



OPEN ACCESS

EDITED BY

Fakhriddin Kushanov,
Academy of Sciences Republic of Uzbekistan
(UzAS), Uzbekistan

REVIEWED BY

Shibo Wang,
University of California, Riverside, United States
Yingpeng Han,
Northeast Agricultural University, China

*CORRESPONDENCE

Anders S. Carlsson,
✉ anders.carlsson@slu.se

RECEIVED 20 May 2024

ACCEPTED 14 August 2024

PUBLISHED 28 August 2024

CITATION

Alemu A, Sundaramoorthy J, Abreha KB,
Enyew M, Geleta M and Carlsson AS (2024)
Developing genomic tools to assist turnip rape
[*Brassica rapa* (L.) subsp. *oleifera* (DC.)
Metzg.] breeding.
Front. Genet. 15:1435474.
doi: 10.3389/fgene.2024.1435474

COPYRIGHT

© 2024 Alemu, Sundaramoorthy, Abreha,
Enyew, Geleta and Carlsson. This is an open-
access article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) 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.

Developing genomic tools to assist turnip rape [*Brassica rapa* (L.) subsp. *oleifera* (DC.) Metzg.] breeding

Admas Alemu¹, Jagadeesh Sundaramoorthy¹, Kibrom B. Abreha¹,
Muluken Enyew^{1,2,3}, Mulatu Geleta¹ and Anders S. Carlsson^{1*}

¹Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden, ²School of Biological Sciences, Washington State University, Pullman, WA, United States, ³Institute of Biotechnology, Addis Ababa University, Addis Ababa, Ethiopia

Introduction: Turnip rape is recognized as an oilseed crop contributing to environmentally sustainable agriculture via integration into crop rotation systems. Despite its various advantages, the crop's cultivation has declined globally due to a relatively low productivity, giving way to other crops. The use of genomic tools could enhance the breeding process and accelerate genetic gains. Therefore, the present research investigated 170 turnip rape accessions representing its global gene pool to identify SNP markers associated nine phenological and agro-morphological traits and estimate the genomic breeding values (GEBVs) of the germplasm through GWAS and genomic prediction analyses, respectively.

Methods: Field trials were conducted at two sites in northern and southern Sweden to obtain the phenotypic data while genotyping was conducted via the genotyping-by-sequencing (GBS) method. The traits studied include days to flowering (DTF) and maturity (DTM), plant height (PH), seed yield (YLD), thousand seed weight (TSW), silique length (SL), number of siliques (NS), number of seeds per silique (SS), and pod shattering resistance (PSHR).

Results and conclusion: Analysis of variance revealed substantial variation among accessions, with significant genotype-by-environment interaction for most traits. A total of 25, 17, 16, 14, 7, 5, 3, and 3 MTAs were identified for TSW, DTF, PH, PSHR, SL, YLD, SS and DTM, respectively. An 80%–20% training-test set genomic prediction analysis was conducted using the ridge regression – BLUP (RR-BLUP) model. The accuracy of genomic prediction for most traits was high, indicating that these tools may assist turnip rape breeders in accelerating genetic gains. The study highlights the potential of genomic tools to significantly advance breeding programs for turnip rape by identifying pivotal SNP markers and effectively estimating genomic breeding values. Future breeding perspectives should focus on leveraging these genomic insights to enhance agronomic traits and productivity, thereby reinstating turnip rape as a competitive and sustainable crop in Sweden and broader global agriculture.

KEYWORDS

turnip rape, agro-morphological traits, genomic prediction, GWAS, QTL, SNPs

Introduction

Brassica rapa L. (syn. *B. campestris*) (AA, $2n = 20$) is the most widely distributed and cultivated member of the genus *Brassica* within the family *Brassicaceae* (Gómez-Campo and Prakash, 1999; Kaur et al., 2021a; Wu et al., 2022). It is one of the earliest domesticated *Brassica* species and wild progenitors of the allotetraploids *Brassica juncea* (AABB, $2n = 36$) and *Brassica napus* (AACC, $2n = 38$) (Gómez-Campo and Prakash, 1999; Chalhoub et al., 2014; Wu et al., 2022). It comprises several subspecies and morphotypes cultivated as leafy vegetables (Broccolotto – *B. rapa* ssp. *oleifera*; Chinese cabbage – ssp. *pekinensis* and Bok choy – ssp. *chinensis*); root vegetable (Turnips – *B. rapa* ssp. *rapa*); oilseed crops (turnip rape – *B. rapa* ssp. *oleifera* and brown sarson – ssp. *dichotoma*) and fodder crops (Fodder turnip – *B. rapa* ssp. *rapifera*) (Gómez-Campo and Prakash, 1999; Shyam et al., 2012; Cartea et al., 2021; Wu et al., 2022). Genomic inferences and anecdotal pieces of evidence suggest the origin of *B. rapa* as European-Central Asian (Qi et al., 2017; Kaur et al., 2021a). Previous genetic and phylogenetic analyses grouped *B. rapa* into five subgroups reflecting the morphology and geographical distributions. These are i) Asian turnip and oilseed cultivars; ii) European turnip and oilseed cultivars; iii) yellow and brown sarson; iv) Chinese cabbage; and v) Pakchoi, choi sum and tatsoi (Bird et al., 2017; Vélez-Gavilán, 2018). Factors including biased gene retention, genetic redundancy and functional diversification are believed to be the major contributors to the widespread of the various morphotypes and subspecies (Lysak et al., 2005; Tang and Lyons, 2012; Kaur et al., 2021a).

Brassica rapa has been cultivated in several parts of the world including Asia and Europe, creating genetically highly diverse subspecies, morphotypes and cultivars. For instance, some leafy vegetables, such as Chinese cabbage (ssp. *pekinensis*), have been believed to be evolved naturally in Central China from Pak choi (ssp. *chinensis*) and turnip rape (ssp. *oleifera*) followed by artificial selection for adaptation to different environments and different purposes (Lee, 1982). Around the 14th century, the huge demand increment for vegetable oil led to the introduction of turnip rape *B. rapa* ssp. *oleifera* in temperate regions of European countries (Reiner et al., 1995; Guo et al., 2014).

Turnip rape (*B. rapa* ssp. *oleifera*) is the third most important Brassica oilseed crop next to *B. napus* and *B. juncea* widely grown in China, Canada, India, and northern Europe (Wu et al., 2022). It has been identified as a rich source of human nutrition as it contains up to 46% oil, 38% proteins and more than 20% carbohydrates in the seeds on a dry weight basis (Bassegio and Zanotto, 2020; Kaur et al., 2021a; Wanasundara et al., 2024). However, turnip rape is currently cultivated in a very small area globally compared to its allotetraploid relatives *B. napus* and *B. juncea* (Kaur et al., 2021a).

Forage and grain fodder crops have been predominantly incorporated into Northern European agriculture via crop rotation. They have shown several benefits by contributing to sustainable and environment-friendly agriculture. Spring turnip rape is a well-adapted and environmentally sustainable oil crop in crop rotations. Spring turnip rape cultivars are widely known because they mature earlier

than oilseed rape cultivars making them ideal for temperate countries in the northern hemisphere. In Sweden, turnip rape has been identified as an environmentally sustainable oilseed crop with the potential to be incorporated into crop rotation systems, primarily in northern Sweden, but it is also a viable alternative for cultivation in the Svealand (Central Sweden) parts of (<https://www.slu.se/en/Collaborative-Centres-and-Projects/grogrund/projekt/rybs—en-flexibel-och-talig-oljegröda-for-sverige/>).

Turnip rape was one of the largest oil crops in terms of area coverage in Sweden for a few years in the 1980s, predominantly grown in the temperate latitudes of the northern hemisphere from 55°N to 65°N of the country. However, in the late 1990s, spring turnip rape cultivation declined, mainly due to the lack of investment in improvement programs leading to the termination of spring turnip rape breeding programs in the country (Persson, 2017). This led to a lack of high-yielding improved varieties with the ability to withstand recurring biotic and abiotic stresses, resulting in a significant reduction in the crop's cultivation area, with only 13,000 ha at present (Jordbruksverket, 2022).

Expanded accessibility along with the cost-effectiveness of high throughput genomic data of various crops revolutionized plant breeding shifting from phenotypic to genomic-based selection. Shortly after the genotyping revolution (starting in the 1980s), a large number of linkage and association mapping studies were conducted in various crops to identify quantitative trait loci (QTL) controlling different traits of interest (Bernardo, 2008). Genome-wide association studies (GWAS) have been extensively used in several Brassicaceae crops including *B. rapa* to identify QTLs or quantitative trait nucleotides (QTNs) using a collection of different historical germplasms of a particular crop species with highly accumulated historical recombination events (Yano et al., 2016; Challa and Neelapu, 2018; Alemu et al., 2021b; Tibbs Cortes et al., 2021; Alemu et al., 2022). The complex nature of QTLs identified via GWAS from diverse germplasm sets makes the approach challenging for direct and immediate genomic-enabled breeding applications. Genome-wide prediction/selection has emerged as a powerful genomic-assisted breeding method by alleviating the limitations of methods that utilize GWAS and biparental QTL mapping-derived markers. Genomic prediction (GP) estimates individual genotypes' breeding values based on overall genetic merits. Hence, GP predicts breeding values [also known as genomic estimated breeding values (GEBVs)] of candidate genotypes instead of identifying QTLs linked to a particular trait. Compared to conventional phenotypic-based selection, genomic prediction-based selection reduces both the cost and required time per breeding cycle and enhances testing efficiency in field evaluations conducted to develop a crop variety in several breeding programs (Alemu et al., 2024).

The objectives of this study were to identify SNP markers associated with nine phenological and agro-morphological traits of turnip rape via GWAS and estimate the breeding values of turnip rape collections in terms of these traits via genomic prediction. To accomplish this 170 turnip rape accessions representing a worldwide collection were assembled and characterized for the nine traits by growing them at two field experimental sites. Genotyping-by-sequencing (GBS) method was used for genotyping accessions to generate SNP markers.

Materials and methods

Plant material

The germplasm panel used in this study comprises 170 turnip rape accessions that include cultivars, breeding populations, landraces, and wild populations originated from at least 26 countries across the globe. Fifty of the 170 accessions were of Swedish origin while the vast majority of the other accessions originated from Italy (22), India (17), Canada (12), Germany (12), China (10), Finland (8), Bangladesh (7), United States (4), Georgia (3), Pakistan (3), Austria (2), Colombia (2), Denmark (2) and Slovakia (2). Australia, Cuba, Iran, Nepal, Poland, Portugal, the Russian Federation, Spain, Tibet, Turkey and United Kingdom were represented by a single accession each. The countries of origin of the remaining three accessions are unknown (Supplementary Table S1). These accessions were procured from several gene banks including the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK – 73 accessions), Plant Gene Resources of Canada (PGRC – 36 accessions), Nordic Genetic Resource Center (NordGen – 33 accessions), Centre for Genetic Resources, the Netherlands (CGN – 8 accessions), the National Plant Germplasm System, United States (NPGS – 5 accessions) and A&A Seed Farms, Canada (A&A Canada – 2 accessions). In addition, Jerrestad Agro AB, Sweden (a turnip rape breeding company) provided 13 turnip rape cultivars (Supplementary Table S1).

Experimental design, phenotypic data collection and phenotypic plasticity analysis

Field trials were conducted in 2021 using 170 accessions at two locations in Sweden: SITES, Lönnstorp (5539'N, 1306'E) and Öjebyn Agro Park (6521'N, 2124'E), representing different agro-ecological zones. The field experiment was laid out in an alpha lattice design comprising three replications and 17 blocks with 10 plots each. The plot size was 3 m² (3 m × 1 m) and seeds were planted in six rows in each plot. Subsequently, thinning was done at the seedling stage to maintain a spacing of 10 cm between plants in each plot. In total, phenotypic data of nine different phenotypic traits of turnip rape were collected from each plot of one or both experimental sites. The data for days to flowering (DTF), plant height (PH), number of siliques (NS), silique length (SL), pod shattering resistance (PSHR), and thousand seed weight (TSW) were collected from both (Öjebyn and Lönnstorp) experimental sites. The phenotypic data for days to maturity (DTM) and seed yield per plot (YLD) were collected only from Lönnstorp while the data for seeds per silique (SS) was collected only from Öjebyn. DTF was recorded for each plot as the number of days taken from sowing to the flowering of half of the plants on the plot. DTM was measured as the number of days that have elapsed from sowing to physiological maturity of plants on the plot as indicated by the yellowing of the siliques. PH was measured as the distance from the base of the aboveground plant to the tip of the main inflorescence at maturity. The PSHR was visually evaluated and recorded as low, intermediate and high at maturity. A total of 1,000 seeds/plot was cleaned, counted with a seed counter and TSW determined. The number of siliques per plant (NS) on each plot was determined based on five representative plants from the middle of the plot at maturity.

