.2
chr11	15,799	3,539	1,207	83	104.5
chr12	47,711	8,991	3,966	125	102.4
Total	449,717	86,286	32,644	1,465	1,489.7

homo-SNPs was discovered in F2 line 147, and the highest ratio (96.02%) was identified in line 163.

The genomic heterozygosity of each *ERV1* inbred F2 line was then analyzed by comparing the sequencing data with the *Nipponbare* reference genome. After excluding abnormal values, sequencing data of 196 lines out of 216 *ERV1* inbred F2 lines were left for in-depth study (**Table S8**). The mean genomic heterozygosity of the 196 lines was approximately 1.68×10^{-4}, of which 30.61% F2 lines (n=60) ranged from 0.16×10^{-4}-0.98×10^{-4}, 32.14% lines (n=63) ranged from 1.0×10^{-4}-1.96×10^{-4}, 23.98% lines (n=47) ranged from 2.04×10^{-4}-2.99×10^{-4}, 11.73% lines (n=23) ranged from 3.00×10^{-4}-3.81×10^{-4}, and 1.53% lines (n=3) ranged from 4.08×10^{-4}-4.57×10^{-4}. The lowest genomic heterozygosity (0.16×10^{-4}) was identified in line 139, and the highest value (4.57×10^{-4}) in line 167.

The spontaneous mutation rate of the *ERV1* inbred F2 lines were further estimated using the comparison results. Only the SNPs that were covered by the sequencing data of over 20 F2 lines

TABLE 4 | Summary of quantitative trait loci (QTL) information of all traits.

Trait	Chromosome	Peak bin	Start	End	Bin length (bp)	LOD	R2 (%)
PH	chr01	bin22	7779967	8059948	279,981	4.1	8.4
PL	chr02	bin263	25718520	25761663	43,143	3.55	7.3
GW	chr04	bin525	19108032	19172986	64,954	3.53	7.3
TN	chr02	bin260	25424655	25516358	91,703	3.9	8
	chr11	bin1207	15862580	16106680	244,100	3.53	6.8

were considered as candidate sites of spontaneous mutation (**Table S8**). After a series of filtering and screening, the average spontaneous mutation rate of the *ERV1* inbred F2 lines was approximately 5.94×10^{-4}, approximately 3.5-fold higher than the mean genomic heterozygosity of the *ERV1* inbred F2 population, of which 1.53% D2 lines (n=3) ranged from 3.91×10^{-4}- 4.98×10^{-4}, 55.61% lines (n=109) ranged from 5.05×10^{-4}-5.99×10^{-4}, 42.35% lines (n=83) ranged from 6.00×10^{-4}-6.97×10^{-4}, and one line had a mutation rate of 7.01×10^{-4}. The lowest spontaneous mutation rate (3.91×10^{-4}) was identified in line 4, and the highest value (7.01×10^{-4}) in line 30.

DISCUSSION

A series of previous studies have documented that SSI could induce extensive genomic variations, and result to phenotypic changes of important traits (Zhou et al., 1983; de la Peña et al., 1987; Zhao et al., 2007). They also experimently confirmed the integration of donor DNA fragments in variant genomes (Liu et al., 2000; Miao et al., 2000; Xing et al., 2004), and revealed that allelic variation of important genes might be the main factor for phenotypic changes of variant plants (Wang et al., 2013; Peng et al., 2016). Nevertheless, these studies mainly based on limited molecular markers such as RFLP, RAPD, AFLP and SSR (Shan et al., 2005; Wang et al., 2010), the extent and distribution pattern of the induced variations in a genome-wide scale and its inheritance patterns were largely obscure.

People used to use introgressive hybrid added with embryo rescue to overcome the cross incompatibility between the wild species of genus *Oryza* and cultivated rice, and to generate interspecific hybrids (Ramanujam, 1937; Nayar, 1973). Notably, the rice variant *ERV1* from SSI method seemed to a hybrid F1 cross between *RH78* and *O. eichingeri*, the latter belongs to the wild species of genus *Oryza*, has 2n=24 chromosomes representing CC genome, and is incompatible with *O. sativa* genome (Brar and Khush, 1997). This proved that SSI method could realize the exchange or transfer of genetic materials without the effects of reproductive isolation. Meanwhile, the values of plant height, panicle length, spikelet number, and length and width of flag leaf of the rice variant *ERV1* in this study are all higher than the rice recipient *RH78* but lower than the donor *O. eichingeri*, and the genomic heterozygosity of *ERV1* was approximately 8-fold higher than *RH78*. This phenomenon was similar to the introgressive hybridization by pollens of alien sexually incompatible species (Hongyan et al., 2009; Dowen et al., 2012), which would provoke genome-wide and extensive genomic changes, and even results in important phenotypic novelties (Wang et al., 2013; Sakai et al., 2013; Schiestl and Johnson, 2013). We also noted that the genomic heterozygosity rapidly decreased from as high as 25.96×10^{-4} in *ERV1* to an average of approximately 1.68×10^{-4} in *ERV1* inbred F2 lines, this was much different from introgressive hybridization, and indicated that the novel phenotypes or traits of these rice variants generated by SSI method could be quickly fixed, and utilized in crop genetic improvements.

Spontaneous mutation is an important factor in genome evolution and phenotypic variation (Zhang et al., 2016). In this study, we identified a total of 5,518 homozygous SNPs in *ERV1* relative to *RH78*, resulting in a spontaneous mutation rate of 0.182×10^{-4} in *ERV1*. The average spontaneous mutation rate of *ERV1* inbred F2 lines was estimated to be approximately 5.94×10^{-4}, significantly increased compared to *ERV1*, and approximately 3.5-fold higher than the mean genomic heterozygosity of the *ERV1* inbred F2 lines. This interesting phenomenon implies a highly unpurified genetic background, and high genetic variability of the *ERV1* genome.

In conclusion, our study presented a method that could quickly and comprehensively improve important agronomic traits of cultivated rice, and preliminarily unveiled the fundamental rules of the genetic variation and inheritance of variation in the early generation *via* multi-omics strategies and comparative genomic analysis. The results enhanced our understanding of SSI methods, and may accelerate our understanding of plant genome evolution, domestication, and speciation in nature.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA658863.

AUTHOR CONTRIBUTIONS

YH, BM, and YX extracted DNA, identified the genotype of F2 population, and wrote the manuscript. YP, DZ and YS performed SNP data analysis and QTL detecting. LT and YL validated some results of the bioinformatics analysis. BZ designed the experiments and supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This research was financially supported by the Natural Science Foundation of Hunan province (2020JJ5399), the National Major Special Project on New Varieties Cultivation for Transgenic Organisms (Grant No. 2016ZX08001-004), the Major Science and Technology Projects of Hunan province (2018NK1020, 2018NK1010), and the 2020 Research Program of Sanya Yazhou Bay Science and Technology City (Grant No.202002006).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.575373/full#supplementary-material

REFERENCES

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., et al. (2005). Cytokinin Oxidase Regulates Rice Grain Production. *Science* 309 (5735), 741–745. doi: 10.1126/science.1113373

Bennetzen, J. L. (2007). Patterns in grass genome evolution. *Curr. Opin. Plant Biol.* 10 (2), 176–181. doi: 10.1016/j.pbi.2007.01.010

Brar, D. S., and Khush, G. S. (1997). Alien introgression in rice. *Plant Mol. Biol.* 35 (1-2), 35–47. doi: 10.1023/A:1005825519998

Chen, L., and Liu, Y. G. (2014). Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* 65 (1), 57–71. doi: 10.1146/annurev-arplant-050213-040119

Chen, Z. J. (2010). Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* 15 (2), 57–71. doi: 10.1016/j.tplants.2009.12.003

Clark, A.G., Eisen, M. B., Smith, D. R., Bergman, C. M., Oliver, B., Markow, T. A., et al. (2007). Evolution of genes and genomes on the Drosophila phylogeny: Drosophila 12 Genomes Consortium. *Nature* 450, 203.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., Blaxter, M. L., et al. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12 (7), 499–510. doi: 10.1038/nrg3012

de la Peña, A., Ouml, H., and Schell, J. (1987). Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nat. (London)* 325, 274–276. doi: 10.1038/325274a0

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., and Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. U. States America* 109 (32), E2183–E2191. doi: 10.1073/pnas.1209329109

Duan, M., Sun, Z., Shu, L., Tan, Y. N., Yuan, D. Y., Sun, X. W., et al. (2013). Genetic analysis of an elite super-hybrid rice parent using high-density SNP markers. *Rice* 6 (1), 1–15. doi: 10.1186/1939-8433-6-21

Fitzgerald, M. A., Mccouch, S. R., and Hall, R. D. (2009). Not just a grain of rice: the quest for quality. *Trends Plant Sci.* 14 (3), 133–139. doi: 10.1016/j.tplants.2008.12.004

Gill, B. S., Friebe, B. R., and White, F. F. (2011). Alien introgressions represent a rich source of genes for crop improvement. *Proc. Natl. Acad. Sci. U.S.A.* 108 (19), 7657–7658. doi: 10.1073/pnas.1104845108

Gu, S., Fang, L., and Xu, X. (2013). Using SOAPaligner for Short Reads Alignment. *Curr. Protoc. Bioinf.* 44, 11.11.1–17. doi: 10.1002/0471250953.bi1111s44

Hongyan, W., Wang, H. Y., Chai, Y., Chu, X. C., Zhao, Y. Y., Wu, Y., et al. (2009). Molecular characterization of a rice mutator-phenotype derived from an incompatible cross-pollination reveals transgenerational mobilization of multiple transposable elements and extensive epigenetic instability. *BMC Plant Biol.* 9, 63. doi: 10.1186/1471-2229-9-63

Hu, Y., Mao, B., Peng, Y., Sun, Y. D., Pan, Y. L, Xia, Y. M., et al. (2014). Deep re-sequencing of a widely used maintainer line of hybrid rice for discovery of DNA polymorphisms and evaluation of genetic diversity. *Mol. Genet. Genomics* 289 (3), 303–315. doi: 10.1007/s00438-013-0807-z

Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., et al. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Res.* 19 (6), 1068–1096. doi: 10.1101/gr.089516.108

Jena, K. K. (2010). The species of the genus Oryza and transfer of useful genes from wild species into cultivated rice, O. sativa. *Breed Sci.* 60 (5), 518–523. doi: 10.1270/jsbbs.60.518

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map (SAM) Format and SAMtools. *Bioinform* 25 (16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Liu, Z., Dong, Y., and Liu, B. (2000). Isolation of Zizania latifolia Species_specific DNA Sequences and Their Utility in Identification of Z. latifolia DNA Introgressed into Rice. *J. Integr. Plant Biol.* 42 (3), 324–326. doi: 10.1016/S0168-9452(99)00223-X

Luo, Z., and Wu, R. (1989). A simple method for the transformation of rice via the pollen-tube pathway. *Plant Mol. Biol. Rep.* 7 (3), 240–240. doi: 10.1007/BF02668692

Miao, J., Zhao, M., and Li, W. (2000). Molecular proof of wheat transformed by total DNA of leymus racemosus. *J. Genet. Genomics* (7), 621–627.

Nayar, M. M. (1973). Origin and cytogenetics of rice. *Adv. Genet.* 17, 153–292. doi: 10.1016/S0065-2660(08)60172-8

Ouyang, S., Liu, J., Jones, K. M., Gansberger, K., Moffat, K., Hill, J., et al. (2005). The map-based sequence of the rice genome. *Nature* 436 (7052), 793–800. doi: 10.1038/nature03895

Paterson, A. H., Freeling, M., Tang, H., and Wang, X. (2010). Insights from the comparison of plant genome sequences. *Annu. Rev. Plant Biol.* 61 (1), 349–372. doi: 10.1146/annurev-arplant-042809-112235

Peng, Y., Hu, Y., Mao, B., Xiang, H. T., Shao, Y., Pan, Y. L., et al. (2016). Genetic analysis for rice grain quality traits in the YVB stable variant line using RAD-seq. *Mol. Genet. Genomics* 291 (1), 297–307. doi: 10.1007/s00438-015-1104-9

Peng, S. (1994). *Evolution of the new plant ideotype for increased yield potentia[C] Breaking the yield barrier. Proceedings of a workshop on rice yield potential in favorable environments, IRRI.*

Ramanujam, S. (1937). Cytogenetical studies in Oryzeae. III. Cytogenetical behaviour of an interspecific hybrid in Oryza. *J. Genet.* 35, 223–258. doi: 10.1007/BF02982352

Sakai, S., Kawakita, A., Ooi, K., and Inoue, T. (2013). Variation in the strength of association among pollination systems and floral traits: Evolutionary changes in the floral traits of Bornean gingers (Zingiberaceae). *Am. J. Bot.* 100 (3), 546–555. doi: 10.3732/ajb.1200359

Scannell, D. R., Byrne, K. P., Gordon, J. L., Wong, S., and Wolge, K. H. (2006). Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* 440, 341–345.

Schiestl, F. P., and Johnson, S. D. (2013). Pollinator-mediated evolution of floral signals. *Trends Ecol. Evol.* 28 (5), 307–315. doi: 10.1016/j.tree.2013.01.019

Shan, X. H., Liu, Z. L., Dong, Z. Y., Wang, Y. M., Chen, Y., Lin, X. Y., et al. (2005). Mobilization ofthe active MITE transposons mPing and Pong in rice by introgressionfrom wild rice (Zizania latifolia Griseb.). *Mol. Biol. Evol.* 22, 976–990. doi: 10.1093/molbev/msi082

Wang, Y. M., Dong, Z. Y., Zhang, Z. J., Lin, X. Y., Shen, Y., Zhou, D. W., et al. (2005). Extensive de Novo genomic variation in rice induced by introgression from wild rice (Zizania latifolia Griseb.). *Genetics* 170 (4), 1945–1956. doi: 10.1534/genetics.105.040964

Wang, N. N., Wang, H. Y., Wang, H., Zhang, D., Wu, Y., Ou, X. F., et al. (2010). Transpositional reactivation of the Dart transposon family in rice lines derived from introgressive hybridization with Zizania latifolia. *BMC Plant Biol.* 10, 190. doi: 10.1186/1471-2229-10-190

Wang, Z. H., Zhang, D., Bai, Y., Zhang, Y. H., Liu, Y., Wu, Y., et al. (2013). Genomewide Variation in an Introgression Line of Rice-Zizania Revealed by Whole-Genome re-Sequencing[J]. *PloS One* 8 (9), e74479. doi: 10.1371/journal.pone.0074479

Wang, C., Hu, S., Gardner, C., and Lübberstedt, T. (2017). Emerging Avenues for Utilization of Exotic Germplasm. *Trends Plant Sci.* 22 (7), 624–663. doi: 10.1016/j.tplants.2017.04.002

Xing, Q., Zhao, B. R., Xu, K., Yang, H. H., Liu, X., Wang, S. W., et al. (2004). Test of agronomic characteristics and amplified fragment length polymorphism analysis of new rice germplasm developed from transformation of genomic DNA of distant relatives. *Plant Mol. Biol. Rep.* 22 (2), 155–164. doi: 10.1007/BF02772722

Yu, J., Hu, S., Wang, J., Wong, G. K. S., Li, S. G., Liu, B., et al. (2001). A draft sequence of the rice (Oryza sativa ssp. indica) genome. *science* 46 (23), 1937–1942. doi: 10.1126/science.1068275

Yuan, L. P., and Virnani, S. S. (1989). *Status of hybrid rice research and development*. International Rice Research Institute, Los Baños, pp. 7–24.

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.* 2 (12), 983–989. doi: 10.1038/35103590

Zhang, Q. J., Zhu, T., Xia, E. H., Shi, C., Liu, Y. L., Zhang, Y., et al. (2014). Rapid diversification of five Oryza AA genomes associated with rice adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 111 (46), E4954. doi: 10.1073/pnas.1418307111

Zhang, J., Chen, L. L., Xing, F., David, A. K., and Zhang, Q. F. (2016). Extensive sequence divergence between the reference genomes of two elite indica rice varieties Zhenshan 97 and Minghui 63. *Proc. Natl. Acad. Sci. U.S.A.* 113 (35), E5163. doi: 10.1073/pnas.1611012113

Zhao, B. R., Xing, Q. H., Xia, H. A., Yang, H. H., Jin, D. M., Liu, X., et al. (2005). DNA Polymorphism Among Yewei B, V20B, and Oryza minuta J. S. Presl. ex C. B. Presl. 植物学报*(英文版) J. Integr. Plant Biol.* 47 (12), 1485–1492. doi: 10.1111/j.1744-7909.2005.00122.x

Zhao, C., Zhao, B., Ren, Y., Tong, W., Wang, J. Q., Zhao, K., et al. (2007). Seeking transformation markers: an analysis of differential tissue proteomes on the rice germplasm generated from transformation of Echinochloa crusgalli genomic DNA. *J. Proteome Res.* 6 (4), 1354–1363. doi: 10.1021/pr0605015

Zhou, G., Weng, J., Zeng, Y., Huang, J. G., and Liu, G. L. (1983). Introduction of exogenous DNA into cotton embryos. *Methods Enzymol.* 101 (6), 433–481. doi: 10.1016/0076-6879(83)01032-0

Zhou, H., Liu, Q., Li, J., Jiang, D. G., Zhou, L. Y., Wu, P., et al. (2012). Photoperiod- and thermo-sensitive genic male sterility in rice are caused by a point mutation

in a novel noncoding RNA that produces a small RNA. *Cell Res.* 22 (4), 649–660. doi: 10.1038/cr.2012.28

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 21 January 2021
doi: 10.3389/fpls.2020.589189

Current Perspectives on Introgression Breeding in Food Legumes

Aditya Pratap[1], Arpita Das[2], Shiv Kumar[3] and Sanjeev Gupta[1]**

[1] ICAR-Indian Institute of Pulses Research, Kanpur, India, [2] Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, India, [3] International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat Office, Rabat, Morocco

Food legumes are important for defeating malnutrition and sustaining agri-food systems globally. Breeding efforts in legume crops have been largely confined to the exploitation of genetic variation available within the primary genepool, resulting in narrow genetic base. Introgression as a breeding scheme has been remarkably successful for an array of inheritance and molecular studies in food legumes. Crop wild relatives (CWRs), landraces, and exotic germplasm offer great potential for introgression of novel variation not only to widen the genetic base of the elite genepool for continuous incremental gains over breeding cycles but also to discover the cryptic genetic variation hitherto unexpressed. CWRs also harbor positive quantitative trait loci (QTLs) for improving agronomic traits. However, for transferring polygenic traits, "specialized population concept" has been advocated for transferring QTLs from CWR into elite backgrounds. Recently, introgression breeding has been successful in developing improved cultivars in chickpea (*Cicer arietinum*), pigeonpea (*Cajanus cajan*), peanut (*Arachis hypogaea*), lentil (*Lens culinaris*), mungbean (*Vigna radiata*), urdbean (*Vigna mungo*), and common bean (*Phaseolus vulgaris*). Successful examples indicated that the usable genetic variation could be exploited by unleashing new gene recombination and hidden variability even in late filial generations. In mungbean alone, distant hybridization has been deployed to develop seven improved commercial cultivars, whereas in urdbean, three such cultivars have been reported. Similarly, in chickpea, three superior cultivars have been developed from crosses between *C. arietinum* and *Cicer reticulatum*. Pigeonpea has benefited the most where different cytoplasmic male sterility genes have been transferred from CWRs, whereas a number of disease-resistant germplasm have also been developed in *Phaseolus*. As vertical gene transfer has resulted in most of the useful gene introgressions of practical importance in food legumes, the horizontal gene transfer through transgenic technology, somatic hybridization, and, more recently, intragenesis also offer promise. The gains through introgression breeding are significant and underline the need of bringing it in the purview of mainstream breeding while deploying tools and techniques to increase the recombination rate in wide crosses and reduce the linkage drag. The resurgence of interest in introgression breeding needs to be capitalized for development of commercial food legume cultivars.

Keywords: pre-breeding, distant hybridization, food legumes, breeding populations, introgression line

OPEN ACCESS

Edited by:
Mariola Plazas, Polytechnic University of Valencia, Spain

Reviewed by:
Bernardo Ordas, Consejo Superior de Investigaciones Científicas (CSIC), Spain
Hamid Khazaei, University of Saskatchewan, Canada

****Correspondence:***
Sanjeev Gupta
saniipr@rediffmail.com
Shiv Kumar
sk.agrawal@cgiar.org

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *30 July 2020*
Accepted: *03 December 2020*
Published: *21 January 2021*

Citation:
Pratap A, Das A, Kumar S and Gupta S (2021) Current Perspectives on Introgression Breeding in Food Legumes. Front. Plant Sci. 11:589189. doi: 10.3389/fpls.2020.589189

INTRODUCTION

Deployment of plant breeding tools has been successful for bolstering crop productivity, harmonizing crop phenology, enhancing nutritional quality, and developing resistance to multiple stresses. This became possible with identification of new combinations of genes and construction of superior populations possessing desirable novel characteristics, which have been exploited for human welfare (Anderson, 1949; Arnold, 1992). Although there are numerous examples for purposeful introgression of advantageous traits into crop varieties as a part of regular plant breeding programs, the extent and impact of either natural or farmer-aided introgression are yet to be ascertained (Jarvis and Hodgkin, 1998, 1999). With almost 20,000 species, legumes are the members of the Fabaceae/Leguminosae, the third largest family of the higher plants, which are ubiquitously present all over the temperate and tropical parts of the world (Polhill and Raven , 1981). Food legumes are important to human and animal life and occupy an important place in the global food supply chain, as well as sustainable agricultural production systems. With high protein content and 15 essential minerals, these are indispensable constituents of the cereal-based vegetarian diets and are grown traditionally with cereals, oilseeds, sugarcane, etc. Food legumes have prominent biological features and an inherent capability to fix atmospheric nitrogen owing to the presence of symbiotic association with *Rhizobium* bacteria in root nodules. Therefore, these crops become an indispensable part of the sustainable agricultsure strategy throughout the world (Chaturvedi et al., 2011).

A quest is on for the search of genes that can impart resistance to biotic and abiotic stresses in different food legumes, as well as to improve the physical and nutritional qualities of grains. The ever-changing climatic conditions led to emergence of new insect-pests and diseases and their biotypes and races, which are becoming a major threat limiting crop production and productivity (Chakraborty and Newton, 2011; Gautam et al., 2013). Broadening the genetic base will provide the needed armor to legume crops against these emerging challenges under climate change. The crop wild relatives (CWRs) are known to possess useful alien alleles and cryptic genetic variation, which are introgressed and expressed in cultivated genepool only when a systematic breeding scheme is put in place (Doyle, 1988; Tanksley and McCouch, 1997; Gupta and Singh, 2009; Pratap and Gupta, 2009). Recent advances in breeding and genomic tools and techniques provide an opportunity to introgress useful alleles left behind in the secondary and tertiary genepool into the elite background useful for legumes breeders. This review illustrates factors affecting wild gene introgression, population development, and success resorting wild gene introgression in cultivated food legumes.

The alien gene transfer in a crop species is paramount when the breeding value of the parental genepool no longer responds to selection, resulting in slow or no genetic gain. Conventionally, it is easier to manipulate desirable genes present within a crop species compared to the alien genes from distant relatives or exotic germplasm. This is because the gene transfer within a species is comparatively easy as there are no crossing barriers, and also it is largely free from linkage drags of unwanted traits. In food legumes, ∼3,700 improved varieties with narrow genetic base form the present varietal portfolio (Kumar et al., 2020), resulting in the genetic uniformity in farmers' fields. These varieties have been developed by the repeated use of a handful of elite germplasm from the primary genepool and therefore resulted in narrow genetic base and limited genetic buffers (Kumar et al., 2004). Introgression of alien genes from CWR offers a viable option to diversify and widen the genetic base of legume varieties, which provide insulation against the vagaries, as well as scope for continuous genetic gains over many breeding cycles (Kumar et al., 2009).

The horizontal gene transfer from wild species and even across different genera has played a significant role in the evolution of eukaryotic genomes (Bock, 2009) as wild species have evolved through different degrees of selection pressure exerted by environmental forces and biotic agents over a long period of time. As a result, these species have acquired many useful genes/alleles imparting adaptation to environmental cues such as extreme temperature, drought, waterlogging, salinity, and mineral toxicity, as well as biotic factors such as diseases, insect-pests, parasitic weeds, etc. Thus, hybridizing wild species with elite germplasm following a proper breeding scheme offers scope for the generation of multitude of pre-bred lines with novel recombination, which can further be utilized in the mainstream breeding for continuous accelerated genetic gains.

