



A Belated Green Revolution for *Cannabis*: Virtual Genetic Resources to Fast-Track Cultivar Development

Matthew T. Welling¹, Tim Shapter^{1,2}, Terry J. Rose¹, Lei Liu¹, Rhia Stanger¹ and Graham J. King^{1*}

¹ Southern Cross Plant Science, Southern Cross University, Lismore, NSW, Australia, ² Ecofibre Industries Operations Pty Ltd, Maleny, QLD, Australia

OPEN ACCESS

Edited by:

Jaime Prohens,
Polytechnic University of Valencia,
Spain

Reviewed by:

Ryan C. Lynch,
University of Colorado Boulder, USA
Giuseppe Mandolino,
Council for Agricultural Research and
Economics, Italy

*Correspondence:

Graham J. King
graham.king@scu.edu.au

Specialty section:

This article was submitted to
Crop Science and Horticulture,
a section of the journal
Frontiers in Plant Science

Received: 29 April 2016

Accepted: 13 July 2016

Published: 29 July 2016

Citation:

Welling MT, Shapter T, Rose TJ, Liu L,
Stanger R and King GJ (2016) A
Belated Green Revolution for
Cannabis: Virtual Genetic Resources
to Fast-Track Cultivar Development.
Front. Plant Sci. 7:1113.
doi: 10.3389/fpls.2016.01113

Cannabis is a predominantly dioecious phenotypically diverse domesticated genus with few if any extant natural populations. International narcotics conventions and associated legislation have constrained the establishment, characterization, and use of *Cannabis* genetic resource collections. This has resulted in the underutilization of genepool variability in cultivar development and has limited the inclusion of secondary genepools associated with genetic improvement strategies of the Green Revolution. The structured screening of *ex situ* germplasm and the exploitation of locally-adapted intraspecific traits is expected to facilitate the genetic improvement of *Cannabis*. However, limited attempts have been made to establish the full extent of genetic resources available for pre-breeding. We present a thorough critical review of *Cannabis ex situ* genetic resources, and discuss recommendations for conservation, pre-breeding characterization, and genetic analysis that will underpin future cultivar development. We consider East Asian germplasm to be a priority for conservation based on the prolonged historical cultivation of *Cannabis* in this region over a range of latitudes, along with the apparent high levels of genetic diversity and relatively low representation in published genetic resource collections. Seed cryopreservation could improve conservation by reducing hybridization and genetic drift that may occur during *Cannabis* germplasm regeneration. Given the unique legal status of *Cannabis*, we propose the establishment of a global virtual core collection based on the collation of consistent and comprehensive provenance meta-data and the adoption of high-throughput DNA sequencing technologies. This would enable representative core collections to be used for systematic phenotyping, and so underpin breeding strategies for the genetic improvement of *Cannabis*.

Keywords: genetic conservation, industrial hemp, marijuana, phenotypic variation, plant breeding, cannabinoids

INTRODUCTION

Cannabis, an erect annual herb and member of the *Cannabaceae* family, is monotypic and characterized by a single species *Cannabis sativa* L. (Small and Cronquist, 1976). Plants are diploid ($2n = 20$) with an estimated haploid genome of ~830 Mb (Van Bakel et al., 2011). The extant genepool is thought to be comprised primarily of domesticated or feral populations, cultivars,

and selections (Small and Cronquist, 1976), with a subset having been subject to steep selective gradients toward phenotypes for specific end-uses (Mandolino and Carboni, 2004; Potter, 2014; Small, 2015a). *Cannabis* has been cultivated in Eurasia over several thousand years (Li, 1974; Bradshaw et al., 1981; Murphy et al., 2011; Herbig and Sirocko, 2013) and has since radiated from this region and been subject to prolonged artificial selective pressures in Africa (Duvall, 2016) and North and South America (Small and Marcus, 2002), and is now cultivated globally (Salentijn et al., 2014). Plants are dioecious and obligate outbred, although some fiber forms are monoecious (Faeti et al., 1996). This has contributed to a high level of hybridization between pre-, post-, and de-domesticated populations (Gilmore et al., 2007), and therefore few if any intact wild populations are thought to exist (Small and Cronquist, 1976).

The Green Revolution saw the productivity of major crops such as wheat (*Triticum* spp.), rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), and millet (*Pennisetum glaucum*) increase by 0.5–1% annually between 1960 and 2000 (Evenson and Gollin, 2003). A large proportion of these gains has been attributed to germplasm improvements initiated by the Consultative Group on International Agricultural Research (CGIAR; Hajjar and Hodgkin, 2007), which facilitated genetic resource exchange, adaptive transfers, and the free flow of materials and knowledge between research groups (Pingali, 2012). However, the genetic improvement of crop-types was not evenly distributed (Evenson and Gollin, 2003). During the latter part of the twentieth century legitimate crop-types of *Cannabis*, commonly referred to as industrial hemp, not only failed to benefit from advances in breeding technologies and genetic resource utilization, but also suffered significant losses in *ex situ* conservation (Clarke, 1998; Small and Marcus, 2002; Ranalli, 2004), with likely significant genetic erosion, as has been documented with the loss of rice wild relatives in China (Gao, 2003).

The major contributing factor to the erosion of *Cannabis* genetic resources and limited genetic improvement of industrial hemp has been the status of *Cannabis* as the source of one of the most widely used illicit drugs (Hall and Degenhardt, 2007). Marijuana refers to an informal subtaxa of *Cannabis* which has been selected specifically for its psychoactivity and propensity to accumulate high levels of delta-9-tetrahydrocannabinol (THC; Hewavitharana et al., 2005). The prohibition on marijuana, which is morphologically indistinguishable from other forms of *Cannabis* (Small and Cronquist, 1976; Small, 2015a), has contributed to the decline of industrial hemp (Small and Marcus, 2003; Piluzza et al., 2013; Small, 2015a). In recent years there has been growing public acceptance that medical cannabis has therapeutic applications, which has been particularly prominent in a number of states of the USA, along with Australia, Canada, Israel, and Uruguay. This growing public tolerance toward *Cannabis* in North and South America (Schuermeyer et al., 2014) and in other parts of the world has led to a number of political and legal reforms in these legislatures (Pardo, 2014), and suggests not only a renewed consumer demand for *Cannabis*-related products, but also an increased requirement for breeders to develop varieties for specific end-use applications.

This deterioration in conservation of industrial hemp genetic resources is not consistent with the intrinsic functionality or agricultural value of the species. The long pericyclic and phloem fibers in the bast fraction of the culm (stem) have high tensile strength, and were used for millennia as the predominant source of fiber for the production of rope (Murphy et al., 2011). They have also been used in the textile and paper industries (Amaducci et al., 2014), as well as in fiber mat thermoplastics for automobile manufacture (Pervaiz and Sain, 2003). The woody core or hurd (shives) within the xylem vascular tissue, comprising up to 70% of the stem biomass (De Meijer, 1994), has also found minor low value use in the production of lime-hemp concrete, a light composite building material (De Bruijn et al., 2009), as well as animal bedding and garden mulch (Pervaiz and Sain, 2003). Achene (seeds) are also suitable for human and animal consumption and contain all essential amino acids (Callaway, 2004), with the exception of lysine, (House et al., 2010) in a readily bioavailable form (Malomo and Aluko, 2015), as well as a balanced omega-6/omega-3 essential polyunsaturated fatty acid (PUFA) ratio (Callaway, 2004; Carvalho et al., 2006; Da Porto et al., 2012).

In addition to the traditional uses in paper and textile industries, a number of novel end-use applications for *Cannabis* are currently being explored. These include cellulosic bioethanol which can be extracted from plant biomass and utilized as a potential transportation biofuel (Finnan and Styles, 2013; Kuglarz et al., 2016). Higher value products being assessed include cellulose nanocrystals, which have been generated from low quality fibers (Li et al., 2010) for the development of high-performance nanocomposites (George and Sabapathi, 2015), while graphitic carbon nanosheets developed from hemp bast fiber precursors have been shown to have ionic-liquid-based supercapacitor properties (Wang et al., 2013).

The phenolic compounds *N*-trans-caffeoyltyramine and cannabisin B (Chen et al., 2012), as well as prenylflavonoids cannflavins A and B (Werz et al., 2014) occurring within seed husks and sprouts, respectively, may have potential nutraceutical applications. In addition, *Cannabis* plants produce a group of pharmacologically active terpenophenolic phytocannabinoids (cannabinoids; ElSohly and Slade, 2005; Appendino et al., 2011). These compounds are currently fueling an emerging cannabinoid-based pharmaceutical industry (Potter, 2014), with botanical extracts being safely and efficiently administered via oromucosal sprays (Izzo et al., 2009) and by electrically-driven vaporizers (Lanz et al., 2016). Other terpenoid-related compounds that are present at high concentrations in *Cannabis* may also contribute to pharmacological entourage effects (Russo, 2011), or have their own specific medicinal applications, such as α -humulene which has antifungal properties potentially beneficial in the treatment of cryptococcosis (Wanas et al., 2016).

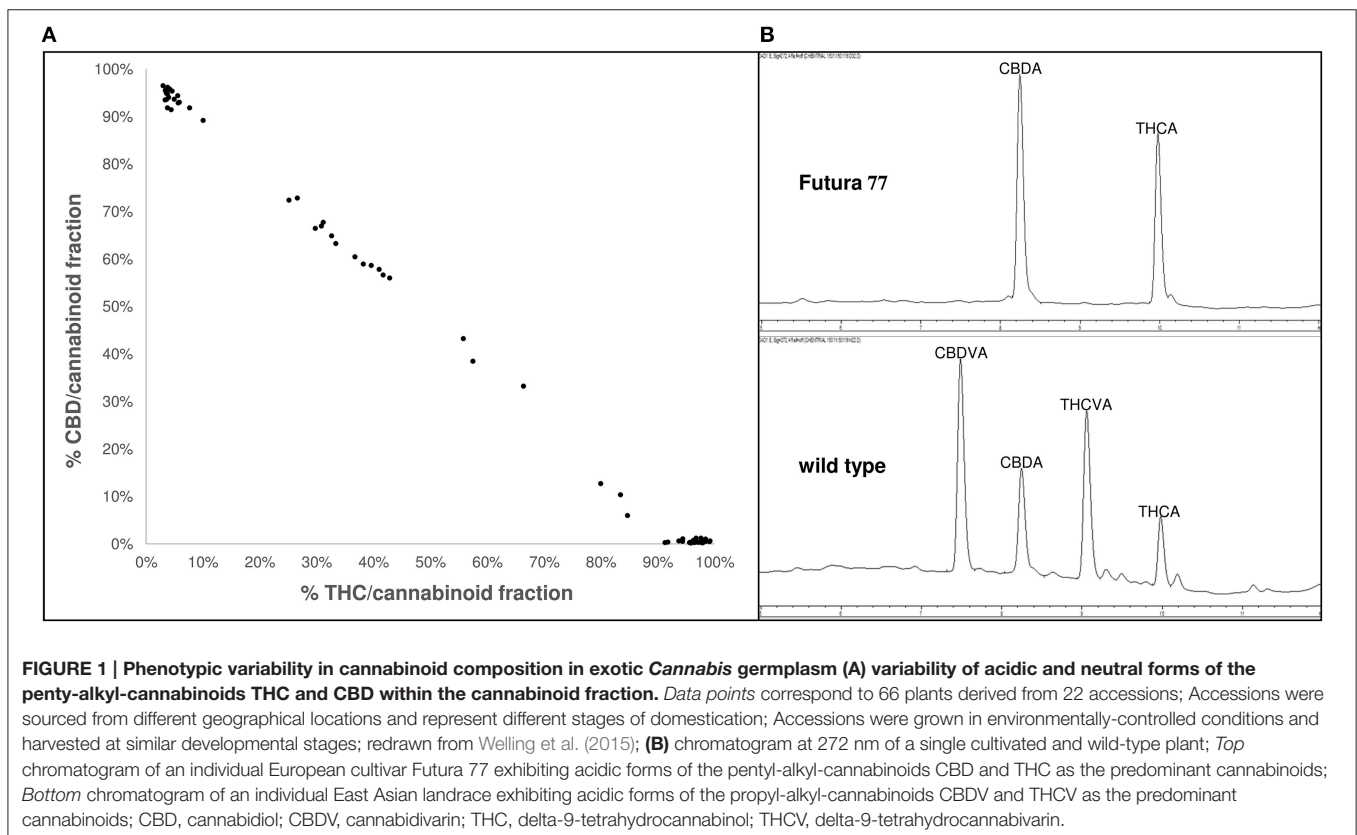
The traditional and still predominantly successful approach to crop genetic improvement has been based on exploiting naturally occurring genetic diversity. Exploitation of genetic diversity can be achieved through the hybridization of elite breeding material with exotic germplasm, defined here as a genetic resource which has not been artificially subject to a high level of selection

for a given targeted trait or growing environment. In the case of heterozygote breeding lines, extensive trait diversity can be uncovered by crossing siblings of the same accession and observing transgressive trait segregation within the progeny. A recurrent introgressive population enrichment approach can then be used to generate elite populations and cultivars with improved performance or novel properties. However, this requires access to structured plant genetic resources (Babic et al., 2015; Scossa et al., 2015). *Ex situ* genetic resources provide a valuable pool of genetic diversity (Hajjar and Hodgkin, 2007) and have been proven to be an essential source of allelic variation for trait improvement (Hajjar and Hodgkin, 2007; Castañeda-Álvarez et al., 2015; Scossa et al., 2015). Wild or underdomesticated germplasm, in the form of either crop wild relatives or locally-adapted landraces, provide a repository to reintroduce allelic diversity into domesticated populations (Vincent et al., 2013).