The number of seeds/silique (SS) and silique length (SL) were counted and measured, respectively, based on 10 representative siliques sampled from the middle part of each plot. Seed yield per plot (YLD) was measured after machine-harvesting the whole plot and cleaning.

Phenotypic data analysis from single and combined environment were evaluated using the Multi-environment Trial Analysis with R (META-R) (Alvarado et al., 2020). The genotypic and environmental variances and adjusted phenotypic mean values [best linear unbiased predictions (BLUPs)] of the studied traits was estimated including the genotypes, locations, blocks and replications. The variances of a single environment and BLUPs were estimated using a linear model:

$$Y_{ijk} = \mu + R_i + B_j(R_i) + G_k + \varepsilon_{ijk}$$

where Y_{ijk} is the target trait phenotypic value of the K_i th genotype in the R_i th replication and $B_j(R_i)$ th block; μ is the intercept; R_i is the i th replicate effect; $B_j(R_i)$ is the effect of j th block nested in i th replicate; G_k is k th genotype effect; and ε_{ijk} is the error associated with the i th replicate, j th block and k th genotype with the assumption of independent and identically distributed normal random variables (iid) with mean zero and variance $\delta^2\varepsilon$.

For the combined environment, ANOVA and BLUP analyses were conducted as follows:

$$Y_{ijkl} = \mu + E_i + R_j(E_i) + B_k(E_iR_i) + G_l + E_i \times G_l + \varepsilon_{ijkl}$$

where E_i is the E_i th environment effect and $E_i \times G_l$ is the effect of genotype-by-environment interaction (G×E).

Trait repeatability (H) was analyzed using the phenotypic values scored across replicates within a location as:

$$H = \frac{\delta^2g}{\delta^2g + \delta^2E/nReps}$$

where δ^2g is the genotypic variance; δ^2E is the estimated error and $nReps$ is the number of replications.

The broad-sense heritability (H^2) of the traits was calculated as:

$$H^2 = \frac{\delta^2g}{\delta^2g + \delta^2ge/nEnv + \delta^2E/(nEnv \times nReps)}$$

where δ^2ge is the G×E variance; $nEnv$ is the number of environments.

The Pearson correlation coefficient analysis conducted in the *metan* package in R (Olivoto et al., 2020).

The Bayesian Finlay-Wilkinson regression model (FWR) was employed to estimate accessions phenotypic plasticity following the two stage ordinary least squares (OLS) method in the *FW* package (Lian and de Los Campos, 2016) with a model:

$$Y_{ij} = \mu + g_i + (b_i + 1)h_j + \varepsilon_{ij}$$

where Y_{ij} is the phenotype of the i th accession measured in the j th environment, μ represents the intercept as the population mean, g_i is the main genetic effect of the i th accession, h_j is the environmental effect of the j th environment, $b_i + 1$ indicates the change in performance of the i th accession per unit change in the environmental effect (h_j), and ε_{ij} is the residual error. Hence, FWR decomposes the phenotype of individual accessions into three components: genotypic main effect, a regression slope, and

residual value. The regression slope, also called linear plasticity, measures the linear response of a genotype to environmental changes relative to other genotypes in the population. The linear plasticity score of accessions were used to identify SNP markers and genomic regions controlling phenotypic plasticity of the six traits measured in both environments.

DNA extraction, sequencing, read alignment and variant discovery

The seeds of 165 accessions were planted in a greenhouse at the Department of Plant Breeding, Swedish University of Agricultural Sciences (SLU), Sweden. Fifty-cell (5 × 10) plastic seedling trays filled with soil were used for planting. For each accession, leaf tissue was sampled from 100 two-week-old seedlings and pooled into a single Eppendorf tube for DNA extraction. Each seedling was represented by a 6 mm disc sampled using a Unicore Punch Kit (Qiagen). Genomic DNA was extracted from each pooled sample of each accession using the Thermo Scientific GeneJET Genomic DNA Purification Kit. Following extraction and purification, a high-quality genomic DNA sample of each accession was transferred to a single well of a 96-well plate. Subsequently, sealed 96-well plates containing the samples were dispatched to LGC Genomics GmbH (Berlin, Germany) for genotyping.

SNP genotyping was employed using the genotyping-by-sequencing (GBS) method at the LGC Genomics GmbH following the method by Elshire et al. (2011) with minor modifications. Briefly, pooled DNA samples of individual accessions were digested with MseI (5'-CAYNN*NNRTG-3', a 5 bp cutter) restriction enzyme based on the recommendation of the experts of the LGC, Biosearch Technologies. An average insert size of 215 bp was produced by the restriction enzyme, suitable for sequencing on Illumina platforms. The restriction fragments were ligated to unique sequence barcode adapters. Equal aliquots of adapter-ligated DNA samples were pooled in a single tube to produce sequencing libraries. The pooled DNA was amplified with the sequencing primer and purified using a QIAGEN PCR purification kit. The purified DNA was quantified and sequenced with 2 × 150 bp mode on the Illumina NextSeq 500/550 v2 and NovaSeq FC NGS platforms.

The in-house Illumina bcl2fastq v2 software package was used for demultiplexing of samples based on the barcode adapters ligated to individual DNA samples followed by verification of restriction site. This was followed by adapter-clipping, discarding of reads containing Ns and whose 5'-ends did not match the restriction enzyme site, trimming the 3'-end of the remaining reads to obtain an average Phred quality score of ≥20 over a window of ten bases, and discarding reads with a final length <20 bp. *B. rapa* reference genome assembly CAAS_Brap_v3.01 (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000309985.2/; 352.8 Mb) was used as a reference for sequence analysis. The quality trimmed reads were aligned against the reference genome using the BWA-MEM software package version 0.7.12 (Li, 2013). A 95.3% mapping rate was obtained between the GBS reads and the reference genome. Then, FreeBayes v1.0.2-16 (Garrison and Marth, 2012) was used for variant discovery and allele count. The alignment of the reads of the 165 turnip rape accessions resulted in the discovery of

653,362 SNPs. Variants were filtered using a GBS-specific rule set. The total number of fully covered SNPs in 91% of the samples with a minimum read count of 100 was 24,856. Further filtering of these SNPs to obtain bi-allelic SNPs with no missing data, 5% minor allele frequency (MAF), and a maximum of 20% samples having both alleles resulted in 2,327 SNPs, which were used for marker-trait association analysis and genomic prediction.

Then, barcode sequences were removed clipped from the read sequences followed by trimming of the reads using the in-house software Trimmomatic-0.33. The SNP polymorphism discovery of filtered reads was made by mapping to the *B. rapa* reference genome assembly (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000309985.2/). The filtering of variants and the minimum read depth was applied according to the GBS-specific ruleset.

Genome-wide association study and linkage disequilibrium analysis

The overall design of the current research is illustrated in Figure 1. Adjusted phenotypic mean values (BLUPs) of the studied traits collected from the 170 turnip rape accessions along with their SNP marker profile were employed for GWAS analysis. The analysis was conducted using Fixed and random model Circulating Probability Unification (FarmCPU) (Liu X. et al., 2016). The model was implemented within the package GAPIT v.3 (Wang and Zhang, 2021) in R environment v.4.3.2 (R Core Team, 2023) using the default “method.bin” parameter as “optimum.” FarmCPU is one of the most robust GWAS models, controlling confounding factors and false positive marker-trait associations (Liu X. et al., 2016). The model applies a genomic kinship matrix (K) developed from pseudo-quantitative trait nucleotides (QTNs) as random effects. The default “method.bin” parameter in FarmCPU applies two functions: i) *bin.size* = *c* (5e5, 5e6, 5e7) used for segmentation of the whole genome into bins with particular kilo base pairs representing the window size to select possible quantitative trait nucleotides (QTNs); and ii) *bin.selection* = *seq* (10, 100, 10) to specify the number of QTNs that could be selected and used in the FarmCPU model as covariates in each loop (Liu X. et al., 2016). The model implements two iterative steps. The first step is the fixed effect model (FEM) while the second step is the random effect model (REM) (Liu X. et al., 2016). In the current GWAS study, the model can be interpreted as follows:

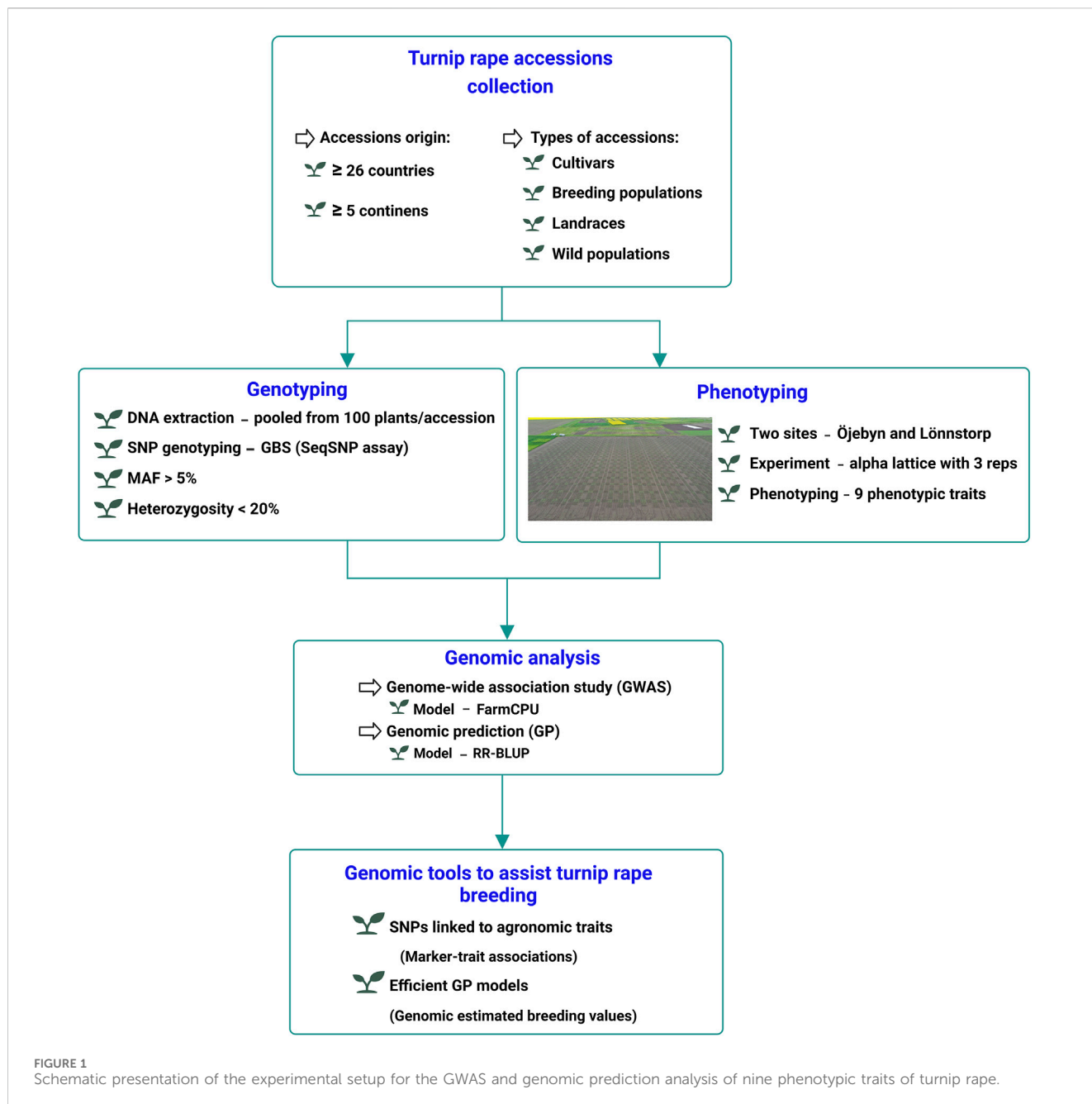
Step 1 (FEM):

$$y_a = M_{a1}Q_1 + M_{a2}Q_2 + \dots + M_{an}Q_n + M_{at}Q_t + S_{at}e_t + \varepsilon_a$$

Step 2 (REM):

$$y_a = J_a + \varepsilon_a$$

where y_a is the BLUPs of the target trait's phenotypic values recorded for the a^{th} accession and ε_a is the residual. In the first step (FEM), $M_{a1}, M_{a2}, \dots, M_{an}$ are genotypes of n pseudo-QTNs; Q_1, Q_2, \dots, Q_n represent the corresponding effects from pseudo-QTNs; S_{at} is a genotype score of a^{th} accession at the t^{th} SNP marker and e_t is the corresponding effect of the t^{th} SNP marker. In the second step (REM), J_a is the total genetic effect of the a^{th} accession



derived from the variance-covariance matrix as $G = 2K\delta_a^2$, where K is the kinship derived from the pseudo-QTNs and δ_a^2 is SNPs genetic variance estimated in the REM step of the model (Liu X. et al., 2016).