GENETIC BOTTLENECKS AND GERMPLASM REDUNDANCY

During evolution and domestication, wild progenitors have graduated to the cultivated forms passing through various genetic modifications and acquiring a combination of traits referred as "domestication syndrome." Nevertheless, the persistence of these species in nature for a long time, largely remaining unattended, might have led to disappearance of many genes/alleles responsible for input response and higher grain yield in legume crops (Jain, 1975). Further, only limited samples of the accessions representing the narrow genetic base of the total diversity might have been brought to the center of domestication leading to the "founder effect" (Ladizinsky, 1985). The history of food legumes matches with human civilization while their evolution took place throughout many different regions of the world (Pratap and Kumar, 2011). However, keeping in view that many food legumes now have their major production base away from the actual center of diversity and also that during their domestication limited sampling might have narrowed down their genetic base, these crops might have started their domestication journey with the "founder effect." For example, limited genetic diversity is reported in soybean outside its center of origin (Shoemaker, 1986; Pratap et al., 2012). Likewise, *Phaseolus*, chickpea, lentil, and pigeonpea also witnessed this bottleneck during their domestication.

While large germplasm repository of food legumes is preserved in different genebanks across the globe, mining

of genetic diversity for use in mainstream breeding remains limited because of the paucity of information on economic traits and the nature of diversity itself (Kumar et al., 2007). This becomes much more alarming when we consider the use of exotic and unadapted germplasm in breeding programs. Further, the large size of germplasm collection, breeders' preference for elite × elite crosses due to obvious advantages of their adaptability to local conditions, presence of cryptic genetic variation, and the linkage drag associated with transferring genes from wild relatives are other factors associated with restricted use of germplasm (Sharma et al., 2016).

Linkage drag is one of the major apprehensions while utilizing exotic and wild species in genetic amelioration of food legumes. In most of the cases, undesirable linkages hinder the transfer of desirable traits into cultivated backgrounds, and breaking such linkages needs dedicated efforts with a larger population and an efficient selection pressure. To overcome the problem of linkage drag, an additional generation of crossing among progenies prior to the selection or recurrent selection program over several generations is recommended. Nonetheless, it is now possible to recover or transfer into the elite germplasm the favorable alleles that were inadvertently left behind during the process of domestication. This can be done more efficiently by deploying molecular maps and integrative quantitative trait locus (QTL) analysis (for details, see Chamarthi et al., 2011) either through constructing introgression libraries that are made up of several introgression lines (ILs) or utilizing advanced-backcross QTL (AB-QTL) analysis. Introgression libraries can be constructed by crossing cultivated parent with wild donor followed by three to four times backcrossing of F_1 with cultivated parent (Kumar et al., 2011). In the past, attempts have been made to develop such libraries in soybean using *Glycine soja*, a wild species (Concibido et al., 2003), and from synthetic tetraploids in peanut (Fonceka et al., 2012). The AB-QTL approach also deploys repeated backcrossing involving elite parent and wild accession with an aim to reduce the number and size of the donor segment transferred through alien introgression. The ultimate objective here is to minimize the effect of linkage drag in such crosses, and advanced backcrossed populations thus derived are further subjected to QTL analysis to identify desirable genes/QTL. Common bean and soybean are the best examples where this approach has been used successfully (Blair et al., 2003; Chaky et al., 2003).

BREEDING POPULATIONS FOR GENE INTROGRESSION

Crop wild relatives are valuable source of novel and cryptic variation for broadening the genetic base of cultivated genepool (Dwivedi et al., 2005; Pratap et al., 2014). CWRs also harbor superior QTLs for improving agronomic and yield attributing traits. However, currently available approaches for introgression are not suitable for polygenic traits because of selection bias against the alien alleles. Moreover, penetrance and expressivity of alien genes and traits when introgressed in the cultivated background are often incomplete and limited, resulting in poor genetic gains. As a result, breeding for introgression of QTLs from CWR to elite background is avoided, and emphasis has been laid upon transfer of oligogenic traits governing stress resistance mostly. Nevertheless, with the advancement in genomic tools and techniques, it has become feasible to identify and target selection for major QTLs from CWRs. For QTL analysis, mostly balanced populations (F_2, BC_1) have been utilized previously where alleles of both wild and elite populations are available in the same frequency, although these populations are easy to develop but are characterized by several drawbacks. Balanced populations have the most complete genetic construction and only allow for analyzing both dominant and additive effects (Wang and Chee, 2010). These populations are temporary and highly heterozygous; thus, it is difficult to use them in replicated yield trials because in every generation of either selfing or backcrossing, the genetic constitution of these populations would change. Moreover, undesirable QTLs from the unadapted wild background could lead to the linkage drag. Further, during the transfer of QTLs, epistatic component augments the complication because it is difficult to detect through statistical inference, often sensitive to environments, is difficult to manipulate, and is likely to be present in balanced populations (Bernardo, 2010). To overcome these difficulties, "specialized population concept" has been advocated for transferring QTLs from CWR into elite backgrounds. For details, please see **Figure 1** The breeding population developed through different methods of gene introgression has been described in **Table 1**.

ADVANCED-BACKCROSS QTL

Advanced-backcross QTL strategy was proposed by Tanksley and Nelson (1996) for concurrently mining and transferring positive QTLs from CWR into elite genepool. It is a kind of inbred backcrossing for transferring complex trait from unadapted genepool to the elite background (Sullivan and Bliss, 1983). In this methodology, QTL analysis is deferred until advanced (BC_2 or BC_3 and so on) generations. This is because in early generation the effects of beneficial QTLs often remain unrecognized because of the presence of epistatic interactions between favorable QTLs and other genes from the donor parent, which might be resolved in later generations, thus allowing possible silencing of the measured QTL effects (Pillen et al., 2003). The common segregating populations (F_2, BC_1, RIL, and DH) generally utilized for QTL analysis have some major drawbacks when involving wild species in introgression program. First, these populations represent a large segment of genes from wild parent, and the QTLs with small effects remain unseen. Second, with these populations, the discovery and further introgression of QTLs through subsequent backcrossing or intercrossing becomes a two-step process, thus becoming a time-consuming affair with mere chance of utilizing QTL information to develop superior cultivar (Tanksley and Nelson, 1996). In AB-QTL strategy, the discovery and further transfer of positive QTLs from unadapted background to elite pool are

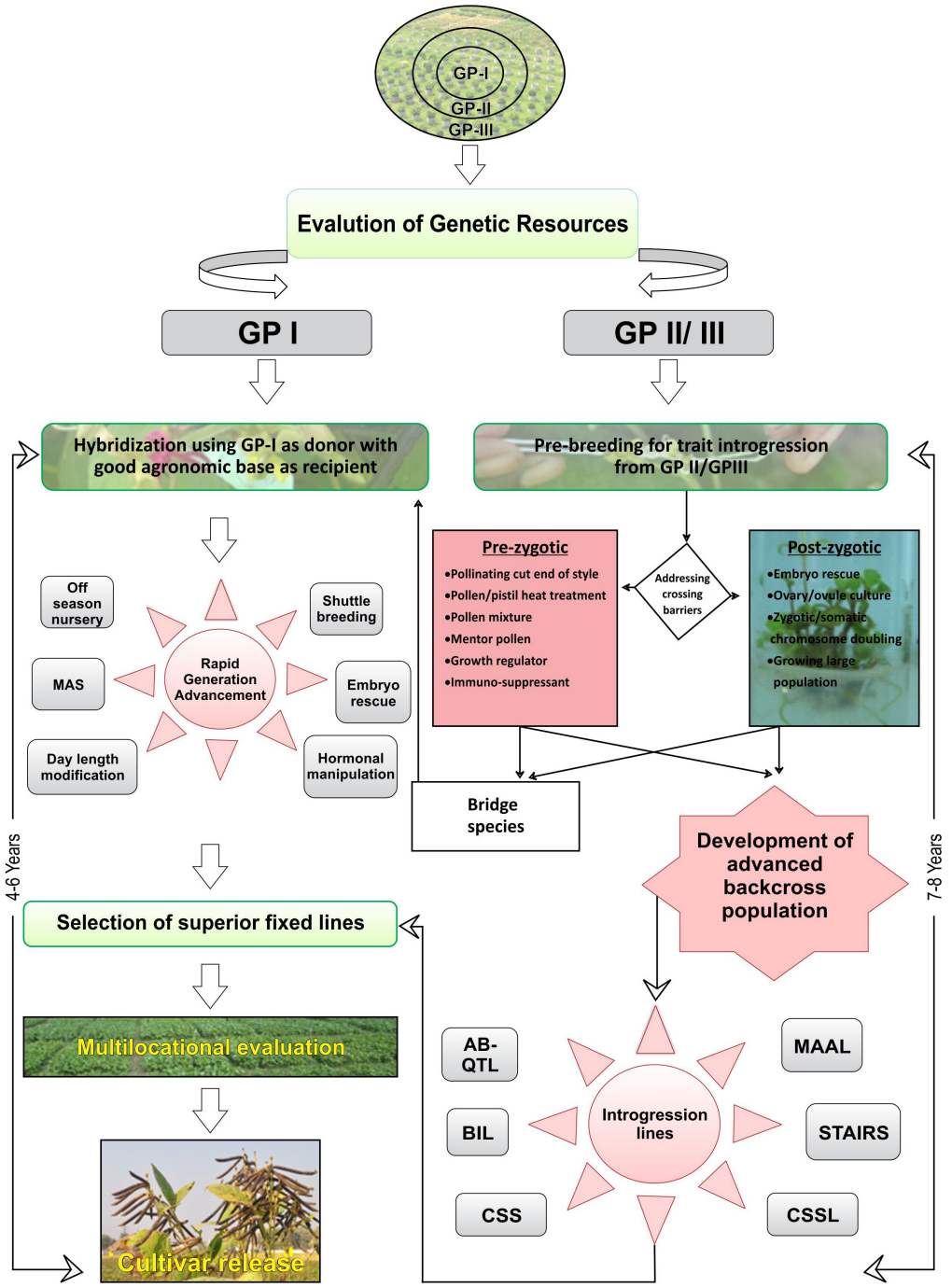

FIGURE 1 | Scheme for gene introgression for improvement of food legumes. **Left**, Hybridization using GP-I as donor with good agronomic base as recipient. **Right**, Pre-breeding for trait introgression using GP-II and/or GP-III.

a single-step process where QTL analysis is performed in later generations to facilitate sound statistical power for detection of QTLs with small effect.

The application of AB-QTL strategy has been tested in several food legumes and recommended as a useful crop breeding venture. An attempt was made to transfer QTLs conferring yield attributing traits from two accessions of *Cicer reticulatum*, "EC556270" and "ILWC21" into *Cicer arietinum* cv. GPF2 by attempting two cross combinations. The respective BC_2F_3 population was tested for yield attributing traits for confirmation of introgression of productivity traits into elite chickpea cultivar (Bhavyasree et al., 2018). Field pea weevil, *Bruchus pisorum*,

TABLE 1 | Examples of AB-QTL analysis for wild QTL introgression in legumes.

Crop	Cross	Population	No. of lines	Marker systems	Trait	References
Pigeonpea	ICPW 68 × ICPL 85010	BC_4F_{12}	138	–	*Phytophthora drechsleri* resistance	Mallikarjuna et al., 2011
	ICPL 87119 × ICPW 12	BC_2F_7	149			
Chickpea	ICC 4958 × (ICC 17264 × IG 69978)	BC_2F_4	1,500	–	Agronomic traits	Sharma et al., 2017
	ICCV 95311 × (IG 72933 × ICC 20192)	BC_2F_3	2,000			
	EC556270 × GPF2	BC_2F_3	52	SSR	Agronomic traits	Bhavyasree et al., 2018
	ILWC21 × GPF2					
Field pea	*P. sativum* cv. Pennant × ATC113 (*P. fulvum*)	BC_1F_3	72	–	Resistance against Pea weevil	Aryamanesh et al., 2012
Common bean	ICA Cerinza (Andean) × G24404 (Colombian)	$BC_2F_{3:5}$	157	SCAR, SSR	Agronomic traits	Blair et al., 2006
	OR 91G (snap bean) × PI 255956 (runner bean)	BC_2F_4	115	AFLP, SSR	Resistance against White mold	Haggard, 2007
	Cerinza × G10022	$BC_2F_{2:5}$	138	SSR	Agronomic traits, Fe and Zn	Blair et al., 2012
Peanut	ICGV 91114 × ISATGR 1212	BC_2F_9	416	DArT	Agronomic traits, as well as biotic stresses	Mallikarjuna et al., 2012
	ICGV 87846 × ISATGR 265-5A	BC_2F_9	579			
	ICGV 87846 × ISATGR 278-18	BC_2F_8	250			
	TMV 2 × ISATGR 121250	BC_2F_8	686			
	Florunner × TxAG-6	BC_3F_6	90	SSR, RFLP	Oil quality	Wilson et al., 2017
	Florunner × TxAG-6	BC_3F_6	233	RFLP	Resistance to root knot nematode	Burow et al., 2014
	Fleur11 × (*A. ipaensis* KG30076 × *A. duranensis V14167*) ^{4x}	BC_2F_1	87: BC_3F_1 and 55: BC_2F_2	SSR	Drought resistance	Fonceka et al., 2012

is a severe menace in cultivated field pea (*Pisum sativum*). An attempt was made to develop AB population (BC_2F_6) by involving *Pisum fulvum* accession "ATC113" as a resistant donor and susceptible *P. sativum* cv. Pennant as recipient (Aryamanesh et al., 2012). Wild beans are very diverse and useful source for enriching genetic variation of cultivated beans with low diversity. AB population ($BC_2F_{3:5}$) was developed in common bean involving a cross between large red-seeded commercial Columbian variety, "ICA Cerinza" as recurrent parent and wild accession "G24404" for detecting QTL toward improvement of agronomic performance. This strategy uncovered 13 QTLs for plant height, yield, and yield-attributing traits along with detection of a QTL for seed size from the wild parent (Blair et al., 2006). Another attempt has been made in common bean to transfer resistance against white mold, caused by *Sclerotinia sclerotiorum*. AB-QTL strategy has been undertaken to identify and transfer QTL conferring resistance to white mold into an interspecific cross of *Phaseolus vulgaris* cv. OR91G and *Phaseolus coccineus* cv. PI255956. A population of 115 BC_2F_4 lines were developed and genotyped using amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers and screened under greenhouse for phenotypic scoring (Haggard, 2007). Wild common beans accumulate high minerals but are not commercially acceptable because of small seed size. AB population was developed for improving the mineral status of the Andean variety "Cerinza," a large red seeded bush bean cultivar with wild genotype "G10022." The $BC_2F_{3:5}$ ILs derived from this cross combination were subjected to multilocation yield trial for contemplating the role of genotype × environment interaction toward the expression of Fe and Zn content in AB population (Blair et al., 2012). The result from the study confirmed that the AB-QTL method was effective for identifying the QTL controlling Fe and Zn content, as well as their transfer into elite background and further evaluation.

Therefore, it is evident that the AB-QTL approach has been successfully applied in numerous legume species for harnessing the favorable alleles from wild into elite background, although in many other crops, *viz*., mungbean, urdbean, *Lathyrus*, and lentil, this approach is yet to be employed to explore its advantages. This strategy has paved the way for identifying more QTLs, precise measurement of the effect of individual loci, and their transfer into the cultivated background. Additionally, the formation of AB-NIL would facilitate further genetic dissection of QTLs and subsequently the map-based cloning of the underlying genes, thus opening a new vista for other legumes also.

INTROGRESSION LINES

Introgression lines are specialized populations derived through advanced backcrossing, which are nearly isogenic to recurrent parent and contain only a small fraction from donor parent (Eshed and Zamir, 1994; Tian et al., 2006). These populations are much more efficient for QTL identification and fine mapping followed by studying QTL × environment interaction due to their homozygous nature as compared to the conventional populations (Wehrhahn and Allard, 1965; Eshed and Zamir, 1995; Yin et al., 2016). The major drawback of ILs is the time taken for their development (Tanksley et al., 1989; Hospital et al., 1992; Hospital and Charcosset, 1997) and high cost for marker evaluation (Ali et al., 2010). However, with

the availability of densely saturated marker systems in some legumes such as chickpea, pigeonpea, soybean, peanut, etc., the foreground as well as background selection become easier. In addition, marker-based selection facilitates detection of non-target introgression in early generation, as well as further elimination from the recipient background for speeding up IL recovery (Frisch and Melchinger, 2005).

The critical factors for reducing the problem of linkage drag in backcross population through marker-assisted selection are the tightness of the linkage between the introgressed genes and the flanking markers and the size of the population, as well as the total duration of the backcross scheme (Hospital, 2001) besides the size of the segment to be transferred. Theoretical explanation given by Hospital (2001) nicely pointed out that presence of functional markers within the genes to be introgressed or tightly linked flanking markers along with three to five generations of backcrossing would be cost-effective to minimize the length of the donor segment. In peanut, pyramiding of nematode resistance and the trait governing high oleic:linoleic acid has been introgressed successfully to develop improved Tifguard variety "Tifguard High O/L" through tightly linked cleaved amplified polymorphic sequences and SSR markers with less linkage drag problem (Chu et al., 2011). Another high oleic acid line was developed in peanut through marker-assisted introgression of two *FAD2* mutant alleles conferring high oleic acid from donor parent "SunOleic95R" into the background of "ICGV 06100" (Bera et al., 2019). In chickpea, two *Fusarium* wilt (FW) resistance ILs, namely, "Annigeri 1" and improved "JG 74," have been developed through marker-assisted backcrossing using "WR 315" as the donor parent (Mannur et al., 2019). Foreground selection was done with TA59, TA96, TR19, TA27, and GA16 markers, whereas background selection was done using SSR markers. Likewise, Pratap et al. (2017) developed improved "Pusa 256" using "Vijay" as the donor parent using TA 37 and TA 110 as the markers for foreground selection. Two parallel marker-assisted introgression programs have been implemented to improve both FW, as well as *Ascochyta* blight (AB) resistance of "C 24" cultivar by introgressing resistant locus of race 1 of FW coupled with two QTL clusters for AB resistance (Varshney et al., 2014).

Besides conventional marker-assisted introgression approaches deployed in food legumes, various IL-based strategies have been recommended like backcross inbred lines (BILs), chromosome segment substitution lines (CSSLs), stepped aligned inbred recombinant strain (STAIRS), etc., for removing background noise and measuring yield associated traits precisely. BILs are characterized by small introgression of segment from donor parent and useful for reducing background noise from donor parent, as well as for mapping interspecific variation (Eshed and Zamir, 1994). In soybean, BIL populations have been developed for overcoming abiotic stresses by mining and introgressing useful QTLs from the donor parent. Water limiting situation is one of the main restraints for soybean production (Sinclair et al., 2007). Earlier reports confirmed that root length and absorption surface area along with root architecture are prime determinants for yield performance under variable moisture regime (Price et al., 2002; Manavalan et al., 2009). As variability regarding root architecture is limited in cultivated soybean, an attempt was made to explore the potential of exotic wild species for broadening the genetic base. BIL mapping population has been developed by crossing *Glycine max* cv. Dunbar (PI 552538) as a recipient with a wild soybean accession "PI326582A" of *G. soja*. BILs have been created to minimize the magnitude of gene introgression from the wild soybean parent "PI326582A" by allowing two generations of backcrossing to produce 296 $BC_2F_{4:5}$ progenies (Manavalan et al., 2015). Genetic linkage map was constructed by using SSR and SNP markers, resulting in the identification of a major QTL (Satt315-I locus) on chromosome 8 that governs root traits and shoot length. It has been observed that, sometimes, the same metabolic pathway governs different stresses in plants (Xiong and Yang, 2003), and it is mostly associated with overlapping QTLs. The reason for this genetic overlapping is due to pleiotropy and linkage disequilibrium (Zhang et al., 2012). In soybean, BIL mapping population was developed by crossing a Chinese variety, "Hongfeng 11" with an American variety, "Harosoy" for mapping QTLs related to drought and low-temperature tolerance during germination. Finally, 12 QTLs were detected that were correlated with drought and low-temperature tolerance and confirmed the mechanism of partial genetic overlap between drought and low-temperature tolerance in soybean (Zhang et al., 2012). This study further validated the effectiveness of using BILs for gene introgression, trait identification, QTL mapping, and gene cloning in legume.

CHROMOSOME SEGMENT SUBSTITUTION LINE

Chromosome segment substitution lines are very robust population for QTL mapping or cloning and gene discovery, as well as for gene pyramiding (Tanksley and Nelson, 1996), and can be developed by deploying AB strategy subsequently by selfing and selection of backcross population with molecular markers. Selection of backcross population with markers leads to identification of individuals carrying the introgressed gene(s) of interest along each chromosome. CSSLs generally exclude non-targeted portion from the donor, which can create background noise due to epistatic interaction (Ali et al., 2010). In CSSLs, each line carries a single defined chromosomal section from the wild donor into the recipient background, unlike BILs, where each line carries several homozygous introgressed segments from donor parent. Unlike BIL, CSSL libraries have been developed to recover the whole genome of donor parent (Ali et al., 2010). These populations can be compared with the individual IL or recurrent parent for finding out significant differences between them. CSSLs are also useful populations for controlling allelic variation and facilitate "breeding by design" (Peleman and van der Voort, 2003; Wei et al., 2010).

This strategy has been used in many crops including wheat (Liu et al., 2006), tomato (Monforte and Tanksley, 2000), rice (Bian et al., 2010), maize (Wang et al., 2007), cotton (Wang et al., 2008), and barley (Matus et al., 2003) for gene discovery and map-based cloning and opens a new vista for exploring

the potential of CSSL population in legumes for detection of genes or QTL explicitly, as well as their pyramiding into elite background. Cultivated soybean (*G. max*) is domesticated from wild *G. soja* (Broich and Palmer, 1980), which harbors useful genes governing large number of pods, richness in protein, adaptability to various biotic and abiotic factors, etc. Previous studies confirmed the versatility of *G. soja* as a useful donor for enriching the genetic diversity of cultivated soybean (Concibido et al., 2003; Li et al., 2008). The problem of linkage drag often circumvents the useful introgression process. Keeping these in mind, attempt has been made to construct CSSL population consisting of 151 lines by involving *G. max* cv. NN1138-2 as female and *G. soja* cv. N24852 as male. Polymorphic SSR markers between the parents were deployed for marker assisted selection (MAS) for easy recovery of CSSL. In this study, four QTLs related with plant height, as well as node numbers per plant, have been identified (Wang et al., 2013). The same CSSL population was used for mining and fine mapping of QTLs underlying seed quality traits including size and shape, as well as other agronomic traits (Wang et al., 2012, 2013; He et al., 2014).

CHROMOSOME SUBSTITUTION STRAIN

Another approach is the construction of chromosome substitution strain containing a large number of lines each carrying a homozygous chromosome with single crossover in such a way that the chromosome contains recurrent genotype at one end and donor genotype at the other end and known as single recombinant lines (SRLs). When the SRLs for each chromosome are sequentially stacked, they reveal a step-like progression, with each successive line having a little more donor chromosome, and constitute STAIRS libraries (Koumproglou et al., 2002). The concept was first applied in *Arabidopsis thaliana* for fine mapping of QTL. Although STAIRS has not yet been explored in legume crops, it is an effective strategy for comparison of genetic differences in the precise region of selected chromosome for QTL analysis, gene mining, and expression studies (Koumproglou et al., 2002). All these ILs can be maintained as an immortal representation in the form of "exotic library" for efficient detection and mapping of QTLs conferring agronomic traits (Zamir, 2001). This library is a permanent resource, which enables the researchers to explore over time and access the data generated for further use. The homozygous lines maintained in the library can be utilized as a parent for crossing with different tester lines to identify the chromosomal segments associated with heterosis. Development of an exotic library will immensely facilitate to counter the problem of linkage drag and precisely examine the phenotypic effect of QTL interaction for better insight into the epistatic effect (Eshed and Zamir, 1996). All these mapping populations along with genomic tools will be valuable for demonstrating the scope of introgression of desirable QTLs from CWR that was hitherto difficult to accomplish. The methodologies described can be extended to legume crops for harnessing the potential of CWRs for broadening the genepool through genomics-assisted genetic enhancement.

POTENTIAL WILD SPECIES FOR ALIEN GENE TRANSFER OF TARGET TRAITS

Most food legumes and their wild relatives (CWRs) are diploid and self-pollinated in nature. Considerable variability exists in wild species for yield contributing traits including number of pods per plant, number of seeds per pod, and seed size, as well as nutritional traits and biotic and abiotic stress resistance. The success of alien gene transfer through distant hybridization generally depends on the ploidy level of the species, pollination behavior of the plant, nature, and direction of the cross and frequency of pollination, which are further influenced by the deployment of appropriate hybridization schemes (Pratap et al., 2015a). Efforts were made to identify potential wild accessions for alien gene introgression in different food legumes by several researchers (**Table 2**).