Germplasm surveys of *Cannabis* have revealed substantial phenotypic diversity in oil content (Kriese et al., 2004; Matthäus et al., 2006; Grigor'ev et al., 2010), cannabinoid (Welling et al., 2015), and xylem composition (De Meijer, 1994). Exploitation of this phenotypic diversity could be used for the development of novel cultivars tailored for specific end-use applications. For instance, phenotypic diversity in the composition of THC and cannabidiol (CBD; **Figure 1A**) as well as other rare-alkyl-cannabinoid homologs (**Figure 1B**; Welling et al., 2015) has the potential to be used to develop breeding lines for the

production of specialized and target-specific botanical drug products (Izzo et al., 2009). Phenotypic diversity in xylem composition (**Figure 2**) could also be exploited to increase xylem and hemicellulose quality to improve the suitability of *Cannabis* as a feedstock for the production of furfural (Brazdauskis et al., 2016), a promising and versatile chemical precursor for a number of biofuels and bio-renewables, including plastics (see review López et al., 2016). However, at present the limiting step for introgression of exotic allelic variants into novel cultivars for these as well as other end-use and agronomic traits (Hillig, 2005a; Tang et al., 2013), is access to well-represented and characterized *Cannabis* genetic resources.

Explicit characterization of *Cannabis* genetic resources as a key stage in cultivar development has been underutilized (Mandolino and Carboni, 2004; Ranalli, 2004; Amaducci et al., 2014; Salentijn et al., 2014), and a more sophisticated use of intraspecific germplasm will be required to advance elite breeding lines able to meet market requirements and a wider range of growing environments worldwide. In this review we describe *Cannabis* genetic resources reported within the literature. Potential gaps in coverage of available *ex situ* collections, as well as recommendations for genetic resource management and the use of germplasm resources in crop improvement programs, are proposed. A novel virtual approach to *ex situ* conservation is outlined to facilitate the exploitation of *Cannabis* genetic resources within the legal constraints surrounding this plant.



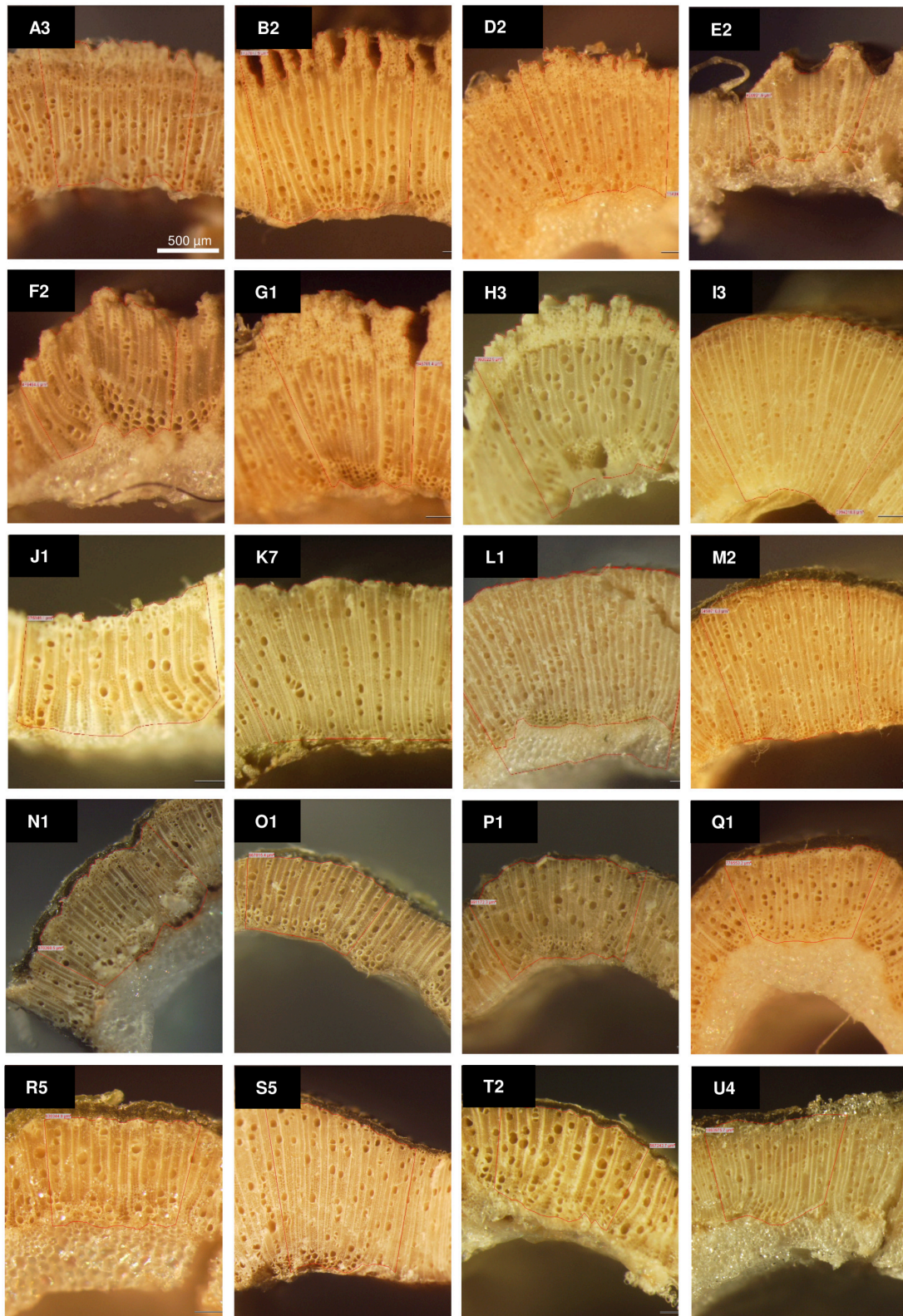


FIGURE 2 | Phenotypic variation in the xylem: phloem ratio for a range of *Cannabis* germplasm from the Ecofibre Global Germplasm Collection. Stereo microscopy images taken from internode 3–4 across a representative sub-set of 20 accessions; Accessions were grown in environmentally-controlled conditions and harvested at similar developmental stages; Letters correspond to accessions and numbers to replicate plant; Scale bar: 500 μ m; redrawn from (Stanger, 2015).

A HYPERVARIABLE SPECIES SHAPED BY DOMESTICATION

Taxonomy and Classification

Patterns of genetic and phenotypic variation to inform prioritization of intraspecific groupings for *ex situ* conservation require an appreciation of current taxonomic nomenclature. *Cannabis* has traditionally been classified as monotypic, characterized by a single species *C. sativa* (Small and Cronquist, 1976). The presence of additional species has been inferred from allozyme variation (Hillig, 2005b) and chemotaxonomic analysis (Hillig and Mahlberg, 2004), and placed in context of Pleistocene climatic events which are believed to have resulted in the isolation and diversification of this genus (Clarke and Merlin, 2013). However, given the lack of validated reports of extant “natural” populations, one has to consider *C. sativa* as a species solely represented by domesticated or post-domesticated germplasm. Thus, claims of additional taxonomic divisions are controversial, on the basis that plants readily inter-cross and physiological and genetic barriers to gene flow appear absent (Small and Cronquist, 1976).

The use of allozyme variation and chemotaxonomic analysis to determine population structure tends to bias non-neutral selection. Moreover, allocating observed variation to specific loci and allelic status requires validation through segregation analysis. Electrophoretic variation in allozyme loci does not always correspond to DNA sequence variation, and may be associated with experimental artifacts (May and Hoelzel, 1992; Olsen et al., 2014), post-translational modification, *cis*-regulatory, or epigenetic processes (Olsen et al., 2014). Although chemotaxonomic differentiation has been inferred on the basis of cannabinoid composition (Hillig and Mahlberg, 2004), this is thought to be governed by only a handful of loci (De Meijer et al., 2009; Weiblen et al., 2015). Thus, the use of allozyme variation and chemotaxonomic analyses performed to date may provide limited information on the genetic composition of diverse accessions, and therefore lend little support for speciation.

Nonetheless, *Cannabis* is polymorphic (Small et al., 2003) and is often differentiated into intraspecific taxonomic groupings on the basis of chemotype, ecotype, crop-type (fiber or drug), and leaflet morphology (Small et al., 2003; Lynch et al., 2015). However, agreement on a practical and workable nomenclature for subspecies differentiation within *Cannabis* is lacking (Hillig, 2005b; Clarke and Merlin, 2015; Small, 2015b), and could be best described as *ad hoc*. The terms “*sativa*” and “*indica*” can be used variously to differentiate between recreational drug varieties, crop-types (Small and Cronquist, 1976), chemotypes (Elzinga et al., 2015), ecotypes, and changes in leaflet morphology (Hillig, 2005b; Clarke and Merlin, 2013). Classification based on the terms “*fiber-type*” and “*drug-type*” to differentiate between subspecies is also problematic, given the widespread changes in the status of *Cannabis* in different legislatures (Pardo, 2014) and the propensity for fiber accessions to exhibit characteristics indicative of both fiber- and drug-type plants (Tipparat et al., 2012; Welling et al., 2015). The use of descriptors associated with a combination of stable chemical and morphological traits,

and/or that incorporate DNA evidence, is more likely to provide practical operational classification.

Diversification through Domestication

Anthropogenic selection focused predominantly on the traits defining industrial hemp, hempseed, pharmacological, and marijuana end-uses has been influential in the phenotypic and chemotypic diversification of *Cannabis* (see review Small, 2015a), and is therefore an important consideration for establishing the coverage and representation of germplasm accessions required for *ex situ* genetic resource conservation. Moreover, detection of population structure reflecting long-term crop use, post-domestication radiation, and eco-geographic adaptation may be valuable for the management of plant genetic resources and breeding programs, as has been demonstrated for the vegetable *Brassica oleracea* (Smith and King, 2000), a species which has been subject to similar patterns of renewed domestication into diverse crop forms and secondary radiation.

It has been suggested that domestication of marijuana has resulted in both qualitative and quantitative increases in THC. Analyses of THC content as a percentage of dried weight (w/w) as high as 39.8% have been recorded from contemporary marijuana preparations (Swift et al., 2013). However, such high levels in content could also be attributed to an improvement in analytical and extraction methodologies (De Backer et al., 2009) and the preferential sampling of cannabinoid-rich floral tissues (buds; Happyana et al., 2013), as opposed to demonstrable (specific) heritable genetic changes. Nonetheless, contemporary fiber and marijuana forms have shown that the latter can not only exhibit a higher overall cannabinoid content than the former, but also a higher proportion of THC within the cannabinoid fraction (Hillig and Mahlberg, 2004; Pacifico et al., 2006; Staginnus et al., 2014; Lynch et al., 2015). However, comprehensive large scale quantitative comparison both within and between representatives from each grouping are lacking.

Transcriptome analysis (Van Bakel et al., 2011; McKernan et al., 2015; Onofri et al., 2015; Weiblen et al., 2015) and QTL mapping (Weiblen et al., 2015) further support the historical selection pressure that has led to chemotypic separation between contemporary fiber and marijuana groupings. Cannabinoids accumulate in plants in their carboxylic acid forms, such as tetrahydrocannabinolic acid (THCA) and cannabidiolic acid (CBDA) which form neutral cannabinoids THC and CBD in a non-enzymatic reaction when exposed to heat (Dussy et al., 2005). The marijuana variety Purple Kush was found to have higher expression levels of genes encoding cannabinoid biosynthetic pathway intermediates than the fiber cultivar Finola, with Purple Kush only expressing functional sequence variants of genes coding for THCA and Finola those for CBDA synthase (Van Bakel et al., 2011). Linkage mapping in 62 F₂ individuals derived from a cross between full-sib inbred contemporary fiber cultivar Carmen and marijuana variety Skunk#1 revealed QTL for THCA and CBDA composition, as well as putative QTL for cannabinoid content, with differences in composition associated with a CBDA synthase locus and loss of CBDA synthase functionality in Skunk#1 (Weiblen et al., 2015).

The occurrence of contemporary fiber cultivars which have relatively high levels of CBDA (Small, 2015a) can be attributed to breeding efforts within France and other European countries toward the middle to latter part of the twentieth century (Amaducci et al., 2014), where techniques such as the Bredemann method, an *in vivo* fiber evaluation method, and the counter selection for THC using marker assisted selection (MAS), were employed (Ranalli, 2004). Fifty one fiber cultivars are currently registered for use within the European Union (Salentijn et al., 2014) and these cultivars have been exported to North America (Small, 2015a) and the Northern provinces of China (Salentijn et al., 2014). Despite the scarcity of published data relating to ancestry of such accessions, it is believed that a large number of contemporary fiber cultivars are descendants from Central Russian and Mediterranean landraces and derivative cross-progenies (De Meijer and van Soest, 1992).