The exploratory and Bonferroni-corrected thresholds were applied to identify SNP markers significantly associated with the studied traits. Here, the most commonly applied exploratory threshold with a P -value ≤ 0.0001 or $[-\log_{10}(P - value) \geq 4]$ was applied to identify nominal marker-trait associations. In addition, the Bonferroni threshold adjusted for multiple marker tests at $P \leq 0.05$ was applied to exclude potential false positive marker-trait associations. It was estimated as Bonferroni threshold = $-\log_{10}(\frac{t}{n})$ where t is the overall false positive threshold (0.05) and n is the number of SNP markers used in the current GWAS analysis (2,328). Therefore, the Bonferroni threshold was calculated as $-\log_{10}(\frac{0.05}{2328}) = 4.67$.

The pairwise linkage disequilibrium (LD) between SNP markers was calculated using r^2 values in TASSEL (Bradbury et al., 2007) in a full matrix size.

The GWAS analysis for phenotypic plasticity was conducted for DTF, TSW, PSHR, NS, PH and SL measured across the two environments.

Genomic prediction and cross-validation analysis

The potential of genomic prediction to assist turnip rape breeding was evaluated using one of the most efficient and widely applied ridge-regression BLUP (RR-BLUP) using the

TABLE 1 Analysis of variance and broad-sense heritability of six phenotypic traits of turnip rape conducted based on data collected from both Öjebyn and Lönnstorp field trials.

Statistics	DTF	PH	NS	SL	PSHR	TSW
σ^2G	2.33	65.51	26	0.08	0.04	0.11
$\sigma^2G \times E$	0.90	0.92	82.68	0.00	0.07	0.07
σ^2E	0.86	66.68	334.75	0.41	0.29	0.03
Mean	35.76	81.34	58.55	5.95	2.48	2.76
CV	2.59	10.04	31.25	10.71	21.84	6.61
H^2	0.80	0.85	0.21	0.55	0.32	0.73
Gen. Sign	***	***	NS	***	*	***
G \times E sign	***	NS	***	NS	***	***

DTF, days to flowering; PH, plant height; NS, number of siliques; SL, silique length; PSHR, pod shattering resistance; TSW, thousand seed weight.

rrBLUP package in R (Endelman, 2011). The RR-BLUP model was implemented using the following mixed linear model:

$$Y = \mu + Z\beta + \varepsilon$$

where Y is the $N \times 1$ vector values of the BLUPs recorded for a studied trait; μ is a vector of trait's overall phenotypic mean; and Z is the $N \times N_m$ SNP marker matrix. N and N_m are the number of accessions and SNP markers, respectively, applied to the current genomic prediction analysis. β is the $N_m \times 1$ vector of random SNP effects calculated using the "mixed.solve" function in the RR-BLUP package following $\beta \sim N(0, I\delta_g^2)$, where I is the identity matrix and δ_g^2 is the genetic variance component contributed by SNP markers; and ε is the $N \times 1$ vector of residuals.

The accuracy of the prediction models was assessed through a cross-validation method using 80% of the accessions as a training set (TRS) and the remaining 20% as a test set (TS). Here, the genomic estimated breeding values (GEBVs) of accessions in the TS were estimated with the genomic prediction models trained on the TRS accessions. Then, Pearson's correlation coefficient between the GEBVs and adjusted phenotypic values (BLUPs) was recorded as genomic prediction ability, which was repeated 500 times for each trait. The prediction accuracy was estimated by dividing the prediction ability by the square root of the trait's broad-sense heritability (Legarra et al., 2008; Chen et al., 2011; Alemu et al., 2021a).

Results

Analysis of phenotypic variance, plasticity, correlation and heritability

Analysis of variance (ANOVA) for phenotypic data was conducted within and across sites (environments). At the Lönnstorp site, genotypic variance was highly significant ($p < 0.0001$) for all traits except NS and SL (Supplementary Table S2). Thousand seed weight (TSW) had the highest repeatability (0.95) followed by PH, DTF and YLD, which had repeatability values of 0.89, 0.82, and 0.7, respectively. However, SL and NS had low repeatability values of 0.21 and 0.22, respectively. At the Öjebyn site, the genotypic variance of all seven studied traits was highly significant. TSW and

DTF had the highest repeatability score (0.94) while SS exhibited the lowest repeatability value (0.47) (Supplementary Table S2).

The combined environment ANOVA was conducted for six of the nine traits with data from the two sites. In this case, DTF, PH, SL and TSW exhibited highly significant variation between accessions (Table 1). The variance due to G \times E was highly significant for DTF, NS, PSHR, and TSW. There was no significant G \times E in PH and SL. Broad-sense heritability (H^2) was high for PH, DTF and TSW with values of 0.85, 0.80, and 0.73, respectively. Other traits had low to moderate broad-sense heritability (Table 1).

Strong correlations, both positive and negative, were evident among the traits examined, as illustrated in Figure 2. The number of days to flowering (DTF) was significantly negatively correlated with yield ($r = -0.55$, $p < 0.001$), PSHR ($r = -0.38$, $p < 0.001$) and SL ($r = -0.25$, $p < 0.01$) and positive correlation with DTM ($r = 0.60$, $p < 0.001$). Seed yield (YLD) demonstrated significant positive correlations with PSHR ($r = 0.50$, $p < 0.001$), PH ($r = 0.45$, $p < 0.001$), NS ($r = 0.28$, $p < 0.001$), and TSW ($r = 0.16$, $p < 0.05$). Moreover, TSW exhibited a significant negative correlation with both NS and seed SS (Figure 2).

Except for PSHR, the estimated phenotypic plasticity of the five traits showed a strong correlation (both positive and negative) with the estimated BLUP values for the combined phenotypic performance of accessions across the two environments. Specifically, DTF and TSW exhibited a negative correlation with phenotypic plasticity, whereas NS, SL, and PH demonstrated positive correlations (Figure 3).

Linkage disequilibrium and marker-trait associations identified via GWAS

A total of 2,706,301 pairwise linkage disequilibrium (LD) comparisons were derived from the full matrix analysis with the currently applied SNP markers. The genome-wide linkage disequilibrium (LD) decay analysis demonstrated a rapid decay of LD. Initially, the maximum LD started at an r^2 value of 0.44 and reached to half LD decay at $r^2 = 0.23$ (Figure 4). Notably, the LD decay was rapid as the r^2 value decreased from 0.3 to 0.1, the corresponding genomic distance increased substantially from 2.74 to 19.84 megabase pairs (Mbp) (Figure 4).

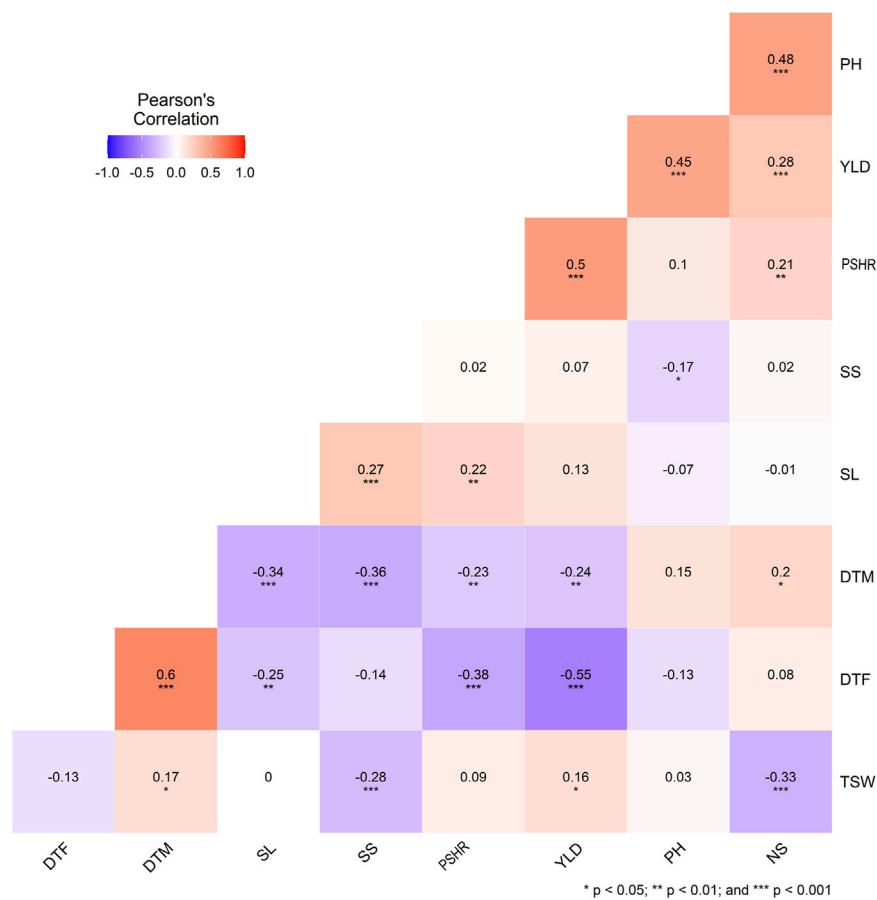


FIGURE 2 Correlation coefficients and significance levels for agronomic traits in 170 turnip rape accessions. DTF, days to flowering; DTM, days to maturity; SL, silique length; SS, number of seeds per silique; PSHR, pod shattering resistance; YLD, yield per plot; PH, plant height; NS, number of silique; TSW, thousand seed weight.

The germplasm panel used in this study comprised accessions from more than 27 countries across five continents (Supplementary Table S1). However, the SNP marker-based kinship and population structure analysis revealed poor population structure, as the vast majority of the accessions were clustered together and there was no clear clustering of accessions based on factors such as country or region of origin (Figure 5).

Several significant marker-trait associations (MTAs) were identified via GWAS analysis for the nine studied traits. In total, 65 SNP markers exhibited significant associations with one or more traits for a single environment and combined environment GWAS with values above the exploratory threshold (Supplementary Table S3). For TSW, 25 MTAs were identified, spanning all chromosomes except chromosome 10 (Chr A10). Nine, eight and five SNPs were significant only for Öjebyn, Lönnstorp and the combined environment, respectively. One SNP marker (*CM020890.1_33244505*) was significant for Öjebyn and the combined environment while another one (*CM020894.1_4615438*) was significant for Lönnstorp and the combined environment. Whereas, one marker (*CM020895.1_12959514*) was significant for Öjebyn, Lönnstorp and the combined environment. The SNP marker *CM020895.1_12959514* (MAF = 0.46) on Chr A08 displayed a stable association with TSW in each

environment as well as the combined environment with levels of significance above the Bonferroni threshold (Table 2; Figure 6A).

Accessions homozygous for the reference allele (A) at this locus had significantly higher TSW compared to accessions homozygous for the alternative allele (G) ($p < 0.01$) for the two sites as well as the combined environment (Figure 6B). The two SNP markers *CM020888.1_2486460* (MAF = 0.08) and *CM020888.1_2486421* (0.07) on Chr A01 with the same physical position (2.49 Mbp; only 39 bp apart) were significant for TSW for Öjebyn and Lönnstorp, respectively, and hence they may refer to the same minor QTL. The two SNP markers on Chr A07, *CM020894.1_4615438* (MAF = 0.05) and *CM020894.1_11913677* (MAF = 0.11) were also significantly associated with TSW (Supplementary Table S3; Figure 6A). *CM020894.1_4615438* may be considered stable, as it was significant for Lönnstorp and the combined environment.