Chickpea is the most important cool season grain legume and offers tremendous opportunities for its genetic improvement through introgression breeding, especially concerning biotic and abiotic stresses. Of the eight annual species, only one wild species, *C. reticulatum*, is readily crossable with the cultivated chickpea (Kumar et al., 2003). The success of hybridization with the remaining annual wild *Cicer* species requires specialized techniques such as the application of growth hormones and embryo rescue techniques (Lulsdorf et al., 2005; Mallikarjuna and Jadhav, 2008). Among the biotic stresses, FW and AB cause maximum damage to the plant and lead to severe yield reduction. FW causes up to 100% yield losses (Sharma et al., 2004; Pratap et al., 2017). AB, caused by *Ascochyta rabiei* usually appears at the reproductive phase, and in severe cases, the entire plant dries up suddenly. Several accessions of *Cicer bijugum*, *Cicer echinospermum*, *Cicer judaicum*, *Cicer chorassanicum*, *Cicer pinnatifidum*, *C. reticulatum*, *Cicer cuneatum*, *Cicer yamashitae*, and *Cicer canariense* have shown high resistance to AB (Haware et al., 1992; Kaiser et al., 1994; Collard et al., 2001; Shah et al., 2005). Simultaneously, many of them also possessed a high degree of resistance to FW (Kaiser et al., 1994; Infantino et al., 1996; Singh K. B. et al., 1998; Collard et al., 2001; Shah et al., 2005; Pande et al., 2006). Some of the accessions belonging to *C. bijugum*, *C. echinospermum*, *C. judaicum*, and *C. reticulatum* were reported to be highly resistant to *Botrytis* gray mold (Pande et al., 2006; Knights et al., 2008; Isenegger et al., 2011; Coyne et al., 2020). Several chickpea CWRs have shown high tolerance to abiotic stresses. For example, tolerance to drought and heat stresses was reported in six *Cicer* species (Toker et al., 2007a; Canci and Toker, 2009; Imtiaz et al., 2011). Likewise, tolerance to cold was reported in *C. bijugum*, *C. echinospermum*, *C. pinnatifidum*, and *C. reticulatum* (Singh et al., 1990; Robertson et al., 1995; Singh K. B. et al., 1998; Toker, 2005; Saeed et al., 2010).

Among *Vigna* crops, the Asiatic *Vigna* have tremendous scope for improvement with respect to yield and yield attributes, biotic and abiotic resistance, and nutritional quality (For review, see

TABLE 2 | Potential of wild species for alien gene transfer in food legumes.

Trait	Species	References
Chickpea		
Ascochyta blight	*C. bijugum* K.H. Rech.	Stamigna et al., 2000; Shah et al., 2005
	C. echinospermum P.H. Davis	Collard et al., 2003; Saeed and Darvishzadeh, 2017
	C. judaicum Boiss.	Stamigna et al., 2000; Shah et al., 2005; Saeed and Darvishzadeh, 2017
	C. pinnatifidum Jaub. & Sp.	Collard et al., 2001; Shah et al., 2005
	C. reticulatum Ladiz.	Shah et al., 2005; Saeed and Darvishzadeh, 2017
	C. yamashitae Kitamura	Shah et al., 2005
	C. canariense S. Guerra & Lewis	Kaiser et al., 1994
Fusarium wilt	*C. bijugum* K.H. Rech., *C. cuneatum* Hochst. Ex Rich, *C. echinospermum* P.H. Davis, *C. judaicum* Boiss., *C. pinnatifidum* Jaub. & Sp.	Singh K. B. et al., 1998
	C. chorassanicum (Bge) M. Pop.	Kaiser et al., 1994
	C. reticulatum Ladiz.	Infantino et al., 1996
	C. canariense S. Guerra & Lewis	Kaiser et al., 1994
Botrytis gray mold	*C. bijugum* K.H. Rech.	Isenegger et al., 2011
	C. echinospermum P.H. Davis	Knights et al., 2008
	C. judaicum Boiss., *C. reticulatum* Ladiz.	Pande et al., 2006
Root lesion nematode (*Pratylenchus thornei*)	*C. reticulatum* and *C. echinospermum*	Reen et al., 2019
Drought	*C. echinospermum* (ILWC 235), *C. oxyodon (C. oxyodon* L-4, L-9)	Saeed and Darvishzadeh, 2017
Phytophthora root rot	*C. echinospermum*	Amalraj et al., 2019
Drought	ICC7571	Kashiwagi et al., 2013
Terminal heat stress	ICC1205 and ICC15614	Devasirvatham et al., 2013
Drought	ICC14778	Krishnamurthy et al., 2013
Drought	*C. yamashitae*	Sharma and Upadhyaya, 2015; Sharma and Upadhyaya, 2019
Cold	*C. bijugum* K.H. Rech.	Toker, 2005
Drought		Imtiaz et al., 2011
Cold	*C. echinospermum* P.H. Davis	Toker, 2005; Saeed et al., 2010
	Drought and heat	Canci and Toker, 2009
Cold	*C. reticulatum* Ladiz	Singh K. B. et al., 1998; Toker, 2005; Saeed et al., 2010
Drought and heat		Canci and Toker, 2009; Imtiaz et al., 2011
Drought and heat	*C. anatolicum* Alef., *C. microphyllum* Benth., *C. montbretii* Jaub. et Sp., *C. oxydon* Boiss. et Hoh., *C. songaricum* Steph. ex DC.	Toker et al., 2007b
Lentil		
Anthracnose	*Lens ervoides, L. lamottei, L. nigricans*	Tullu et al., 2006
Ascochyta blight	*L. ervoides, L. culinaris* ssp. *orientalis, L. odemensis, L. nigricans, L. montbretti*	Bayaa et al., 1994; Tullu et al., 2010
Fusarium wilt	*L. culinaris* ssp. *orientalis, L. ervoides*	Gupta and Sharma, 2006; Singh et al., 2020
Powdery mildew	*L. culinaris* ssp. *orientalis, L. nigricans*	Gupta and Sharma, 2006
Rust	*L. culinaris* ssp. *orientalis, L. ervoides, L. nigricans, L. odemensis*	Gupta and Sharma, 2006
Drought	*L. odemensis, L. ervoides, L. nigricans*	Hamdi and Erskine, 1996; Gupta and Sharma, 2006
Cold	*L. culinaris* ssp. *orientalis*	Hamdi et al., 1996
Orobanche	*Lens ervoides, L. odemensis, L. orientalis*	Fernández-Aparicio et al., 2009
Bruchids	*L. culinaris Medikus* subsp. *orientalis, L. nigricans, L. lamottei*	Laserna-Ruiz et al., 2012
Rust and powdery mildew	*L. orientalis*	Singh et al., 2020
Powdery mildew and *Fusarium* wilt	*L. ervoides*	Singh et al., 2020
Vigna species		
Bruchid	*V. riukinensis, V. reflexo-pilosa*	Tomooka et al., 1992
	V. radiata var. *sublobata*	Miyagi et al., 2004
	V. umbellata	Somta et al., 2006
	V. tenuicaulis	Tomooka et al., 2000
	V. nepalensis	Somta et al., 2008
Powdery mildew	*V. stipulacea*	Tomooka et al., 2006
	V. reflexo-pilosa var. *glabra*	Egawa et al., 1996
Low trypsin inhibitor activity	*V. tenuicaulis*	Konarev et al., (2002)
Chymotrypsin absent	*V. grandiflora*	Konarev et al., (2002)
Heat	*V. aconitifolia*	Tomooka et al., 2001
	V. riukinensis	Egawa et al., 1999

(Continued)

TABLE 2 | Continued

Trait	Species	References
Bean fly resistance	*V. reflexo-pilosa*	Egawa et al., 1996
Resistance to pod bug	*V. unguiculata* ssp. *dekindtiana TVNu 151*	Koona et al., 2002
Resistance to yellow mosaic virus	*V. radiata* var. *sublobata*	Reddy and Singh, 1990; Pal et al., 2000
	V. umbellata, V. trilobata, V. mungo	Pandiyan et al., 2008
Photothermo insensitivity	*V. umbellata, V. glabrescens*	Pratap et al., 2014
Soybean cyst nematode (*Heterodera glycines*)	*V. angularis*	Kushida et al., 2012
Salt stress	*V. luteola, V. marina, V. vexillata* *V. riukiuensis*, *V. trilobata*, *V. vexillata*, *V. marina* subsp. *oblonga*, *V. luteola*, and *V. marina*	Yoshida et al., 2020 Iseki et al., 2016
Salinity stress	Domesticated *V. unguiculata, V. vexillata,* wild *V. luteola, V. marina, V. nakashimae, V. riukiuensis, V. vexillata* var. *macrosperma.*	Van Zonneveld et al., 2020 Harouna et al., 2020
High temperature, salinity	*V. trilobata*	
Dry climate and salinity	*V. vexillata* var. *ovate*	
Dry and seasonally hot climate	*V. monantha, V. aconitifolia, V. aridicola, V. exilis*	
Resistance against Bruchid beetles	*V. radiata* var. *sublobata*	Schafleitner et al., 2016
Bruchids, *Cercospora* leaf spot, powdery mildew and MYMV	TCR 20	Tripathy et al., 2016
Field pea		
Drought tolerance	*P. fulvum*	Naim-Feil et al., 2017
TI1 and TI2 seed protease inhibitors	*P. sativum* subsp. *elatius*	Clemente et al., 2015
Pulse beetle (*Callosobruchus chinensis* L.)	*P. elatius* – AWP 442 and *P. fulvum* – AWP 600, AWP 601	Esen et al., 2019
Rust (*Uromyces* pisi.)	*P. fulvum*	Barilli et al., 2018
PSbMV virus (*Potyvirus*)	*P. fulvum*	Konečná et al., 2014
Powdery mildew	*P. fulvum*	Cobos et al., 2018
Pigeonpea		
Heat, drought	*C. acutifolius, C. cinereus, C. lanceolatus, C. latisepalus*	Khoury et al., 2015
Cold	*C. confertiflorus, C. mollis, C. platycarpus, C. trinervius*	
High precipitation, waterlogging, drought	*C. sericeus, C. lineatus*	
Heat, temperature variation/seasonality, cold	*C. platycarpus, C. scarabaeoides*	
Insect resistance – *Helicoverpa armigera*	*C. scarabaeoides* – IBS 3471	Ngugi-Dawit et al., 2020
Helicoverpa pod borer	*C. scarabaeoides, C. sericeus, C. lineatus, C. acutifolius,* and *C. platycarpus*	Saxena et al., 2018
Pod fly	*C. sericeus*	
Bruchids	*C. scarabaeoides, C. platycarpus,* and *C. acutifolius*	
Water logging	*C. acutifolius*	Hingane et al., 2015; Mallikarjuna et al., 2017
Antinutritional factors, high antioxidant potential	*C. scarabaeoides* (ICP15683/W15)	Sekhon et al., 2017
Common bean		
Nutritional composition and cooking characteristics	*Phaseolus acutifolius*	Porch et al., 2017
Abiotic stresses	*P. acutifolius*	Gujaria-Verma et al., 2016
Drought-tolerant	*P. acutifolius*	Mwale et al., 2020
Abiotic stresses – drought and subzero temperatures	*P. acutifolius* A. (Gray)	Souter et al., 2017
Cow pea		
Heat and salinity	*V. unguiculata* group *sesquipedalis*	Van Zonneveld et al., 2020
Aphid	Wild cowpea relative – line TVNu 1158	Boukar et al., 2019
Peanut		
Late leaf spot pathogen	*Arachis diogoi*	Kumar and Kirti, 2015
Drought tolerance	*A. duranensis*	Lílian et al., 2019

Pratap et al., 2020). Accessions of *Vigna mungo* var. *silvestris* have been identified as a durable source of mungbean yellow mosaic virus (MYMV) resistance (Reddy and Singh, 1993). Likewise, variation for yield contributing traits and MYMV resistance was observed in *V. mungo* var. *silvestris*, *Vigna radiata* var. *sublobata* (Singh, 1990), *Vigna umbellata*, *Vigna trilobata*, and *V. mungo* (Nagaraj et al., 1981; Singh and Dikshit, 2002; Pratap et al., 2018a). "PLN15," a wild accession of *V. radiata* var. *sublobata* possessed a high number of pods per plant and seeds per pod and therefore identified as a potential donor for these traits (Reddy and Singh, 1990). *V. trilobata* and *Vigna stipulacea* was proposed as the candidates for neodomestication for drought tolerance and disease and pest resistance, respectively, by Tomooka et al. (2014), whereas resistance to diseases and pests in both the species was also reported in several other reports (Nagaraj et al., 1981; Chandel et al., 1984; Tomooka et al., 2006; Pandiyan et al., 2008; Gore et al., 2019). Accession "TC1966" of *V. radiata* var. *sublobata* was reported to possess bruchid tolerance (Tomooka et al., 1992). Likewise, resistance to legume pod borers and pod-sucking bugs was reported in *Vigna vexillata* and *Vigna oblongifolia* (Fatokun, 1991). *V. mungo* var. *silvestris* was reported as immune to bruchids (Fujii et al., 1989; Dongre et al., 1996). Many accessions of ricebean (*V. umbellata*) show complete resistance to bruchids and therefore have been identified as useful donors for introgressing bruchid resistance into other *Vigna* species. Further, as it is a cultivated species, gene introgression from ricebean into mungbean and urdbean is comparatively easier. Dense hairs on different parts of the wild cowpea, *V. vexillata*, are reported to impart antixenosis to pod-sucking bugs and pod borer (Oghiakhe et al., 1992; Boukar et al., 2013). Likewise, strength and hardness of the pod wall are also considered to impart pod borer resistance (Oigiangbe et al., 2002). Other wild *Vigna* species with resistance to *Maruca vitrata* and the pod-sucking bugs include *Vigna unguiculata* ssp. *dekindtiana*, *Vigna luteola*, *Vigna oblongifolia*, and *Vigna reticulata*. Pratap et al. (2014) reported *V. umbellata* (accession IC251442) and *Vigna glabrescens* (accession IC251372) as photo- and thermo-period insensitive as these were able to flower and set pods at temperatures as high as 43.9°C and as low as 2.7°C.

Wild *Lens* species have emerged as a great reservoir of useful genes for traits of breeders' interest including resistance to strategically important diseases, insect-pests, and plant parasitic weeds. A high degree of resistance was observed for *Stemphylium* blight in *Lens lamottei* followed by *Lens ervoides* (Podder et al., 2013). Similarly, some accessions of *Lens odemensis* possessed high resistance against *Sitona* weevil followed by *L. ervoides* (El-Bouhssini et al., 2008). Some of the wild accessions with combined resistance to AB and FW or anthracnose diseases have been identified for their use in lentil breeding programs (Bayaa et al., 1995; Tullu et al., 2006). Preliminary screening of *Lens* CWR has indicated drought tolerance in *Lens nigricans*, *L. odemensis*, and *L. ervoides* (Gupta and Sharma, 2006) and cold tolerance in *Lens culinaris* ssp. *orientalis* (Hamdi et al., 1996). Simultaneously, promising donors for yield traits, *viz*., 100-seed weight and pods per plant were observed in *L. lamottei* and *L. culinaris* ssp. *orientalis* (Gupta and Sharma, 2006). Earlier reports have indicated *L. ervoides* as a good source of alleles for plant architectural traits such as phenology, plant growth habit, and biomass besides seed traits (Tullu et al., 2011, 2013; Kumar et al., 2014). Based on the extensive evaluation of global wild *Lens* taxa representing 27 countries, Kumar et al. (2014) observed wide variation for yield attributing traits, as well as resistance to multiple diseases in *L. nigricans* and *L. ervoides*. Nutritional quality traits in wild *Lens* showed significant diversity not only for micronutrients (Sen Gupta et al., 2016; Kumar et al., 2018) but also for prebiotics, RFO, raffinose, and verbascose (Tahir et al., 2011).

GENE INTROGRESSION IN FOOD LEGUMES: SUCCESS

Wild gene introgression as a breeding strategy has been deployed successfully in food legumes for development of improved varieties, pre-bred lines, genetic stocks, mapping populations, and bridge species. Legumes such as chickpea, pigeonpea, lentil, mungbean, urdbean, and peanut have benefited from the wild gene introgression with successful examples of discovery, development, and deployment of useful traits in cultivated genepool. In chickpea, after the first report of successful interspecific crosses between *C. arietinum* and *C. reticulatum* (Ladizinsky and Adler, 1976), attempts have been made for crossing between *C. arietinum* and *C. echinospermum* (Singh and Ocampo, 1993; Pundir and Mengesha, 1995). *C. reticulatum* accession "ILWC119," when involved in hybridization program, led to the development of "ILC10765" and "ILC10766," two cyst nematode-resistant chickpea lines (Malhotra et al., 2002). Singh et al. (2005) and Ramgopal et al. (2012) utilized the diversity of *C. reticulatum* and *C. echinospermum* to transfer useful traits including tolerance to cold and resistance to diseases such as wilt, root rot, and *Botrytis* gray mold into cultivated chickpea. There are reports of successful interspecific crosses between *C. arietinum* and *C. judaicum* (Singh et al., 1999), *C. arietinum* and *C. cuneatum* (Singh and Singh, 1989), *C. arietinum* and *C. pinnatifidum* (Badami et al., 1997; Mallikarjuna and Jadhav, 2008), and *C. arietinum* and *C. bijugum* (Mallikarjuna et al., 2007). Successful introgression of useful genes into cultivated chickpea from these crosses has shown the transferability even from the cross-incompatible wild *Cicer* species.

Successful interspecific hybridization of *P. vulgaris* has been reported to a limited extent with the members of other wild *Phaseolus* species. Interspecific hybrids with *Phaseolus costaricensis* in the secondary genepool were reported by Singh et al. (1997), and consequently, VRW 32 was reported as the first white mold-resistant interspecific breeding line derived from *P. costaricensis*. Congruity backcrossing (CBC) involves recurrent backcrossing to each parent in alternate generations as opposed to the traditional recurrent backcrossing to a single recurrent parent and was first reported as a method to produce fertile intermediate hybrids between *Phaseolus acutifolius* and *P. vulgaris* (Haghighi and Ascher, 1988). This method allows substantial recombination between distant species, and new

phenotypes can arise as a result of CBC (Anderson et al., 1996). Singh et al. (2009) reported the release of white mold-resistant "VCW 54" and "VCW 55" bean germplasm lines that were developed using CBC between the black bean cultivar "ICA Pijao" and the scarlet runner bean accession "G35172." CBC has also been used to transfer traits from wild tepary species *Phaseolus parvifolius* into common bean (Singh S. P. et al., 1998).

Wilkinson (1983) reported a root rot-resistant line "Cornell 2114-12" derived from a cross between common bean and scarlet runner bean lines. Likewise, Miklas et al. (1999) developed the common bacterial blight-resistant bean germplasm lines "ICB-3," "IBC-6," "ICB-8," and "ICB-10," which were derived from an interspecific cross with scarlet runner bean. Beaver et al. (2012) released a bean germplasm line "PR0650-31," which was derived from the cross BAT 93/PI 417662//VAX 6 using wild-type bean germplasm "PI 417662" collected from Jalisco, Mexico, and was resistant to web blight and common bacterial blight. Acosta-Gallegos et al. (2007) developed an inbred backcross population from a cross between G 24423, a wild bean accession from Colombia and "Negro Tacana," a Mexican black bean cultivar. One Bc2F4:7 line from this population was later observed to produce >5,000 kg/ha seed in field trials.

Lentil CWRs have been evaluated extensively to discover and deploy traits of interest into cultivated species. These efforts have led to identification of extra early photoinsensitive (ILWL118 maturing in <90 days) and high micronutrient content (ILWL74 and ILWL80) germplasm. These CWRs have been used extensively in mainstream breeding, resulting in the development of short-duration biofortified pre-bred lines (Kumar et al., 2018). Wide crosses in lentil have also been mined for transgressive segregants for agronomically important traits (Kumar et al., 2011, 2014; Singh et al., 2013). More recently, hybridization of the cultivated lentil with *L. ervoides* using embryo rescue (Tullu et al., 2013) has been reported with successful transfer of resistance to *Orobanche crenata* and anthracnose in cultivated species (Fiala et al., 2009; Tullu et al., 2011). The International Center for Agricultural Research in the Dry Areas has successfully deployed *L. orientalis* and *L. ervoides* for introgression of resistance to key diseases, phenology, biofortification, plant habit, and other important agronomic traits toward the development of pre-bred lines. These pre-bred lines demonstrated >40% yield advantage over the best check (Bakaria) coupled with richness in micronutrients. These pre-bred lines can also fit well in short-season windows of 80–100 days (Kumar et al., 2020). These lines are currently under multilocation testing under the CWR project.

In pigeonpea, despite large visible genetic variation (Yang et al., 2006), the use of wild species in breeding programs has been rather limited to the development of cytoplasmic genic male sterility systems (Saxena et al., 2010). To date, seven cytoplasmic male sterility (CMS) systems have been reported (Saxena et al., 2010), and six of them have been developed from wild relatives belonging to the secondary genepool. The seventh system was developed utilizing *Cajanus platycarpus*, a member from the tertiary genepool (Saxena et al., 2010; Mallikarjuna et al., 2011). The A_1 CMS system derived from *Cajanus sericeus* (Saxena et al., 1996) was not stable at low temperature (<10°C) as the male-sterile plants revert to male fertility (Saxena et al., 2005). However, the presence of pollen shedders in the female line and non-availability of good maintainers did not make it commercially viable for hybrid breeding. The A_2 cytoplasm derived from *Cajanus scarabaeoides* was reported as highly stable (Tikka et al., 1997; Saxena and Kumar, 2003). Although this system is promising with respect to yield, inconsistency was observed in the fertility restoration over diverse environments, which reduced its acceptance for hybrid production. In the A_3 system, the cytoplasm was derived from *Cajanus volubilis* (Wanjari et al., 1999). The A_4 CMS system was developed from *Cajanus cajanifolius*, which is so far the best among different CMS systems developed. This CMS system has a good number of both maintainers and restorers. In A_5 system, the cytoplasm of cultivated species of *Cajanus cajan* was placed along with nuclear genome of a CWR of pigeonpea, *Cajanus acutifolius* (Mallikarjuna and Saxena, 2005) while using *C. cajan* as the female parent. This system also exhibited perfect fertility restoration by cultivated accessions. The A_6 cytoplasm was developed from *Cajanus lineatus* (A_6) in 2002, from one naturally out-crossed plant with erect growth and different morphological traits. This CMS system was observed to be very stable (Saxena et al., 2010) showing perfect fertility restoration by cultivated accessions. The A_7 cytoplasm derived from *C. platycarpus* (A_7) produced good heterosis (Saxena et al., 2010). Four CMS lines, *viz*., "GT 288A," "CMS 67A," "ICRISAT CMS," and "AKCMS 1A," were developed from different wild sources *viz*., *C. sericeus*, *C. scarabaeoides*, and *C. volubilis* (Chauhan et al., 2008).

In *Vigna* species, mungbean × urdbean hybridization has been routinely practiced for mungbean and urdbean improvement programs, and the derivatives from these hybridizations exhibit many desirable features *viz*., resistance to vagaries, both biotic and abiotic, synchronous podding, and non-shattering pods (Pratap et al., 2019). Several traits, such as longer pods, increased seeds number (>10 seeds/pod), and erect plant type, have been transferred from mungbean to urdbean, whereas multiple clusters per peduncle and sympodial pod-bearing habit have been transferred from urdbean into mungbean (Gupta et al., 2004). Similarly, mungbean × ricebean and mungbean × *V. radiata* var. *sublobata* hybridization have also been practiced by many breeders, and progenies were derived which were resistant to MYMV (Verma and Brar, 1996). Singh et al. (2003) produced successful hybrids between *V. radiata* and *V. umbellata* with intermediate morphology and MYMV resistance. Several popular and widely adaptable cultivars have been developed as a result of wild gene introgression in both mungbean and urdbean (**Table 3**). These cultivars show wide adaptation, synchronous maturity, and improved plant architecture in addition to a high degree of resistance to MYMV. Recently, "IPM 312-20" and "Tripura Mung-1" have been developed as a result of mungbean × urdbean hybridization. Likewise, the resultants of mungbean × urdbean crosses were also used further to develop some of the most popular varieties of mungbean. For example, IPM 99-125 was used to develop the most popular pan India variety "IPM 205-7" (Virat) of mungbean (Pratap et al., 2013) which matures in 52–55 days and offers the farmers an excellent choice for cultivation during summer season. Earlier, "IPM 02-3"

TABLE 3 | Commercial cultivars of mungbean and blackgram developed through wild gene introgression.

Crop/variety	Pedigree	Introgression
Mungbean		
Pant Mung-4	T 44 × UPU 2	*V. radiata* × *V. mungo*
HUM-1	BHUM 1 × Pant U-30	*V. radiata* × *V. mungo*
Meha	Pant Mung-2 × AMP 36	*V. radiata* × amphidiploid of (*V. radiata* × *V. mungo*)
Pant Moong-6	Pant Mung-2 × AMP-36	*V. radiata* × amphidiploid of (*V. radiata* × *V. mungo*)
IPM 312-20	IPM 3-1 × SPS 5	*V. radiata* × *V. mungo*
Tripura Mung 1 (TRCM 131)	IPM 99-125 × SPS 5	*V. radiata* × *V. mungo*
Blackgram		
Mash 118	Mungbean × urdbean	*V. radiata* × *V. mungo*
Vamban 7	Vamban-3 × *V. mungo* var. *silvestris*	*V. mungo* × *V. mungo* var. *silvestris*
VBN 6	VBN 1 × *V. mungo* var. *silvestris*	*V. mungo*
TU-40	TU 94-2 × *V. mungo* var. *silvestris*	*V. mungo*
VBG 04-008	Vamban 3 × *V. mungo* var. *silvestris*	*var. silvestris*

was also a highly preferred variety of mungbean by farmers (Singh et al., 2017).