Phylogenetic relationships between domesticated *Cannabis* germplasm have recently been examined using reduced representation DNA sequencing. Genotyping by sequencing (GBS) analysis of 195 accessions using 2894 single nucleotide polymorphisms (SNPs) inferred close relatedness and shared ancestry between contemporary fiber accessions, with the latter forming a separate clade (Lynch et al., 2015). These accessions were also observed to exhibit lower levels of heterozygosity than other intraspecific taxa (Mann-Whitney U-test $p < 0.001$; Lynch et al., 2015), suggesting that recent domestication of fiber traits has resulted in a genetic bottleneck and reduction in allelic diversity. However, these varieties were not well-represented in the sample collection, with only 16 analyzed (Lynch et al., 2015). Moreover, a separate GBS study using 14,031 SNPs in 43 contemporary fiber and 81 marijuana varieties produced results that conflicted with this more recent study, and showed significantly lower levels of heterozygosity in marijuana varieties compared with fiber cultivars (Mann-Whitney U-test $p = 8.64 \times 10^{-14}$; Sawler et al., 2015).

Despite this lack of congruence between GBS analyses, domestication for either industrial hemp or marijuana traits has likely resulted in a loss of genetic and allelic diversity, potentially brought about by changes in breeding systems. Processes such as linkage drag can be associated with complex polygenic flowering QTL (Mace et al., 2013) in relation to latitudinal and environmental adaption (Gao et al., 2014). Regardless of reductions in allelic diversity that may have arisen either from clonal propagation in marijuana (Russo, 2007), or from the propagation of monoecious varieties in industrial hemp (Forapani et al., 2001), it is unclear to what extent contemporary *Cannabis* germplasm deviates from the broader gene pool. Analysis of 45 SNPs in both GBS sample sets reveals an overall limited genetic distance between 22 industrial hemp and 173 marijuana groupings (Lynch et al., 2015). Resequencing and mapping of 30 billion sequence reads from 302 domesticated and wild soybean (*Glycine max*) accessions identified selective sweeps associated with domestication events (Zhou et al., 2015). By comparison, resequencing of various species of *Citrus* has also revealed a complex arrangement of large haplotype blocks and admixture between ancestral and domesticated species (Wu et al., 2014). On completion of a fully annotated *Cannabis* genome

(Van Bakel et al., 2011), it may be possible to resequence diverse germplasm (Scossa et al., 2015) to quantify differences in genetic diversity and to determine the contribution wild ancestors have conferred to contemporary forms. However, access to wild and landrace accessions may be a limiting factor to exploring *Cannabis* phylogeny.

Identifying Centers of Diversity

The region associated with the origin of a species is often considered the center of genetic diversity, although this may not be the case with species that have been subject to prolonged periods of domestication and secondary radiation. This has been observed in the common bean (*Phaseolus vulgaris*), where simple sequence repeat (SSR) genetic marker comparisons between native American germplasm and secondary domesticated groupings revealed a higher level of nuclear diversity in the African germplasm than that of the native collection, suggesting that non-native geographical regions can act as both source and sinks for genetic diversity over different historical periods (Bellucci et al., 2014). Palaeobotanical evidence in the form of pollen deposits and historical accounts would strongly suggest domestication of *Cannabis* in the post-Neolithic era over tens of thousands of years, with subsequent secondary domestication events in non-native geographical regions (Li, 1974; Bradshaw et al., 1981; Small and Marcus, 2002; Murphy et al., 2011; Herbig and Sirocko, 2013; Duvall, 2016).

Based on field observations it has been concluded that *Cannabis* originated within Central Asia (Hillig, 2005b; Russo, 2007) and references therein, although such inferences may not be justified given the domestication and radiation of *Cannabis* throughout Eurasia over several millennia (Li, 1974; Bradshaw et al., 1981; Murphy et al., 2011; Herbig and Sirocko, 2013). Conclusive phylogenetic evidence in support of a specific geographical region is also incomplete, with ruderal or wild populations underrepresented or absent from sample collections (Hillig, 2005b; Lynch et al., 2015). Combining genetic and phenotypic evidence relating to the predominant characteristics of a species in both ruderal and domesticated forms in the context of their allelic eco-geographical distribution, is more likely to define a putative center of diversity, as has been noted in other widely cultivated plant species (Smith and King, 2000).

East Asia appears to be a rich source of genetic diversity within the *Cannabis* gene pool, and a potentially valuable genetic resource both for future phylogenetic analyses and *ex situ* conservation. China is botanically megadiverse (Li, 2008). The Hengduan Mountains in the south west of China have been identified as one of only 35 biodiversity hotspots worldwide (Sloan et al., 2014) and this region encompasses parts of the Yunnan province in which the Yunnan Academy of Agricultural Sciences (YAAS) *Cannabis* germplasm collection is maintained. China also benefits from a latitudinal gradient from ~ 23 – 50° N (Amaducci et al., 2014), with hundreds of *Cannabis* landraces reported to have undergone distinct domestication events along these latitudes in provinces spanning from Hebei in the north west, through to ShanDong, Henan, Guizhou and Yunnan in the south west (Salentijn et al., 2014). Moreover, historical evidence strongly suggests that *Cannabis* has been cultivated in China over

several thousand years. Pottery paintings depicting *Cannabis* are believed to have been produced by the Neolithic Yangshao culture, and *Cannabis* fibers were reportedly utilized in the production of paper during the Han dynasty >1790 years before present (Li, 1974). Excavation of the 2700 year old Yanghai Tombs in Xinjiang-Uyghur Autonomous Region in China has also revealed high THC plant material (Russo et al., 2008), implying that this plant was used within a cultural and potentially medicinal context within early Chinese societies.

A number of genetic markers have been used to determine levels of heterozygosity in Chinese *Cannabis* germplasm (Table 1), with SSR, amplified fragment length polymorphism (AFLP) and randomly amplified polymorphic DNA (RAPD) genetic markers indicating a high level of genetic diversity, with the proportion of polymorphic loci ranging from 75 to 92% (Gilmore et al., 2007; Hu et al., 2012; Gao et al., 2014; Zhang et al., 2014). Analysis of 76 accessions from 26 countries using 12 chloroplast and mitochondrial DNA loci revealed six haplotypes, all of which were located within or adjacent to China (Gilmore et al., 2007). Fifty six loci derived from expressed sequence tag (EST) simple sequence repeat (EST-SSR) markers were tested on a sample collection of 100 varieties from 10 provinces in China and 15 varieties from Europe. Principle coordinate analysis revealed four clusters relating to geographical location and latitude. Interestingly, clusters from Central, Northern, and Southern China had a higher percentage of polymorphic loci than the European cluster (Gao et al., 2014; Table 1), suggesting a higher level of diversity within Chinese germplasm.

Of the 808 accessions reported to have been collected within *ex situ Cannabis* genetic resource collections worldwide, only 58 (7.2%) have origins within China (Figure 3). Moreover, from the 156 accessions listed in the former Centre for Plant Breeding and Reproduction Research (CPRO) germplasm collection, cultivars of predominantly European origin and marijuana varieties contributed to more than 40% of all accessions (De Meijer and van Soest, 1992). Considering the long cultivation history of *Cannabis* in China, and the high density of landrace accessions occurring within a large latitudinal range, sourcing accessions from China and adjacent regions should be a priority for *Cannabis ex situ* conservation, irrespective of whether the center of origin for this species has been fully characterized. Nevertheless, appropriate management of germplasm *ex situ* and systematic characterization of East Asian *Cannabis* genetic resources will be required if the full potential of genepool-enabled crop improvement is to be maximized.

LIMITED COORDINATION OF CANNABIS GENETIC RESOURCES

Limited effort has been devoted to comprehensive characterization of currently available *Cannabis ex situ* germplasm resources. In the latter part of the twentieth century *Cannabis* genetic resource collections in North America were confiscated, or in some instances reported to have been destroyed by regulatory authorities (Small and Marcus, 2002). Those that did remain were subject to funding constraints, resulting in

suboptimal maintenance, and regeneration of seed (Clarke, 1998). However, an increasing number of research groups are currently active in *Cannabis* research (see Supplementary Table), and a concerted effort to characterize *ex situ* genetic resource collections would have significant benefit to the research community and for the continued genetic improvement of this plant.

Ex situ Cannabis genetic resource collections have been partially characterized and reported within peer reviewed conference proceedings and other reports from public sector researchers, as well as in a variety of gray literature and online databases (Table 2). These include reference to collections in Europe from the Research Institute of Crop Production (RICIP; Pavelek, 2006), the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK; Höppner, 2006), and the Istituto Sperimentale per le Colture Industriali (CRA-ISCI; Mandolino et al., 2006). The characterization of three genetic resource collections has been published in refereed scientific journals; the CPRO germplasm collection (De Meijer and van Soest, 1992), the Vavilov Research Institute (VIR; Clarke, 1998) and the Institute of Natural Fibres and Medicinal Plants (INF&MP) gene bank collections (Mankowska and Silska, 2015).

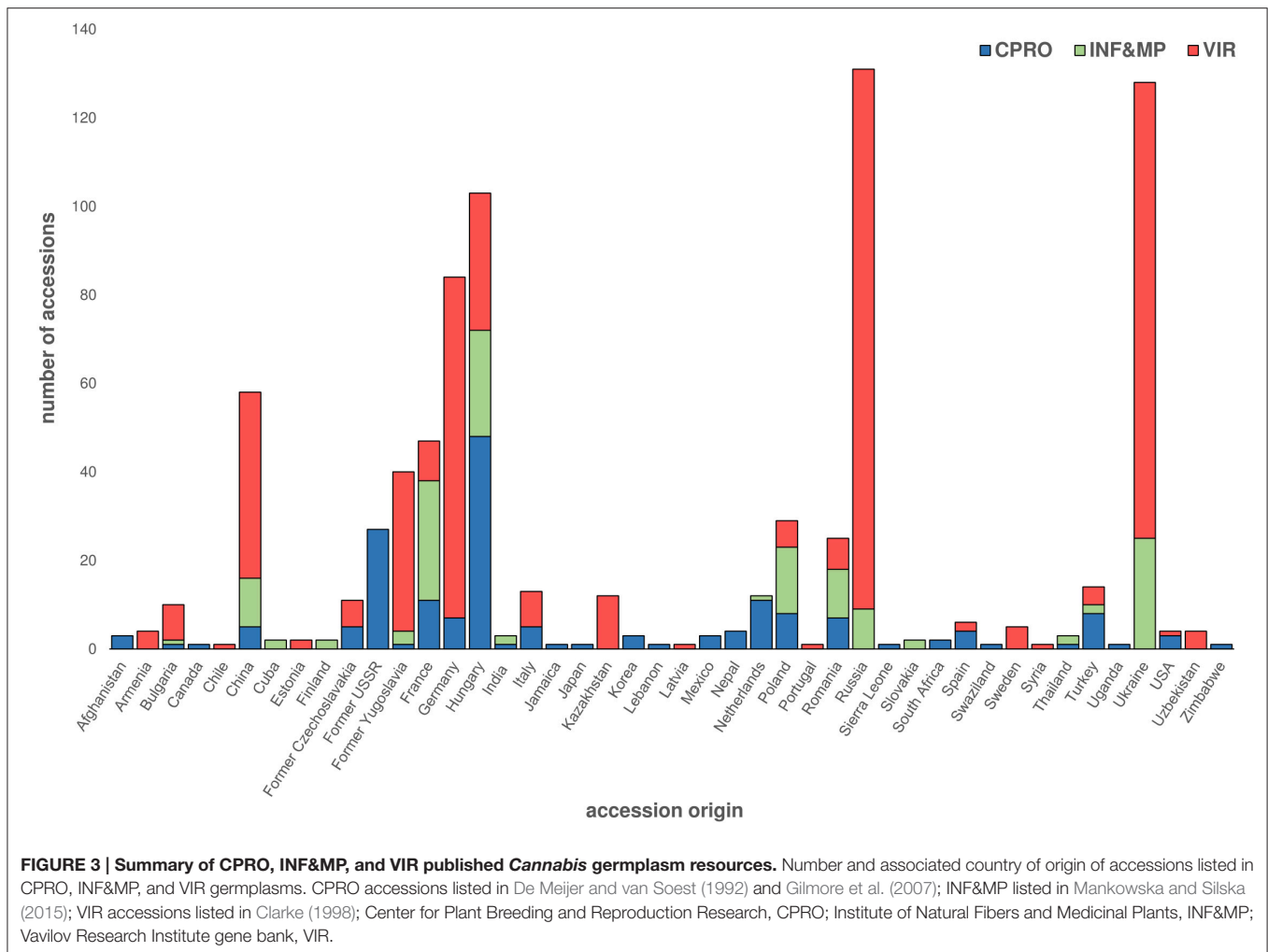
The CPRO (now the Wageningen University) germplasm collection was established in 1988 (De Meijer and van Soest, 1992; Table 2). This is the only published example of a collection which provides passport data for each accession (De Meijer and van Soest, 1992; Clarke, 1998; Mankowska and Silska, 2015). The last comprehensive report in 1992 identified over 156 accessions (De Meijer and van Soest, 1992; Gilmore et al., 2007), comprising of cultivars, varieties, strains or forms, as defined by International Code of Nomenclature for Cultivated Plants (Brickell et al., 2009), as well as landrace, feral, or ruderal accessions. Accessions were sourced from previous genetic resource collections and academic and commercial organizations. Accessions held within this collection originated from more than 22 countries (De Meijer and van Soest, 1992; Gilmore et al., 2007) (Figure 3), with 42.6% originating from the former Union of Soviet Socialist Republics (USSR) and Hungary. The last published entry of a CPRO accession was in 1999 (Gilmore et al., 2007), although this collection is now being used in the EU-led MultiHemp consortium (Salentijn et al., 2014).