DTF is a crucial trait for improvement in turnip rape, particularly in northern temperate regions via shortening days to flowering and maturity. Seventeen MTAs were identified for this trait for the two single environments and the combined environment involving 14 SNP markers across all chromosomes except Chr A05. The SNP marker *CM020896.1_28000790* (MAF = 0.06) on chromosome A09 was significant for Öjebyn and the combined environment with $-\log P$ values of 12.6 and 10,

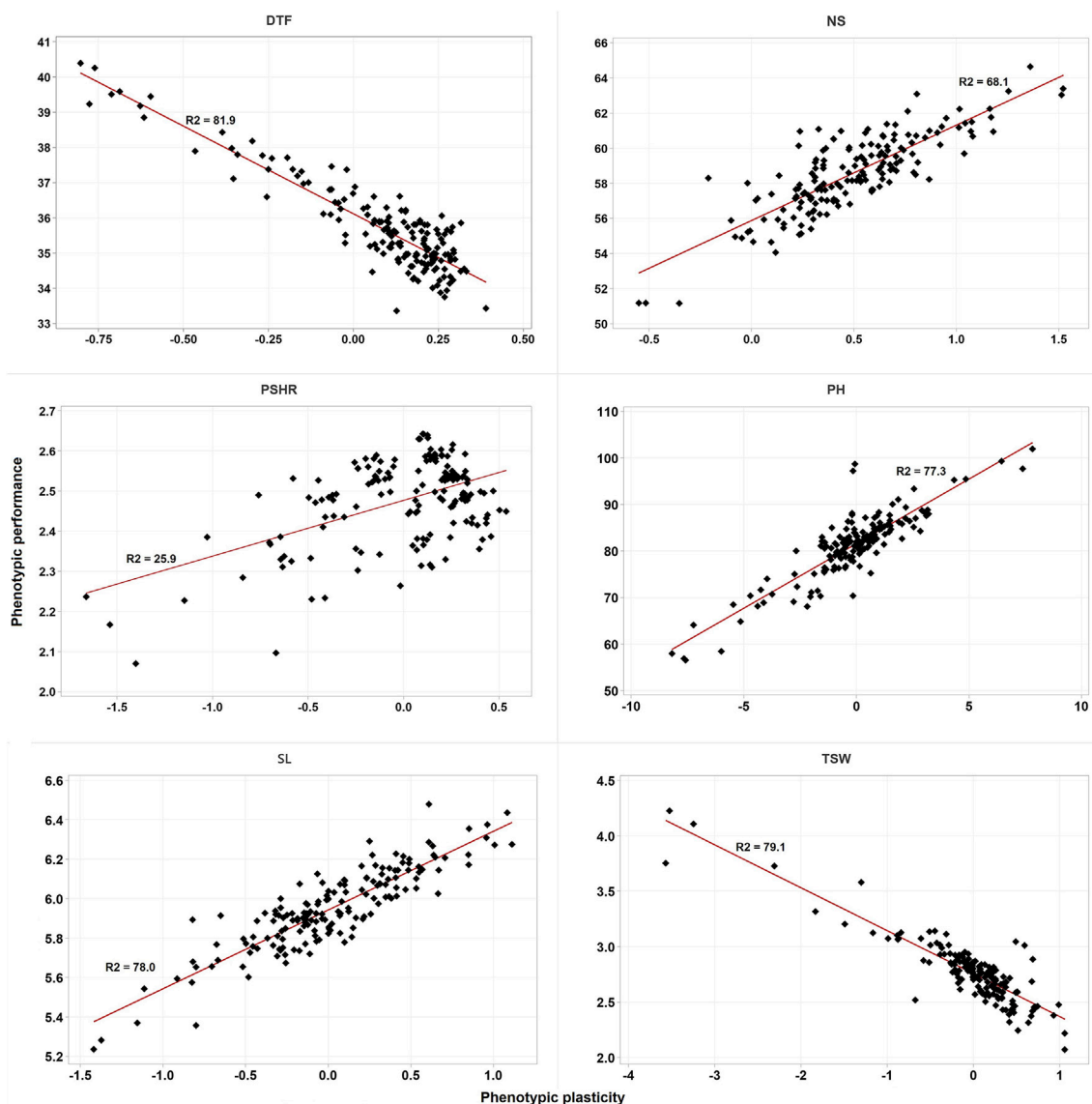


FIGURE 3
Scatter plot of the BLUPs estimated phenotypic performance of accessions against their estimated phenotypic plasticity. DTF, days to flowering; NS, number of siliques; PH, plant height; PSHR, pod shattering resistance; SL, silique length; TSW, thousand seed weight. R² is the coefficient of determination.

respectively (Supplementary Table S3; Figure 7A). The two closely located markers on Chr A07, *CM020894.1_13558395* (13.6 Mb) and *CM020894.1_14019976* (14 Mb) were significantly associated with DTF for Öjebyn and Lönnstorp, respectively (Supplementary Table S3; Figure 7A). The SNP marker *CM020897.1_2008052* (MAF = 0.11) on Chr A10 was significant for Öjebyn, Lönnstorp and the combined environment, and hence is a stable marker for DTF. Analysis of variance revealed that accessions homozygous for allele A flowered earlier than accessions homozygous for allele T for both sites and the combined environment ($p < 0.01$) (Figure 7B).

Silique-related traits, pod resistance (PSHR), silique length (SL), seeds per silique (SS) and number of siliques (NS) were also evaluated for MTA, and several MTAs were identified except for NS. For PSHR, SL and SS, 16, 7 and 3 significant SNP markers,

respectively, were identified across the two sites and the combined environment (Supplementary Table S3). SNP markers significantly associated with PSHR were distributed across all chromosomes except Chr A07. The SNP markers *CM020888.1_28035981* and *CM020890.1_26416338* on chromosomes 1 and 3, respectively, were significantly associated with PSHR for the Lönnstorp site and combined environment.

Other traits evaluated in the current study were yield (YLD), plant height (PH) and days to maturity (DTM). Yield and days to maturity were evaluated only for the Lönnstorp site. Five MTAs were identified for YLD on chromosomes 4, 5, 7, and 9, while 13 MTAs were identified for PH with associated SNPs distributed across all chromosomes except Chrs A04 and A07. For DTM, three MTAs were detected on Chr A02, A04 and A07 (Supplementary Table S3).

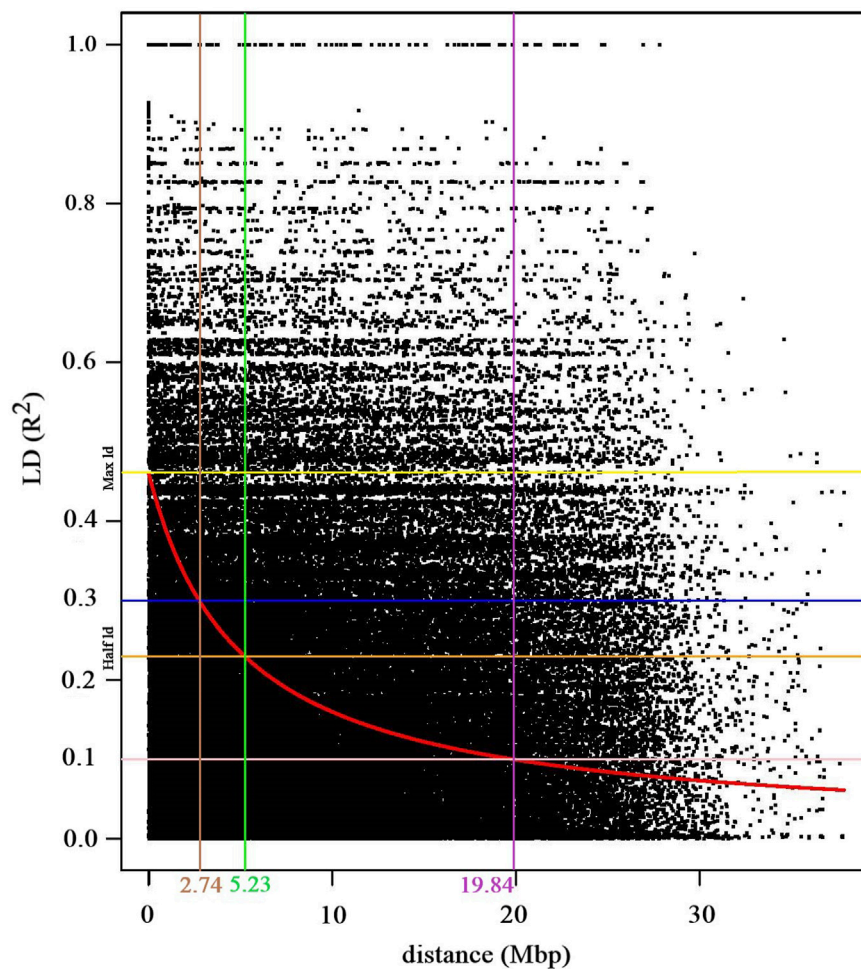


FIGURE 4

Scatter plot depicting the genome-wide linkage disequilibrium (LD) decay curve, which plots pairwise SNPs r^2 values against their physical distances in mega base pairs (Mbp). The red curve line indicates the LD decay trend, which was derived by fitting the smoothing spline regression model (loess). The yellow and orange horizontal lines indicate the pairwise r^2 values at the maximum and half LD decay, respectively. The blue (r^2 , 0.3) and pink (r^2 , 0.1) horizontal lines intersects with the LD decay curve at 2.74 Mbp (brown vertical line) and 5.23 Mbp (pink vertical line), respectively. The green vertical line indicates the distance (5.23 Mbp) where the LD decay curve intersects with the half LD decay.

The GWAS analysis was performed using the phenotypic plasticity values estimated for traits evaluated at two locations (Figure 8). This analysis identified 23 major marker-trait associations (MTAs) for five traits (Supplementary Table S4). Given the limited number of tested environments (two), the majority of these MTAs (15) were also detected on the previous phenotypic BLUP-based GWAS analysis.

Accuracy of the genomic estimated breeding values

The GWAS analysis conducted in this study successfully identified marker-trait associations for eight out of the nine traits examined, demonstrating the effectiveness of the marker set employed. Consequently, we tested the effectiveness of these marker set for genomic prediction of the target traits in order to promote the integration of genomic selection into turnip rape

breeding programs. The GWAS analysis did not detect significant MTAs for the number of siliques (NS) at the significant threshold applied. Interestingly, the genomic-based prediction accuracy for this trait was 0.67, 0.73 and 0.64 for Öjebyn, Lönnstorp and the combined environment, respectively (Table 3). In general, the genomic prediction accuracy for the studied traits was promising (Table 3; Figure 9). For instance, the prediction accuracy for DTF for Öjebyn, Lönnstorp and the combined environment was 0.71, 0.47 and 0.71, respectively. For TSW, it was 0.77, 0.65 and 0.88 for Öjebyn, Lönnstorp and the combined environment, respectively. Moderate to low prediction accuracy was recorded for silique length, 0.52 (Öjebyn site), 0.37 (Lönnstorp site) and 0.48 (combined environment). On the other hand, the accuracy of genomic prediction for pod-shattering (PSHR) ranged from low to high (0.35–0.91) across the two sites and combined environment. Seed yield was evaluated only for the Lönnstorp site, and the genomic prediction accuracy for this trait was interestingly high (0.8) (Table 3).