In urdbean, at least five commercial varieties have been developed and deployed using wild gene introgression. The first such variety was Mash 118 developed from an urdbean × mungbean cross in 2008. This was followed by the development of four more cultivars, *viz*., Vamban 7, VBN 6, TU 40, and VBG04-008 in 2011. Among these, VBG04-008 showed high tolerance to heat stress, making it most popular cultivar in heat-prone environments of South India. Interspecific crosses have also been attempted successfully between *V. umbellata* and its wild relatives. However, the success of crosses with respect to pod set differed with the combination of parents involved in the interspecific crosses (Chen et al., 1983; Bharathi et al., 2006).

EPILOGUE

The narrow genetic base of the elite genepool of food legumes and resultant vulnerability of the existing varieties to climate vagaries and changing insect-pest and disease scenario warrants introgression of novel genes or alleles through hybridization and deployment of more diverse germplasm including exotic lines and CWRs in crop improvement programs. Food legumes being majorly grown by small and marginal farmers are more prone to fluctuations in the soil, water, and climate factors as compared to other crops due to limited resources at disposal of these farmers to counter these challenges. The slow process of natural evolution has been significantly replaced by human interventions of domestication, hybridization, and selection. The transformation of humans from food collector to food producer has witnessed the natural attempts of domestication to a planned and focused crop breeding, which has ultimately concluded into the modern "super domestication." While at one side, this has ensured food and nutritional security to ever-increasing population, on the other side, it has narrowed down the genetic base of the cultivated genepool. Keeping this in view, there is a need to reorient legumes improvement programs in such a way that more diverse sources of yield contributing traits, resistance to stresses, both biotic and abiotic, and seed quality are involved in widening the genetic base of cultivated types. This requires trait discovery and deployment from CWR, exotic germplasm, and landraces in mainstream breeding programs. A huge repository of germplasm and CWR (>7 million germplasm accessions) of different crops is maintained together in more than 1,750 national and international genebanks. This includes >86,000 accessions of chickpea[1], >57,000 of *Phaseolus* (Basavaraja et al., 2020), >43,000 of mungbean (Nair et al., 2013; Gayacharan et al., 2020), >13,500 of pigeonpea (Upadhyaya et al., 2011), >16,000 accessions of cowpea, and so on. Nonetheless, characterization and evaluation data on economically important traits are limited to a smaller set of cultivated genepool. The situation is still worse when it comes to CWRs, which needs to be addressed on priority. The genetic bottlenecks leading to narrowing down the genetic base of food legumes need to be recognized and efforts to be initiated through intensive pre-breeding programs. However, owing to pre- and post-fertilization barriers applicable to distant crosses, special tools and techniques need to be adopted. These include application of growth hormones, using mentor pollen technique, deployment of embryo rescue, and several other methods bypassing these barriers (Pratap et al., 2010, 2018b). Wild gene introgression has yielded dividends in some legume crops such as mungbean, urdbean, pigeonpea, chickpea, lentil, etc., either directly through development of commercial cultivars or indirectly through the development of breeding materials and male sterile lines helping in hybrid variety development. Nonetheless, the advantages gained are still far from the potential of gene introgression, and focused planning and implementation in this direction are needed. Development of ILs, NILs, and specialized experimental populations may help in unleashing the genetic and genomic potential of wild gene introgression in the improvement of cultivated food legumes. These populations when subjected to precise and high throughput phenotyping will provide fast and inexpensive genomic information (Pratap et al., 2015b). Wild genetic resources are enormous, opportunities are tremendous, and challenges are manifold. Thus, the need is to venture into the wild gene introgression approach as a long-term strategy with great patience and care.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work has funding support from the International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat Office, Rabat-Institus for open access publication fees.

[1]www.croptrust.com

REFERENCES

Acosta-Gallegos, J. A., Kelly, J. D., and Gepts, P. (2007). Prebreeding in common bean and use of genetic diversity from wild germplasm. *Crop Sci.* 47, S44-S49.

Ali, M. L., Sanchez, P. L., Yu, S., Lorieux, M., and Eizenga, G. C. (2010). Chromosome segment substitution lines: a powerful tool for the introgression of valuable genes from *Oryza* wild species into cultivated rice (*O. sativa* L.). *Rice* 3, 218–234. doi: 10.1007/s12284-010-9058-3

Amalraj, A., Taylor, J., Bithell, S., Li, Y., Moore, K., Hobson, K., et al. (2019). Mapping resistance to Phytophthora root rot identifies independent loci from cultivated (*Cicer arietinum* L.) and wild (*Cicer echinospermum* PH Davis) chickpea. *Theor. Appl. Genet.* 132, 1017–1033. doi: 10.1007/s00122-018-3256-6

Anderson, E. (1949). *Introgressive hybridization*. New York: Wiley.

Anderson, N.O., Ascher, P.D., and Haghighi, K. (1996). Congruity backcrossing as a means of creating genetic variability in self pollinated crops: seed morphology of *Phaseolus vulgaris* L. and *P.acutifolius* A. Gray Hybrids. *Euphytica* 87, 211–224 doi: 10.1007/bf00023748

Arnold, S. J. (1992). Constraints on phenotypic evolution. *Am. Natural.* 140, S85–S107.

Aryamanesh, N., Byrne, O., Hardie, D. C., Khan, T., Siddique, K. H. M., and Yan, G. (2012). Large-scale density-based screening for pea weevil resistance in advanced backcross lines derived from cultivated field pea (*Pisum sativum*) and *Pisum fulvum*. *Crop Pasture Sci.* 63(7), 612–618. doi: 10.1071/cp12225

Badami, P. S., Mallikarjuna, N., and Moss, J. P. (1997). Interspecific hybridization between *Cicer arietinum* and *C. pinnatifidum*. *Plant Breed.* 116, 393–395. doi: 10.1111/j.1439-0523.1997.tb01019.x

Barilli, E., Cobos, M. J., Carrillo, E., Kilian, A., Carling, J., and Rubiales, D. (2018). A high-density integrated DArTseq SNP-based genetic map of *Pisum fulvum* and identification of QTLs controlling rust resistance. *Front. Plant Sci.* 9:167. doi: 10.3389/fpls.2018.00167

Basavaraja, T., Pratap, A., Dubey, V., Gurumurthy, S., Gupta, S., Singh, N. P. et al. (2020). Molecular and Conventional Breeding Strategies for Improving Biotic Stress Resistance in Common Bean. In: Gosal- S.S. Wani S.H. (eds.) Accelerated Plant Breeding. 3. Switzerland: Springer Nature. 389–421. doi: 10.1007/978-3-030-47306-8_13

Bayaa, B., Erskine, W. and Hamdi, A. (1995). Genet. *Resour. Crop Evol.* 42, 231–235.

Bayaa, B., Erskine, W., and Hamdi, A. (1994). Response of wild lentil to *Ascochyta fabae* f.sp. *lentis* from Syria. *Genet. Resour. Crop Evol.* 41, 61–65. doi: 10.1007/BF00053049

Beaver, J.S., Zapata, M., Alameda, M. and Porch, T.G. (2012). Registration of PR0401-259 and PR0650-31 dry bean germplasm lines. *J. Plant Regist.* 6, 81–84. doi: 10.3198/jpr2011.05.0283crg

Bera, S. K., Kamdar, J. H., Kasundra, S. V., Patel, S. V., Jasani, M. D., Maurya, A. K. et al. (2019). Steady expression of high oleic acid in peanut bred by marker-assisted backcrossing for fatty acid desaturase mutant alleles and its effect on seed germination along with other seedling traits. *PloS One* 14:e0226252. doi: 10.1371/journal.pone.0226252

Bernardo, R. (2010). *Breeding for quantitative traits in plants*. Woodbury, MN: Stemma Press.

Bharathi, A., Vijay Selvraj, K. S., Veerabadhiran, P., and Subba Lakshami, B. (2006). Crosability barriers in mungbean (*Vigna radiata* L. Wilczek): with its wild relatives. *Indian J. Crop Sci.* 1, 120–124.

Bhavyasree, R. K., Singh, S., and Singh, I. (2018). Advanced backcross strategy for alien introgression for productivity enhancing traits in chickpea (*Cicer arietinum* L.). *Legume Res.* 41, 379–383.

Bian, J. M., Jiang, L., Liu, L. L., Wei, X. J., Xiao, Y. H., Zhang LJ, et al. (2010). Construction of a new set of rice chromosome segment substitution lines and identification of food weight and related traits QTLs. *Breed. Sci.* 60, 305–313. doi: 10.1270/jsbbs.60.305

Blair M. W., Herrera A. L., Sandoval T. A., Caldas G. V., Fileppi M., Sparvoli F. et al. (2012). Inheritance of seed phytate and phosphorus levels in common bean *Phaseolus vulgaris* L. and association with newly-mapped candidate genes for the phytic acid pathway. *Mol. Breed.* 30, 1265–1277. doi: 10.1007/s11032-012-9713-z

Blair, M. W., Iriarte, G., and Beebe, S. (2006). QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean× wild common bean (*Phaseolus vulgaris* L.) cross. *Theor. Appl. Genet.* 112, 1149–1163. doi: 10.1007/s00122-006-0217-2

Blair, M. W., Pedraza, F., Buendia, H. F., Gatian-Soils, E., Beebe, S. E., Gepts, P. et al. (2003). Development of a genome-wide anchors microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 107, 1362–1374. doi: 10.1007/s00122-003-1398-6

Bock, R. (2009). The give and take of DNA: horizontal gene transfers in plants. *Trends in Plant Sci.* 15, 11–22. doi: 10.1016/j.tplants.2009.10.001

Boukar, O., Belko, N., Chamarthi, S., Togola, A., Batieno, J., Owusu, E., et al. (2019). Cowpea (*Vigna unguiculata*): genetics, genomics and breeding. *Plant Breed.* 138, 415–424.

Boukar, O., Bhattacharjee, R., Fatokun, C., Kumar, L. P., Gueye, B. (2013). Cowpea. In Singh M Upadhyaya HD Bisht IS (eds) *Genetic And Genomic Resources Of Food Legume Improvement*. Elsevier Inc: London. 137–157.

Broich, S. L., Palmer, R. G. (1980). A cluster analysis of wild and domesticated soybean phenotypes. *Euphytica* 29, 23–32. doi: 10.1007/bf00037246

Burow, M. D., Starr, J. L., Park, C. H., Simpson, C. E., and Paterson, A. H. (2014). Introgression of homeologous quantitative trait loci (QTLs) for resistance to the root-knot nematode [*Meloidogyne arenaria* (Neal) Chitwood] in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.). *Mol. Breed.* 34, 393–406. doi: 10.1007/s11032-014-0042-2

Canci, H., and Toker, C. (2009). Evaluation of annual wild Cicer species for drought and heat resistance under field conditions. *Genet. Resour. Crop Evol.* 56, 1–6. doi: 10.1007/s10722-008-9335-9

Chakraborty, S., and Newton, A.C. (2011). Climate change, plant diseases and food security: an overview. *Plant Pathol.* 60, 2–14. doi: 10.1111/j.1365-3059.2010.02411.x

Chaky, J. M., Specht, J. E., and Cregan, P. B. (2003). Advanced backcross QTL analysis. In A Mating between *Glycine max* and *Glycine soja*. *Plant Anim. Genome Abst.* 2003:545.

Chamarthi, S. K., Kumar, A., Vuong, T. D., Blair, M. W., Gaur, P. M., Nguyen, H. T., et al. (2011). "Trait mapping and molecular breeding," in *Biology and Breeding of Food Legumes*, eds A. Pratap and J. Kumar (Wallingford, UK: CABI), 296. doi: 10.1079/9781845937669.0296

Chandel, K. P. S., Lester, R. N. and Starling, R. J. (1984). The wild ancestors of urd and mungbeans (*Vigna mungo* (L.) Hepper and *V. radiata* (L.) Wilczek). *Botanical J. Linnean Soc.* 89, 85–96. doi: 10.1111/j.1095-8339.1984.tb01002.x

Chaturvedi, S. K., Gupta, D. S., and Jain Rashmi. (2011). "Biology of Food Legumes," in *Biology and Breeding of Food Legumes*, eds A. Pratap and J. Kumar (Oxfordshire: CABI), 35–48. doi: 10.1079/9781845937669.0035

Chauhan, R. M., Panwar, L. D., Patel, P. T., and Tikka, S. B. S. (2008). Identification of heterotic combination of CMS lines and restorers of pigeonpea. *J. Food Legum.* 21, 25–27.

Chen, N. C., Baker, R. L. and Honma, S. (1983). Interspecific crossability among four species of *Vigna* food legumes. *Euphytica* 32, 925–937. doi: 10.1007/bf00042175

Chu, Y., Wu, C. L., Holbrook, C. C., Tillman, B. L., Person, G., Ozias-Akins, P. et al. (2011). Marker-assisted selection to pyramid nematode resistance and the high oleic trait in peanut. *Plant Genome* 4, 110–117. doi: 10.3835/plantgenome2011.01.0001

Clemente, A., Arques, M. C., Dalmais, M., Le Signor, C., Chinoy, C., Olias, R., et al. (2015). Eliminating anti-nutritional plant food proteins: the case of seed protease inhibitors in pea. *PLoS One* 10:e0134634. doi: 10.1371/journal.pone.0134634

Cobos, M. J., Satovic, Z., Rubiales, D., and Fondevilla, S. (2018). Er3 gene, conferring resistance to powdery mildew in pea, is located in pea LGIV. *Euphytica* 214:203.

Collard, B. C. Y., Ades, P. K., Pang, E. C. K., Brouwer, J. B., and Taylor, P. W. J. (2001). Prospecting for sources of resistance to ascochyta blight in wild *Cicer* species. *Aus. J .Plant Pathol.* 30, 271–276.

Collard, B. C. Y., Pang, E. C. K., Ades, P. K., and Taylor, P. W. J. (2003). Preliminary investigation of QTLs associated with seedling resistance to ascochyta blight from *Cicer echinospermum*, a wild relative of chickpea. *Theor. Appl. Genet.* 107, 719–729. doi: 10.1007/s00122-003-1297-x

Concibido, V. C., Vallee, B. L., Mclaird, P., Pineda, N., Meyer, J., Hummel, L. et al. (2003). Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor. Appl. Genet.* 106, 575–582 doi: 10.1007/s00122-002-1071-5

Coyne C. J., Kumar S., van Wettberg, E. J. B., Marques, E., Berger, J. D., Redden, R. J., Noel Ellis, T. H., Brus, J., Zoblatzka, L. Smykal, P. (2020). Potential and limits of exploitation of crop wild relatives for pea, lentil and chickpea improvement. *Legume Sci*. 2:e36. doi: 10.1002.leg3.36

Devasirvatham, V., Gaur, P. M., Mallikarjuna, N., Raju, T. N., Trethowan, R. M., and Tan, D. K. (2013). Reproductive biology of chickpea response to heat stress in the field is associated with the performance in controlled environments. *Field Crops Res.* 142, 9–19. doi: 10.1016/j.fcr.2012.11.011

Dongre, T. K., Pawar, S. E., Thakare, R. G. and Harwalkar, M. R. (1996). Identification of resistant source to cowpea weevil [*Callosobruchus maculatus* (F.)] in *Vigna* sp. and inheritanceof their resistance in black gram (*Vigna mungo* var. *mungo*). *J. Stored Prod. Res.* 32, 201–204. doi: 10.1016/s0022-474x(96)00028-8

Doyle, J. J. (1988). 5S ribosomal gene variation in the soybean and its progenitor. *Theor. Appl. Genet*. 75, 621–624. doi: 10.1007/bf00289130

Dwivedi, S. L., Blair, M. W., Upadhyaya, H. D., Serraj, R., Buhariwalla, H. K., Ortiz, R., et al. (2005). Usinggenomics to exploit grain legume biodiversity in cropimprovement. *Plant Breed. Rev*. 26, 171–357. doi: 10.1002/9780470650325.ch6

Egawa, Y., Bujang, I. B., Chotechuen, S., Tomooka, N., and Tateishi, Y. (1996). Phyloegenetic differentiation of tetraploid *Vigna* species, *V. glabrescens and V. reflexo-pilosa*. *Jpn. Int. Res. Center Agric. Sci.* 3, 49–58.

Egawa, Y., Takeda, H., and Suzuki, K. (1999). Research plan on crop heat tolerance at the crop introduction and cultivation laboratory. *Jpn. Int. Res. Center Agric. Sci. Working Rep.* 14, 103–107.

El-Bouhssini, M., Sarker, A., Erskine, W. and Joubi, A. (2008). First sources of resistance to *sitona* weevil *sitona crinitus* herbst. in wild lens species. *Genet. Resour. Crop Evol*. 55, 1–4. doi: 10.1007/s10722-007-9297-3

Esen, A., Sari, H., Erler, F., Adak, A., Sari, D., Eker, T., et al. (2019). Screening and selection of accessions in the genus *Pisum* L. for resistance to pulse beetle (*Callosobruchus chinensis* L.). *Euphytica* 215:82.

Eshed, Y., and Zamir, D. (1994). A genomic library of *Lycopersicon pennellii* in *L. esculentum:* a tool for fine mapping of genes. *Euphytica* 79, 175–179. doi: 10.1007/BF00022516

Eshed, Y., and Zamir, D. (1995). An introgression line population of *Lycopersiconpennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* 141, 1147–1162.

Eshed, Y., and Zamir, D. (1996). Less than additive epistatic interactions of QTL in tomato. *Genetics* 143, 1807–1817.

Fatokun, C. A. (1991). Wide hybridization in cowpea: problems and prospects. *Euphytica* 54, 137–140.

Fernández-Aparicio, M., Sillero, J. C., and Rubiales, D. (2009). Resistance to broomrape in wild lentils (*Lens* spp.). *Plant Breed.* 128, 266–270. doi: 10.1111/j.1439-0523.2008.01559.x

Fiala, J. V., Tullu, A., Banniza, S., Séguin-Swartz, G., and Vandenberg, A. (2009). Interspecies transfer of resistance to anthracnose in lentil (*Lens culinaris* Medic.). *Crop Sci*. 49, 825–830. doi: 10.2135/cropsci2008.05.0260

Fonceka, D., Tossim, H. A., Rivallan, R., Vignes, H., Faye, I., Ndoye, O. et al. (2012). Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. *BMC Plant Biol.* 12:26. doi: 10.1186/1471-2229-12-26

Frisch, M., and Melchinger, A. (2005). Selection theory for marker-assisted backcrossing. *Genetics* 170, 909–917. doi: 10.1534/genetics.104.035451

Fujii, K., Ishimoto, M., and Kitamura, K. (1989). Pattern of resistance to bean weevil Bruchidae. in *Vigna radiata-mungo-sublobata* complex inform the breeding of new resistant varieties. *Appl. Entomol. Zool.* 24, 126–132. doi: 10.1303/aez.24.126

Gautam, H. R., Bhardwaj, M. L., and Kumar, R. (2013). Climate change and its impact on plant diseases. *Curr. Sci.* 105, 1685–1691.

Gayacharan, C., Archak, S., Gupta, K., Gupta, V., Tyagi, V., Singh, K. et al. (2020). Mungbean genetic resources and utilization. In: Nair, R.M., Schafleitner, R, Lee, S. H. (eds.). *The Mungbean Genome*. Springer: Switzerland. 9–25. doi: 10.1007/978-3-030-20008-4_2

Gore, P. G., Tripathi, K., Pratap, A., Bhat, K. V., Umdale, S. D., Gupta, V. et al. (2019). Delineating taxonomic identity of two closely related Vigna species of section Aconitifoliae: *V. trilobata* (L.) Verdc. and *V. stipulacea* (Lam.) Kuntz in India. *Genet. Resour. Crop Evol.* 66, 1155–1165. doi: 10.1007/s10722-019-00767-9

Gujaria-Verma, N., Ramsay, L., Sharpe, A. G., Sanderson, L. A., Debouck, D. G., Tar'an, B., et al. (2016). Gene-based SNP discovery in tepary bean (*Phaseolus acutifolius*) and common bean (*P. vulgaris*) for diversity analysis and comparative mapping. *BMC Genom.* 17:239. doi: 10.1186/s12864-016-2499-3

Gupta S. and Singh, B. B. (2009). Utilization of potential germplasm and species spectrum in improvement of pulse crops. In: M. Ali, S. Gupta, P.S.Basu and K. Naimuddin (eds). *Legumes For Ecological Sustainability*. Kanpur, India: Indian Institute of Pulse Research. 332–341.

Gupta, D., and Sharma, S. K. (2006). Evaluation of wild *Lens* taxa for agro-morphological traits, fungal diseases and moisture stress in northwestern Indian hills. *Genet. Resour. Crop Evol.* 53, 1233–1241. doi: 10.1007/s10722-005-2932-y

Gupta, S., Kumar S. and Singh, B. B. (2004). Relative Genetic Contributions of Ancestral Lines of Indian Mungbean Cultivars Based on Coefficient of Parentage Analysis. *Indian J. Genet.* 64, 299–302.

Haggard, J. E. (2007). Characterization of physiological resistance to white mold and search for molecular markers linked to resistance via advanced backcross QTL analysis in an interspecific cross between *Phaseolus coccineus* and *P. vulgaris*. *HortSci* 41:973.

Haghighi, K. R., and Ascher, P.D. (1988). Fertile, intermediate hybrids between *Phaseolus vulgaris* and *P. acutifolius* from congruity backcrossing. *Sex. Plant Reprod.* 1, 51–58.

Hamdi, A., and Erskine, W. (1996). Reaction of wild species of the genus *Lens* to drought. *Euphytica* 91, 173–179.

Hamdi, A., Küsmenoglu, I., and Erskine, W. (1996). Sources of winter hardiness in wild lentil. *Genet. Resour. Crop Evol.* 43, 63–67. doi: 10.1007/bf00126942

Harouna, D. V., Venkataramana, P. B., Matemu, A. O., and Ndakidemi, P. A. (2020). Agro-morphological exploration of some unexplored wild *Vigna* legumes for domestication. *Agronomy* 10:111. doi: 10.3390/agronomy10010111

Haware, M. P., Nene, Y. L., Pundir, R. P. S. and Narayana, R. J. (1992). Screening of world chickpea germplasm for resistance to fusarium wilt. *Field Crops Res.* 30, 147–154. doi: 10.1016/0378-4290(92)90063-f

He, Q., Yang, H., Xiang, S., Wang, W., Xing, G., Zhao, T., et al. (2014). QTL mapping for the number of branches and pods using wild chromosome segment substitution lines in soybean [*Glycine max* (L.) Merr.]. *Plant Genet. Resour.* 12:S172.

Hingane, A. J., Saxena, K. B., Patil, S. B., Sultana, R., Srikanth, S., Mallikarjuna, N., et al. (2015). Mechanism of water-logging tolerance in pigeonpea. *Indian J. Genet.* 75, 208–214. doi: 10.5958/0975-6906.2015.00032.2

Hospital, F. (2001). Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs. *Genetics* 158, 1363–1379.

Hospital, F., and Charcosset, A. (1997). Marker-assisted introgression of quantitative trait loci. *Genetics* 147, 1469–1485.

Hospital, F., Chevalet, C., and Mulsant, P. (1992). Using markers in gene introgression breeding programs. *Genetics* 132, 1199–1210.

Imtiaz, M., Malhotra, R. S., and Yadav, S. S. (2011). Genetic adjustment to changing climates: chickpea. (eds). S. S. Yadav, R. Redden, J. L. Hatfield, H. Lotze-Campen, A. Hall. In *Crop Adaptation to Climate Change*. Oxford: John Wiley and Sons. 251–268. doi: 10.1002/9780470960929.ch18

Infantino, A., Porta-Puglia, A., and Singh, K. B. (1996). Screening wild *Cicer* species for esistance to fusarium wilt. *Plant Dis.* 80, 42–44. doi: 10.1094/pd-80-0042

Iseki, K., Takahashi, Y., Muto, C., Naito, K., and Tomooka, N. (2016). Diversity and evolution of salt tolerance in the genus *Vigna*. *PLoS One* 11:e0164711. doi: 10.1371/journal.pone.0164711

Isenegger, D.A., Ford, R. and Taylor, P.W. (2011). Disease reaction of chickpea (*Cicer* spp.) genotypes to Botrytis grey mould (Botrytis cinerea). *Aus. Plant Pathol.* 40:583. doi: 10.1007/s13313-011-0081-7

Jain, H. K. (1975). Breeding for yield and other attributes in food legumes. *Ind. J. Genet*. 35, 169–187.