The VIR collection is considered the largest collection of *Cannabis* genetic resources in Europe (Small and Marcus, 2002; Table 2). The VIR, formerly the All-Union Institute of Plant Industry, established in 1920 by Nikolai Vavilov was the first *ex situ* gene bank of its kind to exist worldwide (Tyagi and Agrawal, 2015). The last published report in 1998 characterized the origin of 493 accessions, with accessions from Ukraine, Russia, and Germany representing 61.3% of the collection (Figure 3; Clarke, 1998). Lack of funding is thought to have significantly compromised regeneration efforts (Khouri et al., 2010) leading to a loss of unique genetic diversity held within this resource (Clarke, 1998; Small and Marcus, 2002).

The INF&MP gene bank is the most recently published genetic resource collection (Mankowska and Silska, 2015; Table 2). A report in 2015 identified 139 accessions of predominantly European Origin, with accessions from France, Hungary, and

TABLE 1 | Evidence for the diversity of East Asian germplasm.

DNA marker	No. loci /bands	No. of accessions	Country of origin of accessions	Province of Chinese accessions	Genetic variability	References
AFLP®	442	49	Europe, China	Anhui, Gansu, Guangxi, Guizhou, Heilongjiang, Inner Mongolia, Jilin, Kunming, Liaoning, Shanxi, Sichuan, Xinjiang, Yunnan	Percentage of polymorphic loci (PPL): 92.1% Yunnan (88.8%) and Heilongjiang (75.6%) populations exhibited the highest PPL	Hu et al., 2012
CpSSR / mtSSR	7	76	Afghanistan, Australia, Canada, China, Former Czechoslovakia, Former East Germany, Former USSR, France, Hungary, India, Italy, Jamaica, Japan, Korea, Lebanon, Nepal, Netherlands, Romania, Sierra Leone, South Africa, Swaziland, Thailand, Turkey, Uganda, USA, Zimbabwe	Not specified	Six organelle haplotypes identified All haplotypes occurred in South and East Asia, either within China or countries adjacent to China	Gilmore et al., 2007
EST-SSR	56	115	China, France, Poland, Ukraine	Anhui, Chongqing, Gansu, Gansu, Hanma, Hebei, Heilongjiang, Henan, Huangzhangma, Jilin, Liaoning, Neimenggu, Ningxia, Shandong, Shanxi, Xinjiang, Yunnan, Zhejiang	PCoA: accessions clustered into four groups; Central, Northern and Southern China, and Europe Germplasm originating from Central China (85.5%), Northern China (76.4%), and Southern China (74.6%) exhibited the highest PPL compared with germplasm originating from Europe (41.8%)	Geo et al., 2014
ISSR and chromosome	183	27	China	Anhui, Chongqing, Hebei, Heilongjiang, Henan, Inner Mongolia, Jiangsu, Jilin, Liaoning, Ningxia, Qianghai, Shanxi, Shandong, Shanxi, Xinjiang, Yunnan, Zhejiang	PPL: 85.8% Five chromosome types $2n = 20 = 14m + 6sm$; $2n = 20 = 20m$; $2n = 20 = 18m + 2sm$; $2n = 20 = 16m + 4sm$; and $2n = 20 = 12m + 8sm$ 3 karyotypes identified	Zhang et al., 2014
RAPD	106	16	China	Anhui, Heilongjiang, Inner Mongolia, Jilin, Liaoning, Shandong, Tibet, Yunnan, Xinjiang	PPL: 74.5%	Tang et al., 2013



the Ukraine contributing to 54.7% of the collection (Mankowska and Silska, 2015; **Figure 3**). However, passport data of individual accessions held within the INF&MP and VIR collections are not listed within the literature, with both articles limiting data to the number of accessions, and their associated country of origin (Clarke, 1998; Mankowska and Silska, 2015). Further, provenance is unclear, since the authors do not clearly define the term origin, which can be interpreted as either the location that the plant exists within its natural state, the location with which the accession was acquired, or in the case of a cultivar, strain, variety or form, the location where the accession was developed or registered.

Reference to other significant genetic resource collections has also been made in the literature, including the YAAS collection in China-Yunnan, comprising of ~350 accessions of mainly East Asian origin (Salentijn et al., 2014), and the Ecofibre Global Germplasm Collection (EFGGC) in Australia which has a diverse collection of accessions from multiple regions within Eurasia (Welling et al., 2015). However, comprehensive published data detailing the extent of genetic resources held within these collections is lacking. In the absence of more

contemporary peer reviewed published data that systematically characterizes genetic resources, along with detailed passport and other metadata, estimates of the extent to which genetic resource collections represent the global allelic diversity of *Cannabis* will be incomplete.

ADDRESSING THE FRAGILITY OF GLOBAL CANNABIS GERmplasm RESOURCES

Maintenance and Regeneration

Regeneration is one of the most costly factors in the conservation of *ex situ* genetic resource collections and requires significant input in the form of labor, resources, and infrastructure (Li and Pritchard, 2009; Khoury et al., 2010). The regeneration of *Cannabis* germplasm is especially problematic as plants are anemophilous (wind pollinated) and dioecious, with male or hermaphrodite plants releasing large amounts of pollen during flowering (Amaducci et al., 2014). It has been estimated that a single plant can produce in excess of 350 million pollen grains (Faegri et al., 1989). Prevention of gene flow and hybridization

TABLE 2 | Summary of *Cannabis* germplasm resources reported within the literature.

<i>Cannabis</i> genetic resource collection	Location	Link	Online database	References
Centre for Plant Breeding and Reproduction Research (CPRO)	Netherlands-Wageningen	http://www.wageningenur.nl/	n/a	De Meijer and van Soest, 1992; De Meijer et al., 1992; Hillig, 2005b; Gilmore et al., 2007
Ecofibre Global Germplasm Collection	Australia, Queensland	http://www.ecofibre.com.au/	n/a	Welling et al., 2015
HortaPharm B.V., Amsterdam, Holland (now merged with GW pharmaceuticals)	Netherlands, Amsterdam	n/a	n/a	Ranalli, 2004; Hillig, 2005b; Datwyler and Weiblen, 2006
Institute of Natural Fibers and Medicinal Plants (INF&MP)	Poland-Poznan	http://archive-ecpgr.cgiar.org/fileadmin/www.ecpgr.cgiar.org/Presentations/Fibre_crops_NL_2006/Collection%20of%20cannabis%20at%20INF.pdf	http://archive-ecpgr.cgiar.org/	Mankowska and Silska, 2015
Istituto Sperimentale per le Colture Industriali (CRA-ISCI)	Italy, Bologna	http://archive-ecpgr.cgiar.org/fileadmin/www.ecpgr.cgiar.org/Presentations/Fibre_crops_NL_2006/Italy%20Country%20report.pdf	http://archive-ecpgr.cgiar.org/	Mandolino et al., 1999, 2006; Pacifico et al., 2006
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)	Germany, Gatersleben	http://archive-ecpgr.cgiar.org/fileadmin/www.ecpgr.cgiar.org/Presentations/Fibre_crops_NL_2006/Germany%20Country%20report%20final.pdf	http://gbis.ipk-gatersleben.de/gbis_i/	Small and Marcus, 2003; Höppner, 2006
National Cannabis Collection (NCC)	Hungary, Tápiószéle	http://archive-ecpgr.cgiar.org/fileadmin/www.ecpgr.cgiar.org/Presentations/Fibre_crops_NL_2006/Hungary%20Country%20report.pdf	http://archive-ecpgr.cgiar.org/	Simon, 2006
Research Institute of Crop Production (RICIP; owned by AGRITEC Ltd)	Prague, Ruzyne	http://archive-ecpgr.cgiar.org/fileadmin/www.ecpgr.cgiar.org/Presentations/Fibre_crops_NL_2006/Czech%20Country%20report.pdf	http://archive-ecpgr.cgiar.org/	Small and Marcus, 2003; Pavelek, 2006
Vavilov Research Institute (VIR)	Russia-St Petersburg	http://www.vir.nw.ru/	http://91.151.189.38/virdb/	Clarke, 1998; Small and Marcus, 2003; Hillig, 2005b
Yunnan Academy of Agricultural Sciences (YAAS)	China, Kunming	http://www.university-directory.eu/China/Yunnan-Academy-of-Agricultural-Sciences.html	n/a	HuaDong and YingFang, 2012; Salentijn et al., 2014

It should be noted that the availability of Cannabis germplasm may vary over time between genetic resource collections and legislatures.

between accessions is therefore an important consideration in *Cannabis ex situ* regeneration and conservation (De Meijer and van Soest, 1992), and it may have contributed to reports of a widely shared *Cannabis* genepool.

Outdoor regeneration can require large areas of land, with distances of up to 5 km required to prevent cross-pollination (De Meijer and van Soest, 1992; Small and Antle, 2003). However, given the long distances *Cannabis* pollen can travel and the sensitivity of pollen distribution to wind velocity (Small and Antle, 2003), as well as the potential for prolonged viability of *Cannabis* pollen in low relative humidities post-anthesis (Bassani et al., 1994), introgression may not necessarily be prevented at these distances. Outdoor regeneration can also be impractical where multiplication of diverse accessions requires specific photoperiods spanning several degrees of latitude (Gao et al., 2014) in order to initiate flowering (Cosentino et al., 2012),

thus limiting multiplication of *Cannabis* germplasm to certain periods throughout the year.

Protected cultivation within pollen secure facilities is a standard alternative to outdoor multiplication. However, these may be costly and also problematic in maintaining genetic diversity. For many plant species, genetic drift has been attributed to the process of regeneration, and associations have been detected between time spent *ex situ* and loss of alleles per locus, gene diversity, and percentage of polymorphic loci (Parzies et al., 2000). For example, loss of *ex situ* genetic diversity has been observed in outcrossing species such as barley (*Hordeum vulgare* L.), maize and common bean genetic resource collections (Parzies et al., 2000; Rice et al., 2006; Negri and Tiranti, 2010), although much of the reduction in genetic diversity appears dependent on the number of parents used in mass pollination (Parzies et al., 2000). Given the spatial limitations associated with

the construction of indoor pollen-secure infrastructure (Negri and Tiranti, 2010), which limits the number of outbreeding individuals for regeneration, a decline in allelic variation may occur with each regeneration cycle, ultimately leading to erosion of genetic variability from the time of acquisition.

Seed Storage

Considering the complexities associated with *Cannabis* germplasm regeneration, much attention has been directed to efficient methods of seed storage. Duration of storage, temperature, and seed moisture content are all variables which can significantly affect the viability of *Cannabis* seeds (Small and Brookes, 2012; Parihar et al., 2014). Common practice for short-term seed storage is 4°C with a moisture content of 6% (De Meijer and van Soest, 1992; Small and Brookes, 2012; Mankowska and Silska, 2015), while for longer periods of storage >3 years, seeds are held at -20°C with ~4% moisture content (De Meijer and van Soest, 1992; Mankowska and Silska, 2015). *Cannabis* seed appears to have orthodox storage behavior, and the ability to withstand periods of up to 66 months after desiccation with minimal effects on seed viability (Small and Brookes, 2012; Parihar et al., 2014). Nevertheless, systematic evidence for the long-term viability of *Cannabis* seed is lacking. Given the variation in seed size amongst *Cannabis* germplasm (Piluzza et al., 2013), it is important to establish optimal periods for seed storage prior to multiplication.

Desiccation and cold storage does not necessarily guarantee seed longevity: Only 61/276 species were found to exhibit a half-life >100 years at these conditions (Walters et al., 2005). Both orthodox and non-orthodox seed types may benefit from cryopreservation conditions (Li and Pritchard, 2009; Michalak et al., 2015; Perullo et al., 2015; Prada et al., 2015). Despite the absence of data supporting the long-term viability of seed storage at temperatures below -180°C, cryopreservation of seed has successfully been demonstrated as a proof of concept in a number of plant species (Li and Pritchard, 2009; Michalak et al., 2015; Perullo et al., 2015; Prada et al., 2015). Short-term cryopreservation of seed of several species, including black poplar (*Populus nigra* L.; Michalak et al., 2015), swamp pink (*Helonias bullata* L.; Perullo et al., 2015), and Barbados nut (*Jatropha curcas* L.; Prada et al., 2015), had no detrimental effect on germination. However, seed water content prior to immersion in liquid nitrogen can significantly determine seed viability (Michalak et al., 2015). Given the expense and propensity for genetic drift associated with *Cannabis* germplasm regeneration, as well as the potential orthodox storage properties of *Cannabis* seed, contemporary methodologies of seed storage should be explored to determine if they are compatible and economically feasible with long-term *Cannabis ex situ* conservation.

