Received: 13 May 2019 | Revised: 26 August 2019 | Accepted: 28 September 2019

DOI: 10.1111/pbr.12771

ORIGINAL ARTICLE

Plant Breeding WILEY

Genome-wide association study for frost tolerance in rapeseed/canola (*Brassica napus*) under simulating freezing conditions

Danielle F. Wrucke | Zahirul I. Talukder | Mukhlesur Rahman

Department of Plant Sciences, North Dakota State University, Fargo, ND, USA

Correspondence
Mukhlesur Rahman, Department of Plant Sciences, North Dakota State University, Fargo, ND, USA
Email: md.m.rahman@ndsu.edu

Funding information
Northern Canola Growers Association, Grant/Award Number: NCGA-2013; U.S. Department of Agriculture-National Institute of Food and Agriculture, Grant/Award Number: ND01581

Communicated by: Rod Snowdon

Abstract

Rapeseed/canola seedlings can be easily damaged by spring frost, which can rupture the cells and kill the plant. Genetic variations for frost tolerance have known to exist within rapeseed/canola gene pool. A genome-wide association study (GWAS) was conducted using 231 diverged rapeseed/canola germplasm to find the significant markers of the freezing tolerance traits. The genotypes were obtained from 21 countries and comprised of spring, winter and semi-winter growth types. The genotypes were evaluated in plant growth chamber under simulated freezing conditions. Highly significant genotypic variation was observed for the freezing tolerance. The best three freezing tolerant germplasms (Rubin, KSU-10, and AR91004) were winter type, while the four most freezing susceptible germplasms (Polo Canada, Prota, Drakkar, and BO-63) were all spring type. No geographical or growth habit type clusters were identified by structure analysis in this mixed population. One QTL was identified that was located on chromosome A02. Six freezing/abiotic stress tolerance genes have been identified in this study.

KEYWORDS
Brassica napus, freezing tolerance, GWAS, simulated conditions

1 | INTRODUCTION

Rapeseed/canola (*Brassica napus* L.) is one of the most important oilseed crops in the world and second to soybean (Foreign Agriculture Services, United States Department of Agriculture, National Agricultural Statistics Service, 2019). It is an amphidiploid (AACC, $2n = 4x = 38$) and evolved through spontaneous interspecific hybridization followed by genome duplication, chromosome fission, fusion and rearrangements between two diploid species, *B. rapa* L. (AA, $2n = 2x = 20$) and *B. oleracea* L. (CC, $2n = 2x = 18$) (Lagercrantz, 1998; Lagercrantz & Lydiate, 1996; U, 1935). The genome size of *B. napus* is about 1,130 Mb, consistent with the combined genome sizes of *B. rapa* and *B. oleracea* (Chalhoub et al., 2014). Three growth habits types are found in canola: spring, winter and semi-winter. Spring type canola is typically grown in North Dakota, because the winter canola is not hardy enough to survive in the harsh winters that are experienced in North Dakota. The spring canola is planted in the spring and harvested during the same growing season (Rahman & McClean, 2013). The winter type mostly grown in Western Europe is seeded in the fall, vernalized during the winter to induce flower in spring and harvested in the summer (Rahman & McClean, 2013). Semi-winter type is not as hardy as the winter type, and a shorter period of vernalization is required to induce flower (Wang et al., 2011). The seed yield of spring canola in the United States is about 1,889 kg/ha, the winter canola in Western Europe is about 3,597 kg/ha, and the semi-winter canola in China is about 1,964 kg, (5 years average from 2013 to 2017; FAOSTAT, Food and Agriculture Organization of the United Nations, accessed on May, 2019, http://www.fao.org/faostat/en/#data/QC). North Dakota is the largest producer

of canola in the United States with about 84% of acreage, 85% of total production, and valued at about $428 million dollars (5 years average from 2014 to 2018; USDA-NASS, accessed of February 2019, https://www.nass.usda.gov).

Early planting is a concern for growers in North Dakota (ND) and Canada because of the potential threat of a freezing temperature shortly after planting. Plant growth and crop production can be severely affected by freezing temperatures, and the whole crop can be destroyed (Chinnusamy, Zhu, & Zhu, 2007). Frost tolerance cultivars would allow growers to plant earlier and utilize early season moisture, compete with early weeds and avoid heat during flowering time, which can reduce seed yield. In North Dakota, usually 25 May is considered as frost-free date, but frost can happen later as well. The cotyledon stage of canola is more susceptible to frost than the three to four leaf stages, which makes canola most susceptible in the early spring when the risk of frost is higher (Fiebelkorn & Rahman, 2016). The severity of the injury depends on many factors, including moisture, growth stage and duration of the temperature. Frost damage can have different symptoms such as plant death, bleaching or wilting of the plant. The loss of water in the leaves resulted in seedling wilting. Bleaching is due to the phytooxidation of leaf pigments (Wilson, 1997). Plant cells can be ruptured during freezing temperature due to ice formation, and thus, the plant can be killed. Artificial freezing simulating conditions can be used to evaluate canola germplasm for frost tolerance year-round and is not dependent on outside environmental conditions.

Genome-wide association study (GWAS) is a powerful method compared with traditional QTL mapping and is often used for quantitative traits (Rafalski, 2002; Sun et al., 2014; Visioni et al., 2013; Waugh, Jannink, Muller, & Ramsay, 2009; Yu & Buckler, 2006). GWAS examines genetic variation throughout the whole genome to find association signals for quantitative traits (Risch & Merikangas, 1996; Zhu et al., 2008). An important concept of GWAS is linkage disequilibrium, which is the non-random association between traits and alleles at different loci (Zhu et al., 2008). A diverse population with many historical recombination is often used for GWAS, and common trait-linked alleles can be identified from the population.

The objective of the study was to screen a wide collection of *B. napus* germplasm accessions for frost tolerance under artificial freezing conditions and to find the genomic regions that control the frost tolerant traits. The study will give an access to select freezing tolerant germplasm for crossing with our elite breeding lines and will allow developing molecular markers for marker-assisted breeding programme.

2 | MATERIALS AND METHODS

2.1 | Plant materials

A diversity panel, consisting of 231 *B. napus* genotypes, was used for association analysis in this study. The germplasms were obtained from the USDA-ARS, Germplasm Resources Information Network (GRIN) (http://www.ars-grin.gov/npgs/searchgrin.html), and were originated from 21 countries on four continents (Table S1, Figure 1). The germplasm panel consisted of three growth habit types: spring, winter and semi-winter. The spring growth type included 95, winter type contained 95, and semi-winter had 41 germplasm (Table S1). Germplasm accessions were self-pollinated in the greenhouse for 4–5 generations before being used in this study.

2.2 | Phenotypic evaluation

The germplasm accessions were grown in 4" × 6" plastic pots in greenhouse using a randomized complete block design with three replications. Each replication had six plants per genotype. The temperature of the greenhouse was maintained at 20°C with a 16-hr photoperiod provided by natural sunlight supplemented with 400 W HPS PL 2000 lights (P.L. Light Systems Inc.). The plants were fertilized using 20–20–20 (N-P-K) liquid fertilizer. After 14 days, the seedlings were moved to a walk-in plant growth chamber (BioCold line of Environmental Rooms, Innovative Laboratory Systems Inc.) for cold acclimation at 4°C with a 12-hr photoperiod for seven days. Light was provided by GE Ecolux F32T8 SP35 Eco (32 W T8) style bulbs (General Electric Company). Artificial freezing conditions were

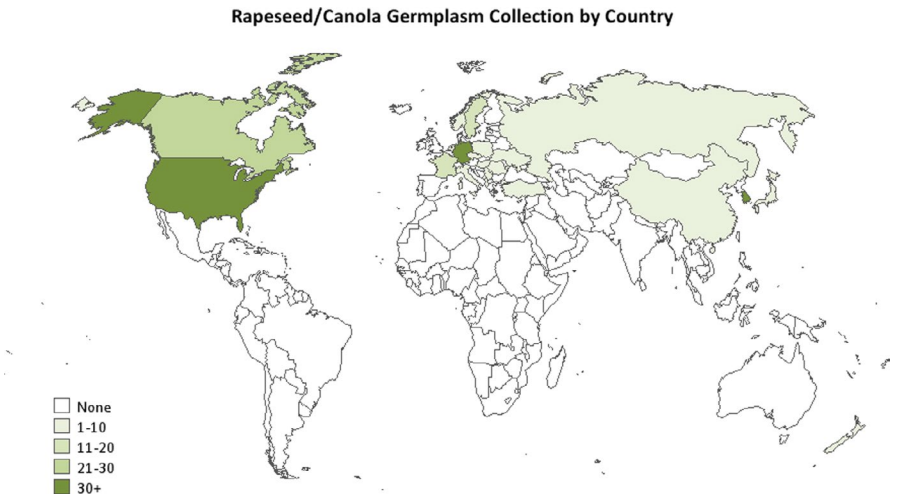

FIGURE 1 Geographical distribution of *B. napus* germplasm accessions used in this study

created in the plant growth chamber where the temperature was slowly lowered from 4°C to − 8°C to reach the treatment temperature. The seedlings were kept at − 8°C for 8.0 hr, and then, the temperature was slowly raised to 4°C. The treated seedlings were kept in the growth chamber at 4°C for 24 hr and then were brought back to the greenhouse for seedling damage scoring. The scoring was conducted three times at three, six and nine days after the freezing treatment. The plants were scored visually for freezing damage using a 0–5 scale (Fiebelkorn & Rahman, 2016), where 0 = dead plant, while 5 = no damage. The same experiment was repeated two times. Thus, 36 (six plants × three reps × two times) data points were generated per germplasm accession. The scoring at three days after freezing treatment was used according to Fiebelkorn and Rahman (2016) for genome-wide association study.