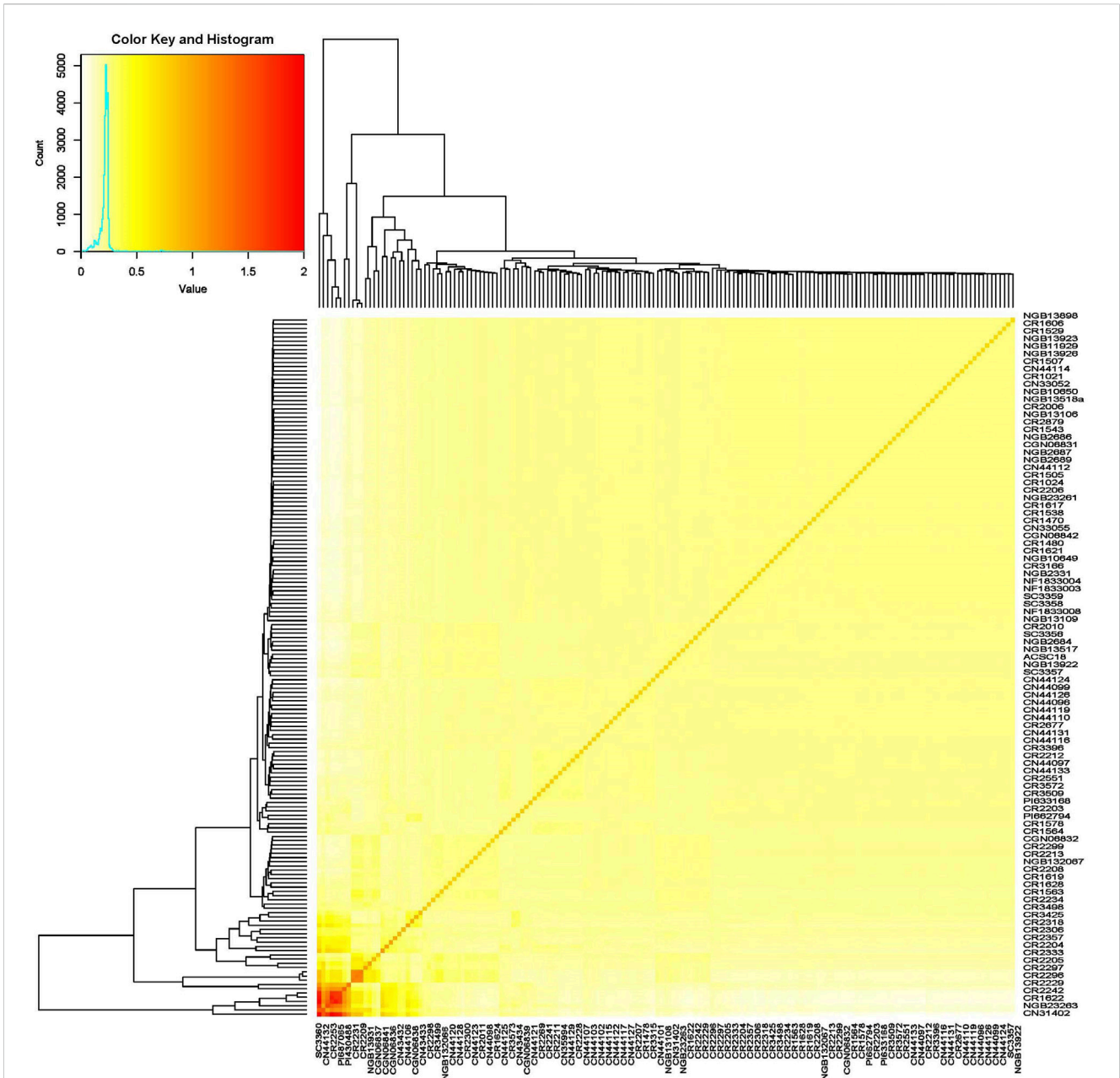


FIGURE 5 Heat map of the turnip rape panel comprising accessions representing its global gene pool based on their SNP marker profile revealing the absence of a strong population structure.

Discussion

Turnip rape was among the largely cultivated edible and industrial oilseed crops globally until the 1970s. However, the crop is currently cultivated in very limited areas in India, Canada and Northern Europe (Kaur et al., 2021b). This is mainly because of the lack of breeding programs to develop cultivars with high yield and resistance/tolerance to recurring biotic and abiotic stresses. The current study was conducted at two locations in Sweden to evaluate the performance of turnip rape accessions for nine phenotypic traits commonly targeted in breeding. The panel comprised accessions of worldwide origin with diverse genetic backgrounds, including

landraces, cultivars, breeding populations and wild populations. The single site-based and combined environment-based analysis of variance revealed highly significant variation among accessions for the majority of studied traits, and hence are suitable for marker-trait association (MTA) analysis via GWAS (Brachi et al., 2011). This facilitates the development of genomic tools that assist turnip rape breeding in order to accelerate genetic gains in target traits. Crop yield, nutritional quality and stress tolerance/resistance are quantitative traits regulated by several genes located in genomic regions referred to as quantitative trait loci (QTLs) (Consortium et al., 2003). These genes/QTLs further interact with each other and the environments where plants are grown, making phenotypic

TABLE 2 List of some SNP markers significantly associated at above-Bonferroni threshold levels with TSW, DTF or PSHR at least in two of the three cases (Öjebyn site, Lönnstorp site and combined environment) together with their SNP position within the corresponding chromosomes and minor allele frequencies (MAF).

Trait	Location	SNP	Chr	Pos (MB)	-log ₁₀ (P values)	MAF
TSW	Lönnstorp	CM020895.1_12959514	8	12.96	16.57	0.46
	Öjebyn				12.04	
	Combined				9.36	
	Öjebyn	CM020890.1_33244505	3	33.24	9.01	0.12
	Combined				5.69	
	Combined	CM020894.1_4615438	7	4.62	8.88	0.05
	Lönnstorp				8.17	
DTF	Combined	CM020896.1_28000790	9	28.00	12.63	0.06
	Öjebyn				10.01	
	Öjebyn	CM020897.1_2008052	10	2.01	6.88	0.11
	Combined				6.70	
	Öjebyn	CM020897.1_2008052	10	2.01	6.88	0.11
	Lönnstorp				6.52	
PSHR	Combined	CM020888.1_28035981	1	28.04	9.30	0.10
	Lönnstorp				5.08	
	Combined	CM020890.1_26416338	3	26.42	5.39	0.19
	Lönnstorp				5.13	

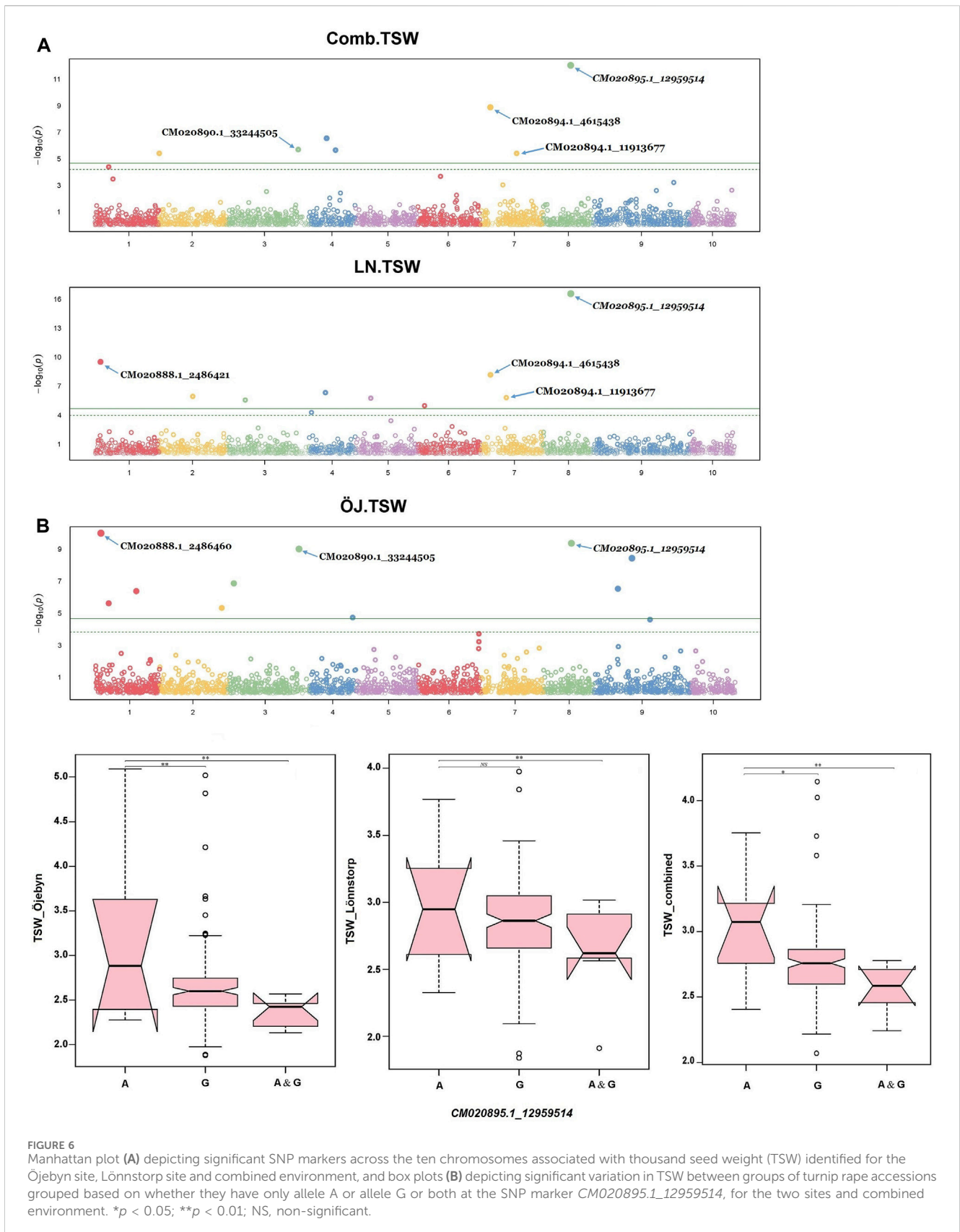
TSW, thousand seed weight; DTF, days to flowering; PSHR, pod shattering resistance.

evaluation challenging. Multi-environment trial evaluation of genetically diverse accessions across different environments allows us to understand genotype stability and performance in these target traits. The current study was designed to assess the phenotypic and genetic variation of turnip rape accessions followed by GWAS and genomic prediction analyses.

The availability of whole genome assemblies of *B. rapa* (<http://brassicadb.cn>) has made it possible to discover genes/QTLs linked to economically valuable traits (Cai et al., 2017; Zhang et al., 2018). The current study has discovered several markers linked to the studied traits. Flowering and maturity time are among the most important traits in crop plants affecting biomass and seed yield as well as adaptation to different growing environments (Weis and Kossler, 2004; Gaudinier and Blackman, 2020). A large variation in flowering and maturity times has been observed among *B. rapa* germplasm leading to winter, semi-winter, and spring growth habits (Kaur et al., 2021b). Days to flowering (DTF) is a quantitatively inherited trait controlled by several genes with minor to major effects. Several of these genes have been identified and characterized in Brassicas, such as *FLOWERING LOCUS T (FT)*, *FRUITFUL (FUL)*, *FLOWERING LOCUS C (FLC)*, *CONSTANS (CO)*, *FRIGIDA (FRI)* and *PHYTOCHROME (PHY)* (Long et al., 2007; Raman et al., 2013; Xiao et al., 2013; Raman et al., 2016; Gao et al., 2017). Previous studies showed that several genes (including those belonging to the same gene families) in Brassica and *Arabidopsis* control flower initiation and flowering (Schiessl et al., 2017). Studies on *B. rapa* showed that these genes are located on Chrs A01, A02, A03, A07 and A10 with a different number of copies (Schranz et al., 2002; Zhang

et al., 2015; Leijten et al., 2018; Li et al., 2018). Similarly, the current GWAS study identified SNP markers significantly associated with DTF on Chrs A01 (single environment), A02 (single environment), A03 (combined environment), A07 (single environment) and A10 (both single and combined environments, and hence stable) (Supplementary Table S3). Li et al. (2018) reported 1,064, 510, and 524 putative orthologs of *Arabidopsis thaliana* flowering-time genes from *B. napus*, *B. rapa*, and *Brassica oleracea*, respectively, including several genes on Chr A09. Likewise, the present study identified the most significant marker-trait association (MTA) for DTF on the same chromosome. Day to maturity (DTM) was only analyzed for the Lönnstorp site and MTAs were identified on Chrs A04, A07 and A02. Similarly, Kaur et al. reported significant marker for this trait on Chr 02 (Kaur et al., 2021b). Zhou et al. (2017) also reported significant markers on Chrs A02, A03, A06, A07 and A09 (corresponding to the *B. rapa* genome) in rapeseed (*B. napus*) using a diverse GWAS panel containing 300 inbred lines.