Few studies have attempted to monitor seed aging in *Cannabis* (Parihar et al., 2014) or associate a specific biochemical, metabolic, or physiological characteristic with seed viability or vigor (Small and Brookes, 2012). However, a number of novel biomarker-based prediction tools have been developed, with the potential to monitor seed aging beyond traditional germination and biochemical tests (Fu et al., 2015). Biomarkers associated with DNA methylation (Rocha et al., 2005) as well as FA (Li et al.,

2007) and endogenous antioxidant metabolism (Revilla et al., 2009) have been found to be reliable at predicting seed aging (Fu et al., 2015). Such tools may contribute toward predicting *Cannabis* seed viability, and be adopted for development of “best practice” methods for long-term *ex situ* seed management.

Passport Data

The long-term value and utility of *ex situ* germplasm is subject to the quality and quantity of data and metadata available for a given genetic resource (Ramírez-Villegas et al., 2010; Endresen et al., 2011). The systematic collation of passport data can aid in identifying duplication which may have occurred both between and within genetic resource collections (Van Hintum and Visser, 1995). Indeed, the recently revised gene bank standards for plant genetic resources recommend the comprehensive assignment of universally accepted passport descriptors, and that the conservation and use of genetic resources should be collated within a suitably designed database (Tyagi and Agrawal, 2015). The adoption of Digital Object Identifiers (DOIs) for datasets and associated entities, linked to an Open Researcher and Contributor ID (ORCID) can also encourage data sharing and reduce management complexities associated with homonyms and synonyms.

Gap methodologies which incorporate sampling representativeness scores and distribution modeling can be utilized to prioritize geographical locations for future *ex situ* conservation. These methods have been successfully applied to intra- and inter-specific taxa of the common bean (Ramírez-Villegas et al., 2010), Pigeon pea (*Cajanus cajan* L.), African cowpea (*Vigna unguiculata* L.; Moray et al., 2014), wheat, barley (Endresen et al., 2011), and tomato (Castañeda-Álvarez et al., 2015). Determining the eco-geographical profile or biome-specific adaption of germplasm with the use of passport data, geographic information system (GIS) technologies (<http://www.diva-gis.org>), and modeling software (Endresen, 2010; Ramírez-Villegas et al., 2010) can also increase the probability of targeting phenotypically relevant accessions within *ex situ* genetic resource collections (Bekessy et al., 2003; Endresen, 2010; Endresen et al., 2011; Jones et al., 2013). However, in the case of *Cannabis* genetic resources, where comprehensive publicly available passport, provenance, and metadata are either absent or deemed low-quality, these approaches to *Cannabis* germplasm characterization and conservation are currently not possible.

A NOVEL APPROACH TO CANNABIS CONSERVATION

Establishing a Virtual Core Collection

Additional approaches have been developed to enhance the management of *ex situ* seed genetic resource collections. These include the establishment of core collections, which represent the range of genetic diversity for a species with minimal repetition (Brown, 1989). Whilst core collections contain a subset of the genetic resources for a species, they do provide an enriched source of diversity, enabling efficient genepool characterization and utilization (Brown, 1989; Odong et al., 2013). We propose

that an international core collection of *Cannabis* be established from existing gene banks comprising up to 400 accessions collected from diverse geographical regions. Based on theoretical studies of natural populations (Lawrence et al., 1995a), in *ex situ* collections this should represent 99% of allelic polymorphisms for alleles that occur at species-wide frequencies >2% (Lawrence et al., 1995b). Having established an initial core collection, more comprehensive empirical testing could employ methods which use pairwise comparisons of allele frequencies of the core to guide the inclusion or exclusion of additional accessions (Thachuk et al., 2009; Odong et al., 2013).

Major barriers to the immediate development of a centralized physical core collection for *Cannabis*, include the lack of publicly accessible seed banks (Sawler et al., 2015) and the limitations on import, export, and transfer of seed between sovereign states that are signatories to the United Nations 1961 Single Convention on Narcotic Drugs and the 1971 Convention on Psychotropic Substances (Nutt, 2015; Pain, 2015). Nevertheless, this need not prevent the generation and exchange of data, nor the development of internationally agreed nomenclature and characterization standards, as well as the establishment of a “virtual” core collection and associated online data repository. Indeed, considering these obstacles, the *in silico* management of genetic resources and development of a commonly shared database containing detailed passport, meta-, and characterization-data, together with a coordinated distribution system, may be a necessity for *ex situ Cannabis* conservation, and even facilitate the commodification of privately managed germplasm collections as a pre-breeding resource.

High Throughput DNA Sequencing

Industrial hemp displays considerable phenotypic plasticity in response to environmental factors (Struik et al., 2000; Amaducci et al., 2014). Thus, characterization of genetic diversity based on morphological characteristics (Jansky et al., 2015) or eco-geographical data alone (Jones et al., 2013) are not always accurate indicators of diversity within genetic resource collections. Neutral DNA markers, such as SSR, RAPD, and AFLP, which may not be subject to selective pressures, have typically been employed to assess the level of genetic diversity and phylogeny between *Cannabis* genetic resources (Mandolino and Carboni, 2004; **Table 1**). However, these do not necessarily reflect quantitative patterns of trait variation, nor can they always be considered a reliable tool in characterizing the functional genetic diversity of a species, nor the extent of genetic structure required for conservation (Bekessy et al., 2003).

High density SNP genotyping offers the promise of characterizing the genomic architecture underlying plant phenotypic variation and the extent of allelic diversification held within germplasm collections (Kilian and Graner, 2012; Scossa et al., 2015; Varshney et al., 2015). This has been demonstrated in other species (Kim and Buell, 2015; Luo, 2015), such as pepper (*Capsicum annuum*; Devran et al., 2015), lupin (*Lupinus angustifolius* L.; Yang et al., 2012), soybean (Kadam et al., 2016), and tomato (Kaminski et al., 2016). High-throughput multi-parallel resequencing would also promote sequence

informed *ex situ* conservation and could potentially eliminate germplasm redundancy (Kilian and Graner, 2012), and may further aid in resolving the plasticity of *Cannabis* traits which appear associated with epigenetic regulation, such as sex differentiation (Razumova et al., 2015).

Using high-throughput DNA sequencing to characterize *Cannabis* genome diversity and to develop markers is currently limited, due to the lack of a chromosome anchored and fully annotated reference genome, which limits sequencing applications (Ekblom and Galindo, 2011; Scossa et al., 2015). Two draft genomes are available for *Cannabis* (<http://www.medicinalgenomics.com>; <http://genome.ccb.utoronto.ca/>; Van Bakel et al., 2011) and these have been characterized through the completion of >195 K scaffolds. *De novo* sequencing is now widely utilized in non-model species, although this approach has a number of biological, computational, and biomolecular challenges specific to plants; (Schatz et al., 2012), such as the relatively high levels of heterozygosity (Gore et al., 2009) and genome complexity (Schnable et al., 2009). Reduced genome representation and complexity sequencing can also be utilized in the absence of a contiguous reference genome (Davey et al., 2011; Kim et al., 2016), although library preparation and choice of restriction enzyme can result in erroneous genotyping due to method induced sequencing errors, low read depth (Kim et al., 2016), uneven marker density and ultimately, missing data (Beissinger et al., 2013; Gardner et al., 2014).

Data Integration

The accumulation of high-throughput DNA sequence data requires systematic management of large datasets (Kilian and Graner, 2012) and analysis to obtain meaningful assemblies (Lange, 2016) that correspond to causative genomic regions and the underlying phenotypic traits of interest (Love et al., 2012). Navigation between data from different realms of genomics, genetics and trait-phenomics also requires a detailed understanding of the underlying genetic and experimental resources. The collation of data, however, can be incorporated into a suite of interconnected online databases such as those outlined for InterStoreDB (www.interstoredb.org/), which provides a web interface between raw sequence data, pairwise alignments and metadata relating to provenance, phenotype, and experimental parameters (Love et al., 2012). This framework incorporates the generic CropStoreDB (<http://www.cropstoredb.org>) database, which is currently being developed by researchers at Southern Cross University in Australia to characterize candidate genes involved in cannabinoid and lignocellulose composition in *Cannabis*.

Cannabis genome and transcriptome data are currently scattered within a number of public depositories (see review Lange, 2016), including the National Centre for Biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov/>; Van Bakel et al., 2011; Lynch et al., 2015; Sawler et al., 2015). Phenotype-specific data have also been made available through databases such as TrichOME (Dai et al., 2010; Lange, 2016), which incorporates ESTs and metabolite data from the trichomes of various aromatic plant species including *Cannabis* (Dai et al.,

2010), although no entries post 2012 have been recorded (Lange, 2016). A coordinated effort by researchers to centralize and normalize these data sets has the potential not only to harbor collaborative analysis, but also accelerate the *in silico* discovery of complex gene-to-trait relationships, ultimately benefiting both the research community and plant breeders alike.

CONCLUSION

Cannabis is a phenotypically and genetically diverse genus which has yet to benefit from the advanced level of breeding applied to other commercial crops. Coordinated and comprehensive conservation and characterization of *ex situ Cannabis* resources holds the promise of preserving gene pool diversity and enabling cultivar development. However, the legal constraints imposed by international narcotics conventions over more than 50 years have been influential in the fractionation and erosion of publicly accessible *Cannabis ex situ* genetic resources. The restrictions on legal exchange of *bona fide* research materials continues to limit the establishment of physical and centralized *ex situ* core collections. Nevertheless, the advent of low cost high-throughput DNA sequencing technologies and user-friendly data and metadata online repositories, make it possible to develop a virtual *Cannabis* core collection to facilitate the *in silico* mining of trait diversity and to guide genetic improvement strategies. This approach not only has the potential to accelerate the introgression and stabilization of commercially valuable genes into desirable germplasm, but may also precipitate a Green Revolution for *Cannabis* and the renewed commercial exploitation of this multi-functional plant species.

REFERENCES

- Amaducci, S., Scordia, D., Liu, F. H., Zhang, Q., Guo, H., Testa, G., et al. (2014). Key cultivation techniques for hemp in Europe and China. *Ind. Crops Prod.* 68, 2–16. doi: 10.1016/j.indcrop.2014.06.041
- Appendino, G., Chianese, G., and Tagliatela-Scafati, O. (2011). Cannabinoids: occurrence and medicinal chemistry. *Curr. Med. Chem.* 18, 1085–1099. doi: 10.2174/092986711794940888
- Babic, V., Vančetović, J., Prodanović, S., Kravić, N., Babić, M., and Anđelković, V. (2015). Numerical classification of Western Balkan drought tolerant maize (*Zea mays* L.) landraces. *J. Agric. Sci. Technol.* 17, 455–468.
- Bassani, M., Pacini, E., and Franchi, G. G. (1994). Humidity stress responses in pollen of anemophilous and entomophilous species. *Grana* 33, 146–150. doi: 10.1080/00173139409428991
- Beissinger, T. M., Hirsch, C. N., Sekhon, R. S., Foerster, J. M., Johnson, J. M., Muttoni, G., et al. (2013). Marker density and read depth for genotyping populations using genotyping-by-sequencing. *Genetics* 193, 1073–1081. doi: 10.1534/genetics.112.147710
- Bekessy, S. A., Ennos, R. A., Burgman, M. A., Newton, A. C., and Ades, P. K. (2003). Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. *Biol. Conserv.* 110, 267–275. doi: 10.1016/S0006-3207(02)00225-2
- Bellucci, E., Bitocchi, E., Rau, D., Rodriguez, M., Biagetti, E., Giardini, A., et al. (2014). “Genomics of origin, domestication and evolution of *Phaseolus vulgaris*,” in *Genomics of Plant Genetic Resources*, eds R. Tuberosa, A. Graner, and E. Frison (Springer), 483–507.
- Bradshaw, R. H. W., Coxon, P., Greig, J. R. A., and Hall, A. R. (1981). New fossil evidence for the past cultivation and processing of hemp (*Cannabis sativa* L.) in Eastern England. *New Phytol.* 89, 503–510. doi: 10.1111/j.1469-8137.1981.tb02331.x

AUTHOR CONTRIBUTIONS

MW: Carried out detailed literature survey and prepared manuscript. TS: Contributed to literature review and background information. TR: Contributed to development of review topic, detailed review, and revision of manuscript. LL: Contributed to organization of manuscripts, along with review and revision. RS: Contributed to generation of image data and review of manuscript. GK: Conceived of review topic, detailed review, and revision of manuscript.

FUNDING

MW, TS, and RS were sponsored by scholarships funded by Ecofibre Industries Operations Pty Ltd. TS is an employee of Ecofibre Industries Operations Pty Ltd.