2.3 | Estimation of relative effect

The relative effect varies between 0 and 1 for the different genotypes. A relative effect closer to 1 means that the genotypes are more tolerant to frost. The overall median is used to calculate the mean rank. The mean rank of the genotypes was also calculated by ranking all the medians from low to high and taking the mean of the ranks (Wrucke, Mamidi, & Rahman, 2018).

2.4 | Statistical analysis

Data were analysed using SAS® 9.3 (SAS Institute Inc.). Medians were used to calculate the one-way analysis of variance (ANOVA) using non-parametric method. The estimated relative effect and the 95% confidence interval were also calculated using SAS® 9.3 (SAS Institute Inc.). Broad-sense heritability (H^2) was estimated on an entry mean basis following Nyquist (1991):

$$H^2 = \sigma_g^2 \left(\sigma_g^2 + \sigma_{ge}^2 / l + \sigma_e^2 / lr \right),$$

where σ_g^2 is the genotypic variance, σ_{ge}^2 is the genotype × environment variance, σ_e^2 is the error variance, r is the number of replications, and l is the number of environments. Data normality was tested using the Shapiro–Wilk normality test using R 3.2.4 (R Core Team, 2016).

2.5 | DNA extraction and SNP genotyping

Total genomic DNA was extracted from leaves of 4-leaf stage seedling per germplasm using DNeasy 96 Plant Kit (Qiagen Inc.) following manufacturer's protocol. Genotyping-by-sequencing (GBS) was conducted at the Institute of Genome Diversity (IGD) at Cornell University. The GBS libraries were prepared and analysed according to Elshire et al. (2011). The enzyme used for digestion was *Ape*KI, and 96 unique barcodes were used to create the GBS libraries. The genome sequencing was done on Illumina GAII sequencer (Illumina Inc.). The sequence tags were aligned using the Burrow–Wheelers Alignment tool bwa-mem (Li, 2013) to the Canola reference genome (Chalhoub et al., 2014), and multi-sample SNP calling was conducted using VarScan (Koboldt et al., 2012). A total of 42,575 SNPs were obtained for a collection of 231 genotypes. Markers with minor allele frequency of less than 5% were removed leaving a final data set of 37,699 SNPs for further analysis. Missing genotypes in the resulting data set were imputed with fastPHASE 1.3 software (Scheet & Stephens, 2006) using default settings.

2.6 | Linkage disequilibrium (LD) decay

Linkage disequilibrium (LD) decay between pairs of SNPs was estimated using squared allele frequency correlations (r^2) in TASSEL v5.0 (Bradbury et al., 2007). LD decay trends among markers were estimated assuming a non-linear mutation recombination–drift model of LD (r^2) versus the pairwise genetic distance (cM) (Hill & Weir, 1988; Remington et al., 2001). We fit this equation into a non-linear regression model using the statistical package R version 3.3.1 (R Core Team, 2016). LD decay was also calculated for each subgenome (A and C) and individual chromosomes. According to Mohammadi et al. (2015), the $r^2 > .2$ was used to identify markers located on the same QTL.

2.7 | Model selection

Principal component analysis was used to control the population structure (Price et al., 2006). TASSEL 5.0 (Bradbury et al., 2007) was used to estimate the principle components (PCs). Cumulative variation that explained 25% and 50% was used in the regression model. To account for relatedness between genotypes, an identify by state matrix was calculated in TASSEL (Zhao et al., 2007). Six models, naïve, PC_3, PC_{29}, Kinship, PC_3 + Kinship and PC_{29} + Kinship, were used. The best model was used based on the lowest mean square difference (MSD). Observed and expected p-values were used to determine the MSD (Mamidi et al., 2011).

2.8 | Association mapping

Genome-wide association analyses were performed in TASSEL v5.0 (Bradbury et al., 2007). Six different GWAS models were tested: (a) a naïve analysis using only the general linear model (GLM), $y = X\alpha + e$; (b) a GLM analysis that accounted for first three principal components (PC_3) as cofactor (GLM + PC_3), $y = X\alpha + P_3\beta_3 + e$; (c) a GLM analysis that accounted for first 29 principal components (PC_{29}) as cofactor (GLM + PC_{29}), ($y = X\alpha + P_{29}\beta_{29} + e$); (d) a mixed linear model analysis that accounted for kinship relatedness as a cofactor (MLM), $y = X\alpha + K\mu + e$; (e) a MLM analysis in which both PC_3 and kinship were used as cofactors (MLM + PC_3), $y = X\alpha + P_3\beta_3 + K\mu + e$, and (f) a MLM analysis in which both PC_{29} and kinship were used as cofactors (MLM + PC_{29}), $y = X\alpha + P_{29}\beta_{29} + K\mu + e$ (Zhang et al., 2010). In the equations, y is the phenotype, X is the SNP genotype matrix, α is the vector of marker effects, K is the relative kinship matrix determined from the marker data, μ is the vector of kinship effects, P is the population structure matrices of the principal components,

β is the vector of population structure effects, and e is the vector of residual effects. $X\alpha$ and $P\beta$ represent fixed effects, and $K\mu$ and e represent random effects. The significant markers were identified based on the p-value of the marker and are within the 0.1 percentile tail of 10,000 bootstraps (Gurung et al., 2014; Mamidi, Lee, Goos, & McClean, 2014). Stepwise regression was conducted on the significant markers to estimate the combined variation (r^2) explained by these significant markers as well as to define the major QTL (Gurung et al., 2014; Mamidi et al., 2014). A Manhattan plot was calculated using the $-\log_{10}(p)$ values using R 3.2.4 (The R Foundation). The observed verses expected $-\log_{10}(p)$ values were plotted using R 3.2.4 (The R Foundation) to make a Q–Q plot.

2.9 | Candidate gene search

The annotations of the selected markers were determined using the TAIR 10 protein database. The selected marker included a 100 kb region sequence on both sides. The selected regions were blasted with the *Arabidopsis thaliana* proteome to find candidate genes. The genome sequence and gene model were used based on Chalhoub et al. (2014). Published literature was used to identify the functions related to frost tolerance or abiotic stress responses for the selected annotations.

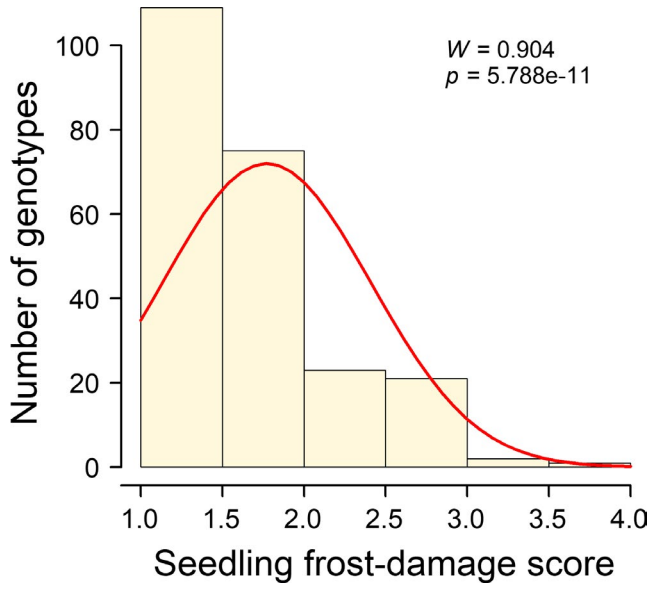

FIGURE 2 Phenotypic distribution of frost induced seedling damage under greenhouse conditions. The seedling damage score is on *x*-axis, and the number of genotypes for the scores is presented on *y*-axis. The curve represents the non-normal distribution of the damage score in the population

3 | RESULTS

3.1 | Phenotypic results

A median of thirty-six readings was considered as phenotypic score of freezing induced seedling damage. A wide variation in damage scores was observed in this study. The overall median score varied from 1.0 to 4.0 with a common median of 1.75 (Table S1). The scores were significantly different at $\alpha = 0.001$ level. The top three freezing tolerant genotypes were Rubin (4.0), KSU-10 (3.5) and AR91004 (3.3), and they belonged to winter type. Eighteen genotypes scored 3.0, of them 11 were winter type, three semi-winter type and four spring type. In contrary, fifty-seven genotypes scored the lowest (1.0), of which 35 were spring type, 11 were semi-winter, and 11 were winter type (Table S1). The broad-sense heritability was 0.54, and the standard error was 0.065. The Shapiro–Wilk normality test indicates a non-normality as the p-value is less than .05 (p = 5.788e-11). The number of genotypes that fit each value was plotted using the frequency of the medians for the study (Figure 2).

3.2 | Estimated relative effect

The estimated relative effect and the confidence interval (95%) were calculated based on the medians and indicated that 95% of the time the genotypes belonged to respective range (Figure 3, and Table S1). A higher mean rank means the genotype performed better, while a lower mean rank means the genotype is more susceptible to frost. The genotypes with higher median, mean rank and estimated relative effect indicated that the accessions have higher frost tolerance potentiality under controlled simulated freezing conditions.

3.3 | Linkage disequilibrium (LD) decay

The LD decay was calculated in the A and C genomes as well as individual chromosomes (A01-C09). The partial squared allele frequency correlation coefficient (r^2) is commonly used as a measure for LD mapping and to quantify and compare LD. In this study $r^2 = 0.2$ was used as the cut-off. The physical distance (kb) for LD decay in the

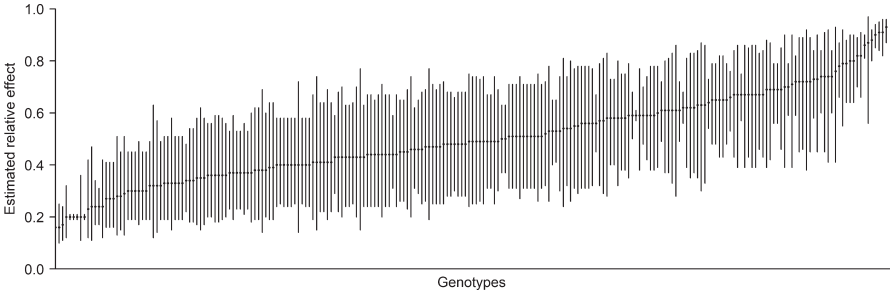

FIGURE 3 Estimated relative effect and 95% confidence interval for the medians from the greenhouse study. Genotypes towards the right side of the graph are closer to 1 and considered more frost tolerant while the genotypes on the left side are closer to 0 and considered less frost tolerant

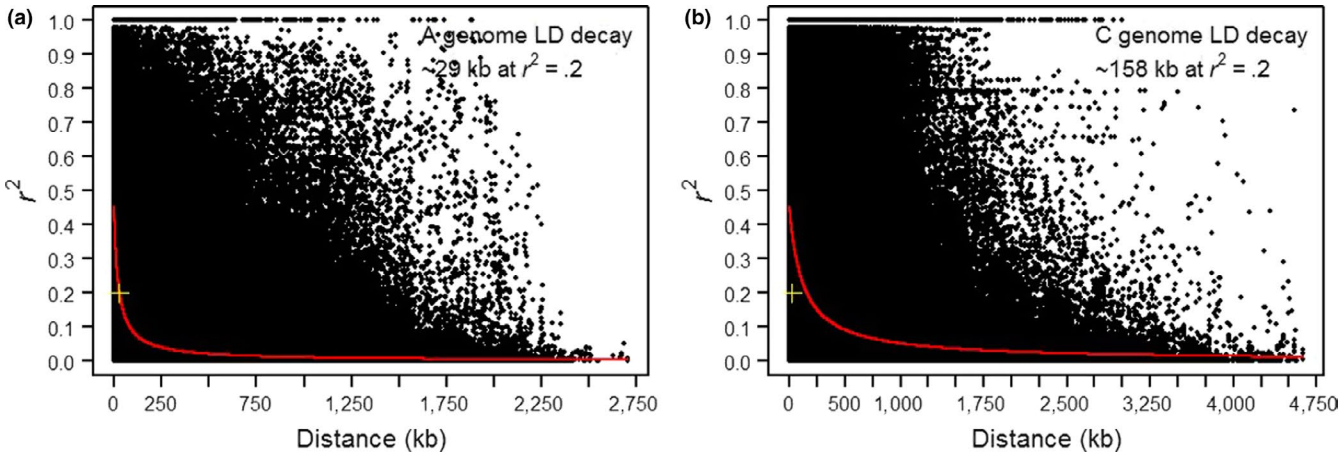

FIGURE 4 Genome-wide LD decay in the A (left) and C (right) subgenomes for 231 *B. napus* accessions. The black dots are the r^2 values of all SNP pairs within chromosomes for each subgenome. The red curve is the non-linear regression trend of r^2 against the SNP physical distance (Mb). The yellow plus sign is the estimated background level of LD (The LD decay of A genome at r^2 = .2 is about 29 kb, LD decay of C genome at r^2 = .2 is about 158 kb)

A subgenome was about 29 kb (Figure 4) and in the C subgenome was about 158 kb (Figure 4). LD decay of each of the 19 chromosomes was calculated using the same cut-off (r^2 = 0.2). The A subgenome had shorter distances in kb for LD decay than the C genome (Figure 5). The LD decay in the A subgenome ranged from 18 to 52 kb (Figure 5: A1 – A10), while the C subgenome ranged from 61 to 364 kb (Figure 5: C1 – C9).

3.4 | Markers, minor allele frequency and genome-wide association scan

The total 42,575 SNPs markers were filtered for minor allele frequencies (<5%) which left a total of 37,699 SNPs for association mapping study. Cumulative variation of 25 and 50% were accounted for using PC_3 and PC_{29}. These were used to control population structure in the model. The first two PCs, PC_1 and PC_2, explained about 13% and 9% of variation, respectively. The PCs clustered the genotypes into three groups (Figure 6). The clusters were random and not by growth habit or geographical locations.

Of the six models tested, the PC_{29} + Kinship (where 50% of the variation is explained and relatedness) model was selected as the best model based on the MSD calculation. Based on 10,000 bootstraps and the 0.1% percentile of the empirical distribution, 38 markers were selected and considered as significant markers (Table 1). The cut-off *p*-value was .001146. One SNP (chrA02_21498544) (Table S2) was selected using stepwise regression located on chromosome A02 (21.5 Mbp) and showed about 3% of the phenotypic variation.

A Manhattan plot was calculated using the $-\log_{10}(p)$ values (Figure 7). The significant markers are above the dashed blue line and are at $-\log_{10}(p)$ = 2.94. Markers present on Ann_rand, Cnn_rand and Unn_rand were removed from the analysis because they were

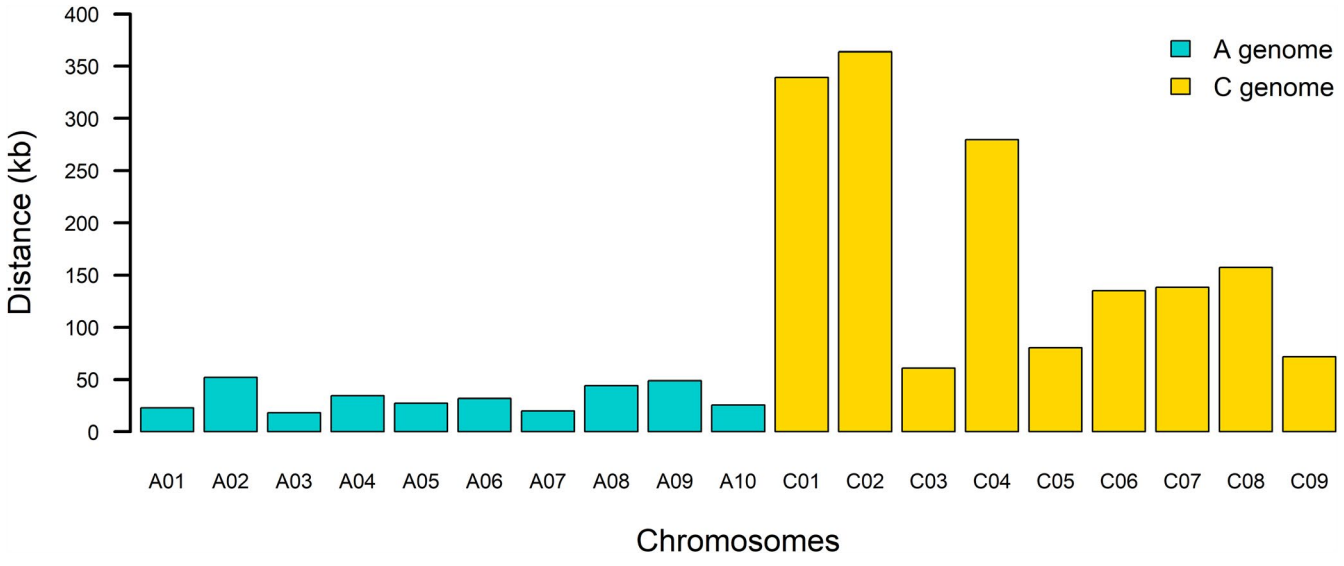

FIGURE 5 Comparison of the average size of LD blocks on the chromosomes of A and C subgenomes

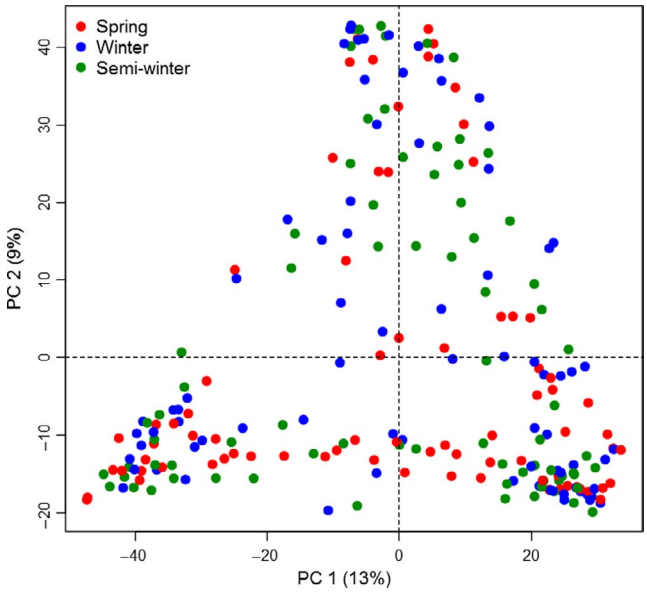

FIGURE 6 PCA graph showing the distribution of the first two principle components using the medians from freezing induced seedling damage. PC_1 explains 13% of the variation, and PC_2 explains 9% of the variation. PC_1 is on the *x*-axis, and PC_2 is on the *y*-axis

not assigned to any chromosome based on the reference literature of *B. napus* genome (Chalhoub et al., 2014).

The Q–Q plot indicates the fitness of the model based on observed and expected $-\log_{10}(p)$ values (Figure 8). The SNP (chrA02_21498544) identified during stepwise regression is highlighted in green and labelled on the plot.

A 100 kb region on each side of the significant marker (chrA02_21498544) identified from stepwise regression was blasted with the *Arabidopsis* proteome (TAIR 10). Genes that were annotated were subjected to a literature search to identify functions and genes that were related to frost or abiotic stress (Table 2). The genes functions ranged from cell wall metabolism and repressed under stress conditions, ABA degradation, DREB2A ubiquitination mediation and transcription factors involved in responses to abiotic stresses.

4 | DISCUSSION

In North Dakota, freezing tolerance traits are important trait for canola and other crop too. Canola is typically grown in the northern part of the state alone with Canadian border and in Canada where spring frost is a concern. Frost tolerance is a difficult trait to study in natural conditions due to unpredictability of the environmental conditions. Therefore, germplasm screening under simulated freezing conditions allowed us to reliably control the temperature, humidity, light and moisture in the plant growth chamber. Strigens et al. (2013) conducted chilling experiments in both the field and the growth chamber and reported that the growth chamber study was reliable to overcome the fluctuating temperature experienced in the field. Various methods can be to study the freezing tolerance in plants, such as plant tissue water content (Brule-Babel & Fowler, 1988), ion leakage from cold-stressed plant cells (Teutonico, Palta, & Osborn, 1993), changes in luminescence (Brzostowicz & Barcikowska, 1987) and electrolytes leakage (Fiebelkorn, Horvath, & Rahman, 2018) for freezing induced phenotypic scoring. These authors used various freezing temperatures on plant leaf samples. However, in this study, we used whole seedling for freezing induced seedling damage score according to our earlier research (Fiebelkorn & Rahman, 2016). We have developed a protocol to screen canola germplasm for freezing tolerance using a smaller freezing chamber (4 cubic foot, model: ESPEC BTU-433, ESPEC North America, Inc.). The suggested protocol for the freezing tolerant screening is cold acclimation of 14-day-old seedlings for 7 days at 4°C followed by freezing treatment at −4°C for 8 hr (Fiebelkorn & Rahman, 2016). This protocol did not work well in the current larger plant growth chamber (960 cubic foot). Therefore, we used the freezing treatment at −8°C for 8 hr for better seedling damage scores. It is logical that the smaller chamber has better temperature control over the larger chamber. However, in the breeding programme larger chamber is needed for large-scale germplasm screening. We used 14-day-old seedlings for the freezing tolerance screening. At this stage, the seedlings are more susceptible to freezing temperature and expected to observe the most variation in seedling damage score in response to different freezing temperatures. Moreover, this is the high-risk stage for frost damage for canola growers in North Dakota. We have repeated the experiment two times, which helped us to correctly phenotype the germplasm and reduce the experimental error, which is the most important parameter for the association study. In this study, we used median scores instead of arithmetic means according to Williams & Delwiche, 1980). The freezing tolerance scores used in this study were categorical data, and therefore, the median of scores was calculated for each replication within the germplasm accessions. Rahman et al. (2016) used median data for blackleg disease scoring, and Zou et al. (2017) also used median data for frost tolerance score in rapeseed/canola.

4.1 | Genome-wide association study (GWAS)

GWAS is a powerful tool to identify marker–trait associations and has been used for many mapping projects (Jia et al., 2012; Li et al., 2011). This method aims to identify common alleles in a diverse population. Traditional QTL mapping can also be used to identify trait-linked markers. However, GWAS is a population-based screening method, which is based on LD in crop species to identify trait–marker relationships that significantly increased the power to localize the QTL (Kraakman, Niks, Berg, Stam, & Eeuwijk, 2004; Meuwissenk & Goddard, 2000). GWAS utilizes a diverse population, which has wider genetic variability generated over many rounds of historical recombination (Hansen, Kraft, Ganestam, Säll, & Nilsson, 2001; Kraakman, Martínez, Mussiraliev, Eeuwijk, & Niks, 2006). In this study, we used a wide collection of diverged 231 germplasm accessions for GWAS. There is no clear consensus on the number of genotypes to be used for GWAS. However, GWAS studies have been conducted in *B. napus* using 49 genetically diverse

TABLE 1 Significant markers for the freezing induced seedling damage under greenhouse conditions selected based on the 0.1 percentile of the Empirical distribution. The reference allele, alternate alleles and the heterozygote had the means calculated for each marker

Marker	Chromosome	Position	Reference allele/ Alternate allele	Reference allele			Alternate allele			Heterozygote (H)		
				Obs	Mean	Effect	Obs	Mean	Effect	Obs	Mean	Effect
chrA01_576522	A01	576522	G/A	140	1.81	0.48	39	1.76	0.02	51	1.65	0
chrA01_22991258	A01	22991258	T/C	75	1.69	−0.21	60	1.72	−0.57	95	1.85	0
chrA02_21498544*	A02	21498544	A/T	194	1.72	0.27	17	2.1	1.17	19	1.91	0
chrA03_20587761	A03	20587761	C/T	140	1.78	−0.3	38	1.86	0.47	52	1.65	0
chrA04_5650012	A04	5650012	G/A	191	1.73	−0.33	19	1.79	0.5	20	2.05	0
chrA04_11707842	A04	11707842	T/G	97	1.83	−0.25	54	1.72	0.25	79	1.71	0
chrA07_10110972	A07	10110972	G/A	130	1.74	0.28	40	1.86	0.7	60	1.75	0
chrA07_13505270	A07	13505270	G/A	190	1.71	−0.31	19	1.95	0.54	21	2.07	0
chrA07_21672105	A07	21672105	G/A	170	1.77	0.55	28	1.74	−0.41	32	1.74	0
chrA07_rand_1125045	A07_rand	1125045	A/G	192	1.72	−0.06	16	1.66	−0.82	22	2.24	0
chrA08_5411182	A08	5411182	G/A	88	1.81	−0.09	52	1.76	0.38	90	1.72	0
chrA09_8152488	A09	8152488	C/A	152	1.79	−0.09	35	1.83	−0.89	43	1.62	0
chrA09_8867241	A09	8867241	T/G	83	1.82	0.58	61	1.7	0.01	86	1.76	0
chrA09_rand_1053698	A09_rand	1053698	A/C	102	1.8	0.04	44	1.74	−0.46	84	1.73	0
chrA10_4985855	A10	4985855	A/C	149	1.79	0.74	34	1.85	0.12	47	1.62	0
chrA10_10273633	A10	10273633	T/G	148	1.73	0.06	38	1.68	1.39	44	1.95	0
chrA10_12339004	A10	12339004	C/T	106	1.76	−0.39	53	1.88	0.29	71	1.68	0
chrAnn_rand_4644594	Ann_rand	4644594	C/T	127	1.69	−0.01	47	1.8	1.02	56	1.91	0
chrC03_4962575	C03	4962575	A/C	100	1.8	0.3	58	1.67	0.87	72	1.79	0
chrC03_58989086	C03	58989086	C/A	155	1.83	−0.27	35	1.65	1.14	40	1.63	0
chrC05_4660254	C05	4660254	A/T	166	1.74	−1.01	30	1.78	−0.61	34	1.87	0
chrC05_12310312	C05	12310312	A/C	150	1.82	0.89	34	1.71	0.43	46	1.63	0
chrC05_17437784	C05	17437784	C/A	142	1.71	0.25	43	1.85	1.11	45	1.86	0
chrC05_31062008	C05	31062008	C/T	93	1.74	0.4	56	1.78	0.37	81	1.78	0
chrC05_42883306	C05	42883306	T/A	93	1.77	0.38	47	1.83	−0.22	90	1.73	0
chrC06_2107349	C06	2107349	A/T	184	1.71	−0.9	20	1.8	−0.68	26	2.09	0
chrC07_42313730	C07	42313730	A/C	147	1.72	na	42	1.84	−0.34	41	1.84	0
chrC09_4883671	C09	4883671	G/A	120	1.69	−0.15	53	1.82	1.34	57	1.88	0
chrC09_31942271	C09	31942271	C/T	125	1.79	0.44	49	1.75	−0.56	56	1.73	0
chrC09_45401975	C09	45401975	A/C	98	1.7	−0.54	52	1.75	−0.61	80	1.85	0
chrCnn_rand_73744305	Cnn_rand	73744305	T/C	122	1.71	−0.28	48	1.78	0.62	60	1.87	0
chrCnn_rand_80593610	Cnn_rand	80593610	T/C	150	1.72	−0.91	39	1.79	−0.37	41	1.91	0

Note: Obs, observations; na, not calculated.
*Marker identified through stepwise regression.

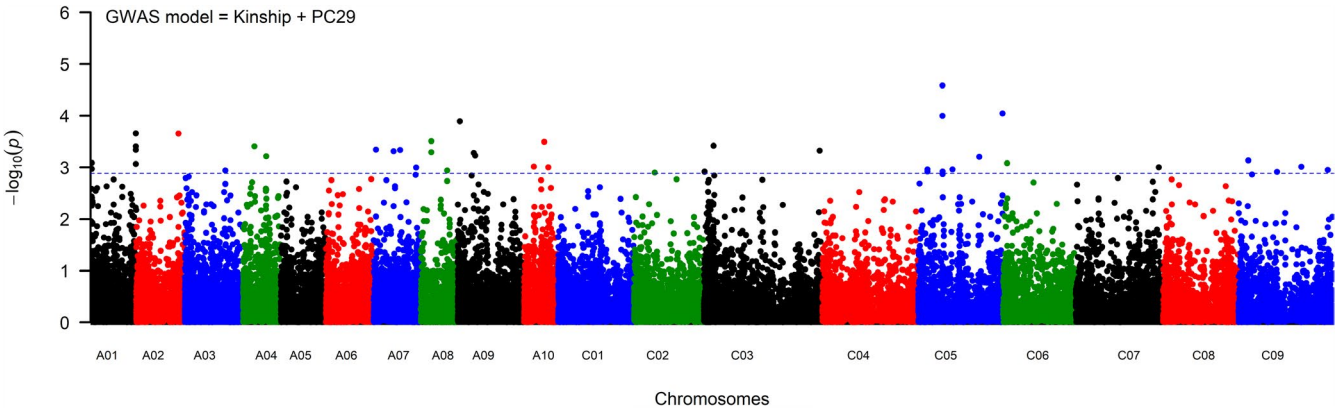

FIGURE 7 Manhattan plot of marker–trait association analysis using Kinship + PC29 model for freezing induced seedling damage score under greenhouse conditions of 231 *B. napus* accessions. The *x*-axis is the chromosomes, and the *y*-axis is the −log10(*p*). The significant SNPs above the horizontal dashed line are chosen through bootstrapping and based on the empirical distribution

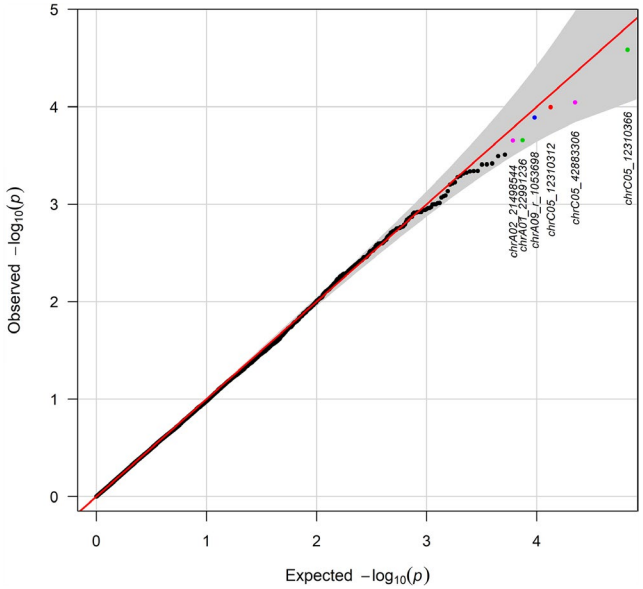

FIGURE 8 Q–Q plot showing the distribution of the observed and expected −log10 *p*-values. The *x*-axis is the expected *p*-values, and the *y*-axis is the observed *p*-values. The red line represents the expected *p*-values with no association existed, and the black line/dots represent observed *p*-values

winter-type germplasm for seed phenolic compounds (Rezaeizad et al., 2011), 89 winter-type accessions were used for six seed quality traits (Gajardo et al., 2015), 139 spring-, semi-winter- and winter-type germplasm were used for blackleg disease (Rahman et al., 2016), 143 spring-, semi-winter- and winter-type accessions were used for branch angle (Liu et al., 2016), 172 *B. napus* germplasm was used for seed oil content (Zou et al., 2010), 248 winter-type germplasm was used for seed germination and early vigour (Hatzig et al., 2015), 405 accessions of winter oilseed, winter fodder, swede, semi-winter, spring, spring fodder and vegetable types were used for agronomic and seed quality traits (Körber et al., 2016), and 540 accessions of winter, semi-winter, spring and vegetable types were used for fatty acid profiles (Qu et al., 2017).

A moderate broad-sense heritability (54%) was identified for the trait. This indicates that the progeny could inherit the trait from the parents. Frost tolerance is believed to be controlled by a large number of genes, which is indicated by the heritability estimate too. In recent years, the invention of newly developed sequencing technologies has dramatically reduced the genotyping cost, which enable the researchers to go for genomic selection in breeding programme. The genomic selection is a powerful tool to estimate for genetic values of high, moderate or low heritable traits (Zhang, Yin, Wang, Yuan, & Liu, 2019). As phenotyping of frost tolerant is a difficult trait that showed a moderate heritability, however, it could be used in genomic selection for freezing tolerance study.

The population structure analysis indicated that three clusters were present in the population which showed an agreement with Hasan et al. (2008), Bus, Körber, Snowdon, and Stich (2011), Qian, Qian, and Snowdon (2014), Raman et al. (2014), Li et al. (2014) and Wang et al. (2014) in *B. napus*. However, the individuals in the subpopulation did not show any growth habit types or geographical distribution in respect to the origin. This mixed clustering of growth habit in *B. napus* showed an agreement with Li et al. (2014) and Wang et al. (2014). In the contrary, growth habit-based clustering in *B. napus* was reported by Bus et al. (2011), Qian et al. (2014) and Raman et al. (2014).

False positives can be observed due to confounding effects during the analysis. False positives are often seen in populations that have structure or relatedness. Different models can be used to account for structure, relatedness or both (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000; Pritchard, Stephens, Rosenberg, & Donnelly, 2000; Yu et al., 2006). Six models, naïve, PC_3, PC_{29}, kinship, kinship + PC_3 and kinship + PC_{29}, were used in this study. The PC models were used to account for structure, kinship was used for relatedness, and PC + Kinship was used for both structure and relatedness. The best model based on the lowest mean square difference (MSD) was PC_{29} + Kinship, which accounts for both structure and relatedness. The empirical

TABLE 2 Candidate genes for freezing/cold tolerance within 100kb region at either side of the significant markers

Brassica gene model	Marker	Gene start (bp)	Gene end (bp)	Marker distance from gene (bp)	A. thaliana equivalent	Gene annotation	Function	Reference
BnaA02g29330D	chrA02_21498544	2,14,53,042	2,14,55,195	43,349	AT5G67150	HXXXD-type acyl-transferase family protein	Involved in cell wall metabolism and repressed under stress conditions	Zhu et al. (2013)
BnaA02g29380D	chrA02_21498544	2,14,78,746	2,14,80,637	17,907	ARALYDRAFT_484649	Cytochrome P450 family protein	Involved in ABA degradation during seed imbibition and dehydration stress	Shinozaki and Yamaguchi-Shinozaki (2007)
BnaA02g29510D	chrA02_21498544	21,546,6	2,15,49,271	48,128	AT2G03090	Expansin	Loosens cell walls and disrupts hydrogen bonds when expression is increased under abiotic stress	Lu et al. (2011)
BnaA02g29490D	chrA02_21498544	2,15,41,176	2,15,42,959	42,632	AAM64792	MYB family transcription factor	Transcription factor involved in responses to abiotic stresses	Hong et al. (2013)
BnaA02g29620D	chrA02_21498544	2,15,90,163	2,15,91,088	91,619	AT3G29020	MYB transcription factor	Involved in responses to biotic and abiotic stresses	Dubos et al. (2010)
BnaA02g29440D	chrA02_21498544	2,15,17,952	2,15,18,817	19,408	AT3G11600	E3 ubiquitin-protein ligase	Involved in DREB2A ubiquitination mediation	Qin et al. (2017)

distribution of *p*-values that was suggested by Mamidi et al. (2014) was used to select the significant markers. QTL peaks for marker-assisted selection (MAS) were identified through stepwise regression, which eliminated the markers those had a minor effect on the phenotype.

In this study, we used a panel of 231 diversified germplasm and 37,699 SNP markers to calculate LD decay distance (r^2 = 0.2) on the A and C subgenomes, which revealed that the mean LD decay was about 5.5 times faster in the A subgenome (average 29 kb, range 18–52 kb) compared with the C subgenome (average 158 kb, range 61–364 kb). The faster LD decay in A subgenome of *B. napus* has been reported by Qian et al. (2014) and Zou et al. (2017). The higher LD decay in A subgenome indicated that the genetic diversity of *B. rapa* (progenitor species of *B. napus*, A subgenome) was higher which has undergone for more recombination. The probable reason for this decay is thought to be that *B. napus* and *B. rapa* are primarily used as oilseed crop, and both species have better cross compatibility resulted more recombination in the chromosomes of A subgenome (Qian et al., 2006). On the other hand, cross compatibility between *B. napus* × *B. oleracea* is difficult (Bennett, Thiagarajah, King, & Rahman, 2008), which limited the genetic diversity of C subgenome of *B. napus*. Due to limited crosses and recombination, it is assumed that there is greater LD and lower genetic diversity available in the C subgenome of *B. napus* (Zou et al., 2017). Usually, the transposon-rich portion of chromosome is recognized as recombination-poor regions (Gorelick, 2003). Chalhoub et al. (2014) conducted the reference genome sequences on *B. napus* and reported that the C subgenome contains more transposon-rich regions compared with the A subgenome, that could also partly justify the difference in LD decay between the A and C subgenomes.

The germplasm accessions used in this study had low LD (Michalak de Jiménez et al., Unpublished). Therefore, we took 100 kb sequence from each side of the significant marker for the candidate gene search (Chalhoub et al., 2014). Gacek et al. (2017) used 50 kb sequence for candidate gene search in *Brassica napus*, Xu et al. (2016) and Zhou et al. (2018) used 100 kb sequence, and Qu et al. (2017) used 200 kb sequence for mining candidate genes. NCBI-BLAST was used to identify genes associated with frost tolerance or abiotic stress. Six potential candidate genes were identified as being related to abiotic stress or frost tolerance. The candidate genes/proteins are annotated as HXXXD-type acyl-transferase family protein that is involved in cell wall metabolism and is repressed under stress conditions (Zhu et al., 2013), the cytochrome P450 family proteins, involved in ABA degradation during dehydration stress and seed imbibition (Shinozaki & Yamaguchi-Shinozaki, 2007), expansin that is involved in loosen cell walls and hydrogen bonding disruption (Lu et al., 2013), the MYB transcription factors (Dubos et al., 2010) or MYB family transcription factors (Hong, Roze, & Linz, 2013), which are involved in abiotic stress response. The last gene was annotated as E3 ubiquitin-protein ligase which is involved in DREB2A (dehydration responsive 2A) ubiquitination mediation (Qin et al., 2008). The DREB genes are involved with dehydration response, which can happen during cold stress. DREB/CBF (C repeat-binding proteins)

factors have been linked to many abiotic stresses, including low temperature stress (Gao, Allard, Byass, Flanagan, & Singh, 2002; Li et al., 2005; Liu et al., 1998; Nakashima et al., 2000; Qin et al., 2004).

Frost tolerance is a complex trait that is controlled by many genes. It is important for the breeding programme to know and understand the trait. Knowing the genes associated with frost tolerance and understanding whether they are positively or negatively related to stress tolerance would also benefit the breeder. Incorporating the positively related genes and minimizing the negatively related genes would help to increase frost tolerance of canola. This study identified six potential genes that are associated with frost tolerance or abiotic stress. Further research is needed to utilize and introgress the genes into commercial cultivars.

ACKNOWLEDGEMENTS

The study was funded by the Northern Canola Growers Association (NCGA-2013) and U.S. Department of Agriculture-National Institute of Food and Agriculture (Hatch Project No. ND01581).

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

AUTHORS CONTRIBUTION

MR designed and supervised this experiment. WFD performed the research and collected phenotypic data. MR, WFD and ZIT conducted data analysis. WFD, MR and ZIT wrote and edited the manuscript. All co-authors reviewed the final version and approved the manuscript before submission.

ORCID

Mukhlesur Rahman https://orcid.org/0000-0003-3355-4803

REFERENCES

Bennett, R. A., Thiagarajah, M. R., King, J. R., & Rahman, M. H. (2008). Interspecific cross of *Brassica oleracea* var. *alboglabra* and *B. napus*: Effects of growth condition and silique age on the efficiency of hybrid production, and inheritance of erucic acid in the self-pollinated backcross generation. *Euphytica*, *164*, 593–601. https://doi.org/10.1007/s10681-008-9788-0

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brule-Babel, A. L., & Fowler, D. B. (1988). Genetic control of cold hardiness and vernalization requirement in winter wheat. *Crop Science*, *28*, 879–884. https://doi.org/10.2135/cropsci1988.0011183X002800060001x

Brzostowicz, A., & Barcikowska, V. (1987). Possibility of frost resistance testing of *Brassica napus* with the help of delayed luminescence intensity. *Cruciferae Newsletter*, *12*, 27.

Bus, A., Körber, N., Snowdon, R. J., & Stich, B. (2011). Patterns of molecular variation in a species wide germplasm set of *Brassica napus*. *Theoretical and Applied Genetics*, *123*, 1413–1423.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., … Wincker, P. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, *345*(6199), 950–953.

Chinnusamy, V., Zhu, J., & Zhu, J.-K. (2007). Cold stress regulation of gene expression in plants. *Trends in Plant Science*, *12*(10), 444–451.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., & Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends in Plant Science*, *15*(10), 573–581. https://doi.org/10.1016/j.tplants.2010.06.005

Elshire, R. J., Glaubitz, J. C., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), e19379.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, *164*, 1567–1587.

Fiebelkorn, D., Horvath, D., & Rahman, M. (2018). Genome-wide association study for electrolyte leakage in rapeseed/canola (*Brassica napus* L.). *Molecular Breeding*, *38*, 129. https://doi.org/10.1007/s11032-018-0892-0

Fiebelkorn, D., & Rahman, M. (2016). Development of a protocol for frost-tolerance evaluation in rapeseed, canola (*Brassica napus* L.). *The Crop Journal*, *4*, 147–152.

Gacek, K., Bayer, P. E., Bartkowiak-Broda, I., Szala, L., Bocianowski, J., Edwards, D., & Batley, J. (2017). Genome-wide association study of genetic control of seed fatty acid biosynthesis in *Brassica napus*. *Frontier Plant Science*, *7*, 2062. https://doi.org/10.3389/fpls.2016.02062

Gajardo, H. A., Wittkop, B., Soto-Cerda, B., Higgins, E. E., Parkin, I. A. P., Snowdon, R. J., … Iniguez-Luy, F. L. (2015). Association mapping of seed quality traits in *Brassica napus* L. using GWAS and candidate QTL approaches. *Molecular Breeding*, *35*, 143.

Gao, M. J., Allard, G., Byass, L., Flanagan, A. M., & Singh, J. (2002). Regulation and characterization of four CBF transcription factors from *Brassica napus*. *Plant Molecular Biology*, *49*, 459–471.

Gorelick, R. (2003). Transposable elements suppress recombination in all meiotic eukaryotes, including automictic ancient asexuals: A reply to Schön and Martens. *Journal of Natural History*, *37*, 903–909. https://doi.org/10.1080/0022293021000007705

Gurung, S., Mamidi, S., Bonman, J. M., Xiong, M., Brown-Guedira, G., & Adhikari, T. B. (2014). Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. *PLoS ONE*, *9*(9), e108179. https://doi.org/10.1371/journal.pone.0108179

Hansen, M., Kraft, T., Ganestam, S., Säll, T., & Nilsson, N.-O. (2001). Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. *Genetical Research*, *77*(1), 61–66. https://doi.org/10.1017/S0016672300004857

Hasan, M., Friedt, W., Pons-Kuhnemann, J., Freitag, N. M., Link, K., & Snowdon, R. J. (2008). Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus ssp. napus*). *Theoretical and Applied Genetics*, *116*, 1035–1049. https://doi.org/10.1007/s00122-008-0733-3

Hatzig, S. V., Frisch, M., Breuer, F., Nesi, N., Ducournau, S., Wagner, M.-H., … Snowdon, R. J. (2015). Genome-wide association mapping unravels the genetic control of seed germination and vigor in *Brassica napus*. *Frontiers in Plant Science*, *6*, 221. https://doi.org/10.3389/fpls.2015.00221

Hill, W. G., & Weir, B. S. (1988). Variances and covariances of squared linkage disequilibria in finite populations. *Theoretical and Population Biology*, *33*, 54–78. https://doi.org/10.1016/0040-5809(88)90004-4

Hong, S.-Y., Roze, L. V., & Linz, J. E. (2013). Oxidative stress-related transcription factors in the regulation of secondary metabolism. *Toxins*, *5*(4), 683–702. https://doi.org/10.3390/toxins5040683

Jia, L., Yan, W., Zhu, C., Agrama, H. A., Jackson, A., Yeater, K., … Wu, D. (2012). Allelic analysis of sheath blight resistance with association mapping in rice. *PLoS ONE*, *7*(3), e32703. https://doi.org/10.1371/journal.pone.0032703

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., … Wilson, R. K. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, *22*, 568–576. https://doi.org/10.1101/gr.129684.111

Körber, N., Bus, A., Li, J., Parkin, I. S. P., Wittkop, B., Snowdon, R. J., & Stich, B. (2016). Agronomic and seed quality traits dissected by genome-wide association mapping in *Brassica napus*. *Frontier of Plant Science*, *7*, 386. https://doi.org/10.3389/fpls.2016.00386

Kraakman, A. T. W., Martínez, F., Mussiraliev, B., van Eeuwijk, F. A., & Niks, R. E. (2006). Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Molecular Breeding*, *17*(1), 41–58. https://doi.org/10.1007/s11032-005-1119-8

Kraakman, A. T. W., Niks, R. E., Van den Berg, P. M. M. M., Stam, P., & Van Eeuwijk, F. A. (2004). Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, *168*, 435–446. https://doi.org/10.1534/genetics.104.026831

Lagercrantz, U. (1998). Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics*, *150*, 1217–1228.

Lagercrantz, U., & Lydiate, D. J. (1996). Comparative genome mapping in Brassica. *Genetics*, *144*, 1903–1910.

Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., … Wu, X. (2014). Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*Brassica napus* L.). *DNA Research*, *21*, 355–367. https://doi.org/10.1093/dnares/dsu002

Li, H. (2013). *Aligning sequence reads, clone sequences, and assembly contigs with BWA-MEM*. arXiv:1303.3997. http://arxiv.org/abs/1303.3997

Li, X. P., Tian, A. G., Luo, G. Z., Gong, Z. Z., Zhang, J. S., & Chen, S. Y. (2005). Soybean DRE-binding transcription factors that are responsive to abiotic stresses. *Theoretical and Applied Genetics*, *110*, 1355–1362. https://doi.org/10.1007/s00122-004-1867-6

Li, X., Yan, W., Agrama, H., Jia, L., Shen, X., Jackson, A., … Wu, D. (2011). Mapping QTLs for improving grain yield using the USDA rice mini-core collection. *Planta*, *234*, 347–361. https://doi.org/10.1007/s00425-011-1405-0

Liu, J., Wang, W., Mei, D., Wang, H., Fu, L., Liu, D., … Hu, Q. (2016). Characterizing variation of branch angle and genome-wide association mapping in rapeseed (*Brassica napus* L.). *Frontier of. Plant Science*, *7*, 21. https://doi.org/10.3389/fpls.2016.00021

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., & Shinozaki, K. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively in *Arabidopsis*. *The Plant Cell*, *10*, 1391–1406. https://doi.org/10.1105/tpc.10.8.1391

Lu, P., Kang, M., Jiang, X., Dai, F., Gao, J., & Zhang, C. (2013). *RhEXPA4*, a rose expansin gene, modulates leaf growth and confers drought and salt tolerance to *Arabidopsis*. *Planta*, *237*, 1547–1559. https://doi.org/10.1007/s00425-013-1867-3

Mamidi, S., Chikara, S., Goos, R. J., Hyten, D. L., Annam, D., Moghaddam, S. M., … McClean, P. E. (2011). Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. *Plant Genome*, *4*, 154–164. https://doi.org/10.3835/plantgenome2011.04.0011

Mamidi, S., Lee, R. K., Goos, R. J., & McClean, P. E. (2014). Genome-wide association studies identifies seven major regions responsible for iron deficiency chlorosis in soybean (*Glycine max*). *PLoS ONE*, *9*(9), e107469. https://doi.org/10.1371/journal.pone.0107469

Meuwissenk, T., & Goddard, M. (2000). Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. *Genetics*, *155*(1), 421–430.

Michalak de Jiménez, M., Mamidi, S., McClean, P. E., & Rahman, M. (Unpublished). *Low level of linkage disequilibrium and population structure in a core collection of Brassica napus discovered using SNPs derived from GBS*.

Mohammadi, M., Blake, T. K., Budde, A. D., Chao, S., Hayes, P. M., Horsley, R. D., … Smith, K. P. (2015). A genome-wide association study of malting quality across eight U.S. barley breeding programs. *Theoretical Applied Genetics*, *128*, 705–721. https://doi.org/10.1007/s00122-015-2465-5

Nakashima, K., Shinwari, Z. K., Sakuma, Y., Seki, M., Miura, S., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2000). Organization and expression of two *Arabidopsis* DREB2 genes encoding DRE-binding proteins involved in dehydration- and high-salinity-responsive gene expression. *Plant Molecular Biology*, *42*, 657–665.

Nyquist, W. E., & Baker, R. J. (1991). Estimation of heritability and prediction of selection response in plant populations. *Critical Reviews in Plant Sciences*, *10*, 235–322. https://doi.org/10.1080/07352689109382313

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Pritchard, J. K., Stephens, M., Rosenberg, N. A., & Donnelly, P. (2000). Association mapping in structured populations. *American Journal of Human Genetics*, *676*, 170–181. https://doi.org/10.1086/302959

Qian, L., Qian, W., & Snowdon, R. J. (2014). Sub-genomic selection patterns as a signature of breeding in the allopolyploid *Brassica napus* genome. *BMC Genomics*, *15*, 1170. https://doi.org/10.1186/1471-2164-15-1170

Qian, W., Meng, J., Li, M., Frauen, M., Sass, O., Noack, J., & Jung, C. (2006). Introgression of genomic components from Chinese *Brassica rapa* contributes to widening the genetic diversity in rapeseed (*B. napus* L.), with emphasis on the evolution of Chinese rapeseed. *Theoretical and Applied Genetics*, *113*, 49–54. https://doi.org/10.1007/s00122-006-0269-3

Qin, F., Sakuma, Y., Li, J., Liu, Q., Li, Y. Q., Shinozaki, K., & Yamaguchi-Shinozaki, K. Y. (2004). Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant Cell Physiology*, *45*, 1042–1052. https://doi.org/10.1093/pcp/pch118

Qin, F., Sakuma, Y., Phan Tran, L.-S., Maruyama, K., Kidokoro, S., Fujita, Y., & …Yamaguchi-Shinozaki, K., (2008). *Arabidopsis* DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress–responsive gene expression. *The Plant Cell*, *20*, 1693–1707. https://doi.org/10.1105/tpc.107.057380

Qu, C., Jia, L., Fu, F., Zhao, H., Lu, K., Wei, L., & … Li, J. (2017). Genome-wide association mapping and Identification of candidate genes for fatty acid composition in *Brassica napus* L. using SNP markers. *BMC Genomics*, *18*, 232.

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology*, *5*, 94–100. https://doi.org/10.1016/S1369-5266(02)00240-6

Rahman, M., Mamidi, S., del Rio, L., Ross, A., Kadir, M. M., Rahaman, M. M., & Arifuzzaman, M. (2016). Association mapping in *Brassica napus* (L.) accessions identifies a major QTL for blackleg disease resistance on chromosome A01. *Molecular Breeding*, *36*, 90. https://doi.org/10.1007/s11032-016-0513-8

Rahman, M., & McClean, P. E. (2013). Genetic analysis on flowering time and root system in *Brassica napus* L. *Crop Science*, *53*, 141–147. https://doi.org/10.2135/cropsci2012.02.0095

Raman, H., Dalton-Morgan, J., Diffey, S., Raman, R., Alamery, S., Edwards, D., & Batley, J. (2014). SNP markers-based map construction and genome-wide linkage analysis in *Brassica napus*. *Plant Biotechnology Journal*, *12*, 851–860.

Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., … Buckler, E. S. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 11479–11484. https://doi.org/10.1073/pnas.201394398

Rezaeizad, A., Wittkop, B., Snowdon, R., Hasan, M., Mohammadi, V., Zali, A., & Friedt, W. (2011). Identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping using SSR markers. *Euphytica*, *177*, 335–342. https://doi.org/10.1007/s10681-010-0231-y

Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, *273*, 1516–1517. https://doi.org/10.1126/science.273.5281.1516

Scheet, P., & Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics*, *78*, 629–644. https://doi.org/10.1086/502802

Shinozaki, K., & Yamaguchi-Shinozaki, K. (2007). Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany*, *58*(2), 221–227. https://doi.org/10.1093/jxb/erl164

Strigens, A., Freitag, N. M., Gilbert, X., Grieder, C., Riedelsheimer, C., Schrag, T. A., … Melchinger, A. E. (2013). Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chamber and field experiments. *Plant, Cell, and Environment*, *36*, 1871–1887. https://doi.org/10.1111/pce.12096

Sun, J., Guo, N., Lei, J., Li, L., Hu, G., & Xing, H. (2014). Association mapping for partial resistance to *Phytophthora sojae* in soybean (*Glycine max* L.). *Journal of Genetics*, *2*, 355–363.

Teutonico, A., Palta, P., & Osborn, C. (1993). In vitro freezing tolerance in relation to winter survival of rapeseed cultivars. *Crop Science*, *3*, 103–107. https://doi.org/10.2135/cropsci1993.0011183X003300010019x

U, N. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *Brassica napus* and peculiar mode of fertilization. *Japanese Journal of Botany*, *7*, 389–452.

United States Department of Agriculture, National Agricultural Statistics Service (2019). Quick Stat. Retrieved from https://quickstats.nass.usda.gov/

Visioni, A., Tondelli, A., Francia, E., Pswarayi, A., Malosetti, M., Russell, J., & … Comadran, J. (2013). Genome-wide association mapping of frost tolerance in barley (*Hordeum vulgare* L.). *BMC Genomics*, *14*, 424–437.

Wang, N., Li, F., Chen, B., Xu, K., Yan, G., Qiao, J., … Wu, X. (2014). Genome-wide investigation of genetic changes during modern breeding of *Brassica napus*. *Theoretical and Applied Genetics*, *127*, 1817–1829. https://doi.org/10.1007/s00122-014-2343-6

Wang, N., Qian, W., Suppanz, I., Wei, L., Mao, B., Long, Y., … Jung, C. (2011). Flowering time variation in oilseed rape (Brassica napus L.) is associated with allelic variation in the FRIGIDA homologue BnaA. FRI.a. *Journal of Experimental Botany*, *62*, 5641–5658.

Waugh, R., Jannink, J. L., Muller, K., & Ramsay, L. (2009). The emergence of whole genome association scans in barley. *Current Opinion in Plant Biology*, *12*, 1–5. https://doi.org/10.1016/j.pbi.2008.12.007

Williams, P. H., & Delwiche, P. A. (1980). *Screening for resistance to blackleg of crucifers in the seedling stage*. Pages 164-170 in: Proc. Eucarpia Cruciferae 1979 Conference. Wageningen, The Netherlands.

Wilson, J. M. (1997). Mechanisms of chilling resistance in plants. In A. S. Basra, & R. K. Basra (Eds.), *Mechanisms of environmental stress resistance in plants* (pp. 111–122). Amsterdam, The Netherlands: Harwood Academic Publishers.

Wrucke, D. F., Mamidi, S., & Rahman, M. (2018). Genome-wide association study for frost tolerance in canola (*Brassica napus* L.) under field conditions. *Journal of Plant Biochemistry and Biotechnology*, *28*(2), 211–222. https://doi.org/10.1007/s13562-018-0472-8

Xu, L., Hu, K., Zhang, Z., Guan, C., Chen, S., Hua, W., … Fu, T. (2016). Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). *DNA Research*, *23*, 43–52.

Yu, J. M., & Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, *17*, 155–160. https://doi.org/10.1016/j.copbio.2006.02.003

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., … Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, *38*, 203–208. https://doi.org/10.1038/ng1702

Zhang, H., Yin, L., Wang, M., Yuan, X., & Liu, X. (2019). Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. *Frontier Genetics*, *10*, 189. https://doi.org/10.3389/fgene.2019.00189

Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., … Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics*, *42*, 355–360. https://doi.org/10.1038/ng.546

Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., … Nordborg, M. (2007). An Arabidopsis example of association mapping in structured samples. *PLoS Genetics*, *3*, e4. https://doi.org/10.1371/journal.pgen.0030004

Zhou, Q. H., Han, D. P., Mason, A. S., Zhou, C., Zheng, W., Li, Y. Z., … Huang, Y. J. (2018). Earliness traits in rapeseed (*Brassica napus*): SNP loci and candidate genes identified by genome-wide association analysis. *DNA Research*, *25*, 229–244. https://doi.org/10.1093/dnares/dsx052

Zhu, J., Jeong, J. C., Zhu, Y., Sokolchik, I., Miyazaki, S., Zhu, J.-K., … Bressan, R. A. (2008). Involvement of *Arabidopsis* HOS15 in histone deacetylation and cold tolerance. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(12), 4945–4950. https://doi.org/10.1073/pnas.0801029105

Zhu, Y.-N., Shi, D.-Q., Ruan, M.-B., Zhang, L.-L., Meng, Z.-H., Liu, J., & Yang, W.-C. (2013). Transcriptome analysis reveals crosstalk of responsive genes to multiple abiotic stresses in Cotton (*Gossypium hirsutum* L.). *PLoS ONE*, *8*(11), 1–13. https://doi.org/10.1371/journal.pone.0080218

Zou, J., Hu, D., Mason, A. S., Shen, X., Wang, X., Wang, N., … Meng, J. (2017). Genetic changes in a novel breeding population of Brassica napus synthesized from hundreds of crosses between B. rapa and B. carinata. *Plant Biotechnology Journal*, *16*, 507–519.

Zou, J., Jiang, C., Cao, Z., Li, R., Long, Y., Chen, S., & Meng, J. (2010). Association mapping of seed oil content in *Brassica napus* and comparison with quantitative trait loci identified from linkage mapping. *Genome*, *53*(11), 908–916.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Wrucke DF, Talukder ZI, Rahman M. Genome-wide association study for frost tolerance in rapeseed/canola (*Brassica napus*) under simulating freezing conditions. *Plant Breed.* 2019;00:1–12. https://doi.org/10.1111/pbr.12771