Jarvis, D. I. and Hodgkin, T. (1998). Wild relatives and crop cultivars: conserving the connection. (eds). N. Zencirci, Z. Kaya, Y. Anikster and W. T. Adams. In: *The Proceedings of an International Symposium on in Situ Conservation of Plant Genetic Diversity*. pp. 73–80. Turkey: Central Research Institute for Field Crops.

Jarvis, D. I. and Hodgkin, T. (1999). Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Mol. Ecol.* 8, S159–S173.

Kaiser, W. J., Alcala-Jimenez, A. R., Hervas-Vargas, A., Trapero-Casas, J. L., and Jimenez-Diaz, R. M. (1994). Screening of wild *Cicer* species for resistance to race 0 and 5 of *Fusarium oxysporum* f. sp. *ciceris. Plant Dis.* 78, 962–967.

Kashiwagi, J., Krishnamurthy, L., Gaur, P. M., Upadhyaya, H. D., Varshney, R. K., and Tobita, S. (2013). Traits of relevance to improve yield under terminal drought stress in chickpea (*C. arietinum* L.). *Field Crops Res.* 145, 88–95. doi: 10.1016/j.fcr.2013.02.011

Khoury, C. K., Castañeda-Alvarez, N. P., Achicanoy, H. A., Sosa, C. C., Bernau, V., Kassa, M. T., et al. (2015). Crop wild relatives of pigeonpea [*Cajanus cajan* (L.) Millsp.]: distributions, *ex situ* conservation status, and potential genetic resources for abiotic stress tolerance. *Biol. Conserv.* 184, 259–270. doi: 10.1016/j.biocon.2015.01.032

Knights, E. J., Southwell, R. J., Schwinghamer, M. W., and Harden, S. (2008). Resistance to *Phytophthora medicaginis* Hansen and Maxwell in wild *Cicer* species and its use in breeding root rot resistant chickpea (*Cicer arietinum* L.). *Aus. J. Agric. Res.* 59, 383–387. doi: 10.1071/ar07175

Koona, P., Osisanya, E. O., Jackai, L. E. N., Tamo, M., and Kham, R. H. (2002). Resistance in accessions of cowpea to the coreid pod bug *Clavigralla tomentosicillis* (Hemiptera: coreidae). *J. Econ Entomol.* 95, 1281–1288. doi: 10.1603/0022-0493-95.6.1281

Konarev, A. V., Tomooka, N., and Vaughan, D. A. (2002). Proteinase inhibitor polymorphism in the genus *Vigna* Savi subgenus Ceratotropis and its biosystematic implications. *Euphytica* 123, 165–177.

Konečná, E., Šafářová, D., Navratil, M., Hanáček, P., Coyne, C., Flavell, A., et al. (2014). Geographical gradient of the eIF4E alleles conferring resistance to potyviruses in pea (*Pisum*) germplasm. *PLoS One* 9:e90394. doi: 10.1371/journal.pone.0090394

Koumproglou, R., Wilkes, T. M., Townson, P., Wang, X. Y., Beynon, J., Pooni, H. S. et al. (2002). STAIRS: a new genetic resource for functional genomic studies of Arabidopsis. *Plant J.* 31, 355–364. doi: 10.1046/j.1365-313x.2002.01353.x

Krishnamurthy, L., Kashiwagi, J., Upadhyaya, H. D., Gowda, C. L. L., Gaur, P. M., Singh, S., et al. (2013). Partitioning coefficient—A trait that contributes to drought tolerance in chickpea. *Field Crops Res.* 149, 354–365. doi: 10.1016/j.fcr.2013.05.022

Kumar, D., and Kirti, P. B. (2015). Transcriptomic and proteomic analyses of resistant host responses in *Arachis diogoi* challenged with late leaf spot pathogen, *Phaeoisariopsis personata*. *PLoS One* 10:e0117559. doi: 10.1371/journal.pone.0117559

Kumar, N. P., Pandiyan, M. and Veerabadhiran, P. (2007). Pre-fertilization barriers in *Vigna radiata* × *Vigna umbellata*. *Plant Arch.* 7, 377–380.

Kumar, J., Srivastava, E., Singh, M., Mahto, D., Pratap, A., and And Kumar, S. (2014). "Lentil," in *Alien Gene Transfer in Crop Plants, Volume 2: Achievements and Impacts*, eds A. Pratap and J. Kumar (New York: Springer), 191–206.

Kumar, S., Choudhary, A. K., Rana, K. S., Sarker, A., and Singh, M. (2018). Biofortification potential of global wild annual lentil core collection. *PLoS One* 13:e0191122. doi: 10.1371/journal.pone.0191122

Kumar S., Imtiaz M., Gupta S., and Pratap A. (eds.). (2011). Distant hybridization and alien gene introgression. Biology and Breeding of Food Legumes. Oxfordshire, UK: CABI. 81–110.

Kumar, S., Gupta, P., Choukri, S., and Siddique, K. H. M. (2020). Efficient breeding of pulse crops. (eds). S. S. Gosal and S. H. Wani. In: *Accelerated Plant Breeding, Volume 3-Food Legumes*. Switzerland: Springer Nature.

Kumar, S., Gupta, S., Chandra, S., and Singh, B. B. (2004). How wide is the genetic base of pulse crops? In: M. Ali, B. B. Singh, S. Kumar, and V. Dhar (eds), *Pulses in New Perspective*. Kanpur, India: Indian Institute of Pulse Research. 211—221.

Kumar, S., Rakshit S. and Gupta S. (2003). Genetics and cytogenetics of chickpea. In: M. Ali, S. Kumar, and N. B. Singh (eds). *Chickpea Research in India*. Kanpur, India: Indian Institute of Pulse Research. 31—37.

Kumar, S., Sandhu, J. S., Gupta, S. K., Dikshit, H. K., and Gupta, S. (2009). Broadening genetic base of pulse crops: present status and future strategies. In: M.Ali, S. Gupta, P.S.Basu and K. Naimuddin (eds), *Legumes for Ecological Sustainability*. Kanpur, India: Indian Institute of Pulse Research. 332–341.

Kushida, A., Tazawa, A., Aoyama, S., and Tomooka, N. (2012). Novel sources of resistance to the soybean cyst nematode (*Heterodera glycines*) found in wild relatives of azuki bean (*Vigna angularis*) and their characteristics of resistance. *Genet. Resour. Crop Evol.* 60, 985–994. doi: 10.1007/s10722-012-9895-6

Ladizinsky, G. (1985). The genetics of hard seed coat in the genus Lens. *Euphytica* 34, 539–543 doi: 10.1007/bf00022952

Ladizinsky, G., and Adler, A. (1976). Genetic relationships among the annual species of *Cicer* L. *Theor. Appl. Genet.* 48, 197–203. doi: 10.1007/bf00527371

Laserna-Ruiz, I., De-Los-Mozos-Pascual, M., Santana-Meridas O., Sanchez-Vioque R., and Rodrıguez-Conde M. F. (2012). Screening and selection of lentil (*Lens* Miller) germplasm resistant to seed bruchids (*Bruchus* spp.). *Euphytica* 188, 153–162. doi: 10.1007/s10681-012-0752-7

Li, D. D., Pfeiffer, T. W., and Cornelius, P. L. (2008). Soybean QTL for yield and yield components associated with *Glycine soja* alleles. *Crop Sci.* 48, 571–581 doi: 10.2135/cropsci2007.06.0361

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M. W. (2008). Control of final seed and organ size by the DA1 gene family in *Arabidopsis thaliana*. *Genes Dev.* 22, 1331–1336. doi: 10.1101/gad.463608

Lílian, S. T., Carmo Andressa, C. Q., Martins Cinthia, C. C., Martins Mário, A. S., Passos Luciano, P., Silva Ana, C. G., et al. (2019). Comparative proteomics and gene expression analysis in *Arachis duranensis* reveal stress response proteins associated to drought tolerance. *J. Proteomics* 192, 299–310. doi: 10.1016/j.jprot.2018.09.011

Liu, S., Zhou, R., Dong, Y., Li, P., and Jia, J. (2006). Development, utilization of introgression lines using a synthetic wheat as donor. *Theor. Appl. Genet.* 112, 1360–1373. doi: 10.1007/s00122-006-0238-x

Lulsdorf, M., Mallikarjuna, N., Clarke, H., and Tar'An, B. (2005). "*Finding Solutions For Interspecific Hybridization Problems In Chickpea (Cicer arietinum L.),*". in 4th *International Food Legumes Research Conference*. (New Delhi: IFLRC).

Mallikarjuna, N., and Jadhav, D. R. (2008). Techniques to produce hybrids between *Cicer arietinum* L. x *Cicer pinnatifidum* Jaub. *Ind. J. Genet.* 68, 398–405.

Mallikarjuna, N., and Saxena, K. B. (2005). A new cytoplasmic nuclear male-sterility system derived from cultivated pigeonpea cytoplasm. *Euphytica* 142, 143–148. doi: 10.1007/s10681-005-1199-x

Mallikarjuna, N., Jadhav, D. R., Nagamani, V., Amudhavalli, C., and Hoisington, D. A. (2007). Progress in interspecific hybridization between *Cicer arietinum* and wild species C. *bijugum*. *J.SAT Agric. Res.* 5, 1–2. doi: 10.1007/978-3-319-23534-9_1

Mallikarjuna, N., Saxena, R. K., Gowda, M. B., and Varshney, R. K. (2017). *Wide Crossing Technology for Pigeonpea Improvement the Pigeonpea Genome*. Chamsford: Springer, 31–39.

Malhotra, R. S., Singh, K. B., Di Vito, M., Greco, N., and Saxena, M. C. (2002). Registration of ILC 10765 and ILC 10766 chickpea germplasm lines resistant to cyst nematode. *Crop Sci.* 42, 1756–1756. doi: 10.2135/cropsci2002.1756

Mallikarjuna, N., Senapathy, S., Jadhav, D. R., Saxena, K. B., Sharma, H. C., Upadhyaya, H. D. et al. (2011). Progress in the utilization of *Cajanus platycarpus* (Benth.) Maesen in pigeonpea improvement. *Plant Breed.* 130, 507–514 doi: 10.1111/j.1439-0523.2011.01870.x

Mallikarjuna, N., Senthilvel, S., and Hoisington, D. (2012). Development of new sources of tetraploid Arachis to broaden the genetic base of cultivated groundnut (*Arachis hypogaea* L.). *Genet. Resour. Crop. Evol.* 58, 889–907. doi: 10.1007/s10722-010-9627-8

Manavalan, L. P., Guttikonda, S. K., Tran, L. P., and Nguyen, H. T. (2009). Physiological and molecular approaches to improve drought resistance in soybean. *Plant Cell. Physiol.* 50, 1260–1276. doi: 10.1093/pcp/pcp082

Manavalan, L.P., Prince, S. J., Musket, T. A., Chaky, J., Deshmukh, R., Vuong, T. D., et al. (2015). Identification of novel QTL governing root architectural traits in an interspecific soybean population. *PLoS One* 10:e0120490. doi: 10.1371/journal.pone.0120490

Mannur, D. M., Babbar, A., Thudi, M., Sabbavarapu, M. M., Roorkiwal, M., Sharanabasappa, B. Y. et al. (2019). Super Annigeri 1 and improved JG 74: two Fusarium wilt-resistant introgression lines developed using marker-assisted backcrossing approach in chickpea (*Cicer arietinum* L.). *Mol. Breed.* 39:2.

Matus, I., Corey, A., Filchkin, T., Hayes, P. M., Vales, M. I., Kling, J., et al. (2003). Development and characterization of recombinant chromosome substitution

lines (RCSLs) using *Hordeum vulgare* subsp. *vulgare* background. *Genome* 46, 1010–1023. doi: 10.1139/g03-080

Miklas, P.N., Zapata, M., Beaver, J.S., and Grafton, K.F. (1999). Registration of four dry bean germplasms resistant to common bacterial blight: ICB-3, ICB-6, ICB-8 and ICB-10. *Crop Sci.* 39:594. doi: 10.2135/cropsci1999.0011183x003900020065x

Miyagi, M., Humphry, M. E., Ma, Z.Y., Lambrides, C. J., Bateson, M., Liu, C. J. et al. (2004). Construction of bacterial artificial chromosome libraries and their application in developing PCR-based markers closely linked to a major locus conditioning bruchid resistance in mungbean (*Vigna radiata* L. Wilczek). *Theor. Appl. Genet.* 110, 151–156. doi: 10.1007/s00122-004-1821-7

Monforte, A. J., and Tanksley, S. D. (2000). Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: a tool for gene mapping and gene discovery. *Genome* 43, 803–813. doi: 10.1139/gen-43-5-803

Mwale, S. E., Shimelis, H., Mafongoya, P., and Mashilo, J. (2020). Breeding tepary bean (*Phaseolus acutifolius*) for drought adaptation: a review. *Plant Breed.* 139, 821–833. doi: 10.1111/pbr.12806

Nagaraj, N. C., Muniyappa, V., Satyan, B. A., Shanmugam, N., Jayarajan, R., and Vidhyasekaran, P. (1981). *Resistance source for mungbean yellow mosaic virus. In Proceedings of the National Seminar on Disease Resistance in Crop Plants*. 69–72.

Naim-Feil, E., Toren, M., Aubert, G., Rubinstein, M., Rosen, A., Eshed, R., et al. (2017). Drought response and genetic diversity in *Pisum fulvum*, a wild relative of domesticated pea. *Crop Sci.* 57, 1145–1159. doi: 10.2135/cropsci2016.10.0880

Nair, R. M., Schafleitner, R., Kenyon, L. Srinivasan, R., Easdown, W, Ebert A.W. et al. (2013). Gnetic improvement of mungbean. *SABRAO J. Breed. Genet.* 44, 177–190.

Ngugi-Dawit, A., Hoang, T. M. L., Williams, B., Higgins, T. J., and Mundree, S. G. (2020). A wild *Cajanus scarabaeoides* (L.), Thouars, IBS 3471, for improved insect-resistance in cultivated pigeonpea. *Agronomy* 10:5.

Oghiakhe, S., Jackai, L. E. N., and Makanjuola, W. A. (1992). A rapid visual field screening technique for resistance of cowpea (*Vigna unguiculata*) to the legume pod borer *Maruca testulalis* (Lepidoptera: Pyralidae). *Bull. Entomol. Res.* 82, 507–512. doi: 10.1017/s0007485300042589

Oigiangbe, O. N., Jackai, L. E. N., Ewete, F. K., Hughes, J., and Lajide, L. (2002). Reduced consumption and use of pods of *Vigna* species (Leguminosae) by *Maruca vitrata* (Lepidoptera: Pyralidae). *African Entomol.* 10(2), 333–340.

Pal, S. S., Singh, J. J., and Singh, I. (2000). Transfer of YMV resistance in cultivar SML32 of *Vigna radiata* from other related *Vigna* species. *Plant Dis. Res.* 15, 67–69.

Pande, S., Ramgopal, D., Kishore, G. K., Mallikarjuna, N., Sharma, M., Pathak, M. et al. (2006). Evaluation of wild *Cicer* species for resistance to ascochyta blight and botrytis gray mold in controlled environment at icrisat, patancheru, india. *J. SAT Agric. Res.* 13:25-26.

Pandiyan, M., Ramamoorthi, N., Ganesh, S. K., Jebraj, S., Pagarajan, P., Balasubramanian, P. et al. (2008). Broadening the genetic base and introgression of MYMY resistance and yield improvement through unexplored genes from wild relatives in mungbean. *Plant Mutation Rep.* 2, 33–38.

Peleman, J. D., and van der Voort, J. R. (2003). Breeding by design. *Trends Plant Sci.* 8, 330–334.

Pillen, K., Zacharias, A., and Léon, J. (2003). Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 107, 340–352. doi: 10.1007/s00122-003-1253-9

Podder, R., Banniza, S., and Vandenberg, A. (2013). Screening of wild and cultivated lentil germplasm for resistance to stemphylium blight. *Plant Genet. Resour.* 11:26. doi: 10.1017/s1479262112000329

Porch, T. G., Cichy, K., Wang, W., Brick, M., Beaver, J. S., Santana-Morant, D., et al. (2017). Nutritional composition and cooking characteristics of tepary bean (*Phaseolus acutifolius* Gray) in comparison with common bean (*Phaseolus vulgaris* L.). *Genetic Resour. Crop Evol.* 64, 935–953. doi: 10.1007/s10722-016-0413-0

Polhill, R. M., and Raven P. H. (1981). *Advances in Legume Systematics, Part 1 and 2*. Kew, UK: Royal Botanical Garden.

Pratap, A. Gupta, S. K., Kumar, J., and Solanki, R. K. (2012). Soybean. In: Gupta S.K. (Ed.) *Technological Innovations In Major World Oil Crops, Vol.I Breeding*. Germany: Springer Science + Business Media. 293–321.

Pratap, A., and Gupta, S. K. (2009). *Biotechnological Interventions in Host Plant Resistance*, eds R. Peshin and A. K. Dhawan in Integrated Pest Management: Innovation, Dissemination and Impact. Dordrecht: Springer Publishers. 183–207. doi: 10.1007/978-1-4020-8992-3_8

Pratap, A., Gupta, D. S., Singh, B. B., and Kumar, S. (2013). IPM 205-7 (IC0589309-IC0589310; INGR11043-INGR11044), a mungbean (*Vigna radiata* (L.) Wilczek) germplasm with Super Early Maturity. *Indian J. Plant Genet. Resour*. 26, 89–90.

Pratap, A., Gupta, S., Basu, P. S., Tomar, R., Dubey, S., Rathore, M., et al. (2019). "Towards development of climate smart mungbean: challenges and opportunities," in *Genomic Designing of Climate-Smart Pulse Crops*, ed. C. Kole (Chamsford, UK: Springer), 235–264. doi: 10.1007/978-3-319-96932-9_5

Pratap, A., and Kumar, J. (2011). History origin and evolution. (eds). A. Pratap, J. Kumar. In *Biology and Breeding of Food Legumes*. Oxfordshire: CABI.

Pratap, A., Basu, P. S., Gupta, S., Malviya, N., Rajan, N., Tomar, R. et al. (2014). Identification and characterization of sources for photo- and thermo-insensitivity in *Vigna* species. *Plant Breed.* 133, 756–764. doi: 10.1111/pbr.12215

Pratap, A., Chaturvedi, S.K., Tomar, R., Rajan, N., Malviya, N., Thudi, M. et al. (2017). Marker-assisted introgression of resistance to Fusarium wilt race 2 in Pusa 256, an elite cultivar of *desi* chickpea. *Mol. Genet. Gen.* 292, 1237–1245. doi: 10.1007/s00438-017-1343-z

Pratap, A., Choudhary, A. K., and Kumar, J. (2010). *In vitro* techniques towards genetic enhancement of food legumes-a review. *J. Food Legum.* 23, 169–185.

Pratap, A., Douglas, C., Prajapati, U., Kumari, G., War, A. R., Tomar, R. et al. (2020). Breeding progress and future challenges: Biotic stresses. In: Nair, R.M., Schafleitner, R and Lee, S. H (Eds.) *The Mungbean Genome*. Switzerland: Springer Nature. 55–80. doi: 10.1007/978-3-030-20008-4_5

Pratap, A., Gupta, S., Malviya, N., Rajan, N., Tomar, R., Latha, M. et al. (2015a). Genome scanning of asiatic *vigna* species for discerning population genetic structure based on microsatellite variation. *Mol. Breed.* 35:178.

Pratap, A., Malviya, N., Gupta, S., Tomar, R., Pandey, V. R., Prajapati, U. et al. (2018a). Field characterization of endemic wild *Vigna* accessions collected from biodiversity hotspots of India to identify promising genotypes for multiple agronomic and adaptive traits. *Legume Res.* 41, 490–499.

Pratap, A., Prajapati, U., Singh, C. M., Gupta, S., Rathore, M., Malviya, N. et al. (2018b). Potential, constraints and applications of *in vitro* methods in improving food legumes. *Plant Breed.* 137, 235–249. doi: 10.1111/pbr.12590

Pratap, A., Tomar, R., Kumar, J., Pandey, V. R., Menhandi, S., Katiyar, P. K. et al. (2015b). High throughput plant phenotyping platforms. In: Kumar, J., Pratap, A. and Kumar, S. (Eds.). *Phenomics Of Crop Plants: Trends, Options And Limitations*. New Delhi: Springer (India) Pvt. Ltd. 285–296 doi: 10.1007/978-81-322-2226-2_19

Price, A. H., Townend, J., Jones, M. P., Audebert, A., and Courtois, B. (2002). Mapping QTLs associated with drought avoidance in upland rice grown in the philippines and west africa. *Plant Mol. Biol.* 48, 683–695.

Pundir, R. P. S., and Mengesha, M. H. (1995). Cross compatibility between chickpea and its wild relative, *Cicer echinospermum* Davis. *Euphytica* 83, 241–245. doi: 10.1007/bf01678136

Ramgopal, D., Srivastava, R. K., Pande, S., Rathore, A., Jadhav, D. R., Sharma, M., et al. (2012). Introgression of Botrytis gray mold resistance gene from *Cicer reticulatum* (bgmr1cr) and *C. echinospermum* (bgmr1ce) to chickpea. *Plant Genet. Resour.* 11, 212–216. doi: 10.1017/s1479262112000470

Reddy, K. R., and Singh, D. P. (1990). Inheritance of resistance to mungbean yellow mosaic virus in blackgram. *Madras Agric. J.* 80, 199–201.

Reddy, M. V., and Singh, K. B. (1993). Rate reducing resistance *to Ascochyta* blight in chickpeas. *Plant Dis.* 77, 231–233. doi: 10.1094/pd-77-0231

Reen, R., Mumford, M., and Thompson, J. P. (2019). Novel sources of resistance to root-lesion nematode (*Pratylenchus thornei*) in a new collection of wild Cicer species (*C. reticulatum* and *C. echinospermum*) to improve resistance in cultivated chickpea *C. arietinum*. *Phytopathology* 109, 1270–1279. doi: 10.1094/phyto-02-19-0047-r

Robertson, L. D., Singh, K. B., and Ocampo, B. (1995). *A Catalog of Annual Cicer Species*. Aleppo: ICARDA.

Saeed, A., Darvishzadeh, R., Hovsepyan, H., and Asatryan, A. (2010). Tolerance to freezing stress in *Cicer* accessions under controlled and field conditions. *African J. Biotech.* 18, 2618–2626.

Saeed, A., and Darvishzadeh, R. (2017). Association analysis of biotic and abiotic stresses resistance in chickpea (*Cicer* s) using AFLP markers. *Biotechnol. Equip.* 31, 698–708. doi: 10.1080/13102818.2017.1333455

Saxena, K. B., and Kumar, R. V. (2003). Development of a cytoplasmic nuclear male-sterility system in pigeonpea using *C. scarabaeoides* (L.) Thouars. *Ind. J. Genet. Plant Breed.* 63, 225–229.

Saxena, K. B., Choudhary, A. K., Saxena, R. K., and Varshney, R. K. (2018). Breeding pigeonpea cultivars for intercropping: synthesis and strategies. *Breed. Sci.* 68, 159–167. doi: 10.1270/jsbbs.17105

Saxena, K. B., Kumar, R. V., Dalvi, V. A., Mallikarjuna, N., Gowda, C. L. L., Singh, B. B., et al. (2005). Hybrid breeding in food legumes– a success story of pigeonpea. (eds). M. C. Khairwal and H. K. Jain. in *Proceedings Of The International Food Legumes Research Conference*. (New Delhi: IFLRC-IV).

Saxena, K. B., Singh, L., Kumar, R. V., and Rao, A. N. (1996). "Development of cytoplasmic-genic male-sterility (CMS) system in pigeonpea at ICRISAT Asia Center," in *Proceedings of the Working Group on Cytoplasmic-Genic Male-Sterility (CMS) in Pigeonpea*, (Andhra Pradesh: ICRISAT Center), 9–10.

Saxena, K. B., Sultana, R., Mallikarjuna, N., Saxena, R. K., Kumar, R. V., Sawargaonkar, S. L., et al. (2010). Male-sterility systems in pigeonpea and their role in enhancing yield. *Plant Breed.* 129, 125–134. doi: 10.1111/j.1439-0523.2009.01752.x

Schafleitner, R., Huang, S. M., Chu, S. H., Yen, J. Y., Lin, C. Y., Yan, M. R., et al. (2016). Identification of single nucleotide polymorphism markers associated with resistance to bruchids (*Callosobruchus s*) in wild mungbean (*Vigna radiata* var. *sublobata*) and cultivated *V. radiata* through genotyping by sequencing and quantitative trait locus analysis. *BMC Plant Biol.* 16:159. doi: 10.1186/s12870-016-0847-8

Sekhon, J., Grewal, S. K., Singh, I., and Kaur, J. (2017). Evaluation of nutritional quality and antioxidant potential of pigeonpea genotypes. *J. Food Sci. Tech.* 54, 3598–3611. doi: 10.1007/s13197-017-2818-y

Sen Gupta, D., Thavarajah, D., McGee, R. J., Coyne, C. J., Kumar, S., and Thavarajah, P. (2016). Genetic diversity among cultivated and wild lentils for iron, zinc, copper, calcium and magnesium concentrations. *Austr. J. Crop Sci.* 10:1381.