ACKNOWLEDGMENTS

The authors would like to thank Phil Warner (EIO) for providing his wealth of knowledge on *Cannabis* genetic resources. The authors would further like to acknowledge the material and financial support generously provided by EIO and for granting the authors access to the Ecofibre Global *Cannabis sativa* L. Germplasm Collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.01113>

- Brazdauskas, P., Paze, A., Rizhikovs, J., Puke, M., Meile, K., Vedernikovs, N., et al. (2016). Effect of aluminium sulphate-catalysed hydrolysis process on furfural yield and cellulose degradation of *Cannabis sativa* L. shives. *Biomass Bioener.* 89, 98–104. doi: 10.1016/j.biombioe.2016.01.016
- Brickell, C. D., Alexander, C., David, J., Hettterscheid, W., Leslie, A., Malécot, V., et al. (2009). *International Code of Nomenclature for Cultivated Plants*. Leuven: International Society for Horticultural Science.
- Brown, A. (1989). Core collections: a practical approach to genetic resources management. *Genome* 31, 818–824. doi: 10.1139/g89-144
- Callaway, J. (2004). Hempseed as a nutritional resource: an overview. *Euphytica* 140, 65–72. doi: 10.1007/s10681-004-4811-6
- Carvalho, I. S., Miranda, I., and Pereira, H. (2006). Evaluation of oil composition of some crops suitable for human nutrition. *Ind. Crops Prod.* 24, 75–78. doi: 10.1016/j.indcrop.2006.03.005
- Castañeda-Álvarez, N. P., De Haan, S., Juárez, H., Khoury, C. K., Achicanoy, H. A., Sosa, C. C., et al. (2015). *Ex Situ* conservation priorities for the wild relatives of Potato (*Solanum* L. Section Petota). *PLoS ONE* 10:e122599. doi: 10.1371/journal.pone.0122599
- Chen, T., He, J., Zhang, J., Li, X., Zhang, H., Hao, J., et al. (2012). The isolation and identification of two compounds with predominant radical scavenging activity in hempseed (seed of *Cannabis sativa* L.). *Food Chem.* 134, 1030–1037. doi: 10.1016/j.foodchem.2012.03.009
- Clarke, R. C. (1998). Maintenance of *Cannabis* germplasm in the Vavilov Research Institute gene bank-five year report. *J. Int. Hemp Assoc.* 5, 75–79.
- Clarke, R. C., and Merlin, M. D. (2015). Letter to the Editor: Small, Ernest. 2015. Evolution and classification of *Cannabis sativa* (marijuana, hemp) in relation to human utilization. *Botanical Review* 81(3), 189–294. *Bot. Rev.* 81, 295–305. doi: 10.1007/s12229-015-9158-2

- Clarke, R., and Merlin, M. (2013). *Cannabis: Evolution and Ethnobotany*. Los Angeles, CA: University of California Press.
- Cosentino, S. L., Testa, G., Scordia, D., and Copani, V. (2012). Sowing time and prediction of flowering of different hemp (*Cannabis sativa* L.) genotypes in southern Europe. *Ind. Crops Prod.* 37, 20–33. doi: 10.1016/j.indcrop.2011.11.017
- Dai, X., Wang, G., Yang, D. S., Tang, Y., Broun, P., Marks, M. D., et al. (2010). TrichOME: a comparative omics database for plant trichomes. *Plant Physiol.* 152, 44–54. doi: 10.1104/pp.109.145813
- Da Porto, C., Decorti, D., and Tubaro, F. (2012). Fatty acid composition and oxidation stability of hemp (*Cannabis sativa* L.) seed oil extracted by supercritical carbon dioxide. *Ind. Crops Prod.* 36, 401–404. doi: 10.1016/j.indcrop.2011.09.015
- Datwyler, S. L., and Weiblen, G. D. (2006). Genetic variation in hemp and marijuana (*Cannabis sativa* L.) according to amplified fragment length polymorphisms. *J. Forensic Sci.* 51, 371–375. doi: 10.1111/j.1556-4029.2006.00061.x
- 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
- De Backer, B., Debrus, B., Lebrun, P., Theunis, L., Dubois, N., Decock, L., et al. (2009). Innovative development and validation of an HPLC/DAD method for the qualitative and quantitative determination of major cannabinoids in cannabis plant material. *J. Chromatogr. B* 877, 4115–4124. doi: 10.1016/j.jchromb.2009.11.004
- De Bruijn, P. B., Jeppsson, K.-H., Sandin, K., and Nilsson, C. (2009). Mechanical properties of lime-hemp concrete containing shives and fibres. *Biosyst. Eng.* 103, 474–479. doi: 10.1016/j.biosystemseng.2009.02.005
- De Meijer, E. P. M. (1994). Variation of *Cannabis* with reference to stem quality for paper pulp production. *Ind. Crops Prod.* 3, 201–211. doi: 10.1016/0926-6690(94)90067-1
- De Meijer, E. P. M., Hammond, K. M., and Sutton, A. (2009). The inheritance of chemical phenotype in *Cannabis sativa* L. (IV): cannabinoid-free plants. *Euphytica* 168, 95–112. doi: 10.1007/s10681-009-9894-7
- De Meijer, E. P. M., Van der Kamp, H. J., and Van Eeuwijk, F. A. (1992). Characterisation of *Cannabis* accessions with regard to cannabinoid content in relation to other plant characters. *Euphytica* 62, 187–200. doi: 10.1007/BF00041753
- De Meijer, E. P. M., and van Soest, L. J. M. (1992). The CPRO *Cannabis* germplasm collection. *Euphytica* 62, 201–211. doi: 10.1007/BF00041754
- Devran, Z., Kahveci, E., Özkaynak, E., Studholme, D. J., and Tör, M. (2015). Development of molecular markers tightly linked to Pvr4 gene in pepper using next-generation sequencing. *Mol. Breed.* 35, 101. doi: 10.1007/s11032-015-0294-5
- Dussy, F. E., Hamberg, C., Luginbühl, M., Schwerzmann, T., and Briellmann, T. A. (2005). Isolation of Δ^9 -THCA-A from hemp and analytical aspects concerning the determination of Δ^9 -THC in *Cannabis* products. *Forensic Sci. Int.* 149, 3–10. doi: 10.1016/j.forsciint.2004.05.015
- Duvall, C. S. (2016). Drug laws, bioprospecting and the agricultural heritage of *Cannabis* in Africa. *Space Polity* 20, 10–25. doi: 10.1080/13562576.2016.1138674
- Eklom, R., and Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107, 1–15. doi: 10.1038/hdy.2010.152
- ElSohly, M. A., and Slade, D. (2005). Chemical constituents of marijuana: the complex mixture of natural cannabinoids. *Life Sci.* 78, 539–548. doi: 10.1016/j.lfs.2005.09.011
- Elzinga, S., Fischechick, J., Podkolinski, R., and Raber, J. (2015). Cannabinoids and terpenes as chemotaxonomic markers in *Cannabis*. *Nat. Prod. Chem. Res.* 3:181. doi: 10.4172/2329-6836.1000181
- Endresen, D. T. F. (2010). Predictive association between trait data and ecogeographic data for Nordic barley landraces. *Crop Sci.* 50, 2418–2430. doi: 10.2135/cropsci2010.03.0174
- Endresen, D. T. F., Street, K., Mackay, M., Bari, A., and De Pauw, E. (2011). Predictive association between biotic stress traits and eco-geographic data for wheat and barley landraces. *Crop Sci.* 51, 2036–2055. doi: 10.2135/cropsci2010.12.0717
- Evenson, R. E., and Gollin, D. (2003). Assessing the impact of the Green Revolution, 1960 to 2000. *Science* 300, 758–762. doi: 10.1126/science.1078710
- Faegri, K., Kaland, P. E., and Krzywinski, K. (1989). *Textbook of Pollen Analysis*. Chichester: Wiley
- Faeti, V., Mandolino, G., and Ranalli, P. (1996). Genetic diversity of *Cannabis sativa* germplasm based on RAPD markers. *Plant Breed.* 115, 367–370. doi: 10.1111/j.1439-0523.1996.tb00935.x
- Finnan, J., and Styles, D. (2013). Hemp: a more sustainable annual energy crop for climate and energy policy. *Energy Policy* 58, 152–162. doi: 10.1016/j.enpol.2013.02.046
- Forapani, S., Carboni, A., Paoletti, C., Moliterni, V. M. C., Ranalli, P., and Mandolino, G. (2001). Comparison of hemp varieties using random amplified polymorphic DNA markers. *Crop Sci.* 41, 1682–1689. doi: 10.2135/cropsci2001.1682
- Fu, Y. B., Ahmed, Z., and Diederichsen, A. (2015). Towards a better monitoring of seed ageing under *ex situ* seed conservation. *Conserv. Physiol.* 3:26. doi: 10.1093/conphys/cov026
- Gao, C., Xin, P., Cheng, C., Tang, Q., Chen, P., Wang, C., et al. (2014). Diversity analysis in *Cannabis sativa* based on large-scale development of expressed sequence tag-derived simple sequence repeat markers. *PLoS ONE* 9:e110638. doi: 10.1371/journal.pone.0110638
- Gao, L.-Z. (2003). The conservation of Chinese rice biodiversity: genetic erosion, ethnobotany and prospects. *Genet. Resour. Crop Evol.* 50, 17–32. doi: 10.1023/A:1022933230689
- Gardner, K. M., Brown, P., Cooke, T. F., Cann, S., Costa, F., Bustamante, C., et al. (2014). Fast and cost-effective genetic mapping in apple using next-generation sequencing. *G3: Genes Genom. Genet.* 4, 1681–1687. doi: 10.1534/g3.114.011023
- George, J., and Sabapathi, S. (2015). Cellulose nanocrystals: synthesis, functional properties, and applications. *Nanotechnol. Sci. Appl.* 8, 45–54. doi: 10.2147/NSA.S64386
- Gilmore, S., Peakall, R., and Robertson, J. (2007). Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of *Cannabis sativa*. *Forensic Sci. Int.* 172, 179–190. doi: 10.1016/j.forsciint.2006.10.025
- Gore, M. A., Chia, J.-M., Elshire, R. J., Sun, Q., Ersoz, E. S., Hurwitz, B. L., et al. (2009). A first-generation haplotype map of maize. *Science* 326, 1115–1117. doi: 10.1126/science.1177837
- Grigor'ev, S., Shelenga, T., Baturin, V., and Sarana, Y. V. (2010). Biochemical characteristics of hemp seeds from various regions of Russia. *Russ. Agric. Sci.* 36, 262–264. doi: 10.3103/S1068367410040105
- Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156, 1–13. doi: 10.1007/s10681-007-9363-0
- Hall, W., and Degenhardt, L. (2007). Prevalence and correlates of cannabis use in developed and developing countries. *Curr. Opin. Psychiatry* 20, 393–397. doi: 10.1097/YCO.0b013e32812144cc
- Happyana, N., Agnolet, S., Muntendam, R., Van Dam, A., Schneider, B., and Kayser, O. (2013). Analysis of cannabinoids in laser-microdissected trichomes of medicinal *Cannabis sativa* using LCMS and cryogenic NMR. *Phytochemistry* 87, 51–59. doi: 10.1016/j.phytochem.2012.11.001
- Herbig, C., and Sirocko, F. (2013). Palaeobotanical evidence for agricultural activities in the Eifel region during the Holocene: plant macro-remain and pollen analyses from sediments of three maar lakes in the Quaternary Westeifel Volcanic Field (Germany, Rheinland-Pfalz). *Veg. Hist. Archaeobot.* 22, 447–462. doi: 10.1007/s00334-012-0387-6
- Hewavitharana, A. K., Golding, G., Tempany, G., King, G., and Holling, N. (2005). Quantitative GC-MS analysis of Δ^9 -tetrahydrocannabinol in fiber hemp varieties. *J. Anal. Toxicol.* 29, 258–261. doi: 10.1093/jat/29.4.258
- Hillig, K. W. (2005a). A combined analysis of agronomic traits and allozyme allele frequencies for 69 *Cannabis* accessions. *J. Ind. Hemp* 10, 17–30. doi: 10.1300/J237v10n01_03
- Hillig, K. W. (2005b). Genetic evidence for speciation in *Cannabis* (*Cannabaceae*). *Genet. Resour. Crop Evol.* 52, 161–180. doi: 10.1007/s10722-003-4452-y
- Hillig, K. W., and Mahlberg, P. G. (2004). A chemotaxonomic analysis of cannabinoid variation in *Cannabis* (*Cannabaceae*). *Am. J. Bot.* 91, 966–975. doi: 10.3732/ajb.91.6.966