The GWAS studies carried out thus far in Brassicas for yield and related traits focused on *B. napus*. Therefore, the use of a different reference genome complicates the comparison of precise chromosome positions of identified QTLs with our study. In this study, 25 MTAs were detected for thousand seed weight (TSW) distributed across all *B. rapa* chromosomes, except Chr A10. Several previous researchers have identified QTLs for TSW on the A genome of *B. rapa* in line with the present study. In this study, Chrs A07, A08 and A09 contain highly significant QTLs for this trait, in line with previous studies (Yang et al., 2012; Li et al., 2014; Lu et al., 2017; Khan et al., 2019; Shen et al., 2019; Shi et al., 2019). The



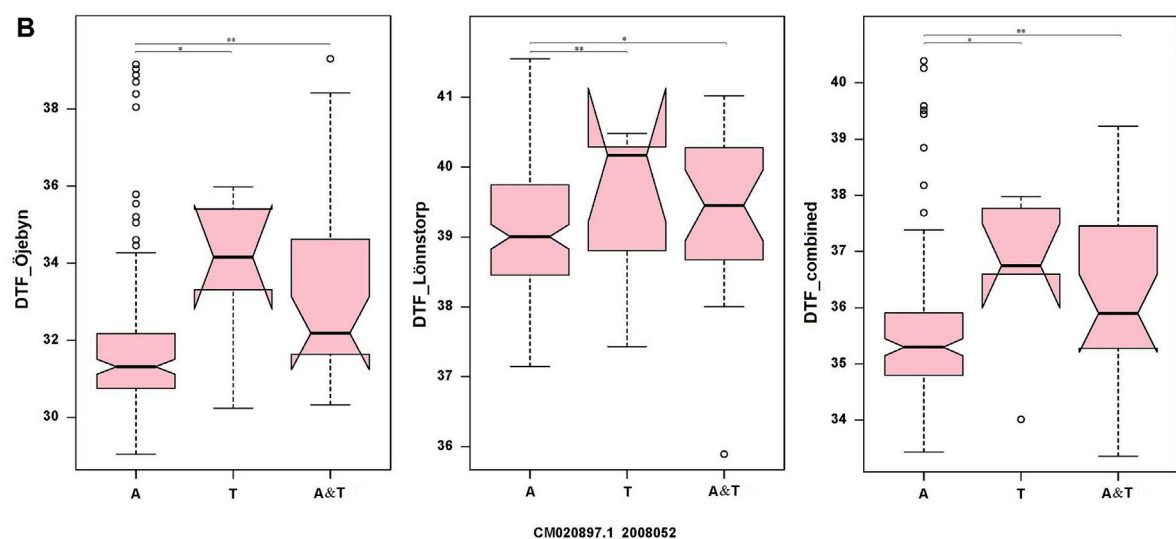
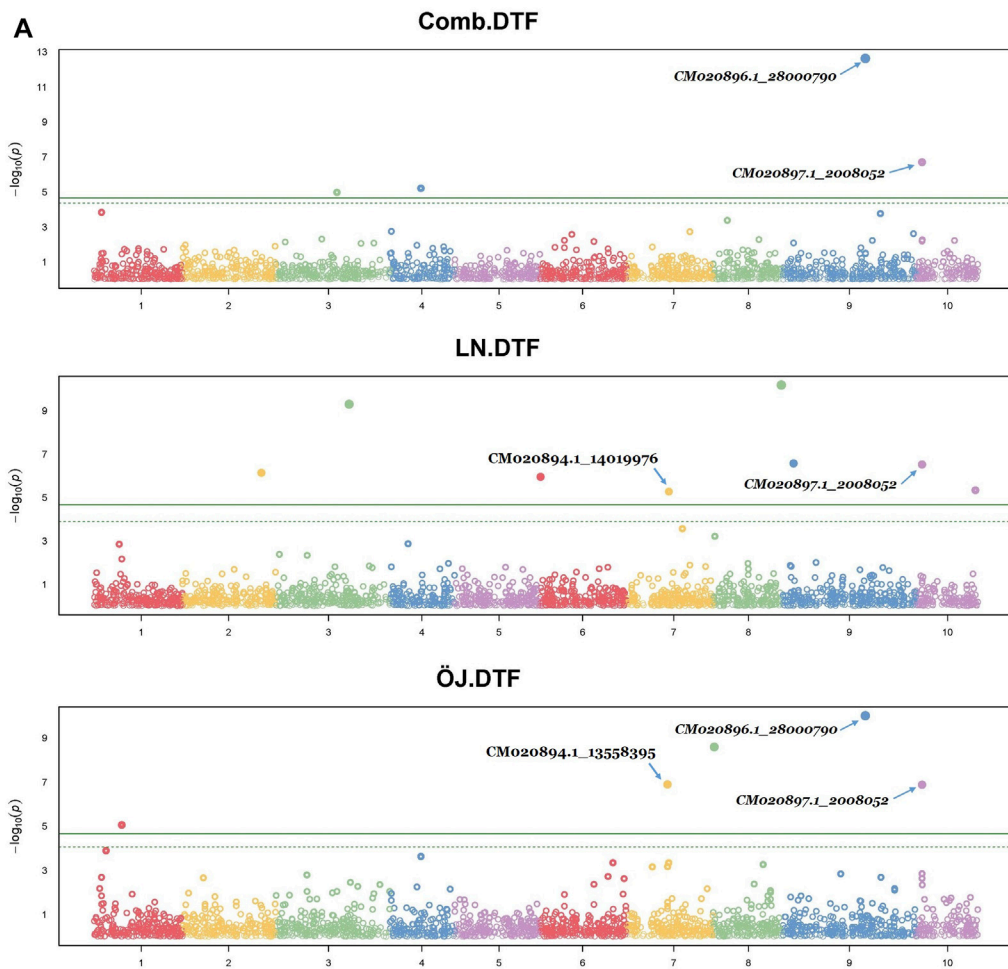
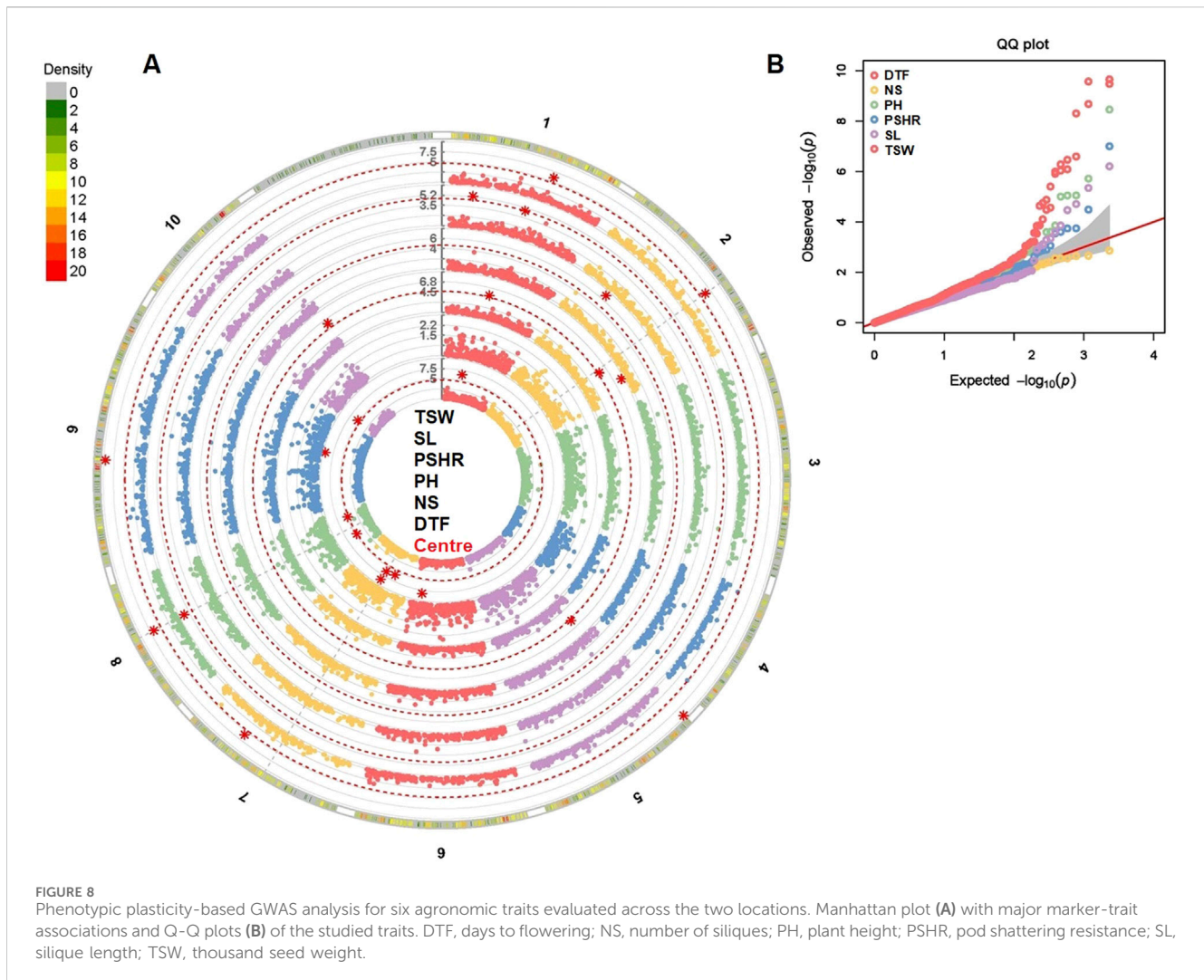


FIGURE 7 Manhattan plot (A) depicting significant SNP markers across the ten chromosomes associated with days to flowering (DTF) identified for the Öjebyn site, Lönnstorp site and combined environment, and box plots (B) depicting significant variation in TSW between groups of turnip rape accessions grouped based on whether they have only allele A or allele T or both at the SNP marker *CM020895.1_12959514*, for the two sites and combined environment. * $p < 0.05$; ** $p < 0.01$.



current study unveiled that all chromosomes except Chrs A04 and A07 carry significant SNP markers associated with PH. Similarly, Li et al. (2016) reported QTLs located on Chrs A03 and A05, while Zheng et al. (2017) reported QTLs on Chrs A02, A08, and A09, in their studies comprising 472 and 333 rapeseed accessions, respectively. Kaur et al. (2021b) reported PH MTAs on Chrs A01, A03, 05, A06, A08, A09 and 10 based on a panel of 195 inbred lines, representing natural and derived forms of *B. rapa* evaluated at two locations.

Regarding silique-related traits, our research focused on exploring silique length (SL), silique number (NS), number of seeds per silique (SS) and pod shattering (PSHR), which are major traits to improve in turnip rape breeding to enhance seed yield. Majority of previous studies on these traits are on *B. napus*, among Brassica crops, and hence we could only make chromosome-level comparisons. The GWAS analysis identified 16, 7 and 3 MTAs for PSHR, SL and SS, respectively, but none for NS. The MTAs for PSHR were identified on all chromosomes except on Chr A07. Similarly, Raman et al. (2014) identified PSHR related QTLs on Chrs A03, A07, and A09 in a study on 126 double haploid lines of *B. napus*. In agreement with our finding, several previous researches also reported MTAs/QTLs for PSHR on Brassica A genome

chromosomes (Liu J. et al., 2016; Kaur et al., 2020; Raman et al., 2023). The current study identified SNP markers significantly associated with SL on Chrs A01, A03, A06, A08 and A09. Similarly, Dong et al. (2018) identified several markers associated with SL on Chrs A03, A08 and A09 from a GWAS study using 157 rapeseed inbred lines. In this study, a SNP marker with significant association with SL (combined environment) was found on Chr A09. This chromosome was recognized as a source of major QTLs for SL in several previous studies (Yang et al., 2012; Fu et al., 2015; Shi et al., 2019; Wang et al., 2019; Wang et al., 2021). We identified three MTAs for SS (Öjebyn site) on Chr A05 (two) and Chr A10. For YLD, five significant markers were identified (Lönstorp site) on Chrs A04, A05, A07 and A09 (two). Previous studies have similarly identified SS related QTL/MTA on chromosomes 5 and 10 (Ding et al., 2012; Yang et al., 2016; Xin et al., 2021). In agreement with our study, a study on 124 recombinant inbred lines of *B. napus* revealed YLD-related QTLs on Chrs A04 and A05 (Ding et al., 2012).

Genomic prediction stands out as a valuable tool for accelerating genetic gains in crop breeding efforts (Alemu et al., 2024). This method has emerged as a preferred approach for polygenic quantitatively inherited traits, especially considering the

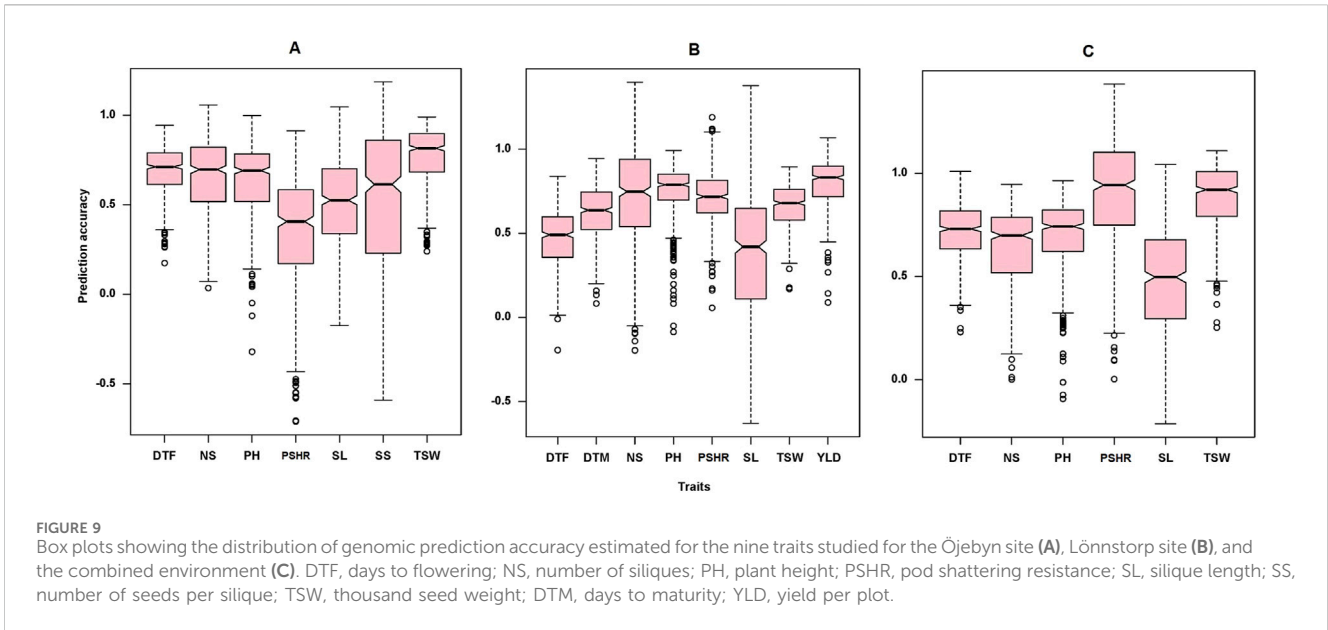


TABLE 3 Average predictive ability and prediction accuracy of the model used for the analysis of nine phenotypic traits for the two sites (Öjebyn and Lönnpstorp) separately and for the combined environment.

Location	Traits	Average predictive ability	Average prediction accuracy
Öjebyn	DTF	0.69	0.71
	NS	0.51	0.67
	PH	0.51	0.63
	PSHR	0.29	0.35
	SL	0.37	0.52
	SS	0.36	0.53
	TSW	0.75	0.77
Lönnpstorp	DTF	0.43	0.47
	DTM	0.55	0.63
	NS	0.34	0.73
	PH	0.7	0.75
	PSHR	0.44	0.71
	SL	0.17	0.37
	TSW	0.59	0.65
	YLD	0.67	0.8
Two locations combined	DTF	0.64	0.71
	NS	0.29	0.64
	PH	0.65	0.7
	PSHR	0.52	0.91
	SL	0.36	0.48
	TSW	0.75	0.88

DTF, days to flowering; NS, number of siliques; PH, plant height; PSHR, pod shattering resistance; SL, silique length; SS, number of seeds per silique; TSW, thousand seed weight; DTM, days to maturity; YLD, yield per plot.

challenges in breeding for such traits (Alemu et al., 2024). The current study explored the potential of genomic prediction in turnip rape breeding through a cross-validation study. The abundance of significantly associated SNP markers identified for the majority of studied traits in this study poses a challenge to their immediate application as selection and breeding tools in turnip rape. This observation prompted us to explore the potential of genomic prediction, which involves estimating the overall merit of individual accessions for improvement of a particular trait, rather than solely identifying specific MTAs. We applied an 80%–20% training-test set cross-validation approach to evaluate the accuracy of genomic prediction. The analysis provided promising results, with high prediction accuracy for the majority of targeted traits. No significant MTAs were detected via GWAS for the number of siliques per plant. However, higher genomic prediction accuracy was recorded for this trait, with values of 0.67, 0.73, and 0.63 for Öjebyn, Lönnstorp, and the combined environments, respectively. Similar to the GWAS studies, the majority of previous genomic prediction studies on Brassicas focused on estimating the GEBVs of hybrid *B. napus* (Würschum et al., 2014; Wei et al., 2016; Yang et al., 2016; Szwarc et al., 2023). Thus, the present study provides an interesting insight into the potential applications of genomic prediction in the breeding of turnip rape.

In summary, our study investigated the utilization of genomic tools to enhance turnip rape breeding. We evaluated a panel of 170 accessions representing a global turnip rape gene pool via field trials at two distinct locations, followed by the identification of MTAs and GEBVs through GWAS and genomic prediction studies, respectively. A total of 65 SNP markers were identified as significantly associated with eight of the nine traits studied. Notably, the genomic prediction accuracy for most targeted traits was high, suggesting the potential of genomics-driven breeding in expediting genetic gains in turnip rape.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material. The SNP markers are deposited at the following link: https://figshare.com/articles/dataset/Turnip_rape_SNP_marker_MAF0_05Hetrozygosity20_/25847716.

Author contributions

AA: Data curation, Formal Analysis, Methodology, Software, Visualization, Writing–original draft. JS: Investigation, Methodology, Writing–review and editing. KA: Investigation, Methodology, Writing–review and editing. ME: Formal Analysis,

Investigation, Methodology, Writing–review and editing. MG: Conceptualization, Investigation, Project administration, Writing–review and editing. AC: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was made possible by the funding from SLU Grogrund Center for Breeding of Food Crops, in a project “Rybs (*B. rapa* subsp. *oleifera*) – en flexibel och tålig oljegröda för Sverige” (#ltv.2019.1.1 .1-1041). SLU LTV faculty supported the growth facilities.

Acknowledgments

We thank Ferdi Agil and Inez Falk for their assistance in the field trial and greenhouse work. We also appreciate Mirela Beganovic, Kaye Bateman, and Marina Kuzmenkova for their assistance in sampling plant material. We also thank the greenhouse personnel at LTV faculty for their support during the experiments.

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1435474/full#supplementary-material>

References

- Alemu, A., Åstrand, J., Montesinos-López, O. A., Isidro Y Sánchez, J., Fernández-González, J., Tadesse, W., et al. (2024). Genomic selection in plant breeding: key factors shaping two decades of progress. *Mol. Plant*. 17, 552–578. doi:10.1016/j.molp.2024.03.007
- Alemu, A., Brantestam, A. K., and Chawade, A. (2022). Unraveling the genetic basis of key agronomic traits of wrinkled vining pea (*Pisum sativum* L.) for sustainable production. *Front. Plant Sci.* 13, 844450. doi:10.3389/fpls.2022.844450

- Alemu, A., Brazauskas, G., Gaikpa, D. S., Henriksson, T., Islamov, B., Jørgensen, L. N., et al. (2021a). Genome-wide association analysis and genomic prediction for adult-plant resistance to septoria tritici blotch and powdery mildew in winter wheat. *Front. Genet.* 12 (12), 661742. doi:10.3389/fgene.2021.661742
- Alemu, A., Suliman, S., Hagra, A., Thabet, S., Al-Abdallat, A., Abdelmula, A. A., et al. (2021b). Multi-model genome-wide association and genomic prediction analysis of 16 agronomic, physiological and quality related traits in ICARDA spring wheat. *Euphytica* 217, 205. doi:10.1007/s10681-021-02933-6
- Alvarado, G., Rodríguez, F. M., Pacheco, A., Burguenio, J., Crossa, J., Vargas, M., et al. (2021). META-R: a software to analyze data from multi-environment plant breeding trials. *Crop J.* 8, 745–756. doi:10.1016/j.cj.2020.03.010
- Bassegio, D., and Zanotto, M. D. (2020). Growth, yield, and oil content of Brassica species under Brazilian tropical conditions. *Bragantia* 79, 203–212. doi:10.1590/1678-4499.20190411
- Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci.* 48, 1649–1664. doi:10.2135/cropsci2008.03.0131
- Bird, K. A., An, H., Gazave, E., Gore, M. A., Pires, J. C., Robertson, L. D., et al. (2017). Population structure and phylogenetic relationships in a diverse panel of Brassica rapa L. *Front. Plant Sci.* 8, 321. doi:10.3389/fpls.2017.00321
- Brachi, B., Morris, G. P., and Borevitz, J. O. (2011). Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol.* 12, 232. doi:10.1186/gb-2011-12-10-232
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinform.* 23, 2633–2635. doi:10.1093/bioinformatics/btm308
- Cai, C., Wang, X., Liu, B., Wu, J., Liang, J., Cui, Y., et al. (2017). Brassica rapa genome 2.0: a reference upgrade through sequence Re-assembly and gene Re-annotation. *Mol. Plant* 10, 649–651. doi:10.1016/j.molp.2016.11.008
- Cartea, M. E., Cámara-Martos, F., Obregón, S., Badenes-Pérez, F. R., and De Haro, A. (2020). Advances in breeding in vegetable Brassica rapa crops. *Brassica Breed. Biotechnol.*, 1–18. doi:10.5772/intechopen.95769
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., et al. (2014). Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. *Science* 345, 950–953. doi:10.1126/science.1253435
- Challa, S., and Neelapu, N. R. R. (2018). "Chapter 9 - genome-wide association studies (GWAS) for abiotic stress tolerance in plants," in *Biochemical, Physiological and Molecular Avenues for Combating Abiotic Stress Tolerance in Plants*. Editor S.H. Wani (Amsterdam, Netherlands: Academic Press), 135–150.
- Chen, C., Misztal, I., Aguilar, I., Tsuruta, S., Meuwissen, T., Aggrey, S., et al. (2011). Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: an example using broiler chickens. *J. animal Sci.* 89, 23–28. doi:10.2527/jas.2010-3071
- Consortium, C. T., Angel, J. M., Avner, P., Bachmanov, A. A., Belknap, J. K., Bennett, B., et al. (2003). The nature and identification of quantitative trait loci: a community's view. *Nat. Rev. Genet.* 4, 911–916. doi:10.1038/nrg1206
- Ding, G., Zhao, Z., Liao, Y., Hu, Y., Shi, L., Long, Y., et al. (2012). Quantitative trait loci for seed yield and yield-related traits, and their responses to reduced phosphorus supply in Brassica napus. *Ann. Bot.* 109, 747–759. doi:10.1093/aob/mcr323
- Dong, H., Tan, C., Li, Y., He, Y., Wei, S., Cui, Y., et al. (2018). Genome-wide association study reveals both overlapping and independent genetic loci to control seed weight and silique length in Brassica napus. *Front. Plant Sci.* 9, 921. doi:10.3389/fpls.2018.00921
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6, e19379. doi:10.1371/journal.pone.0019379
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. doi:10.3835/plantgenome2011.08.0024
- Fu, Y., Wei, D., Dong, H., He, Y., Cui, Y., Mei, J., et al. (2015). Comparative quantitative trait loci for silique length and seed weight in Brassica napus. *Sci. Rep.* 5, 14407. doi:10.1038/srep14407
- Gao, B., Liu, B., Li, S., Liang, J., Cheng, F., Wang, X., et al. (2017). Genome-wide association studies for flowering time in brassica rapa. *Sci. Agric. Sin.* 50, 3375–3385. doi:10.3864/j.issn.0578-1752.2017.17.012
- Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv.* 1207, 3907. doi:10.48550/arXiv.1207.3907
- Gaudinier, A., and Blackman, B. K. (2020). Evolutionary processes from the perspective of flowering time diversity. *New Phytol.* 225, 1883–1898. doi:10.1111/nph.16205
- Gómez-Campo, C., and Prakash, S. (1999). "2 Origin and domestication," in *Developments in plant genetics and breeding*. Editor C. Gómez-Campo (Amsterdam, Netherlands: Academic press), 33–58. doi:10.1111/nph.16205
- Guo, Y., Chen, S., Li, Z., and Cowling, W. A. (2014). Center of origin and centers of diversity in an ancient crop, Brassica rapa (turnip rape). *J. Hered.* 105, 555–565. doi:10.1093/jhered/esu021
- Jordbruksverket (2022). The Swedish agency for agriculture's official statistics.
- Kaur, J., Akhtar, J., Goyal, A., Kaur, N., Kaur, S., Mittal, M., et al. (2020). Genome wide association mapping and candidate gene analysis for pod shatter resistance in Brassica juncea and its progenitor species. *Mol. Biol. Rep.* 47, 2963–2974. doi:10.1007/s11033-020-05384-9
- Kaur, S., Akhtar, J., Kaur, H., Atri, C., Mittal, M., Goyal, A., et al. (2021a). Genome-wide association mapping for key seed metabolites using a large panel of natural and derived forms of Brassica rapa L. *Industrial Crops Prod.* 159, 113073. doi:10.1016/j.indcrop.2020.113073
- Kaur, S., Atri, C., Akhtar, J., Mittal, M., Kaur, R., and Banga, S. S. (2021b). Genetics of days to flowering, maturity and plant height in natural and derived forms of Brassica rapa L. *Theor. Appl. Genet.* 134, 473–487. doi:10.1007/s00122-020-03707-9
- Khan, S. U., Yangmiao, J., Liu, S., Zhang, K., Khan, M. H. U., Zhai, Y., et al. (2019). Genome-wide association studies in the genetic dissection of ovule number, seed number, and seed weight in Brassica napus L. *Industrial Crops Prod.* 142, 111877. doi:10.1016/j.indcrop.2019.111877
- Lee, S.-H. (1982). Vegetable crops growing in China. *Sci. Hortic.* 17, 201–209. doi:10.1016/0304-4238(82)90041-3
- Legarra, A., Robert-Granié, C., Manfredi, E., and Elsen, J.-M. (2008). Performance of genomic selection in mice. *Genetics* 180, 611–618. doi:10.1534/genetics.108.088575
- Leijten, W., Koes, R., Roobeek, I., and Frugis, G. (2018). Translating flowering time from *Arabidopsis thaliana* to Brassicaceae and Asteraceae crop species. *Plants* 7, 111. doi:10.3390/plants7040111
- Li, F., Chen, B., Xu, K., Gao, G., Yan, G., Qiao, J., et al. (2016). A genome-wide association study of plant height and primary branch number in rapeseed (Brassica napus). *Plant Sci.* 242, 169–177. doi:10.1016/j.plantsci.2015.05.012
- Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., et al. (2014). Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (Brassica napus L.). *DNA Res.* 21, 355–367. doi:10.1093/dnares/dsu002
- Li, H., Fan, Y., Yu, J., Chai, L., Zhang, J., Jiang, J., et al. (2018). Genome-wide identification of flowering-time genes in Brassica species and reveals a correlation between selective pressure and expression patterns of vernalization-pathway genes in Brassica napus. *Int. J. Mol. Sci.* 19, 3632. doi:10.3390/ijms19113632
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv.* 1303, 3997. doi:10.48550/arXiv.1303.3997
- Lian, L., and de Los Campos, G. (2016). FW: an R package for Finlay–Wilkinson regression that incorporates genomic/pedigree information and covariance structures between environments. *G3 Genes, Genomes, Genet.* 6 (3), 589–597. doi:10.1534/g3.115.026328
- Liu, J., Wang, J., Wang, H., Wang, W., Zhou, R., Mei, D., et al. (2016a). Multigenic control of pod shattering resistance in Chinese rapeseed germplasm revealed by genome-wide association and linkage analyses. *Front. Plant Sci.* 7, 1058. doi:10.3389/fpls.2016.01058
- Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016b). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12, e1005767. doi:10.1371/journal.pgen.1005767
- Long, Y., Shi, J., Qiu, D., Li, R., Zhang, C., Wang, J., et al. (2007). Flowering time quantitative trait loci analysis of oilseed Brassica in multiple environments and genomewide alignment with Arabidopsis. *Genetics* 177, 2433–2444. doi:10.1534/genetics.107.080705
- Lu, K., Peng, L., Xu, X., Qu, C., Zhang, K., Liu, L., et al. (2017). Genome-wide association and transcriptome analyses reveal candidate genes underlying yield-determining traits in Brassica napus. *Front. Plant Sci.* 8, 206. doi:10.3389/fpls.2017.00206
- Lysak, M. A., Koch, M. A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe Brassicaceae. *Genome Res.* 15, 516–525. doi:10.1101/gr.3531105
- Persson, (2017). Breeding Brassica rapa in scandinavia, GCIRC bulletins. Available at: https://www.gcirc.org/fileadmin/documents/Bulletins/B28/2.Poster_Persson_Brassica_rapa_oK.pdf (Accessed March, 2024).
- Qi, X., An, H., Ragsdale, A. P., Hall, T. E., Gutenkunst, R. N., Chris Pires, J., et al. (2017). Genomic inferences of domestication events are corroborated by written records in Brassica rapa. *Mol. Ecol.* 26, 3373–3388. doi:10.1111/mec.14131
- Raman, H., Raman, R., Coombes, N., Song, J., Prangnell, R., Bandaranayake, C., et al. (2016). Genome-wide association analyses reveal complex genetic architecture underlying natural variation for flowering time in canola. *Plant, Cell and Environ.* 39, 1228–1239. doi:10.1111/pce.12644
- Raman, H., Raman, R., Eckermann, P., Coombes, N., Manoli, S., Zou, X., et al. (2013). Genetic and physical mapping of flowering time loci in canola (Brassica napus L.). *Theor. Appl. Genet.* 126, 119–132. doi:10.1007/s00122-012-1966-8
- Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J., Coombes, N., et al. (2014). Genome-wide delineation of natural variation for pod shatter resistance in Brassica napus. *PLoS One* 9, e101673. doi:10.1371/journal.pone.0101673
- Raman, H., Raman, R., Sharma, N., Cui, X., Mcvittie, B., Qiu, Y., et al. (2023). Novel quantitative trait loci from an interspecific Brassica rapa derivative improve pod shatter resistance in Brassica napus. *Front. Plant Sci.* 14, 1233996. doi:10.3389/fpls.2023.1233996

- R Core Team (2023). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reiner, H., Holzner, W., and Ebermann, R. (1995). The development of turnip-type and oilseed-type Brassica rapa crops from the wild-type in Europe—An overview of botanical, historical and linguistic facts. *Rapeseed Today Tomorrow* 4, 1066–1069.
- Schiessl, S. V., Huettel, B., Kuehn, D., and Snowdon, R. J. (2017). Flowering time gene variation in Brassica species shows evolutionary principles. *Front. Plant Sci.* 8, 1742. doi:10.3389/fpls.2017.01742
- Schranz, M. E., Quijada, P., Sung, S.-B., Lukens, L., Amasino, R., and Osborn, T. C. (2002). Characterization and effects of the replicated flowering time gene FLC in Brassica rapa. *Genetics* 162, 1457–1468. doi:10.1093/genetics/162.3.1457
- Shen, W., Qin, P., Yan, M., Li, B., Wu, Z., Wen, J., et al. (2019). Fine mapping of a silique length- and seed weight-related gene in Brassica napus. *Theor. Appl. Genet.* 132, 2985–2996. doi:10.1007/s00122-019-03400-6
- Shi, L., Song, J., Guo, C., Wang, B., Guan, Z., Yang, P., et al. (2019). A CACTA-like transposable element in the upstream region of BnaA9. CYP 78A9 acts as an enhancer to increase silique length and seed weight in rapeseed. *Plant J.* 98, 524–539. doi:10.1111/tpj.14236
- Shyam, P., Wu, X., and Bhat, S. (2012). History, evolution, and domestication of Brassica crops. *Plant Breed. Rev.* 35, 19–84.
- Szwarc, J., Niemann, J., Bocianowski, J., Kaczmarek, J., Doğu, M. Z., and Nowicka, A. (2023). Improving the selection efficiency of breeding material within interspecific Brassicaceae hybrids with genomic prediction and phenotyping. *Agriculture* 13, 962. doi:10.3390/agriculture13050962
- Tang, H., and Lyons, E. (2012). Unleashing the genome of Brassica rapa. *Front. plant Sci.* 3, 172. doi:10.3389/fpls.2012.00172
- Tibbs Cortes, L., Zhang, Z., and Yu, J. (2021). Status and prospects of genome-wide association studies in plants. *Plant Genome* 14 (14), e20077. doi:10.1002/tpg2.20077
- Vélez-Gavilán, J. (2018). *Brassica rapa (field mustard)*. International: CAB International publisher name.
- Wanasundara, J. P., Kapel, R., and Albe-Slabi, S. (2024). "Proteins from canola/rapeseed—current status," in *Sustainable protein sources*. Editor(s) S. R. Nadathur and P. D. Janitha, Wanasundara, laurie scanlin, *sustainable protein sources*, Academic Press, Pages 285–304. doi:10.1016/B978-0-12-802778-3.00018-4
- Wang, H., Zaman, Q. U., Liu, J., Wang, W., Hao, M., Cheng, H., et al. (2019). QTL and candidate gene identification for silique length based on high-dense genetic map in Brassica napus L. *Front. plant Sci.* 10, 490218. doi:10.3389/fpls.2019.01579
- Wang, J., Fan, Y., Mao, L., Qu, C., Lu, K., Li, J., et al. (2021). Genome-wide association study and transcriptome analysis dissect the genetic control of silique length in Brassica napus L. *Biotechnol. Biofuels* 14, 214–24. doi:10.1186/s13068-021-02064-z
- Wang, J., and Zhang, Z. (2021). GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genom. Proteom. Bioinf.* 19, 629–640. doi:10.1016/j.gpb.2021.08.005
- Wei, L., Jian, H., Lu, K., Filardo, F., Yin, N., Liu, L., et al. (2016). Genome-wide association analysis and differential expression analysis of resistance to Sclerotinia stem rot in Brassica napus. *Plant Biotechnol. J.* 14, 1368–1380. doi:10.1111/pbi.12501
- Weis, A. E., and Kossler, T. M. (2004). Genetic variation in flowering time induces phenological assortative mating: quantitative genetic methods applied to Brassica rapa. *Am. J. Bot.* 91, 825–836. doi:10.3732/ajb.91.6.825
- Wu, J., Xu, X.-D., Liu, L., Ma, L., Pu, Y., Wang, W., et al. (2022). A chromosome level genome assembly of a winter turnip rape (Brassica rapa L.) to explore the genetic basis of cold tolerance. *Front. Plant Sci.* 13, 936958. doi:10.3389/fpls.2022.936958
- Würschum, T., Abel, S., and Zhao, Y. (2014). Potential of genomic selection in rapeseed (*B. L.*) breeding. *Plant Breed.* 133, 45–51. doi:10.1111/pbr.12137
- Xiao, D., Zhao, J. J., Hou, X. L., Basnet, R. K., Carpio, D. P., Zhang, N. W., et al. (2013). The Brassica rapa FLC homologue FLC2 is a key regulator of flowering time, identified through transcriptional co-expression networks. *J. Exp. Bot.* 64, 4503–4516. doi:10.1093/jxb/ert264
- Xin, S., Dong, H., Yang, L., Huang, D., Zheng, F., Cui, Y., et al. (2021). Both overlapping and independent loci underlie seed number per pod and seed weight in Brassica napus by comparative quantitative trait loci analysis. *Mol. Breed.* 41, 52. doi:10.1007/s11032-021-01232-1
- Yang, P., Shu, C., Chen, L., Xu, J., Wu, J., and Liu, K. (2012). Identification of a major QTL for silique length and seed weight in oilseed rape (Brassica napus L.). *Theor. Appl. Genet.* 125, 285–296. doi:10.1007/s00122-012-1833-7
- Yang, Y., Shi, J., Wang, X., Liu, G., and Wang, H. (2016). Genetic architecture and mechanism of seed number per pod in rapeseed: elucidated through linkage and near-isogenic line analysis. *Sci. Rep.* 6, 24124. doi:10.1038/srep24124
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P.-C., Hu, L., et al. (2016). Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* 48, 927–934. doi:10.1038/ng.3596
- Zhang, L., Cai, X., Wu, J., Liu, M., Grob, S., Cheng, F., et al. (2018). Improved Brassica rapa reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Hortic. Res.* 5, 50. doi:10.1038/s41438-018-0071-9
- Zhang, X., Meng, L., Liu, B., Hu, Y., Cheng, F., Liang, J., et al. (2015). A transposon insertion in flowering locust is associated with delayed flowering in Brassica rapa. *Plant Sci.* 241, 211–220. doi:10.1016/j.plantsci.2015.10.007
- Zheng, M., Peng, C., Liu, H., Tang, M., Yang, H., Li, X., et al. (2017). Genome-wide association study reveals candidate genes for control of plant height, branch initiation height and branch number in rapeseed (Brassica napus L.). *Front. Plant Sci.* 8, 1246. doi:10.3389/fpls.2017.01246
- Zhou, Q., Han, D., Mason, A. S., Zhou, C., Zheng, W., Li, Y., et al. (2017). Earliness traits in rapeseed (Brassica napus): SNP loci and candidate genes identified by genome-wide association analysis. *DNA Res.* 25, 229–244. doi:10.1093/dnares/dsx052