Shah, T. M., Hassan, M., Haq, M. A., Atta, B. M., Alam, S. S., Ali, H. et al. (2005). Evaluation of *Cicer* species for resistance to Ascochyta blight. *Pakistan J. Bot.* 37, 431–438.

Sharma, S., and Upadhyaya, H. D. (2015). Vernalization and photoperiod response in annual wild *Cicer* species and cultivated chickpea. *Crop Sci.* 55, 2393–2400. doi: 10.2135/cropsci2014.09.0598

Sharma, S., and Upadhyaya, H. D. (2019). Photoperiod response of annual wild *Cicer* species and cultivated chickpea on phenology, growth, and yield traits. *Crop Sci.* 59, 632–639. doi: 10.2135/cropsci2018.07.0438

Sharma, K. D., Winter, P., Kahl, G., and Muehlbauer, F. J. (2004). Molecular mapping of *Fusarium oxysporum* f. sp. *ciceris* race 3 resistance gene in chickpea. *Theor. Appl. Genet.* 108, 1243–1248. doi: 10.1007/s00122-003-1561-0

Sharma, R., Rawat, V., and Suresh, C. G. (2017). Genome-wide identification and tissue-specific expression analysis of nucleotide binding site-leucine rich repeat gene family in *Cicer arietinum* (*kabuli* chickpea). *Genomics Data* 14, 24–31. doi: 10.1016/j.gdata.2017.08.004

Sharma, S. S., Islam, M. A., Malik, A. A., Kumar, K., Negi, M. S., Tripathi, S. B. et al. (2016). Seed traits, fatty acid profile and genetic diversity assessment in *Pongmia pinnata* (L.) Pierre germplasm. *Physiol. Mol. Biol. Plants* 22, 193–205. doi: 10.1007/s12298-016-0356-0

Shoemaker, R. C. (1986). Chloroplast DNA variation in the genus *Glycine* subgenus *soja*. *J. Hered.* 77, 26–30. doi: 10.1093/oxfordjournals.jhered.a110161

Sinclair, T. R., Purcell, L. C., King, C. A., Sneller, C. H., Chen, P., Vadez, V., et al. (2007). Drought tolerance and yield increase of soybean resulting from improved symbiotic N fixation. *Field Crops Res.*101, 68–71. doi: 10.1016/j.fcr.2006.09.010

Singh, M., Rana, M. K., Kumar, K., Bisht, I. S., Dutta, M., Gautam, N. K., et al. (2013). Broadening the genetic base of lentil cultivars through inter-sub-specific and interspecific crosses of Lens taxa.ă*Plant Breed.* 132, 667–675. doi: 10.1111/pbr.12089

Singh, B. B., and Dikshit, H. K. (2002). Possibilities and limitations of interspecific hybridization involving green gram (*Phaseolus radiatus*) and black gram (*Phaseolus mungo*). *Indian J. Agril. Sci.* 72, 676–678.

Singh, D. P. (1990). Distant hybridization in genus *Vigna*: a review. *Ind. J. Genet. Plant Breed.* 50, 268–276.

Singh, K. B., and Ocampo, B. (1993). Interspecific hybridization in annual *Cicer* species. *J. Genet. Breed*. 47, 199–204.

Singh, K. B., Malhotra, R. S., and Saxena, M. C. (1990). Sources for tolerance to cold in *Cicer* species. *Crop Sci.* 30, 1136–1138. doi: 10.2135/cropsci1990.0011183x003000050036x

Singh, K. B., Ocampo, B., and Robertson, L. D. (1998). Diversity for abiotic and biotic stress resistance in the wild annual Cicer species. *Gene. Res. Crop Evol.* 45, 9–17. doi: 10.1023/A:1008620002136

Singh, K. P., Monika, Sareen, P. K., and Kumar, A. (2003). Interspecific hybridization studies in *Vigna radiata* L. Wilczek and *Vigna umbellata* L. *Nat. J. Pl. Improve.* 5, 16–18.

Singh, M., Kumar, S., Basandrai, A. K., Basandrai, D., Malhotra, N., Saxena, D. R. et al. (2020). Evaluation and identification of wild lentil accessions for enhancing genetic gains of cultivated varieties. *PLoS One* 15:e0229554. doi: 10.1371/journal.pone.0229554

Singh, M., Mishra, J. S., and Bhatt, B. P. (2017). Effect of integrated nutrient management on production potential and quality of summer mungbean (*Vigna radiata* L.). *J. Krishi Vigyan* 5, 39–45. doi: 10.5958/2349-4433.2017.00009.5

Singh, N. P., Singh, A., Asthana, A. N., and Singh, A. (1999). Studies on inter-specific crossability barriers in chickpea. *Ind. J. Pulses Res.*12, 13–19.

Singh, R. P., and Singh, B. D. (1989). Recovery of rare interspecific hybrids of gram *Cicer arieti*num × *C. cuneatum* L. through tissue culture. *Curr. Sci.* 58, 874–876.

Singh, S. P., Debouck, D. G., and Roca, W. W. (1998). Interspecific hybridization between *Phaseolus vulgaris* L. and *P. parvifolius* Freytag. *Annu. Rep. Bean Improv. Coop.* 4, 7–8.

Singh, S., Gumber, R. K., Joshi, N., and Singh, K. (2005). Introgression from wild *Cicer reticulatum* to cultivated chickpea for productivity and disease resistance. *Plant Breed.* 124, 477–480. doi: 10.1111/j.1439-0523.2005.01146.x

Singh, S.P., Debouck, D.G., and Roca, W.W. (1997). Successful interspecific hybridization between *Phaseouls vulgaris* L. and *P. costaricensis* Freytag & Debouck. *Annu. Rep. Bean Improv. Coop.* 40, 40–41.

Singh, S.P., Terán, H., and Beaver, J.S. (2009). Scarlet runner bean germplasm accessions G 35006 and G 35172 possess resistance to multiple diseases of common bean. *Annu. Rep. Bean Improv. Coop.* 52, 22–23.

Somta, C., Somta, P., Tomooka, N., Ooi, P. C., Vaughan, D. A., and Srinives, P. (2008). Characterization of new sources of mungbean (*Vigna radiata* (L.) Wilczek) resistance to bruchids, *Callosobruchus* s (Coleoptera: Bruchidae). *J. Stored Products Res.* 44, 316–321. doi: 10.1016/j.jspr.2008.04.002

Somta, P., Kaga, A., Tomooka, N, Kashiwaba K., Isemura T., Chaitieng B. et al. (2006). Development of *an* interspecific vigna linkage map between *Vigna umbellata* (Thunb.) Ohwi & Ohashi and *V. nakashimae* (Ohwi) Ohwi & Ohashi and its use in analysis of bruchid resistance and comparative genomics. *Plant Breed.*125, 77–84. doi: 10.1111/j.1439-0523.2006.01123.x

Souter, J. R., Gurusamy, V., Porch, T. G., and Bett, K. E. (2017). Successful introgression of abiotic stress tolerance from wild tepary bean to common bean. *Crop Sci.* 57, 1160–1171. doi: 10.2135/cropsci2016.10.0851

Stamigna, C., Crino, P., and Saccordo, F. (2000). Wild relatives of chickpea: multiple disease resistance and problems to introgression in the cultigen. *J. Genet. Breed.* 54, 213–219.

Sullivan, J. G., and Bliss, F. A. (1983). Expression of enhanced seed protein content in inbred backcross lines of common bean. *J. Am. Soc. Hortic. Sci.* 108, 787–791.

Tahir, M., Lindeboom, N., Båga, M., Vandenberg, A., and Chibbar, R. (2011). Composition and correlation between major seed constituents in selected lentil (*Lens culinaris*. Medik) genotypes. *Can. J. Plant Sci.* 91, 825–835. doi: 10.4141/cjps2011-010

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps, unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

Tanksley, S. D., and Nelson, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92, 191–203. doi: 10.1007/s001220050114

Tanksley, S. D., Young, N. D., Paterson, A. H., and Bonierbale, M. W. (1989). RFLP mapping in plant breeding: new tools for an old science. *Biotechnology* 7, 257–264. doi: 10.1038/nbt0389-257

Tian, F., De, J. L., Fu, Q., Zuo, F. Z., Yong, C. F., Xiang, K. W., et al. (2006). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theor. Appl. Genet.* 112, 570–580. doi: 10.1007/s00122-005-0165-2

Tikka, S. B. S., Parmar, L. D., and Chauhan, R. M. (1997). First record of cytoplasmic-genic male sterility system in pigeonpea (*Cajanus cajan* (L.) Millsp.) through wide hybridization. *Gujarat Agric. Univ. Res. J.* 22, 160–162.

Toker, C. (2005). Preliminary screening and selection for cold tolerance in annual wild *Cicer* species. *Genetic. Res. Crop. Evol.* 52, 1–5. doi: 10.1007/s10722-005-1743-5

Toker, C., Canci, H., and Yildirim, T. (2007a). Evaluation of perennial wild *Cicer* species for drought resistance. *Genetic. Res. Crop. Evol.* 54, 1781–1786. doi: 10.1007/s10722-006-9197-y

Toker, C., Lluch, C., Tejera, N. A., Serraj, R., and Siddique, K. H. M. (2007b). Abiotic Stresses. (eds). S.S. Yadav, R. Redden, W. Chen, B. Sharma. In *Chickpea Breeding and Management*. 474–496. Wallingford: CAB International.

Tomooka, N., Kaga, A., and Vaughan, D. A. (2006). The Asian *Vigna* (*Vigna* subgenus Ceratotropis) biodiversity and evolution. In: Sharma AK, Sharma A (eds). *Plant genome: biodiversity and evolution. Part C Phanerogams (angiosperms-dicotyledons). 1*. Enfield: Science Publishers. 87–126.

Tomooka, N., Kashiwaba, K., Vaughan, D. A., Ishimoto, M., and Egawa, Y. (2000). The effectiveness of evaluating wild species: searching for sources of resistance to bruchid beetle in the genus *Vigna* subgenus Ceratotropis. *Euphytica* 115, 27–41.

Tomooka, N., Lairungruang, C., Nakeeraks, P., Egawa, Y., and Thavarasook, C. (1992). Development of bruchid resistantmungbean using wild mungbean germplasm in Thailand. *Plant Breed.* 109, 60–66. doi: 10.1111/j.1439-0523.1992.tb00151.x

Tomooka, N., Naito, K., Kaga, A., Sakai, H., Isemura, T., OgisoTanaka, E. et al. (2014). Evolution, domestication and neo-domestication of the genus Vigna. *Plant Genet. Resour. Charact. Util.* 12, 168–171.

Tomooka, N., Vaughan, D. A., Xu R. Q., Kashiwaba, K., and Kaga, A. (2001). Japanese native *Vigna* genetic resources. *Jpn. Agric. Res. Q.* 35, 1–9. doi: 10.6090/jarq.35.1

Tripathy, S. K., Nayak, P. K., Lenka, D., Swain, D., Baisakh, B., Mohanty, P., et al. (2016). Morphological diversity of local land races and wild forms of mungbean. *Legume Res.* 39, 485–493.

Tullu, A., Banniza, S., B., Tar'an, T., Warkentin and A., Vandenberg (2010). Sources of resistance to ascochyta blight in wild species of lentil (*Lens culinaris* Medik). *Genetic Resour. Crop Evol.* 57, 1053–1063. doi: 10.1007/s10722-010-9547-7

Tullu, A., Bett, K., Banniza, S., Vail, S., and Vandenberg, A. (2013). Widening the genetic base of cultivated lentil through hybridization of *Lens culinaris* 'Eston' and *L. ervoides* accession IG 72815. *Can. J. Plant Sci.* 93, 1037–1047. doi: 10.4141/cjps2013-072

Tullu, A., Buchwaldt, L., Lulsdorf, M., Banniza, S., Barlow, B., Slinkard, A. E. et al. (2006). Sources of resistance to anthracnose (*Colletotrichum truncatum*) in wild *Lens* species. *Genetic. Res. Crop. Evol.* 53, 111–119. doi: 10.1007/s10722-004-1586-5

Tullu, A., Diederichsen, A., Suvorova, G., and Vandenberg, A. (2011). Genetic and genomic resources of lentil: status, use and prospects. *Plant Genetic. Res. Crop. Evol.* 9, 19–29. doi: 10.1017/s1479262110000353

Upadhyaya, H.D., Reddy, K. N. Sharma, S., Varshney, R. K., Bhattacharjee, R., Singh, S. et al. (2011). Pigeonpea composite collection for enhanced utilization of germplasm in crop improvement programs. *Plant Genet. Res.* 9, 97–100. doi: 10.1017/s1479262110000419

Van Zonneveld, M., Rakha, M., yee Tan, S., Chou, Y. Y., Chang, C. H., Yen, J. Y., et al. (2020). Mapping patterns of abiotic and biotic stress resilience uncovers conservation gaps and breeding potential of *Vigna* wild relatives. *Scientific Rep.* 10, 1–11. doi: 10.9755/ejfa.v27i1.17852

Varshney, R. K., Mohan, S. M., Gaur, P. M., Chamarthi, S. K., Singh, V. K., Srinivasan, S. et al. (2014). Marker-assisted backcrossing to introgress resistance to Fusarium wilt race 1 and Ascochyta blight in C 214, an elite cultivar of chickpea. *Plant Gen.* 7, 1–11.

Verma, M. M., and Brar, J. S. (1996). Breeding approaches for increasing yield potential of mungbean. eds A. N. Asthana, D. H. Kim. In *Recent Advances in Mungbean Research*. Kanpur, India: Indian Society of Pulses Research and Development. 102–123.

Wang, B., and Chee, P. W. (2010). Application of advanced backcross quantitative trait locus (QTL) analysis in crop improvement. *J. Plant Breed. crop Sci.* 2, 221–232.

Wang, L. Q., Zhao, Y. F., Xue, Y. D., Zhang, Z. X., Zheng, Y. L., Chen, J. T. et al. (2007). Development and evaluation of two link-up single segment introgression lines (SSILs) in Zea mays. *Acta Agron. Sin.* 33, 663–668.

Wang, P., Ding, Y. Z., Lu, Q. X., Guo, W. Z., and Zhang, T. Z. (2008). Development of Gossypium barbadense chromosome segment substitution lines in the genetic standard line TM-1 of *Gossypium hirsutum*. *Chin. Sci. Bull.* 53, 1512–1517 doi: 10.1007/s11434-008-0220-x

Wang, W. B., He, Q. Y., Yang, H. Y., Xiang, S. H., Xing, G. N., Zhao, T. J. et al. (2012). Detection of wild segments associated with number of branches on main stem and leafstalk angle in soybean. *Sci. Agric. Sin.* 45, 4749–4758.

Wang, W. B., He, Q. Y., Yang, H. Y., Xiang, S. H., Zhao, T. J., Xing, G. N. et al. (2013). Identification of wild segments associated with stem termination, pod color, and seed coat color in soybean. *Acta Agron. Sin.* 39, 1155–1163 doi: 10.3724/sp.j.1006.2013.01155

Wanjari, K. B., Patil, A. N., Manapure, P., Manjayya, J. G., and Patel, M. (1999). Cytoplasmic male sterility in pigeonpea with cytoplasm from *Cajanus volubilis*. *Annals. Plant Physiol.* 13, 170–174.

Wehrhahn, C., and Allard, R. (1965). The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. *Genetics* 51, 109–119.

Wei, X. J., Liu, L. L., Xu, J. F., Jiang, L., Zhang, W. W., Wang, J. K., et al. (2010). Breeding strategies for optimumheading date using genotypic information in rice. *Mol. Breed.* 25, 287–298. doi: 10.1007/s11032-009-9332-5

Wilkinson, R.E. (1983). Incorporation of *Phaseolus coccineus* germplasm may facilitate production of high yielding *P. vulgaris* lines. *Annu. Rep. Bean Improv. Coop.* 26, 28–29.

Wilson, J. N., Chopra, R., Baring, M. R., Selvaraj, M. G., Simpson, C. E., Chagoya, J., et al. (2017). Advanced backcross quantitative trait loci (QTL) analysis of oil concentration and oil quality traits in peanut (*Arachis hypogaea* L.). *Tropical. Plant Biol.* 10, 1–17. doi: 10.1007/s12042-016-9180-5

Xiong, L. Z., and Yang, Y. N. (2003). Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acidinducible mitogen-activated protein kinase. *Plant Cell* 15, 745–759 doi: 10.1105/tpc.008714

Yang, S., Pang, W., Ash, G., Harper, J., Carling, J., Wenzel, P. et al. (2006). Low level of genetic diversity in cultivated pigeonpea compared to its wild relatives is revealed by diversity arrays technology. *Thoer. Appl. Genet.* 113, 585–595. doi: 10.1007/s00122-006-0317-z

Yin, X., Struik, P., Gu, J., and Wang, H. (2016). "Modelling QTL-trait-crop relationships: past experiences and future prospects,". (eds). X. Yin and C. Stuik.in *Crop Systems Biology*. (Berlin: Springer).

Yoshida, J., Tomooka, N., Yee Khaing, T., Shantha, P. G. S., Naito, H., Matsuda, Y. et al. (2020). Unique responses of three highly salt-tolerant wild Vigna species against salt stress. *Plant Prod. Sci.* 23, 1–15. doi: 10.1080/1343943x.2019.1698968

Zamir, D. (2001). Improving plant breeding with exotic geneticlibraries. *Nat. Rev. Genet*. 2, 983–989. doi: 10.1038/35103590

Zhang, L., Luo, J., Hao, M., Zhang, L., Yuan, Z., Yan, Z. et al. (2012). Genetic map of *Triticum turgidum* based on a hexaploid wheat population without genetic recombination for D genome. *BMC Genet.* 13:69. doi: 10.1186/1471-2156-13-69

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 June 2021
doi: 10.3389/fpls.2021.544854

Dynamic Programming for Resource Allocation in Multi-Allelic Trait Introgression

Ye Han[1]*, *John N. Cameron*[2], *Lizhi Wang*[1]*, *Hieu Pham*[1†] and *William D. Beavis*[2]

[1] Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, IA, United States, [2] Department of Agronomy, Iowa State University, Ames, IA, United States

Trait introgression is a complex process that plant breeders use to introduce desirable alleles from one variety or species to another. Two of the major types of decisions that must be made during this sophisticated and uncertain workflow are: parental selection and resource allocation. We formulated the trait introgression problem as an engineering process and proposed a Markov Decision Processes (MDP) model to optimize the resource allocation procedure. The efficiency of the MDP model was compared with static resource allocation strategies and their trade-offs among budget, deadline, and probability of success are demonstrated. Simulation results suggest that dynamic resource allocation strategies from the MDP model significantly improve the efficiency of the trait introgression by allocating the right amount of resources according to the genetic outcome of previous generations.

Keywords: dynamic programming, resource allocation, Markov decision processes, plant breeding, multi-allelic trait introgression

OPEN ACCESS

Edited by:
Jaime Prohens,
Universitat Politécnica de Valéncia, Spain

Reviewed by:
Dean Podlich,
Corteva Agriscience™, United States
Mahmoud Said,
Academy of Sciences of the Czech Republic, Czechia

***Correspondence:**
Ye Han
yehan1989.work@gmail.com
Lizhi Wang
lzwang@iastate.edu

†***Present address:***
Hieu Pham,
College of Business, The University of Alabama in Huntsville, Huntsville, AL, United States

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *23 March 2020*
Accepted: *21 May 2021*
Published: *18 June 2021*

Citation:
Han Y, Cameron JN, Wang L, Pham H and Beavis WD (2021) Dynamic Programming for Resource Allocation in Multi-Allelic Trait Introgression. Front. Plant Sci. 12:544854. doi: 10.3389/fpls.2021.544854

1. INTRODUCTION

Plant breeding has been defined as the art and science of producing desired characteristics through artificial selection (Poehlman, 2013). Practiced since the beginning of civilizations, plant breeders in the twentieth century made enormous changes to important agronomic traits, e.g., grain yield and pest resistance, of cereal crops (Duvick, 1994; Rincker et al., 2014). It is the plant breeder's job to identify new, genetically-superior crop varieties by "testing" the varieties in multiple environments, then selecting those that perform the best. The intention of this process is to breed specific varieties so that certain phenotypic traits (such as yield, height, weight, pest resistance, etc.) of two individuals can be carried over into its offspring. Historically, identifying the best varieties has been done by trial and error, with breeders testing their experimental varieties in a diverse set of locations and measuring their performance, then selecting the varieties that display the desired characteristics. However, analogously to two humans having children, not all traits can be seen in each child. Due to the inherent randomness in the plant breeding system, this process can take many years to produce the ideal variety and is inefficient, simply due to the number of potential combinations to create and test.

Methods for discovery of genetic variants (alleles) associated with specific phenotypic variants have been developed over the last 25 years and are now routinely applied using "omics" technologies in forward and reverse genetics approaches. These technological advancements have the potential to shorten the time-period required for the integration of desired traits. Because the genetic variants associated with phenotypic variability are distributed unevenly throughout a germplasm collections

and breeding populations, it is challenging to combine the most desirable alleles to create improved cultivars. Traditionally, the transfer of a single desirable allele from an inferior cultivar to a superior cultivar is routinely accomplished using marker assisted breeding strategies (Visscher et al., 1996; Frisch et al., 1999; Frisch and Melchinger, 2005; Peng et al., 2014). However, recent developments have demonstrated that the efficiency of these routine processes can be doubled by reframing the objective using principles from operations research (Cameron et al., 2017; Sun et al., 2017; Moeinizade et al., 2019; Xu et al., 2019).

The more complex challenge of aggregating sets consisting of multiple alleles into cultivars with predictable adaptive trait phenotypes will require a specialized breeding strategy to rapidly transfer multiple desirable genetic alleles from a donor individual to an elite recipient individual. In the vernacular of the plant breeder, this is known as multi-allelic trait introgression (MATI) process. The MATI process can be regarded as a decision making system, of which the components are in uncertain states due to the stochastic nature of genetic reconstruction during crop mating. In this process, the plant breeder has the obligation to obtain the available genotypic and phenotypic information, decide parents to breed, allocate resources and fulfill goals. Hospital et al. (2000) demonstrated via simulation that the marker assisted-selection, such as the Marker-based Truncation Selection (MTS) and the QTL Complementation Selection (QCS) could drastically improve the efficiency of parents selection. Recently, De Beukelaer et al. (2015) adapted optimization concepts with heuristics approaches to design a modern and advanced algorithm to solve the gene pyramiding problem. In order to accurately depict this decision making system and optimize the MATI process, a set of mathematical transformations and formulations have been proposed to frame the MATI process as an engineering system (Han et al., 2017). An algorithmic process with mathematical definitions was designed and parental selection was addressed as a key procedure, which can affect the result dramatically. A new metric called the Predicted Cross Value (PCV) with the assistance of genetic markers for parental selection was proposed. The PCV was defined as a quantification metric for any pair of selected breeding parents. Using the metric of PCV, significant improvements with respect to minimizing the cost and amount of time required for successful trait introgression were demonstrated as well as the great potential for further research on MATI process.

As pointed out in Han et al. (2017) and Cameron et al. (2017), in addition to parental selection, resource allocation also plays a crucial role in improving the efficiency of the MATI process. Hospital et al. (2000) discussed similar simulations with fixed population size in each generation but different selection intensity or the number of parents selected. Herein, we expand our discussion on the decision making problem of resource allocation for MATI and improve the breeding strategy by dynamically adjusting the population size for each generation. Resources allocation, as the major topic of this paper, means intelligently determining the population size during the introgression process to efficiently and effectively utilize the resources. Because of the dynamic and uncertain states of the system, we apply the Markov decision processes (MDP) model to frame MATI processes. The MDP model is a technique for solving stochastic sequential decision making problems (Puterman, 2014). The MDP model has been proved to make contributions to various practical decision making projects, such as optimal replacement policy for a motion picture exhibitor (Swami et al., 2001) or the vehicle mix decision in emergency medical service systems (Chong et al., 2015), which share many similarities with MATI processes.

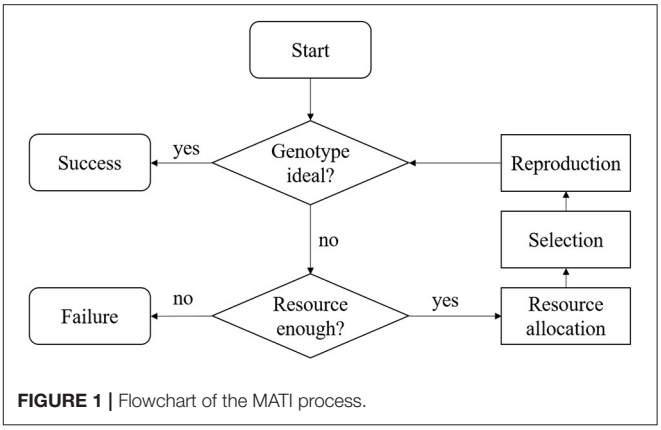

FIGURE 1 | Flowchart of the MATI process.