- Höppner, F. K., F (2006). "Status report on the German collections of flax and hemp," in *First Meeting of the Working Group on Fibre Crops* (Wageningen: The European Cooperative Programme for Plant Genetic Resources).
- House, J. D., Neufeld, J., and Leson, G. (2010). Evaluating the quality of protein from hemp seed (*Cannabis sativa* L.) products through the use of the protein digestibility-corrected amino acid score method. *J. Agric. Food Chem.* 58, 11801–11807. doi: 10.1021/jf102636b
- Hu, Z. G., Guo, H. Y., Hu, X. L., Chen, X., Liu, X. Y., Guo, M. B., et al. (2012). Genetic diversity research of hemp (*Cannabis sativa* L.) cultivar based on AFLP analysis. *J. Plant Genet. Resour.* 13, 555–561.
- HuaDong, W., and YingFang, W. (2012). Survey on the germplasm resources of *Cannabis sativa* L. *Med. Plant* 3, 11–14.
- Izzo, A. A., Borrelli, F., Capasso, R., Di Marzo, V., and Mechoulam, R. (2009). Non-psychoactive plant cannabinoids: new therapeutic opportunities from an ancient herb. *Trends Pharmacol. Sci.* 30, 515–527. doi: 10.1016/j.tips.2009.07.006
- Jansky, S. H., Dawson, J., and Spooner, D. M. (2015). How do we address the disconnect between genetic and morphological diversity in germplasm collections? *Am. J. Bot.* 102, 1213–1215. doi: 10.3732/ajb.1500203
- Jones, H., Gosman, N., Horsnell, R., Rose, G., Everest, L., Bentley, A., et al. (2013). Strategy for exploiting exotic germplasm using genetic, morphological, and environmental diversity: the *Aegilops tauschii* Coss. example. *Theor. Appl. Genet.* 126, 1793–1808. doi: 10.1007/s00122-013-2093-x
- Kadam, S., Vuong, T. D., Qiu, D., Meinhardt, C. G., Song, L., Deshmukh, R., et al. (2016). Genomic-assisted phylogenetic analysis and marker development for next generation soybean cyst nematode resistance breeding. *Plant Sci.* 242, 342–350. doi: 10.1016/j.plantsci.2015.08.015
- Kaminski, K. P., Kørup, K., Andersen, M. N., Sønderkær, M., Andersen, M. S., Kirk, H. G., et al. (2016). Next generation sequencing bulk segregant analysis of potato support that differential flux into the cholesterol and stigmasteryl metabolite pools is important for steroidal glycoalkaloid content. *Potato Res.* 59, 81–97. doi: 10.1007/s11540-015-9314-4
- Khoury, C., Laliberté, B., and Guarino, L. (2010). Trends in *ex situ* conservation of plant genetic resources: a review of global crop and regional conservation strategies. *Genet. Resour. Crop Evol.* 57, 625–639. doi: 10.1007/s10722-010-9534-z
- Kilian, B., and Graner, A. (2012). NGS technologies for analyzing germplasm diversity in genebanks. *Brief. Funct. Genomics* 11, 38–50. doi: 10.1093/bfgp/eln046
- Kim, C., Guo, H., Kong, W., Chandnani, R., Shuang, L.-S., and Paterson, A. H. (2016). Application of genotyping by sequencing technology to a variety of crop breeding programs. *Plant Sci.* 242, 14–22. doi: 10.1016/j.plantsci.2015.04.016
- Kim, J., and Buell, C. R. (2015). A revolution in plant metabolism: genome-enabled pathway discovery. *Plant Physiol.* 169, 1532–1539. doi: 10.1104/pp.15.00976
- Kriese, U., Schumann, E., Weber, W. E., Beyer, M., Brühl, L., and Matthäus, B. (2004). Oil content, tocopherol composition and fatty acid patterns of the seeds of 51 *Cannabis sativa* L. genotypes. *Euphytica* 137, 339–351. doi: 10.1023/B:EUPH.0000040473.23941.76
- Kuglarz, M., Alvarado-Morales, M., Karakashev, D., and Angelidaki, I. (2016). Integrated production of cellulose bioethanol and succinic acid from industrial hemp in a biorefinery concept. *Bioresour. Technol.* 200, 639–647. doi: 10.1016/j.biortech.2015.10.081
- Lange, B. M. (2016). Online resources for gene discovery and biochemical research with aromatic and medicinal plants. *Phytochem. Rev.* 15, 1–22. doi: 10.1007/s11101-015-9450-0
- Lanz, C., Mattsson, J., Soydaner, U., and Brenneisen, R. (2016). Medicinal *Cannabis*: *in vitro* validation of vaporizers for the smoke-free inhalation of *Cannabis*. *PLoS ONE* 11:e0147286. doi: 10.1371/journal.pone.0147286
- Lawrence, M., Marshall, D., and Davies, P. (1995a). Genetics of genetic conservation. I. sample size when collecting germplasm. *Euphytica* 84, 89–99. doi: 10.1007/BF01677945
- Lawrence, M., Marshall, D., and Davies, P. (1995b). Genetics of genetic conservation. II. sample size when collecting seed of cross-pollinating species and the information that can be obtained from the evaluation of material held in gene banks. *Euphytica* 84, 101–107. doi: 10.1007/BF01677946
- Li, D. Z. (2008). Floristics and plant biogeography in China. *J. Integr. Plant Biol.* 50, 771–777. doi: 10.1111/j.1744-7909.2008.00711.x
- Li, D.-Z., and Pritchard, H. W. (2009). The science and economics of *ex situ* plant conservation. *Trends Plant Sci.* 14, 614–621. doi: 10.1016/j.tplants.2009.09.005
- Li, H.-L. (1974). The origin and use of cannabis in eastern asia linguistic-cultural implications. *Econ. Bot.* 28, 293–301. doi: 10.1007/BF02861426
- Li, J., Zhang, Y., Yu, Z., Wang, Y., Yang, Y., Liu, Z., et al. (2007). Superior storage stability in low lipoxygenase maize varieties. *J. Stored Prod. Res.* 43, 530–534. doi: 10.1016/j.jspr.2006.09.005
- Li, S.-Y., Stuart, J. D., Li, Y., and Parnas, R. S. (2010). The feasibility of converting *Cannabis sativa* L. oil into biodiesel. *Bioresour. Technol.* 101, 8457–8460. doi: 10.1016/j.biortech.2010.05.064
- López, R. M., Maireles-Torres, P., Ojeda, M., Sadaba, I., and Granados, M. L. (2016). Furfural: a renewable and versatile platform molecule for the synthesis of chemicals and fuels. *Energy Environ. Sci.* 9, 1144–1189. doi: 10.1039/C5EE02666K
- Love, C. G., Andongabo, A. E., Wang, J., Carion, P. W., Rawlings, C. J., and King, G. J. (2012). InterStoreDB: a generic integration resource for genetic and genomic data^F. *J. Integr. Plant Biol.* 54, 345–355. doi: 10.1111/j.1744-7909.2012.01120.x
- Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* 24, 31–38. doi: 10.1016/j.pbi.2015.01.006
- Lynch, R. C., Vergara, D., Tittes, S., White, K., Schwartz, C., Gibbs, M. J., et al. (2015). Genomic and chemical diversity in *Cannabis*. *BioRxiv*. doi: 10.1101/034314
- Mace, E., Hunt, C., and Jordan, D. (2013). Supermodels: sorghum and maize provide mutual insight into the genetics of flowering time. *Theor. Appl. Genet.* 126, 1377–1395. doi: 10.1007/s00122-013-2059-z
- Malomo, S. A., and Aluko, R. E. (2015). Conversion of a low protein hemp seed meal into a functional protein concentrate through enzymatic digestion of fibre coupled with membrane ultrafiltration. *Innov. Food Sci. Emerg. Technol.* 31, 151–159. doi: 10.1016/j.ifset.2015.08.004
- Mandolino, G., and Carboni, A. (2004). Potential of marker-assisted selection in hemp genetic improvement. *Euphytica* 140, 107–120. doi: 10.1007/s10681-004-4759-6
- Mandolino, G., Carboni, A., Forapani, S., Faeti, V., and Ranalli, P. (1999). Identification of DNA markers linked to the male sex in dioecious hemp (*Cannabis sativa* L.). *Theor. Appl. Genet.* 98, 86–92. doi: 10.1007/s001220051043
- Mandolino, M. C. A., Fontana, F., Maestrini, C., and Fila, G. (2006). "Hemp and flax germplasm collections at CRA-ISCI," in *First Meeting of the Working Group on Fibre Crops* (Wageningen: The European Cooperative Programme for Plant Genetic Resources).
- Mankowska, G., and Silska, G. (2015). Genetic resources of *Cannabis sativa* L. in the collection of the gene bank at INF&MP in Poznan. *J. Nat. Fibers* 12, 332–340. doi: 10.1080/15440478.2014.928246
- Matthäus, B., Schumann, E., Brühl, L., and Kriese, U. (2006). Hempseed oil— influence of the genotype on the composition in a two-year study. *J. Ind. Hemp* 10, 45–65. doi: 10.1300/J237v10n02_05
- May, B., and Hoelzel, A. (1992). "Starch gel electrophoresis of allozymes," in *Molecular genetic analysis of populations: a practical approach*, ed A. Hoelzel (Oxford: Oxford University Press), 1–27.
- McKernan, K. J., Helbert, Y., Tadigotla, V., McLaughlin, S., Spangler, J., Zhang, L., et al. (2015). Single molecule sequencing of THCA synthase reveals copy number variation in modern drug-type *Cannabis sativa* L. *BioRxiv*. doi: 10.1101/028654
- Michalak, M., Plitta, B. P., Tylkowski, T., Chmielarz, P., and Suszka, J. (2015). Desiccation tolerance and cryopreservation of seeds of black poplar (*Populus nigra* L.), a disappearing tree species in Europe. *Eur. J. For. Res.* 134, 53–60. doi: 10.1007/s10342-014-0832-4
- Moray, C., Game, E., and Maxted, N. (2014). Prioritising *in situ* conservation of crop resources: a case study of African cowpea (*Vigna unguiculata*). *Sci. Rep.* 4:5247. doi: 10.1038/srep05247
- Murphy, T. M., Ben-Yehuda, N., Taylor, R. E., and Southon, J. R. (2011). Hemp in ancient rope and fabric from the Christmas Cave in Israel: talmudic background and DNA sequence identification. *J. Archaeol. Sci.* 38, 2579–2588. doi: 10.1016/j.jas.2011.05.004
- Negri, V., and Tiranti, B. (2010). Effectiveness of *in situ* and *ex situ* conservation of crop diversity. *What a Phaseolus vulgaris L. landrace case study can tell us. Genetica* 138, 985–998. doi: 10.1007/s10709-010-9485-5
- Nutt, D. (2015). Illegal drugs laws: clearing a 50-year-old obstacle to research. *PLoS Biol.* 13:e1002047. doi: 10.1371/journal.pbio.1002047

- Odong, T., Jansen, J., Van Eeuwijk, F., and van Hintum, T. J. (2013). Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. *Theor. Appl. Genet.* 126, 289–305. doi: 10.1007/s00122-012-1971-y
- Olsen, M. T., Pampoulie, C., Daniëlsdóttir, A. K., Lidh, E., Bérubé, M., Vikingsson, G. A., et al. (2014). Fin whale MDH-1 and MPI allozyme variation is not reflected in the corresponding DNA sequences. *Ecol. Evol.* 4, 1787–1803. doi: 10.1002/ece3.1046
- Onofri, C., de Meijer, E. P. M., and Mandolino, G. (2015). Sequence heterogeneity of cannabidiolic- and tetrahydrocannabinolic acid-synthase in *Cannabis sativa* L. and its relationship with chemical phenotype. *Phytochemistry* 116, 57–68. doi: 10.1016/j.phytochem.2015.03.006
- Pacifico, D., Miselli, F., Micheler, M., Carboni, A., Ranalli, P., and Mandolino, G. (2006). Genetics and marker-assisted selection of the chemotype in *Cannabis sativa* L. *Mol. Breed.* 17, 257–268. doi: 10.1007/s11032-005-5681-x
- Pain, S. (2015). A potted history. *Nature* 525, S10–S11. doi: 10.1038/525S10a
- Pardo, B. (2014). Cannabis policy reforms in the Americas: a comparative analysis of Colorado, Washington, and Uruguay. *Int. J. Drug Policy* 25, 727–735. doi: 10.1016/j.drugpo.2014.05.010
- Parihar, S., Dadlani, M., Lal, S., Tonapi, V., Nautiyal, P., and Basu, S. (2014). Effect of seed moisture content and storage temperature on seed longevity of hemp (*Cannabis sativa*). *Indian J. Agri. Sci.* 84, 1303–1309.
- Parzies, H., Spoor, W., and Ennos, R. (2000). Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in ex situ gene banks. *Heredity* 84, 476–486. doi: 10.1046/j.1365-2540.2000.00705.x
- Pavelek, M. (2006). “Czech flax (*Linum usitatissimum* L.) and hemp (*Cannabis sativa*) collections – data for country report,” in *First Meeting of the Working Group on Fibre Crops* (Wageningen: The European Cooperative Programme for Plant Genetic Resources).
- Perullo, N., Determann, R., Cruse-Sanders, J., and Pullman, G. (2015). Seed cryopreservation and micropropagation of the critically endangered species swamp pink (*Helonias bullata* L.). *In Vitro Cell. Dev. Biol. Plant* 51, 284–293. doi: 10.1007/s11627-015-9679-8
- Pervaiz, M., and Sain, M. M. (2003). Carbon storage potential in natural fiber composites. *Resour. Cons. Recycling* 39, 325–340. doi: 10.1016/S0921-3449(02)00173-8
- Piluzza, G., Delogu, G., Cabras, A., Marceddu, S., and Bullitta, S. (2013). Differentiation between fiber and drug types of hemp (*Cannabis sativa* L.) from a collection of wild and domesticated accessions. *Genet. Resour. Crop Evol.* 60, 2331–2342. doi: 10.1007/s10722-013-0001-5
- Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12302–12308. doi: 10.1073/pnas.0912953109
- Potter, D. J. (2014). A review of the cultivation and processing of cannabis (*Cannabis sativa* L.) for production of prescription medicines in the UK. *Drug Test. Anal.* 6, 31–38. doi: 10.1002/dta.1531
- Prada, J. A., Aguilar, M. E., Abdelnour-Esquivel, A., and Engelmann, F. (2015). Cryopreservation of seeds and embryos of *Jatropha curcas* L. *Am. J. Plant Sci.* 6, 172–180. doi: 10.4236/ajps.2015.61020
- Ramirez-Villegas, J., Khoury, C., Jarvis, A., Debouck, D. G., and Guarino, L. (2010). A gap analysis methodology for collecting crop gene pools: a case study with *Phaseolus beans*. *PLoS ONE* 5:e13497. doi: 10.1371/journal.pone.0013497
- Ranalli, P. (2004). Current status and future scenarios of hemp breeding. *Euphytica* 140, 121–131. doi: 10.1007/s10681-004-4760-0
- Razumova, O. V., Alexandrov, O. S., Divashuk, M. G., Sukhorada, T. I., and Karlov, G. I. (2015). Molecular cytogenetic analysis of monoecious hemp (*Cannabis sativa* L.) cultivars reveals its karyotype variations and sex chromosomes constitution. *Protoplasma* 253, 895–901 doi: 10.1007/s00709-015-0851-0
- Revilla, P., Butrón, A., Rodríguez, V., Malvar, R., and Ordás, A. (2009). Identification of genes related to germination in aged maize seed by screening natural variability. *J. Exp. Bot.* 60, 4151–4157. doi: 10.1093/jxb/erp249
- Rice, E. B., Smith, M. E., Mitchell, S. E., and Kresovich, S. (2006). Conservation and change: a comparison of *in situ* and *ex situ* conservation of Jala maize germplasm. *Crop Sci.* 46, 428–436. doi: 10.2135/cropsci2005.06-0116
- Rocha, P. S., Sheikh, M., Melchiorre, R., Fagard, M., Boutet, S., Loach, R., et al. (2005). The Arabidopsis homology-dependent gene silencing₁ gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. *Plant Cell Rep.* 17, 404–417. doi: 10.1105/tpc.104.028332
- Russo, E. B. (2007). History of cannabis and its preparations in saga, science, and sobriquet. *Chem. Biodivers.* 4, 1614–1648. doi: 10.1002/cbdv.200790144
- Russo, E. B. (2011). Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. *Br. J. Pharmacol.* 163, 1344–1364. doi: 10.1111/j.1476-5381.2011.01238.x
- Russo, E. B., Jiang, H.-E., Li, X., Sutton, A., Carboni, A., del Bianco, F., et al. (2008). Phytochemical and genetic analyses of ancient cannabis from Central Asia. *J. Exp. Bot.* 59, 4171–4182. doi: 10.1093/jxb/ern260
- Salentijn, E. M. J., Zhang, Q., Amaducci, S., Yang, M., and Trindade, L. M. (2014). New developments in fiber hemp (*Cannabis sativa* L.) breeding. *Ind. Crops Prod.* 68, 32–41. doi: 10.1016/j.indcrop.2014.08.011
- Sawler, J., Stout, J. M., Gardner, K. M., Hudson, D., Vidmar, J., Butler, L., et al. (2015). The genetic structure of marijuana and hemp. *PLoS ONE* 10:e0133292. doi: 10.1371/journal.pone.0133292
- Schatz, M. C., Witkowski, J., and McCombie, W. R. (2012). Current challenges in de novo plant genome sequencing and assembly. *Genome Biol.* 13:243. doi: 10.1186/gb-2012-13-4-243
- Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534
- Schuermeyer, J., Salomonsen-Sautel, S., Price, R. K., Balan, S., Thurstone, C., Min, S.-J., et al. (2014). Temporal trends in marijuana attitudes, availability and use in Colorado compared to non-medical marijuana states: 2003–11. *Drug Alcohol Depend.* 140, 145–155. doi: 10.1016/j.drugalcdep.2014.04.016
- Scossa, F., Brotman, Y., de Abreu e Lima, F., Willmitzer, L., Nikoloski, Z., Tohge, T., et al. (2015). Genomics-based strategies for the use of natural variation in the improvement of crop metabolism. *Plant Sci.* 242, 47–64. doi: 10.1016/j.plantsci.2015.05.021
- Simon, A. (2006). “The National Cannabis Collection in Hungary,” in *First Meeting of the Working Group on Fibre Crops* (Wageningen: The European Cooperative Programme for Plant Genetic Resources).
- Sloan, S., Jenkins, C. N., Joppa, L. N., Gaveau, D. L., and Laurance, W. F. (2014). Remaining natural vegetation in the global biodiversity hotspots. *Biol. Cons.* 177, 12–24. doi: 10.1016/j.biocon.2014.05.027
- Small, E. (2015a). Evolution and classification of *Cannabis sativa* (marijuana, hemp) in relation to human utilization. *Bot. Rev.* 81, 189–294. doi: 10.1007/s12229-015-9157-3
- Small, E. (2015b). Response to the erroneous critique of my *Cannabis* monograph by RC Clarke and MD Merlin. *Bot. Rev.* 81, 306–316. doi: 10.1007/s12229-015-9159-1
- Small, E., and Antle, T. (2003). A preliminary study of pollen dispersal in *Cannabis sativa* in relation to wind direction. *J. Ind. Hemp* 8, 37–50. doi: 10.1300/J237v08n02_03
- Small, E., and Brookes, B. (2012). Temperature and moisture content for storage maintenance of germination capacity of seeds of industrial hemp, marijuana, and ditchweed forms of *Cannabis sativa*. *J. Nat. Fibers* 9, 240–255. doi: 10.1080/15440478.2012.737179
- Small, E., and Cronquist, A. (1976). A practical and natural taxonomy for *Cannabis*. *Taxon* 25, 405–435. doi: 10.2307/1220524
- Small, E., and Marcus, D. (2002). “Hemp: a new crop with new uses for North America,” in *Trends in New Crops and New Uses*, eds J. Janick and A. Whipkey (Alexandria, VA: ASHS Press), 284–326.
- Small, E., and Marcus, D. (2003). Tetrahydrocannabinol levels in hemp (*Cannabis sativa*) germplasm resources. *Econ. Bot.* 57, 545–558. doi: 10.1663/0013-0001(2003)057[0545:TLIHCS]2.0.CO;2
- Small, E., Pocock, T., and Cavers, P. B. (2003). The biology of Canadian weeds. 119. *Cannabis sativa* L. *Can. J. Plant Sci.* 83, 217–237. doi: 10.4141/P02-021
- Smith, L. B., and King, G. J. (2000). The distribution of BoCAL-a alleles in *Brassica oleracea* is consistent with a genetic model for curd development and domestication of the cauliflower. *Mol. Breed.* 6, 603–613. doi: 10.1023/A:1011370525688
- Staginnus, C., Zörntlein, S., and de Meijer, E. (2014). A PCR marker linked to a THCA synthase polymorphism is a reliable tool to discriminate potentially THC-rich plants of *Cannabis sativa* L. *J. Forensic Sci.* 59, 919–926. doi: 10.1111/1556-4029.12448

- Stanger, R. (2015). *Characterisation of a Global Cannabis sativa Germplasm Collection: Quantifying Phenotypic Variation*. Honours dissertation, Southern Cross University, Lismore, NSW.
- Struik, P. C., Amaducci, S., Bullard, M. J., Stutterheim, N. C., Venturi, G., and Cromack, H. T. H. (2000). Agronomy of fibre hemp (*Cannabis sativa* L.) in Europe. *Ind. Crops Prod.* 11, 107–118. doi: 10.1016/S0926-6690(99)00048-5
- Swift, W., Wong, A., Li, K. M., Arnold, J. C., and McGregor, I. S. (2013). Analysis of cannabis seizures in NSW, Australia: cannabis potency and cannabinoid profile. *PLoS ONE* 8:e70052. doi: 10.1371/journal.pone.0070052
- Tang, Z., Chen, X., Zhang, Y., Guo, H., and Yang, M. (2013). Genetic Diversity analysis of wild Cannabis in China based on morphological characters and RAPD markers. *J. West China For. Sci.* 42, 61–66.
- Thachuk, C., Crossa, J., Franco, J., Dreisigacker, S., Warburton, M., and Davenport, G. F. (2009). Core hunter: an algorithm for sampling genetic resources based on multiple genetic measures. *BMC Bioinform.* 10:243. doi: 10.1186/1471-2105-10-243
- Tipparat, P., Natakankitkul, S., Chamnivikaipong, P., and Chutiwat, S. (2012). Characteristics of cannabinoids composition of *Cannabis* plants grown in Northern Thailand and its forensic application. *Forensic Sci. Int.* 215, 164–170. doi: 10.1016/j.forsciint.2011.05.006
- Tyagi, R. K., and Agrawal, A. (2015). Revised genebank standards for management of plant genetic resources. *Indian J. Agri. Sci.* 85, 157–165.
- Van Bakel, H., Stout, J., Cote, A., Tallon, C., Sharpe, A., Hughes, T., et al. (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12:R102. doi: 10.1186/gb-2011-12-10-r102
- Van Hintum, T. J. L., and Visser, D. L. (1995). Duplication within and between germplasm collections. I. Identifying duplication on the basis of passport data. *Genet. Resour. Crop Evol.* 42, 135–145. doi: 10.1007/BF02539517
- Varshney, R. K., Kudapa, H., Pazhamala, L., Chitikineni, A., Thudi, M., Bohra, A., et al. (2015). Translational genomics in agriculture: some examples in grain legumes. *Crit. Rev. Plant Sci.* 34, 169–194. doi: 10.1080/07352689.2014.897909
- Vincent, H., Wiersema, J., Kell, S., Fielder, H., Dobbie, S., Castañeda-Álvarez, N. P., et al. (2013). A prioritized crop wild relative inventory to help underpin global food security. *Biol. Conserv.* 167, 265–275. doi: 10.1016/j.biocon.2013.08.011
- Walters, C., Wheeler, L. M., and Grotenhuis, J. M. (2005). Longevity of seeds stored in a genebank: species characteristics. *Seed Sci. Res.* 15, 1–20. doi: 10.1079/SSR2004195
- Wanas, A. S., Radwan, M. M., Mehmedic, Z., Jacob, M., Khan, I. A., and Elsohly, M. A. (2016). Antifungal activity of the volatiles of high potency *Cannabis sativa* L. against *Cryptococcus neoformans*. *Rec. Nat. Prod.* 10, 214–220.
- Wang, H., Xu, Z., Kohandehghan, A., Li, Z., Cui, K., Tan, X., et al. (2013). Interconnected carbon nanosheets derived from hemp for ultrafast supercapacitors with high energy. *ACS Nano* 7, 5131–5141. doi: 10.1021/nn400731g
- Weiblen, G. D., Wenger, J. P., Craft, K. J., ElSohly, M. A., Mehmedic, Z., Treiber, E. L., et al. (2015). Gene duplication and divergence affecting drug content in *Cannabis sativa*. *New Phytol.* 208, 1241–1250. doi: 10.1111/nph.13562
- Welling, M. T., Liu, L., Shapter, T., Raymond, C. A., and King, G. J. (2015). Characterisation of cannabinoid composition in a diverse *Cannabis sativa* L. germplasm collection. *Euphytica* 208, 463–475. doi: 10.1007/s10681-015-1585-y
- Werz, O., Seegers, J., Schaible, A. M., Weinigel, C., Barz, D., Koeberle, A., et al. (2014). Cannflavins from hemp sprouts, a novel cannabinoid-free hemp food product, target microsomal prostaglandin E2 synthase-1 and 5-lipoxygenase. *PharmaNutrition* 2, 53–60. doi: 10.1016/j.phanu.2014.05.001
- Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat. Biotechnol.* 32, 656–662. doi: 10.1038/nbt.2906
- Yang, H., Tao, Y., Zheng, Z., Li, C., Sweetingham, M. W., and Howieson, J. G. (2012). Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius* L. *BMC Genomics* 13:318. doi: 10.1186/1471-2164-13-318
- Zhang, L., Chang, Y., Zhang, X., Guan, F., Yuan, H., Yu, Y., et al. (2014). Analysis of the genetic diversity of Chinese native *Cannabis sativa* cultivars by using ISSR and chromosome markers. *Genet. Mol. Res.* 13, 10490–10500. doi: 10.4238/2014.december.12.10
- Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 408–414. doi: 10.1038/nbt.3096

Conflict of Interest Statement: 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.

Copyright © 2016 Welling, Shapter, Rose, Liu, Stanger and King. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.