2. MATERIALS AND METHODS

In this section, we cast the MATI process with resource allocation as a Markov decision process model and present a dynamic programming method to solve it. The general idea of this MDP framework is to dynamically simulate and optimize the parent selection, meiosis, gamete production and crossing and other key steps during the trait introgression process. During the simulation, mathematical analysis is applied to adjust parameters to derive the optimal or near optimal decisions. This section covers the flowchart of this engineering process, the necessary mathematical formulations, the detailed discussion on the resource allocation challenge and the MDP model to solve the model.

2.1. The MATI Process

The work flow for the MATI process is presented in **Figure 1**. We summarized the MATI process into three steps with two checking points. The three steps are: resources allocation, selection and reproduction, and the two checking points check the available resources and the population genotype.

- The MATI process begins with the **"Start"** step, in which at least one elite recipient individual and one donor individual are available. In most annual crops, both elite and donor individuals are homozygous throughout their genomes. The majority of alleles in the donor are undesirable, but it does have desirable versions of alleles that the elite individual is lacking at several loci. The goal of this process is to achieve an ideal individual inheriting all the desirable alleles from both donor and elite individuals within the provided resources.

- In the **"Genotype ideal?"** check box, the genotypic information of current progeny is screened to check if the ideal individual was produced. If the ideal individual was sampled, the entire process is considered as a **"Success."**
- Otherwise, the process flows to the **"Resource enough?"** check box. This step involves the resources assessment and the process continues if the remaining resources are adequate. Usually, the resource consists of budget and time. A breeding process is associated with different terms of cost, such as genotyping assays, crossing, growing the crops, and labor. Some costs are fixed, while others are proportional to the number of crosses made or progeny produced. In practice, there may be a total budget constraint for the cost through the entire breeding project. In addition to the cost, the breeding project is often bounded by a deadline, which shall be regarded as a time resource limit.
- In the step **"Resource allocation,"** the decision maker needs to observe the current status of the breeding project and allocate the resources based on policies. For commercial breeding projects, there is revenue associated with the ideal individual when delivered to the market. Hence, for resource allocation, the decision maker needs to consider revenue with the cost.
- When the process reaches the **"Selection"** step, two breeding parents are selected based on a provided selection metric.
- In the **"Reproduction"** step, the selected breeding parents are mated to produce a new generation of progeny and the process flows back to the check box **"Genotype ideal?"** In this MATI process, we assume that the breeding parents would be retained for the next one generation.

2.2. Mathematical Formulations for the MATI Process

According to the flowchart, we design a mathematical algorithmic engineering process for simulating the MATI process, in which some steps can be optimized such as **"Resource allocation"** and **"Selection."** For the **"Selection"** step, random selection, genomic estimated breeding value (GEBV) (Meuwissen et al., 2001), optimal haploid value (OHV) (Daetwyler et al., 2015) and the newly designed predicted cross value (PCV) (Han et al., 2017) are possible metrics for determining the optimal breeding parents for the next generation. For the **"Resource allocation"** step, the remainder of the paper will discuss how to apply dynamic programming model to improve the efficiency. First, we define some major steps in the MATI process.

Definition 2.1. (Han et al., 2017) "We define the `Reproduce` function, $X = \texttt{Reproduce}(L^1, L^2, f, K)$, as follows. Its input parameters include two binary matrices $L^1, L^2 \in \mathbb{B}^{N\times 2}$, a vector $f \in [0, 0.5]^{N-1}$, and a positive integer number K. Its output is a three-dimensional matrix $X \in \mathbb{B}^{N\times 2\times K}$, representing a random population of K progeny."

The `Reproduce` function is defined the same way as the one in Han et al. (2017). We use a binary matrix with dimension of $N \times 2$ to represent the genotype of a diploid individual with N loci where "0" represents undesirable alleles and "1" represents desirable alleles at each of the loci. In the function L_1 and L_2 are the selected breeding parents. The output X of the function represents the genotype of all the progeny produced by the breeding parents, whose element $X_{i,1,k}$ with $i \in \{1, 2, \ldots, N\}, k \in \{1, 2, \ldots, K\}$ represents the allele on the ith row (locus) of the first set ('2' on the second dimension of X representing the second set) chromosome of the kth progeny in the population. The vector $f \in [0, 0.5]^{N-1}$ represents the recombination frequency, which reveals the inheritance characteristics of gene reconstruction. The parameter K in the function decides the number of progeny to produce. In the `Reproduce` function, we assume that the recombination is independent and only related to the recombination frequency.

Definition 2.2. We define the `Selection` function, $[k_1, k_2] = \texttt{Selection}(X)$, as follows. Its input parameter includes a three-dimensional binary matrix $X \in \mathbb{B}^{N\times 2\times K}$ representing a candidate population. Its output includes two integers, $k_1, k_2 \in \mathbb{Z}$ indicating the indexes of selected parents.

The `Reproduce` function and the `Selection` function utilize matrices to represent the information and population genotype. With the information of recombination frequencies, such functions could cast the introgression process into mathematical formulas to be programmed in computer simulation.

Definition 2.3. We define the `Reward` function, $\texttt{Reward}(K, X, t, T) = \texttt{Revenue}(X, t, T) - \texttt{Cost}(K)$, as follows. Its input parameters include a positive integer K representing the progeny number, a three-dimensional binary matrix $X \in \mathbb{B}^{N\times 2\times K}$ representing a candidate population, a non negative integer t representing the current generation number and a non negative integer T representing a deadline. Its output is a reward consisting of the revenue from population X at generation t given deadline T and the cost for producing K progeny.

Definition 2.4. We define the `Allocation` function,

$$K^t = \texttt{Allocation}(T, t, f, P^t, B^t, \texttt{Reward}),$$

as follows. Its input parameters include a positive integer T representing the deadline, a non negative integer t representing the current generation number, a vector $f \in [0, 0.5]^{N-1}$ representing the recombination frequency, a three-dimensional binary matrix $P^t \in \mathbb{B}^{N\times 2\times K^{t-1}}$ ($t \geq 1$ and $K^0 = 2$) representing the candidate breeding population for the current generation (produced by generation $t-1$), a positive number B^t representing the current available budget and the `Reward` function. Its output K^t is a non negative integer representing the number of progeny to produce for generation t. Note that if K^t equals 0 with $t \leq T$ and $B^t > 0$, the project fails.

The `Reward` function describes the estimated value of certain genotype under assumptions, in relation to current generation and the deadline. This function serves as a measure of quality. Together with the `Reward` function, the `allocation` function describes the resources allocation step mathematically. This function determines the population size to produce at a certain generation according to the genetic quality and the time and budget resources left.

With the definitions for three major steps in Flowchart (**Figure 1**), the definition for simulating the entire MATI process is proposed as follows.

Definition 2.5. We define the `MATI` function, $T_s = \texttt{MATI}(P^0, f, B, \texttt{Reward}, T)$, as follows. Its input parameters include a three-dimensional binary matrix $P^0 \in \mathbb{B}^{N\times 2\times 2}$ representing the initial breeding population, a vector $f \in [0, 0.5]^{N-1}$ representing the recombination frequency, a positive integer B representing the total budget, a `Reward` function and a positive integer T representing the deadline. Its output T_s, is the number of generations the process takes to finish the breeding process, which is determined through the following steps.

Step 0 (Initialization) Set $t = 0$ and go to Step 1.
Step 1 (Genotype check)
If $\max_k \left\{ \sum_{i=1}^{N} (P^t_{i,1,k} + P^t_{i,2,k}) \right\} = 2N$
RETURN: $T_s = t$.
Else Go to Step 2.
Step 2 (Resource check and resource allocation)
$K^t = \texttt{Allocation}(T, t, f, P^t, B^t, \texttt{Reward})$.
If $K^t = 0$ or $t > T$
RETURN: $T_s = \infty$.
Else Go to Step 3.
Step 3 (Selection) Obtain $[k_1^t, k_2^t] = \texttt{Selection}(P^t)$ and go to step 4.
Step 4 (Reproduction) Obtain $P^{t+1} = \texttt{Reproduction}(P^t_{:,:,k_1^t}, P^t_{:,:,k_2^t}, f, K^t)$, update $t \leftarrow t + 1$ and $B^{t+1} \leftarrow B^t - \texttt{Cost}(K^t)$, then go to Step 1.

The intuition of the MATI function is as follows:

- Step 0: Initialization;
- Step 1: Check if current population contains the ideal progeny; if it does, return the current generation; otherwise go to the next step;
- Step 2: Check current available time and budget resources and determine the number of progeny to produce; if no resources are left or current time is beyond the deadline, return failure; otherwise go to step 3;
- Step 3: Select the best pair of breeding parents from the current population;
- Step 4: Reproduction step with the determined breeding parents and the number of progeny to produce; Update available resources accordingly; Go back to step 1.

2.3. Resource Allocation in the MATI Process

In this section, we propose the problem definition for the resource allocation step in the MATI process, which is related to designing the `Allocation` function in the `MATI` function. The resource allocation problem for the MATI process is a dynamic decision making problem. The plant breeder needs to determine how many progeny to produce according to the current generation number, the deadline, the budgets remaining from the total budget, the cost and revenue function and the available progeny at the beginning of each generation. This decision is a key factor affecting the MATI process because it determines the number of offspring produced in each generation as well as the cost and revenue.

Herein, we give some intuitive explanations for the resource allocation problem statement. In each generation, producing more progeny can increase the cost but also the probability of obtaining a more promising genotype. The offspring's genotype and the amount of time together determine the revenue of a project. Generally speaking, the earlier a new genotypically designed product (i.e., offspring) can be delivered to the market, the more market share and revenue a company may attain. Hence, *designing the policy for resource allocation (i.e., how many progeny to produce at each generation) to maximize the expected net present value at the beginning of a breeding project* is regarded as the general problem statement of the resource allocation problem in MATI process.

We frame the resource allocation problem as a dynamic programming problem. Based on the previous discussion, the state describing the status of a breeding project shall consist of genotypic indicators and the budget information. Using metrics like MTS score, QCS score (Hospital et al., 2000) or PCV (Han et al., 2017), we can convert genotypic information into a number and use an interval to cover a group of progeny. Associated with the budget, the state is denoted as a combination of available budget and the metric interval for certain genotypes. By carefully designing the metric intervals, we can make the state space discrete and small enough to enumerate and cover all potential progeny genotypes.

The action that the breeder needs to take is to determine the number of progeny to produce at each state after the evaluation of the available population genotypes, which contains the potential breeding parents for the next generation. This action determines the cost. Meanwhile, different actions affect the probabilities of transitioning among states, which are stored in the transition probabilities matrix. In addition, reaching a specific state at a certain generation will generate revenue. Based on the breeder's estimation, the revenue may not only be decided by the state, but also determined by the current generation number and deadline. There will be a decision policy describing a series of actions to optimize the expected revenue of the breeding project.

In such manners, with a discount factor, the objective of a breeding project can be formulated as determining the optimal policy to maximize the expected net present value in terms of rewards subjected to the deadline and budget. In mathematical formulations, the objective of this resource allocation problem can be stated as:

$$\max_\pi \mathbb{E}_s^\pi \{\textstyle\sum_{t=0}^{T} \lambda^t r_t(a, s, T)\},$$

where, s represents the state; a represents the action; T represents the deadline; r represents the reward function; λ represents the discount factor and π represents the decision policy.

2.4. A Markov Decision Processes Model for Resource Allocation

The dynamic programming structure of the MATI process makes Markov decision processes (MDP) an appropriate approach for solving the stochastic decision making problem. In this section, we formulate an MDP model with finite horizon to identify the optimal resource allocation strategy, which is applied in the `Allocation` function of the described process.

An MDP model consists of five major components including decision epochs, states, actions, transition probabilities and rewards. The detailed notations for these components are as follows.

Decision epoches: We define the **decision epoch** as the beginning of each breeding generation, denoted as $\{1, 2, 3, \ldots, T\}$ and T is the deadline of a breeding project. Decisions like parental selection, resource allocation, etc., are made at each decision epoch. We assume the MATI process generally has a specified deadline, which implies that the MDP model has a finite horizon.

States: For any given sample of progeny P, we define a function $V(P)$ to measure the progress in the MATI process, which takes the values within the interval $[V(P^0), V(P^{\text{Ideal}})]$, with P^0 and P^{Ideal} denoting the original sample of progeny and a sample that includes an ideal individual (with all alleles being desirable). Various definitions of breeding values or parents selection metrics, such as MTS score, QCS score (Hospital et al., 2000) or PCV (Han et al., 2017), could be used for this function. Due to the enormous space of all possible samples of progenies, there is potentially a large number of possible values for V. For computational tractability, as illustrated in **Figure 2**, we group all possible V values into a small number of intervals $m_0, m_1, m_2, \ldots, m_{G-1}, m_G$, where G is a predetermined integer. In the figure, m_0 is a single value representing the initial population and m_G is another single value representing the final ideal progeny. The intermediate population is represented by each metric interval.

where (m_g, b) is a 2-tuple. In the 2-tuple, m_g represents the metric interval indicating the genotype status and b represents the remaining budget for the breeding project. In the definition, B represents the total budget at the beginning of the process. The design of metric intervals is associated with the preference of the decision maker and shall not be fixed. We will propose one possible approach in the case study section for designing the metric interval. With such state space definition, the initial state is (m_0, B)

Actions: The **action space** is denoted as $A = \bigcup_{s\in S} A_s = \{0, 1, 2, \ldots, a^{\max}\}$ representing the number of progeny to produce at each decision epoch. The maximum number of progeny that can be produced is set as $a^{\max}$ for each generation determined by the reproductive biology of the plant species. In the remainder of this paper, action a is used to substitute K in the algorithmic process for `Allocation` function.

Transition Probabilities: In the MDP model, we use $W^a_{i,j}$ to denote the transition probability from interval m_i in one generation to m_j in the next generation under action a. One fact of our MDP model is that once the intervals are determined, W^a only depends on the action a and is stationary at different epochs. According to the assumption that the breeding parents are retained to generate a new sample of progeny for the subsequent generation, the process either advances to the next interval or stays in the same one but never moves backwards, i.e., $W^a_{i,j} = 0$ if $j < i$. The matrix W^a could be estimated by simulations recording the information of action, the progeny produced at each generation and the hierarchical kinship information of mating. With the W^a matrix, we are ready to define the transition probabilities matrix, which consists of the probability of transitioning from one state s to another state s' under action a, i.e., $P_t(s'|s, a)$.

Definition 2.6. *Given action a, the* ***transition probabilities matrix*** *can be defined as a partitioned matrix M^a as follows:*

$$
M^a = \begin{array}{c} \\ S_B \\ S_{B-1} \\ S_{B-2} \\ \vdots \\ S_{a+1} \\ S_a \\ S_{a-1} \\ \vdots \\ S_1 \\ \textit{failure} \\ \textit{success} \end{array}
\begin{array}{c}
\begin{array}{ccccccccccc} S_B^\top & S_{B-1}^\top & \ldots & S_{B-a}^\top & S_{B-a-1}^\top & S_{B-a-2}^\top & \ldots & S_1^\top & \textit{failure} & \textit{success} \end{array} \\
\begin{pmatrix}
0 & 0 & \ldots & \bar{W}^a & 0 & 0 & \ldots & 0 & 0 & \hat{W}^a \\
0 & 0 & \ldots & 0 & \bar{W}^a & 0 & \ldots & 0 & 0 & \hat{W}^a \\
0 & 0 & \ldots & 0 & 0 & \bar{W}^a & \ldots & 0 & 0 & \hat{W}^a \\
\vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & \bar{W}^a & 0 & \hat{W}^a \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1-\hat{W}^a & \hat{W}^a \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\
\vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 1
\end{pmatrix}
\end{array}
$$

where $\bar{W}^a = W^a_{1:G-1,1:G-1}$, $\hat{W}^a = W^a_{1:G-1,G}$ *and* $S_b = [(m_1, b), (m_2, b), \ldots, (m_{G-2}, b), (m_{G-1}, b)]^\top$ *is a vector representing* $G-1$ *states. Here,* $P_t(s'|s, a) = M^a_{s,s'}, \forall t < T$.

Next we define the **state space** S as:

$$S = (m_g, b) \cup \{\text{failure}\} \cup \{\text{success}\},$$

and

$$g \in \{1, 2, \ldots, G-1\}, b \in \{1, 2, \ldots, B-1, B\},$$

In the definition of the transition probabilities matrix, the matrix $\bar{W}^a$ represents a sub-matrix containing all the transition probabilities from states group S_b to states group S_{b-a} under

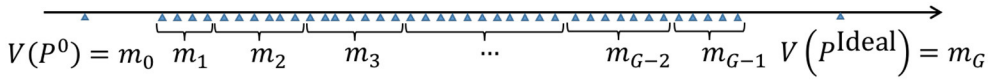

FIGURE 2 | Genotype indicator.

action a. The vector $\hat{W}^a$ represents a sub-vector containing all the transition probabilities from states group S_b to success under action a. Each single value of the transition probability between state s and s' under action a, which is $P_t(s'|s,a)$, is equal to each single element in the matrix $M^a_{s,s'}$.

Rewards: For an MDP model, the reward $r_t(s,a)$ received at epoch t is decided by the state $s \in S$ and action $a \in A_s$, which can be either positive or negative. In our MDP model for the MATI process, the **reward** is defined as $r_t(a,s,T) = -C(a) + R_t(s,T)$, where $C(a)$ is the cost function for producing a progeny and $R_t(s,T)$ is the revenue function at epoch t associated with state s and deadline T.

Our finite horizon MDP model can be efficiently solved by the backwards induction method, which is introduced as follows.

The Backward Induction Algorithm: (Puterman, 2014)
Step 1. Set $t = T$ and $u^*_T(s) = r_T(s)$ for all $s \in S$.
Step 2. Set $t \leftarrow t-1$ for t and compute $u^*_t(s_t)$ for each $s_t \in S$ by

$$u^*_t(s_t) = \max_{a \in A_{s_t}} \{r_t(a, s_t, T) + \lambda \sum_{s' \in S} P_t(s'|s_t, a) u^*_{t+1}(s')\}. \quad (1)$$

and

$$A^*_{s_t,t} = \arg\max_{a \in A_{s_t}} \{r_t(a, s_t, T) + \lambda \sum_{s' \in S} P_t(s'|s_t, a) u^*_{t+1}(s')\}. \quad (2)$$

Step 3. If $t = 1$, stop. Otherwise return to step 2.
We use $\pi = (d_1, d_2, \ldots, d_{T-1})$ to denote a policy, where $d_t : S \rightarrow A_s$ is the decision rule prescribing the procedure for action selection in each state at epoch t. $r_t(a_t, s_t, T)$ denotes the random reward received at epoch $t < T$ and $r_T(s_T)$ denotes the terminal reward. $v^\pi_T(s_1)$ denotes the expected total reward over the decision making horizon if policy π is selected and the system is in state s_1 at the first decision epoch. With the discount factor $\lambda \in [0,1)$, the expected total discounted reward will be

$$v^\pi_T(s_1) = E^\pi_{s_1}\{\sum_{t=1}^{T-1} \lambda^{t-1} r_t(a_t, s_t, T) + \lambda^{T-1} r_T(s_T)\}.$$

And the total expected reward obtained by using policy π at epochs $t, t+1, \ldots, T-1$ will be

$$u^\pi_t(s_t) = E^\pi_{s_t}\{\sum_{n=t}^{T-1} \lambda^{n-1} r_n(a_n, s_n, T) + \lambda^{T-1} r_T(s_T)\},$$

and $u^\pi_T(s_T) = r_T(s_T)$.

Suppose u^*_t, $t = 1, \ldots, T$ and $A^*_{s_t,t}$, $t = 1, \ldots, T-1$ satisfy equation (1) and (2). Let $d^*_t(s_t) \in A^*_{s_t,t}$ for all $s_t \in S$, $t = 1, \ldots, T-1$ and let $\pi^* = (d^*_1, \ldots, d^*_{T-1})$. Then, π^* is the optimal policy and satisfies

$$v^{\pi^*}_T(s) = \sup_\pi v^\pi_T(s), s \in S$$

and

$$u^{\pi^*}_t(s_t) = u^*_t(s_t), s_t \in S \text{ for } t = 1, \ldots, T.$$

3. RESULTS

This section introduces a simulation-based case study for the MDP model to solve the resource allocation problem in MATI process. In this case study, we propose a budget, time and probability of success criteria to assess a breeding strategy. We also discuss how the budget is allocated throughout the process and how to find the most cost-efficient total budget. For purposes of illustrations, we compare static budget allocation strategies and a dynamic budget allocation strategy. All the simulations and case studies are implemented in MATLAB/Octave.

3.1. Simulation Setup

We consider a hypothetical project for a case study with the same data structure as the simulation example 1 in Han et al. (2017). As stated in this paper, "We simulated a polygenic trait consisting of 100 markers that are responsible for genetic variability in the trait. The locations of the marker are distributed as uniform random variables among 10 simulated linkage groups. Each linkage group has from 8 to 12 markers. The recipient and donor are homozygous at all QTL. The recipient has desirable markers at 93 loci, while the donor has desirable markers at the remaining 7. For reference, the recipient has undesirable alleles at C1M4, C1M6, C2M9, C3M1, C5M4, C6M3, and C6M8, where C*i*M*j* denotes the *j*th marker in chromosome *i*. Recombination frequencies used in the simulation are given in the **Supplementary Materials**. The value shown for column C*i* and row M*j* is the recombination frequency between the corresponding marker pairs. The value for adjacent chromosomes is 0.5, in accordance with the principle of independent assortment of chromosomes." In addition to the genotypic information, **Table 1** contains all the parameters for establishing the MDP model. This example represents a realistic plant breeding problem, in which, for instance, 7 disease resistance alleles from a low yield donor need to be introduced to a high yield but disease susceptible recipient. The other 93 markers are used to ensure a high recovery rate of background genes to maintain the favorable agronomic traits of the recipient.

Herein, we introduce one possible way to construct the intervals for state space. In order to estimate the intervals, we run 100 preliminary simulations for each possible non-zero action $a \in \{100, 200, \ldots, 1000\}$.

Preliminary Simulation:
Step 1 Let P^0 denote the initial population and L^{E}, L^{D} denote the elite recipient and donor individuals, respectively, where $P^0_{:,:,1} = L^{\text{E}}$ and $P^0_{:,:,2} = L^{\text{D}}$.
Step 2 Set $G = 0$, which represents the current largest terminal generation number.

Step 3 Set $m_0 = \mathsf{PCV}(L^{\mathrm{E}}, L^{\mathrm{D}}, f)$, in which f represents the recombination frequency.

Step 4

For $a = 100:100:1000$

  For $n = 1:100$

    $g = 0$

    While $\max_k \left\{ \sum_{i=1}^{N} (P^g_{i,1,k} + P^g_{i,2,k}) \right\} < 2N$

      $[k_1^g, k_2^g] = \arg\max_{k_1,k_2} \{\mathsf{PCV}(P^g_{:,:,k_1}, P^g_{:,:,k_2}, f)\}$

      $p_g^{n,a} = \mathsf{PCV}(P^g_{:,:,k_1^g}, P^g_{:,:,k_2^g}, f)$

      $P^{g+1} = \mathtt{reproduce}(P^g_{:,:,k_1^g}, P^g_{:,:,k_2^g}, f, a)$

      $g = g + 1$

    $G = \max(G, g)$

The intuition of this preliminary simulation is as follows:

- Step 1: Initiate the starting population with the donor and elite recipient; Herein, $P^0_{:,:,1} = L^{\mathrm{E}}$ denotes that in the 3-dimensional matrix P^0, all elements in the first and second dimensions are equal to the elite recipient L^{E}, respectively; The ":" represents all elements in a dimension;
- Step 2: Initiate the current largest generation to achieve the ideal target, which is 0;
- Step 3: Initiate the starting metric point as the PCV value of the donor and elite recipient, with the given recombination frequency;
- Step 4: The major simulation step, simulates the effects of different actions (i.e., different population sizes per generation), on the largest number of generations needed to achieve the ideal target.

In this preliminary simulation, we update the G and record the $p_g^{n,a}$ for each simulation run. Then, we construct the state space based on the G and each $p_g^{n,a}$. Since F1 will be the only possible outcome after generation 1, we set $m_1 = p_1^{n,a}, \forall n, a$. Similarly, for the last generation G, m_G will be the PCV value of the ideal individual, which means $m_G = p_G^{n,a} = \mathsf{PCV}(L^{\mathrm{Ideal}}, L^{\mathrm{Ideal}}, f)$. After the preliminary simulations, we define the interval m_g as $m_g = [\min_{n,a}(p_g^{n,a}), \min_{n,a}(p_{g+1}^{n,a})]$ where $2 \leq g \leq$ G-1, $n \in \{1, \ldots, 100\}$, $a \in \{100, 200, \ldots, 1000\}$. The state space construction will be trivial based on the definition.

Next, we need to estimate the matrix W^a for the transition probabilities between each state. First, for any given p, we can trace back the unique interval that p belongs to, based on the preliminary simulation. We use an indicator function $m_k = \mathtt{Interval}(p)$ to represent this procedure. Meanwhile, we use another matrix $N^a \in \mathbb{I}^{G \times G}$ to record the number of simulation runs, which lead to the transition between two intervals under action a.

For $a = 100:100:1000$

  For $n = 1:100$

    $g = 1$

    While $p_g^{n,a} < m_G$

      $m_{k_1} = \mathtt{Interval}(p_g^{n,a})$

      $m_{k_2} = \mathtt{Interval}(p_{g+1}^{n,a})$

      $N^a_{k_1,k_2} = N^a_{k_1,k_2} + 1$

      $g = g + 1$

$W^a_{i,j} = \frac{N^a_{i,j}}{\sum_j N^a_{i,j}}$

The procedures above introduce how to derive each element in the matrix N^a and how to calculate the transition matrix W^a based on N^a.

3.2. Simulation Results

We demonstrate the effectiveness of the dynamic programming method for resource allocation by summarizing the results from the simulation experiments.

3.2.1. Tradeoff Among Cost, Time, and Probability of Success

We first ran the simulation with varying levels of total budget for a static budget per generation and presented the results in the CTP framework associated with each total budget value in **Figure 3**. The horizontal axis shows the total budget (cost) for the MATI process, the vertical axis represents the stacked histogram of probabilities, and different layers with distinct colors indicate the number of generations (time) it takes to successfully complete the process. For example, when the total budget is $11,000, the project can successfully finish in 6, 7 or 8 generations with probability about 2, 20, or 44%, respectively. This project also has about 34% probability to fail. The figure also demonstrates the diminishing effect of increased budget to the performance of the process. From a commercial breeding perspective, this would enable an organization to estimate the cost and time-length needed for successful creating the desired progeny.

3.2.2. Comparison With Static Resource Allocation Strategies

We demonstrate the improvement of optimal dynamic resource allocation over the static resource allocation using two random simulations, which are summarized in the following tables of figures, **Tables 2**, **3**. **Table 2** shows the result simulated using the static strategy with $K^t = 400$ for each generation t, whereas **Table 3** shows the result from the MDP model. In both tables,

TABLE 1 | Parameters.

Parameter	Value	Interpretation
$a^{\max}$	1,000	maximum progeny number for one generation
A	$\{0, 100, 200, \ldots, 900, 1,000\}$	action space
$C(a)$	$10a$	cost function
$R_t(s, T)$	$2,000,000 - 100,000t$	nominal market value (revenue) function
$r_t(s, T)$	$R_t(s, T)\mathcal{I}(s = success)\mathcal{I}(t \leq T)$	reward function
T	8	deadline (in number of generations)
B	$11,000, $12,000, …, or $80,000	budget scenarios

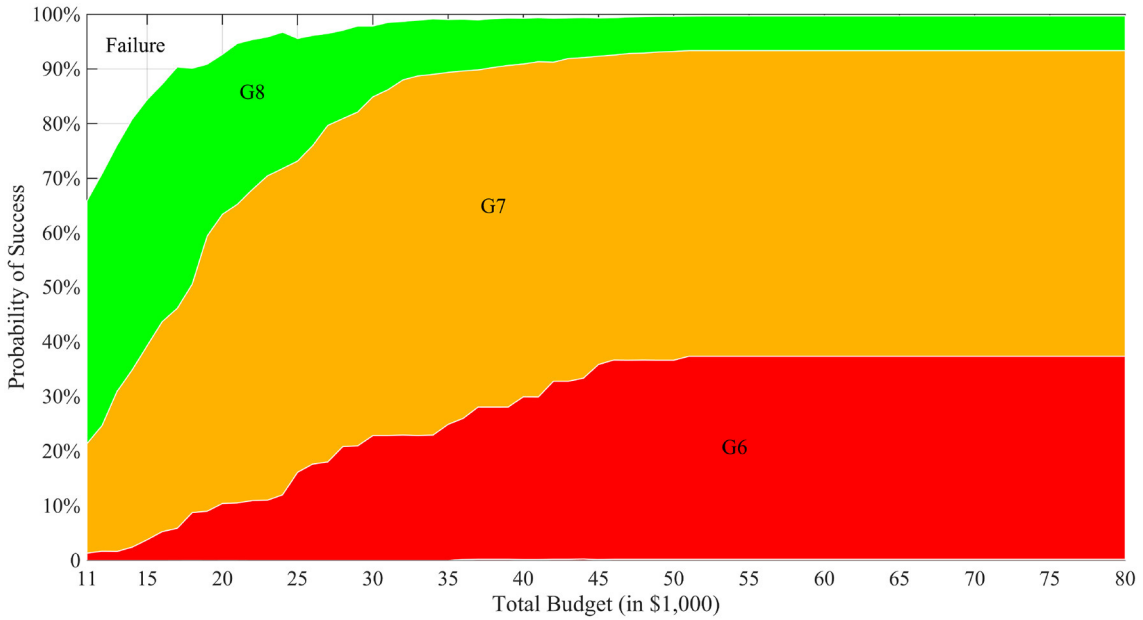

FIGURE 3 | CTP graph with $T = 8$. In the figure, the horizontal axis is different total budget scenarios of the breeding project and the vertical axis represents a stacked histogram of the probabilities of reaching success at different generations. In the figure, "G*X*" label means that the breeding process successfully finishes in *X* generations and "Failure" means no ideal individual is produced when the budget or the time is depleted.

the first column is the generation number. In the second column, at each generation, all the progeny produced in the simulation are put abreast to each other to form a wide rectangle and the width of the rectangle reflects the sample size. Here we use gray pixels to represent the desirable alleles whereas black pixels to represent the undesirable alleles. Those individuals highlighted by white are the selected breeding parents and several ideal individuals are produced at the last generations. The third column of each table is the base 10 logarithm of PCV values of the selected breeding parents. The fundamental difference between these two resource allocation strategies is that the MDP model allows the decision maker to dynamically allocate the resources based on the outcomes from the previous generation. As a result, for the same amount of the total budget, the dynamic approach was able to produce the ideal progeny in the seventh generation, whereas the static strategy required an extra generation.

Figure 4 compares static and dynamic resource allocation strategies with respect to the CTP criteria for a fixed total budget of $32,000. We considered seven different static strategies, in which a fixed number of progeny (ranging from 100 to 700 with an increment of 100) are produced in each generation. A total of 500 simulation repetitions were conducted for the seven static strategies and the dynamic strategy, and the histograms of the terminal generations are compared in the figure. When a small number of progeny are produced, the static strategy takes more time resources to complete the project; when a large number of progeny are produced, on the other hand, the static strategy risks depleting the total budget before successful completion. For instance, the 600-strategy produces 600 progeny in each of generations 1–5 and only 200 progeny in generation 6 with a fixed total budget of $32,000. For such strategy, the success rate of achieving the ideal target in generation 6 is <5%. In contrast, the dynamic strategy has the flexibility to adjust the amount of resource allocation based on the outcome of the previous generation and is more likely to achieve successful completion within a shorter amount of time.

TABLE 2 | Generations 2–8 of one random simulation run with fixed budget allocation.

G	Breeding population	$\log_{10}$ PCV
2		–11.69
3		–6.89
4		–4.98
5		–2.76
6		–1.38
7		–0.25
8		0

3.2.3. Optimal Total Budget and Budget Allocation

Figure 5 enables plant breeders to determine the optimal total budget for the MATI project based on cost-benefit analysis. The blue curve represents a regression line on the estimated total

TABLE 3 | Generation 2–7 of one random simulation run with MDP based budget allocation.

G	Breeding population	$\log_{10}$ PCV
2		–13.45
3		–9.27
4		–4.94
5		–1.77
6		–0.55
7		0

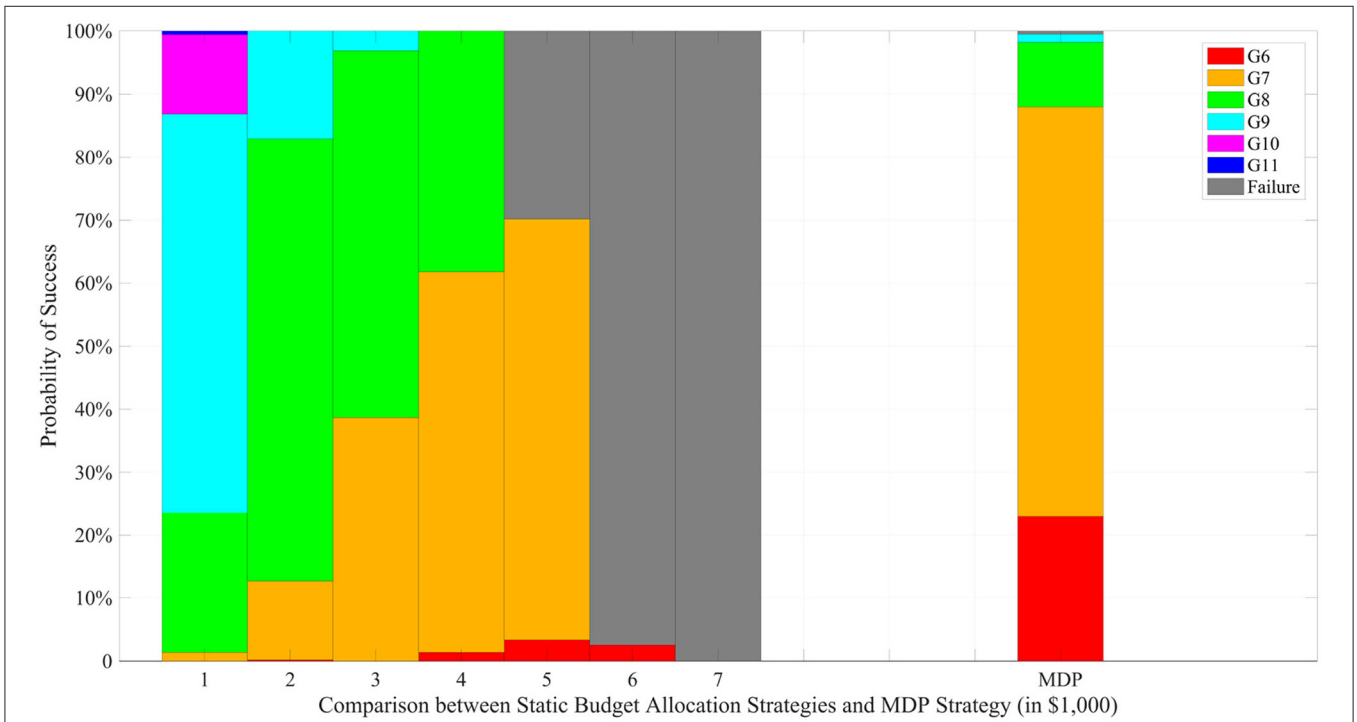

FIGURE 4 | Comparison under a fixed total budget of $32,000. The left 7 stacked bars represent the static budget allocation strategies with different progeny number per generation while the last bar represents the MDP based strategy.

revenue, referring to the blue axis on the left. The red curve represents a regression line on the estimated marginal return, which is the derivative of the total revenue, referring to the red axis on the right. This red curve illustrates the relation between the investment on the total budget and the relative gain on the total revenue. The optimal total budget, approximately $32,000, is achieved where the marginal revenue intersects with $1000, which is the unit increment of the total budget. Before the optimal budget, every extra unit total budget investment brings more return on the total revenue. However, after this point, the increment on the total revenue is comparatively less with the unit total budget increment.

Figure 6 breaks down the cost allocation to different generations for varying levels of total budget. When total budget is less than the optimal level, the model tends to allocate unproportionately higher percentages of budget to early generations, in order to produce enough progeny and preserve genetic diversity for future genetic gains. When total budget exceeds the optimal level, resource allocation to different generations becomes stable. Meanwhile, the model tends to allocate relative more resources on generation 2 and 3 to push the process to succeed in generation 5. However, if it is not finished in 5 generations, the model allocates a second push in generation 6 to pursue a quick success. In general, the model

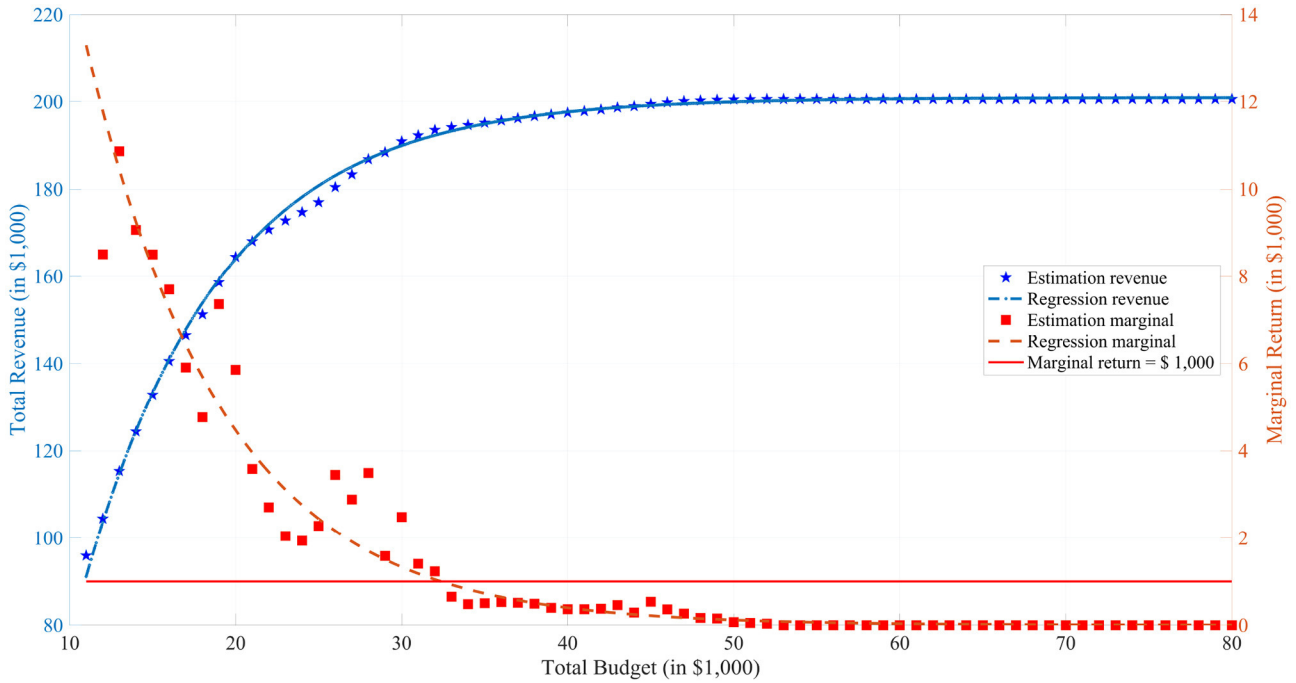

FIGURE 5 | Profits and Budgets. In the figure, the blue pentagrams represent the estimation results from simulations and the blue curve represents a nonlinear regression with model $y = a_1 + a_2 \times \exp(a_3 x)$ for the estimation. The red squares represent the difference between the adjacent estimations and the red curve represents the derivative of the expected total revenue curve. The red horizontal line is the marginal return is equal to one unit increment of total budget, which is $1,000.

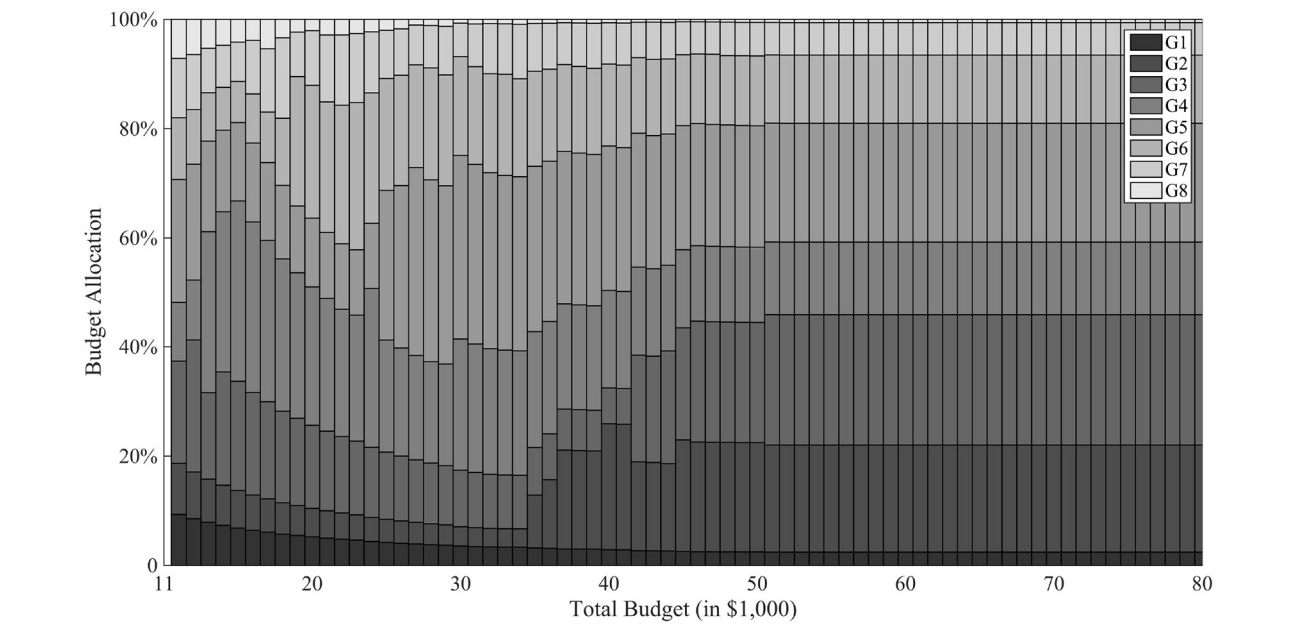

FIGURE 6 | Budget allocation with $T = 8$. In the figure, the horizontal axis is different total budget scenarios of the breeding project and the vertical axis represents the proportion of budget allocated in different generations. Different gray scale are used for different generations.

focuses on dynamic balance of both budget and time resources. After G1, the model tends to allocate higher budget in G2 and G3 to create variability; G4 requires less budget but a little time for favorable recombinations to happen; G5 gives a final push for the "lucky" progeny to succeed in G6 and subsequent generations.

4. CONCLUSIONS

In this paper, we addressed the issue of optimal resource allocation in a MATI process using a Markov decision process model, and made connections to the importance of optimizing this process for a commercial organization. Simulation experiments suggested that the proposed dynamic resource allocation method greatly improves the efficiency of the MATI process. Due to the assumptions made in the problem definition and model construction, the proposed model is by no means the best possible solution to the proposed problem, but this can be seen as a potential efficiency improvement on the traditional MATI process. Future research effort is needed to explore other definitions of the state space and action space to further improve the effectiveness of the model.

Estimating the cost and revenue function is a possible research topic for further discussion, as well. Plant breeding organizations have their own forecasting models about the market value of a certain genotype as well as its revenue associated with time when it is delivered to the market. Thus, the research on the discussion about cost and revenue functions may reveal more economic discoveries about the trait introgression problem and inspire further analysis.

Another fruitful research topic will be applying more advanced artificial intelligence techniques into such research problems. In our model, simplifying assumptions were made to reduce the problem dimension to a relatively small scale with only a few actions and states and finite time horizon. However, as studied in Hospital et al. (2000), different selection intensity or the number of parents selected for each generation could make this resources allocation challenge more comprehensive and complex. At the same time, relaxing the problem to allow multiple donors is challenging. Also, the assumption on independent crossovers could be changed for a more comprehensive analysis. At the same time, it would be a meaningful followup study to relate and compare with the gene-stacking algorithm in De Beukelaer et al. (2015), in which the population size was determined by a statistical formula. In order to solve such problems under fewer assumptions and higher dimensions, more powerful modeling and solution techniques, such as reinforcement learning will be necessary to deal with the uncertainty and complexity of the MATI process to discover more efficient strategies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YH and LW conceived of the presented idea and developed the theory and performed the computations. JC and WB provided guidance on modifications according to domain knowledge. JC, HP, and WB verified the analytical methods and results. All authors discussed the results and contributed to the final manuscript.

FUNDING

LW was partially supported by the United States Department of Agriculture under a NIFA program (grant number 2017-67007-26175/accession number 1011702), by the National Science Foundation under LEAP HI and GOALI programs (grant number 1830478) and an EAGER program (grant number 1842097), and by the Plant Sciences Institute at Iowa State University.

ACKNOWLEDGMENTS

This research was conducted as a part of YH's Ph.D. dissertation (Han, 2017). We especially thank the committee members, Prof. Sarah Ryan, Prof. Dan Nordman, and Prof. Mingyi Hong for their time, effort, support and guidance through the journey. Also, we would like to thank the department of Industrial and Manufacturing Systems Engineering and the department of Agronomy of Iowa State University for all the supports.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.544854/full#supplementary-material

REFERENCES

Cameron, J. N., Han, Y., Wang, L., and Beavis, W. D. (2017). Systematic design for trait introgression projects. *Theor. Appl. Genet.* 130, 1993–2004. doi: 10.1007/s00122-017-2938-9

Chong, K. C., Henderson, S. G., and Lewis, M. E. (2015). The vehicle mix decision in emergency medical service systems. *Manufact. Serv. Operat. Manag.* 18, 347–360. doi: 10.1287/msom.2015.0555

Daetwyler, H. D., Hayden, M. J., Spangenberg, G. C., and Hayes, B. J. (2015). Selection on optimal haploid value increases genetic gain and preserves more genetic diversity relative to genomic selection. *Genetics* 200, 1341–1348. doi: 10.1534/genetics.115.178038

De Beukelaer, H., De Meyer, G., and Fack, V. (2015). Heuristic exploitation of genetic structure in marker-assisted gene pyramiding problems. *BMC Genet.* 16:1. doi: 10.1186/s12863-014-0154-z

Duvick, D. (1994). *Congresso Nacional De Milho E Sorgo*, Maize Breeding: Past, Present and Future, 170–179.

Frisch, M., Bohn, M., and Melchinger, A. E. (1999). Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Sci.* 39, 1295–1301. doi: 10.2135/cropsci1999.3951295x

Frisch, M., and Melchinger, A. E. (2005). Selection theory for marker-assisted backcrossing. *Genetics* 170, 909–917. doi: 10.1534/genetics.104.035451

Han, Y. (2017). "Operations research applications in multi-allelic trait introgression in plant breeding," in *Graduate Theses and Dissertations*. Ames: Iowa State University, 16280.

Han, Y., Cameron, J. N., Wang, L., and Beavis, W. D. (2017). The predicted cross value for genetic introgression of multiple alleles. *Genetics* 205, 1409–1423. doi: 10.1534/genetics.116.197095

Hospital, F., Goldringer, I., and Openshaw, S. (2000). Efficient marker-based recurrent selection for multiple quantitative trait loci. *Genet. Res.* 75, 357–368. doi: 10.1017/S0016672300004511

Meuwissen, T., Hayes, B., and Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Moeinizade, S., Hu, G., Wang, L., and Schnable, P. S. (2019). Optimizing selection and mating in genomic selection with a look-ahead approach: An operations research framework. *G3* 9, 2123–2133. doi: 10.1534/g3.118.200842

Peng, T., Sun, X., and Mumm, R. H. (2014). Optimized breeding strategies for multiple trait integration: I. Minimizing linkage drag in single event introgression. *Mol. Breed.* 33, 89–104. doi: 10.1007/s11032-013-9936-7

Poehlman, J. M. (2013). *Breeding Field Crops*. Oxford: Springer Science & Business Media.

Puterman, M. L. (2014). *Markov decision processes: discrete stochastic dynamic programming*. Hoboken: John Wiley & Sons.

Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Cianzio, S. R., et al. (2014). Genetic improvement of us soybean in maturity groups ii, iii, and iv. *Crop Sci.* 54, 1419–1432. doi: 10.2135/cropsci2013.10.0665

Sun, Y., Shang, J., Liu, J.-X., Li, S., and Zheng, C.-H. (2017). epiaco-a method for identifying epistasis based on ant colony optimization algorithm. *Biodata Min.* 10, 23. doi: 10.1186/s13040-017-0143-7

Swami, S., Puterman, M. L., and Weinberg, C. B. (2001). Play it again, Sam? optimal replacement policies for a motion picture exhibitor. *Manufact. Serv. Operat. Manag.* 3, 369–386. doi: 10.1287/msom.3.4.369.9967

Visscher, P. M., Haley, C. S., and Thompson, R. (1996). Marker-assisted introgression in backcross breeding programs. *Genetics* 144, 1923–1932. doi: 10.1093/genetics/144.4.1923

Xu, Y., Liu, X., Fu, J., Wang, H., Wang, J., Huang, C., et al. (2019). Enhancing genetic gain through genomic selection: from livestock to plants. *Plant Commun.* 1:100005. doi: 10.1016/j.xplc.2019.100005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership