Association mapping of autumn-seeded rye (Secale cereale L.) reveals genetic linkages between genes controlling winter hardiness and plant development

Monica Båga1, Hirbod Bahrami2, Jamie Larsen2, Bernd Hackauf3, Robert J. Graf4, Andre Laroche4 & Ravindra N. Chibbar1

Winter field survival (WFS) in autumn-seeded winter cereals is a complex trait associated with low temperature tolerance (LTT), prostrate growth habit (PGH), and final leaf number (FLN). WFS and the three sub-traints were analyzed by a genome-wide association study of 96 rye (Secale cereale L.) genotypes of different origins and winter-hardiness levels. A total of 10,244 single nucleotide polymorphism (SNP) markers were identified by genotyping by sequencing and 259 marker-trait associations (MTAs; p < 0.01) were revealed by association mapping. The ten most significant SNPs (p < 1.49e−04) associated with WFS corresponded to nine strong candidate genes: Inducer of CBF Expression 1 (ICE1), Cold-regulated 413-Plasma Membrane Protein 1 (COR413-PM1), Ice Recrystalization Inhibition Protein 1 (IRIP1), Jasmonate-resistant 1 (JAR1), BIPP2C1-like protein phosphatase, Chloroplast Unusual Positioning Protein-1 (CHUP1), FRIGIDA-like 4 (FRL4-like) protein, Chalcone Synthase 2 (CHS2), and Phenylalanine Ammonia-lyase 8 (PAL8). Seven of the candidate genes were also significant for one or several of the sub-traits supporting the hypothesis that WFS, LTT, FLN, and PGH are genetically interlinked. The winter-hardy rye genotypes generally carried additional allele variants for the strong candidate genes, which suggested allele diversity was a major contributor to cold acclimation efficiency and consistent high WFS under varying field conditions.

Rye (Secale cereale L.) is a member of the Triticeae tribe within the Pooidae subfamily, which belongs to the Poaceae grass family. Similar to the Triticeae crops barley (Hordeum vulgare L.) and bread wheat (Triticum aestivum L.), rye is mainly cultivated for its grain, but can also be grown for its biomass used as forage or production of bioenergy. The main rye producing areas are located in North America, China, Russia, and Europe, where the grain is used for production of bread, alcoholic beverages, and feed1.

The Triticeae genomes are likely derived from a common progenitor genome which explains the high sequence synteny between the the seven chromosomes of the diploid rye R genome (2n = 2x = 14), diploid H genome of barley and the three genomes (A, B, D) of hexaploid wheat. Since the wheat and rye lineages diverged, several translocations have occurred within the rye genome as revealed by transcript mapping and genome sequence assemblies2–4. The haplotype diversity in rye is high as it is an open-pollinated species, whereas current wheat and barley cultivars have lower genetic diversity due to self-pollination and extensive breeding during the last century. Although some genetic erosion has occurred in improved rye germplasm, the landraces and wild rye lines display a relatively high allelic variation that is available for mining5.

Rye can serve as a good model for winter hardiness studies in Triticeae species, as many autumn-seeded rye varieties developed for northern latitudes have the highest winter-hardiness among cereal crops1. Most of the temperate grasses are inherently chilling tolerant, but acquire frost tolerance following weeks of acclimation to low but non-freezing temperatures6. For autumn-seeded cereals, the cold acclimation process occurs prior to

1Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK S7N 5A8, Canada. 2Harrow Research and Development Centre, Agriculture and Agri-Food Canada, Harrow, ON N0R 1G0, Canada. 3Institute for Breeding Research on Agricultural Crops, Julius Kühn-Institut, 18190 Sanitz, Germany. 4Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, AB T1J 4B1, Canada. 50email: ravi.chibbar@usask.ca
winter and is characterized by two major physiological processes: (a) accumulation of increased freezing tolerance and (b) fulfillment of vernalization requirement causing shift from vegetative to reproductive growth at the shoot apical meristem (SAM). Upon vernalization saturation, further development of floral organs is paused until new growth from the crown tissue is stimulated by inductive photoperiod and temperature in the spring.

Cold acclimation initiates when plants integrate and respond to environmental signals received from gradual reductions in photoperiod, daily temperature, light intensity, and red/far-red ratio in incoming light. An early start and long acclimation period result in higher winter hardness for winter cereals, and thus, the timing of floral transition impacts the total amount of accumulated freezing tolerance. Cold acclimated cereals further increase their freezing tolerance when exposed to slightly below 0 °C temperatures, which stimulate a second hardening process. Transcriptional analyses show different sets of genes are induced during the above-zero and sub-zero cold hardening processes.

Winter-hardy plants undergo a multitude of developmental adjustments during cold acclimation to build up frost resistance. The morphological changes observed in cereals generally include a switch to prostrate growth habit (PGH), compact growth, strengthening of cell walls, changes in membrane structures, and increased number of leaf initials formed at SAM. The biochemical changes involve production of various cryoprotectants, antioxidants, and antifreeze proteins to aid protection against future frost damage. An enhanced photosynthetic performance is observed for the most winter-hardy cereals and results in higher biomass production and reduced susceptibility to photoinhibition as compared to tender genotypes. Through the modification of photosynthesis, winter-hardy cereals can effectively increase their carbon pools in crown tissues to be retrieved during deacclimation incidents throughout winter and regrowth in the spring. Cold acclimation also causes epigenetic changes at the DNA and histone levels with effects on gene activities. Overall, cold responses are coordinated by an extensive cross-talk between cold-induced Ca²⁺, reactive oxygen species, phytohormone, and light signaling pathways, of which retrograde/anterograde stress signaling between plastids and nucleus are suggested to play a central role.

The Vernalization 1 (VRN1) and Frost resistance-2 (FR-2) loci on homoeologous group 5 chromosomes are the major determinants for freezing tolerance in Triticeae species. Allelic differences at VRN1, Fr-A2 and their interactions relate to cold acclimation efficiency and duration. In addition, VRN1 and FR-2 interactions with genes controlling plant development is required for optimal cold acclimation. Large clusters of AP2/ERF transcription factor genes denoted C-repeat Binding Factors (CBFs) are present at FR loci, where variations in CBF expression levels, copy numbers, and haplotypes affect freezing tolerance and WFS levels in hexaploid wheat. Cold-induced CBFs bind to C-repeat/dehydration-responsive elements present in promoters of many cold-regulated (COR) genes and thereby induce their expression during cold stress. Besides the CBF-COR regulon, additional pathways respond to cold stress as demonstrated by triple CBF mutants in Arabidopsis showing reduced frost tolerance but retained cold-response for several COR genes.

In a previous study, we analyzed WFS in a rye population of 96 genotypes of different origins and winter hardness levels and showed WFS was almost entirely determined by the cold acclimation process. In addition, the developmental traits FLN and PGH were highly correlated with both WFS and LTT (r = 0.59–0.80; p < 0.001). Estimations of heritability for WFS, FLN and PGH revealed moderate to high values (h² = 0.45 to 0.81), which allowed us in this study to perform a genome-wide association study (GWAS) with the aim to find genes influencing WFS and associated traits. GWAS is useful for identification of candidate genes for traits showing variation in a population of unrelated genotypes. The genetic mapping technology has greatly benefited from recently developed next generation sequencing and genotyping by sequencing (GBS) technologies, which allow identification of large numbers of single nucleotide polymorphism (SNP) markers. Rye is well suited for association mapping as it’s out-crossing nature generates a low linkage disequilibrium (LD). In addition, a near full-length annotated genome (6.74 Gb of estimated 7.9 Gb) sequence is available for genome mapping. The association study of rye genotype panel revealed several candidate genes were shared between WFS and one or more of the three associated traits; thus, suggesting the traits were genetically interconnected. The characteristics of nine strong candidate genes affecting WFS are discussed.

Materials and methods

Plant materials. Rye (Secale cereale L.) accessions (96 genotypes) used for association mapping are listed in Table S1 and were phenotyped in a separate study. The genotypes were maintained through generations by crossing five plants per pollination bag. All experimental research and field studies on plants (cultivated or wild) including collection of plant material complied with the institutional and funding agencies, national and international guidelines and legislation. The permission to collect seeds and plant material was granted under the material transfer agreements with international gene banks and Universities.

Field trials to determine WFS. Five years (2014–2019) of field trials were conducted at the University of Saskatchewan Experimental Farm, Saskatoon, Saskatchewan, Canada (52° 10′ N, 106° 30′ W, and 457 m altitude) as previously reported. Briefly, the genotypes were seeded in early September with each genotype randomly placed in three 3.6 m rows, 100 seeds/row, and 30 cm row spacing. For each genotype and trial, the germination frequency was determined 6 weeks upon seeding and survival frequency was determined following the spring. The genotypes survival frequencies from the trials were used to calculated the Best Linear Unbiased Estimates (BLUEs) for WFS for each genotype using the statistical analysis Multi Environment Trial Analysis with R (META-R) software version 6.04 (CIMMYT Research Data & Software Repository Network, Mexico). Based on the BLUE scores, WFS for the population was divided into five groups: very high (64.5–92.5%; 19 genotypes), high (56.7–64.3%; 20 genotypes), moderate (46.2–56.6%; 19 genotypes), low (30.5–43.7%; 19 genotypes) and very low (0–25.2%; 19 genotypes) WFS (Table S1).
Freeze tests to determine LTT. Freeze tests of cold acclimated plants were performed by a step wise decrease in temperature using an EPZ-4H test chamber (ESPEC North America Inc., Hudsonville, MI, USA) as described. The freezing temperature at which 50% of the plants regrew two weeks upon exposure to freezing was noted as the LT50 temperature. Negative LT50 values represented the LTT score.

Assessment of plant developmental traits. The two developmental traits PGH and FLN were analyzed in four separate trials with five plants per genotype in each trial as previously described. PGH was rated upon completion of cold-acclimation in growth chamber. The visual PGH rating was based on three types of growth habits: (1) erect, (2) intermediate, and (3) prostrate. To determine FLN, plant leaves were labeled numerically as they developed from the primary stem. The flag leaf number was recorded as FLN. BLUE scores for the two developmental traits were calculated to determine the overall FLN and PGH scores for each genotype.

DNA extraction and GBS analysis. Leaf tissue (~130 mg) was collected from five two-week-old plants of each genotype and ground to a fine powder in liquid nitrogen using a mortar and pestle. DNA was extracted from the homogenized leaf material using a standard hexadecetyltrimethylammonium bromide-based method. The concentration of isolated DNA was determined using a Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen, Molecular Probes, Eugene, OR, USA). Uniform presence of high molecular weight DNA in the samples was confirmed by agarose (0.8%) gel electrophoresis. Normalized DNA preparations at 10 ng μL⁻¹ (20 μL in total) were subjected to library preparations and GBS analysis, a service provided by Plateforme d'analyses Génomiques of the Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, Canada. Reduced-representation libraries were prepared from PstI andMspI double-digested DNA as described using an Ion CHEF System, Hi-Q reagents, and P1 V3 chips (ThermoFisher Scientific, Life Technologies Inc., ON, Canada) for the library and chip preparations. Sequencing analysis was done on an Ion Proton Sequencer (ThermoFisher Scientific, Life Technologies Inc., Waltham, MA, USA).

SNP calling. Raw FASTQ files obtained from GBS analyses of the 96 rye genotypes were processed by Trimmomatic version 0.36 with a four-base window setting and a minimum average quality score of 16. Barcodes, adaptors, and reads less than 36 bp were eliminated from the sequences before alignment to scaffold sequences of rye inbred line Lo7–version 2 using Bowtie 2.0 version 2.2.6 software. Duplicates were removed using the rmdp function and SNP variants including insertions/deletions (InDels) were detected by Sequence Alignment/Map tools version 0.1.19. The markers were filtered using the Variant Call Format tool version 0.1.14, a minor allele frequency >5%, and a minimum reading depth of 6.0. A total of 10,244 SNP markers were identified and numbered according to matching scaffold of rye inbred line Lo7–version 2. The selected markers contained no missing data. BLASTn analyses against the 2021 rye Lo7 v1 genome assembly were done using default settings and e-value threshold set to ≤1e−70. SNP marker distances and squared correlations of allele frequencies (r²) for 9547 mapped markers for the population were determined by the TASSEL version 5.2.20 software. The generated data was displayed as linkage disequilibrium (LD) decay plots using the RStudio package version 3.5.1 software. LD value at r² = 0.20 was estimated from each plot’s trend line.

Analyses of population structure. The genetic relationship among the 96 rye genotypes was analyzed using a model-based clustering method implemented by the STRUCTURE version 2.3.4 software. The analysis was made using genetic information for all 10,244 SNP markers. Based on an admixed model and allele frequencies correlated, the number of subpopulations (K) ranging from 1 to 9 were tested with burn-in time set to 50,000 and Markov Chain Monte Carlo replication number to 100,000. A total of 20 runs were performed for each data set to compute the log likelihood for each simulated K value. The STRU CTU RE HARVESTER version 0.6.94 and Markov Chain Monte Carlo replication number to 100,000. A total of 20 runs were performed for each data set to compute the log likelihood for each simulated K value. The STRU CTU RE HARVESTER version 0.6.94, kinship data, genotype, and phenotype data for the rye population were used to determine the optimal number of sub-populations according to the ΔK method. The result was displayed as bi-plots.

Marker-trait association (MTA) analysis. Phenotypic data for WFS, LTT, FLN, and PGH of the rye population was determined by replicated trials in a previous study. The population structure obtained from STRUCTURE version 2.3.4, kinship data, genotype, and phenotype data for the rye population were used to determine p and R2 (explained variation) for MTAs by the mixed linear model available in TASSEL version 5.2.20 software. The minor allele frequency was set to >5%. Significant SNP variants were initially tested based on a false discovery rate (FDR)-adjusted p value of 0.05 following a step-wise procedure, and lowest adjusted p values (threshold p <1.49e−04) was identified after a Bonferroni correction. To illustrate the relationships between the allele variations and trait values for the 96 genotypes, phenotype and genotype data for the population was subjected to a principal component analysis (PCA) using RStudio package version 3.5.1 software. The evolutionary relationships among the rye genotypes were calculated to determine the overall FLN and PGH scores for each genotype.

Prediction of candidate genes. The marker sequences carrying SNP of interest underwent BLASTN searches (default settings) against the Secale cereale Lo7 version 1 pseudomolecules hosted by Graingenes (https://wheat.pw.usda.gov/blast/) and the IPK Galaxy Blast Suite (https://galaxy-web.ipk-gatersleben.de) databases to identify matches to the latest assembly of rye Lo7 genome and its annotated genes. BLAST searches were also made against the National Center for Biotechnology Information database (https://blast.ncbi.nlm.nih.gov/blast/).
Results
Structure analysis of GBS data generated for rye population. The GBS analysis generated a total of 357.1 million reads with an average read length of 108 nucleotides. Alignment of the processed sequences to scaffold sequences of rye inbred line Lo7 genome version 2 produced 252,158 SNP markers, of which 10,244 SNP markers with no missing data remained upon data filtering. BLASTn analyses of the marker sequences to the 2021 rye genome assembly5 assigned a total of 9,547 markers (93.2%) to chromosome positions. The markers were relatively evenly distributed on the genome, with an average of 1,364 markers per chromosome (Table S2). Estimations of LD yielded values between 3428 and 5153 bp for the individual chromosomes and an average of 4266 bp for the whole genome (Table S2; Fig. S1). With relatively low LD values, the mapping data was considered suitable for association mapping.

The SNP data was analyzed by PCA, which showed the first and second coordinates explained 3.6% and 2.3% of the total genetic variation, respectively (Fig. 1). Three clusters were distinguished from the PCA plot, where the largest cluster (II) with 63 genotypes contained several genotypes with the highest WFS scores, but also relatively cold-sensitive genotypes such as Petkus and Carsten (Fig. 1; Table S1). Like cluster II, cluster I with 28 genotypes included several very winter-hardy genotypes intermingled with tender genotypes. Cluster III consisted mainly of perennial types with relatively low WFS scores. The existence of three subpopulations was also supported by a STRUCTURE HARVESTER analysis, which deducted the maximum value for Delta K to three (Fig. S2). A neighbor-joining phylogenetic tree constructed from the SNP data repositioned the Canadian genotypes Puma, Cougar and Gazelle in PCA cluster I to branch II of the phylogenetic tree (Fig. S3; Table S1). The tree also indicated division of branch II into two sub-branches, where Carsten was placed in branch Ia and Petkus in branch Ib. A separation of Carsten and Petkus into different gene pools can be explained by the use of these genotypes in different breeding programs64,65.

Identification of SNP markers associated with developmental and cold tolerance traits. Association mapping using SNP and trait data for rye population led to identification of 189 SNP marker variations that were significantly (p < 0.01) associated with one or several of the WFS, LTT, FLN, and PGH trait values (Table S3). In total, 259 MTAs were identified, where 98 markers associated with WFS, 26 with LTT, 75 with FLN, and 60 with PGH (Table 1). LTT shared most markers with WFS (16 out of 26; 61.5%), followed by FLN (26 out of 75; 34.7%), and PGH (18 out of 60; 30.0%), thus indicating the four traits may be genetically linked with each other.

The SNP markers were divided into two groups based on their significance level (Table S3). Group 1 contained the ten most significant markers, which were distinguished after a Bonferroni correction of mapping data generating a new and more stringent significant p value (p < 1.49e−4; FDR < 0.05) (Table 2). A total of 22 MTAs was identified for the group 1 markers, which matched nine candidate genes. The second group included 179 markers (p < 0.001), of which 81 matched genes (Table S3). Of the 90 genes associated with MTAs, 32 amino acid substitutions were predicted from the SNP variations. All but five of the 189 SNP markers could be placed on de novo assembled Lo7 physical map of rye4 (Fig. 4). Below is a presentation of the nine strongest candidate genes listed in Table 2.

BIPP2C1-like protein phosphatase. The most significant marker identified in the study (Xuos526258) was associated with all four traits, WFS, LTT, FLN, and PGH (Table 2) and matched SECCE7Rv1G0469720.1, a PPM-type protein phosphatase 2C gene (PP2C) located on chromosome 7R (112,789 Mb; Fig. 4). The predicted rye phosphatase showed highest sequence identity (84.1%) to benzothiadiazole-induced protein phosphatase 2C1 (BIPP2C1) from Triticum dicoccoides (XP_037426047). BIPP2C1 proteases are members of the Mg2+/Mn2+-dependent Ser/Thr phosphatase family subclade K in Arabidopsis, rice, and wheat2,4,6. No function has been assigned for wheat BIPP2C164, whereas clade K member OsBIPP2C1/PP45 from rice is cold-induced65, protects against abiotic and biotic stresses66, and has a negative role in abscisic acid signaling66. The T/G SNP variation for rye BIPP2C1 caused a V185 to G185 change within a glycine-rich region located outside of the catalytic domain for the phosphatase (data not shown). All 38 rye genotypes with low or very low WFS encoded only the G185 variant for the phosphatase (Table S4), and this allele variant was also associated with low LTT values, low LTT and FLN values (Fig. 2; Table S4).

Ice recrystallization inhibition protein 1 (IRIP1). The high-significance marker Xuos530120 associated with WFS, LTT, and PGH (Table 2) was found to match SECCE7Rv1G0458910.1 located on chromosome 7R (23,476 Mb; Fig. 4). The targeted gene codes for an ice recrystallization inhibition protein 1 (IRIP1), which belongs to a group of proteins only produced by Pooidae species67. IRIPs carry a N-terminal leucine-rich repeat (LRR) domain, similar to those of phytosulfokine receptor tyrosine kinases, and an ice-binding domain at the
C-terminal end. During exposure to freezing, IRIPs present in the apoplastic space bind to small ice crystals and thereby inhibit further ice recrystallization and frost damage of the plasma membrane. The proteins may also prevent bacteria from serving as ice-nucleators in the extra-cellular space. The Xuos530120 SNP encoded

**Figure 1.** Principal component analysis (PCA) score plot. Plot generated from analysis of 10,244 SNP markers identified in rye population of 96 accessions. The x axis represents the eigenvalue for principal component 1 (PC1) and the y axis for PC2. Genotypes with highest WFS are encircled and arrows refer to less cold hardy Petkus and Carsten, respectively. Data for the individual genotypes is listed in Table S1.
a G36/R36 variation within one of two preserved cysteine loops preceding the LRR-like region, which may be proteolytically separated from the ice-binding domain in the apoplastic space as suggested by studies in *Brachypodium distachyon* 69. Thus, the amino acid substitution for the rye IRIP was not expected to influence protein binding to ice crystals, but could affect the function of the N-terminal region (Fig. S4). Almost all genotypes (36 out of 38) with low and very low WFS encoded only the R36 IRIP variant (Table S4), whereas the highest WFS, LTT, and FLN values were favored by the G36 IRIP variant (Fig. 2). Many additional IRIP variants may be produced from the 7R locus as the reference genome carries at least 13 IRIP-like coding regions (> 80% nucleotide sequence identity) arranged in tandem with SECCE7Rv1G0458910.

Cold-regulated 413 plasma membrane (COR413-PM)-like. One common SNP for WFS and FLN was carried by marker sequence Xuos613978 (Table 2), which showed high sequence identity to two *COR413* genes, SECCE5Rv1G0354870.1 (89.1%) and SECCE5Rv1G0354880.1 (89.2%), respectively. The two *COR* genes are arranged in tandem ~ 11 Mb distal of the vernalization locus (*ScVRN1*) on chromosome 5R (727,383 and 727,388 Mb; Fig. 4). *COR413* genes are expressed during cold stress and their expression levels relate to accumulated freezing tolerance in cereals and *Arabidopsis* during cold acclimation71,72. The rye *COR413* proteins encoded from 5R locus were from BLAST and SignalP 5.0 algorithm analyses predicted to be targeted to the plasma membrane (Fig. S4). Additional predictions by TMHMM algorithm supported the N-terminal end of COR413-PM encoded by SECCE5Rv1G0354870.1 was positioned inside of the plasma membrane, whereas the N-terminal end of COR413-PM encoded from SECCE5Rv1G0354880.1 was positioned outside of the plasma membrane (Fig. S4). The Xuos613978 SNP caused a T30 to A30 variation proximal of the first predicted transmembrane region on COR413-PM. The SNP variation may involve a possible phosphorylation site at T30 accord-

### Table 1. Total number and shared SNP markers between traits.

| Trait | Total number of SNPs associated with trait | Number of markers shared |
|-------|-------------------------------------------|--------------------------|
|       |                                           | WFS  LTT  FLN  PGH        |
| WFS   | 98                                        | 16  26  18               |
| LTT   | 26                                        | 9   9                     |
| FLN   | 75                                        |                       13 |
| PGH   | 60                                        |                       |

### Table 2. Most significant MTAs identified in the study.

| Marker ID | SNP | SNP position (bp) | P value | R2 | MTA Corresponding gene; protein | Amino acid change |
|-----------|-----|-------------------|---------|----|---------------------------------|------------------|
| Xuos526258 | T/G | Chr7:112,789,055  | 1.10E−05| 20.7 | Xuos526258_WFS | V/G SECC7Rv1G0469720.1;Benzothiadiazole-induced protein phosphatase 2C1 (BIPP2C1) |
| Xuos530120 | G/C | Chr7:23,476,102   | 5.66E−05| 14.5 | Xuos530120_WFS | G/R SECC7Rv1G0458910.1;Ice recrystallization inhibition protein 1 (IRIP1) |
| Xuos615052 | T/A | Chr3:937,715,414  | 7.94E−05| 14.7 | Xuos615052_WFS | L/H SECE3Rv1G0209310.1;Inducer of CBF expression 1 (ICE1) |
| Xuos519455 | T/A | Chr6:434,208,045  | 1.10E−04| 13.1 | Xuos519455_WFS | F/I SECC6Rv1G0399250.1;Phenylalanine ammonia lyase 8 (PAL8) |
| Xuos619798 | A/G | Chr5:727,383,264  | 1.11E−05| 15.5 | Xuos619798_WFS | T/A SECC7Rv1G0354870.1, SECC7Rv1G0354880.1;Cold-regulated plasma membrane 413 protein (COR413-PM) |
| Xuos751999 | A/C | Chr1:716,318,047  | 1.20E−04| 15.3 | Xuos751999_WFS | N/T SECE1Rv1G0061510.1;Jasmonat E−resistant 1 (JAR1) |
| Xuos76228a | C/G | Chr2:933,364,125  | 1.39E−04| 14.5 | Xuos76228a_WFS | Q/E SECCE2Rv1G0140890.1;Chalcone synthase 2 (CHS2) |
| Xuos76228b | C/T | Chr2:933,364,201  | 1.44E−04| 13.4 | Xuos76228b_WFS | A/V SECCE2Rv1G0140890.1;Chalcone synthase 2 (CHS2) |
| Xuos372616 | C/A | Chr5:755,939,404  | 1.47E−04| 14.8 | Xuos372616_WFS | None (G/G) FRIGIDA-like 4 (FRL4-like) |

EXECUTED 1 G36/R36 variation within one of two preserved cysteine loops preceding the LRR-like region, which may be proteolytically separated from the ice-binding domain in the apoplastic space as suggested by studies in *Brachypodium distachyon* 69. Thus, the amino acid substitution for the rye IRIP was not expected to influence protein binding to ice crystals, but could affect the function of the N-terminal region (Fig. S4). Almost all genotypes (36 out of 38) with low and very low WFS encoded only the R36 IRIP variant (Table S4), whereas the highest WFS, LTT, and FLN values were favored by the G36 IRIP variant (Fig. 2). Many additional IRIP variants may be produced from the 7R locus as the reference genome carries at least 13 IRIP-like coding regions (> 80% nucleotide sequence identity) arranged in tandem with SECC7Rv1G0458910.1
ing to NetPhos 3.1 analysis. Genotypes with low or very low WFS encoded only the A30 variant for COR413-PM, which was also associated with low FLN (Fig. 2; Table S4). High WFS and FLN values were associated with the T30 modification for COR413-PM (Fig. 2; Table S4).

**Inducer of CBF expression 1 (ICE1).** Marker Xuos615052 associated with WFS, LTT, FLN and PGH (Table 2) coincided with an *Inducer of CBF Expression 1 (ICE1)* gene (SECCE3Rv1G0209310.1) positioned on chromosome 3R (937,715 Mb; Fig. 4). Like the related ICE2, ICE1 is a MYC-like basic helix-loop-helix (bHLH)
transcription factor, which binds to MYC-recognition sequences present in CBF promoter regions and thereby stimulate cold-induced CBF expression. Induced CBFs then activate cold-dependent and absic acid-independent expression of COR genes conferring cold tolerance in plants. SECCER1v1G0061510.1 showed 97% sequence identity to orthologous TaICE41 on chromosomes 3D in hexaploid wheat, which contributes to increased freezing tolerance when overexpressed in Arabidopsis. The Xuos615052 SNP caused a L283/H283 variation at the end of the DNA-binding bHLH domain and start of a ZIP dimerization domain based on domain structures predicted for ICE proteins. Similar to the rye H283 ICE1 variant, a basic residue (K) is present in the activation domain of systemic defense; thus, JA is involved in several stress responses. The more significant SNP for JAR1 (Xuos75199) caused an N11/T11 variation, of which the T11 variant was common among genotypes with low or very low WFS (29 out of 38 genotypes; Table S4) and associated with low FLN values (Fig. 2). JAR1 encoded by genotypes with high and very high WFS generally carried alleles coding for both (N11/T11) variants (Fig. 2; Table S4).

**Jasmonate-resistant 1 (JAR1).** Two markers, Xuos75199 (Table 2) and Xuos75046 (Table S3), both associated with WFS and FLN, targeted the same gene, (SECCE1Rv1G0061510.1) on chromosome 1R (716,318 Mb; Fig. 4). The SECCER1v1G0061510.1 transcript showed strong sequence identity (95%) to an Aegilops tauschii subsp. tauschii, Jasmonate-resistant 1 (JAR1) mRNA isoform X3. JAR1 is a cytosolic auxin-inducible enzyme catalyzing reversible conversion of Jasmonic Acid (JA) to JA-Ile, which is the active growth regulator in JA signaling. Both JA and methyl jasmonate (MeJA) can be converted to JA-Ile and participate in the activation of systemic defense; thus, JA is involved in several stress responses. The more significant SNP for JAR1 (Xuos75199) caused an N11/T11 variation, of which the T11 variant was common among genotypes with low or very low WFS (29 out of 38 genotypes; Table S4) and associated with low FLN values (Fig. 2). JAR1 encoded by genotypes with high and very high WFS generally carried alleles coding for both (N11/T11) variants (Fig. 2; Table S4).

**FRIGIDA-like 4 (FRL4-like).** Marker Xuos372616 associated with WFS corresponded to SECCE5Rv1G0358510.1 located on chromosome 5R (755,939 Mb; Table 2; Fig. 4) and coding for a FRIGIDA-like protein based on BLASTp searches. High sequence identities were also noted to a C-terminal FRIGIDA-like domains carried by 5′-3′ exoribonuclease 4 (XRN4) proteins encoded by Brachypodium distachyon (XP_024313537), Oryza sativa (XP_025879723), and Tritium urartu (EM554435). Interestingly, SECCER5Rv1G0358510.1 located 1,343 bp upstream and oriented in the same direction as SECCER5Rv1G0358510.1 codes for a XRN4-like protein without the FRIGIDA-like domain. As evidence is lacking for a rye transcript coding for both XRN4 and FRIGIDA-like domains, we assume the gene targeted by Xuos372616 only encoded the FRIGIDA-like protein. Amino acid alignment to FRIGIDA-like proteins characterized in Arabidopsis revealed the rye FRIGIDA-like protein had highest similarity (59–61%; data not shown) to subfamily IV represented by AtFRL4a and AtFRL4b. AtFRL4b is co-regulated with genes encoding chromatin remodeling factors and proteins involved in floral transition and maize FRL4 also participates in a chromatin remodeling network involving MADS box factors controlling floral development. However, no information regarding FRL4 action among Triticeae species is available. For the rye FRL4-like protein, the Xuos372616 SNP did not alter the encoded amino acid (A283). Homozygous A/A alleles at SNP site dominated among genotypes with low or very low WFS (29 out of 38 genotypes; Table S4) and associated with low FLN values (Fig. 2). JAR1 encoded by genotypes with high and very high WFS generally carried alleles coding for both (N11/T11) variants (Fig. 2; Table S4).

**Chloroplast unusual positioning protein-1 (CHUP1).** Two of the markers identified in the study coincided with genes controlling chloroplast movement. Marker Xuos2264 associated with WFS, FLN, and PCH was of higher significance (Table 2) and corresponded to SECCER1v1G0013960.1 gene positioned on chromosome 1R (109,131 Mb; Fig. 4). This gene codes for a Chloroplast Unusual Positioning Protein-1 (CHUP1), which mediates blue light-regulated translocation of chloroplasts to specific sites on the plasma membrane depending on the intensity of incoming light. WEB1 has a role in blue-light-induced reorganization of cp-actin filaments during the avoidance response and is associated with the speed chloroplast move. The identification of highly significant SNPs targeting two different genes involved in chloroplast relocation suggest positioning of chloroplasts could be an important factor for WFS in rye.

**Phenylalanine Ammonia-lyase 8 (PAL8) and Chalcone Synthase 2 (CHS2).** Among the highly significant markers for WFS were markers for genes encoding enzymes in the phenylpropanoid pathway. A Phenylalanine Ammonia-lyase 8 (PAL8) gene (SECCER6Rv1G0399250.1) located on chromosome 6R (434,208 Mb; Fig. 4; Table 2) was targeted by marker Xuos519455 and a Chalcone Synthase 2 (CHS2) gene (SECCER2v1G0140890.1) by two markers (Xuos76228a and Xuos76228b) on chromosome 2R (933,364 Mb; Fig. 4; Table 2). PAL is the first enzyme in the general phenylpropanoid pathway and catalyzes channeling of L-Phe from the primary metabo-
lism to synthesis of trans-cinnamic acid. This enzymatic step constitutes the starting point for numerous polyphenol compounds such as flavonoids, stilbenes, and anthocyanins in plants\textsuperscript{85}. At the Xos519453 SNP site for PAL8, a F291/I291 variation was encoded. The alternative I291 variant was very common in genotypes with low and very low WFS (35 out of 38), whereas the reference F291 allele was primarily encoded by the most winter-hardy genotypes (Fig. 2; Table S4). CHS2 acts downstream of PAL and diverts compounds away from the general phenylpropanoid pathway and towards flavonoids by catalyzing condensation of three malonyl-CoA molecules with one 4-coumaroyl CoA molecule to form naringenin chalcone. The Xos76228a SNP caused a Q79/E79 variation and Xos76228b SNP an A104/V104 variation for CHS2 predicted from SECCE2Rv1G0140890.1. The reference Q79 and A104 residues were favored by genotypes with low and very low WFS (34 out of 38 genotypes; Fig. 2; Table S4), whereas the alternative E79 and V104 residues were most common among genotypes with very high WFS (Fig. 2; Table S4).

### Variation for CBF genes showed low significance for WFS.

In this study, three alternative markers for WFS matched CBF genes at FR-R2 on chromosome 5R, but at lower significance level (p < 0.01; Table S3) than those listed in Table 2. The genes targeted by markers Xos639733, Xos390086, and Xos370362 were CBFIIId-19 (SECCE5Rv1G0340540.1), CBFIVa-2.2 (SECCE5Rv1G0340460.1), and CBFIIId-6.2 (SECCE5Rv1G0340660.1), respectively (5R:614.963–616.449 Mb; Fig. 4). SNP variations for CBFIIId-19 were synonymous (D90/D90), whereas SNP variation for CBFIVa-2.2 caused an H113/P113 change within encoded AP2 DNA-binding region and immediately prior to signature motif CMIII-1 (Fig. S4). SNP variation for CBFIIId-6.2 resulted in a L224 to H224 variation within CBFIIId-6.2 C-terminal region (Fig. S4). A preference for the alternative alleles for CBFIVa-2.2 (P113), and CBFIIId-6 (H113) was found for genotypes with low or very low WFS (data not shown). Besides CBFs at FR-A2, Xos328096 marker associated with FLN matched SECCE4Rv1G0267870.1 encoding CBFII-5 from chromosome 4R (810.923 Mb; Fig. 4). The SNP at CBFII-5 caused an R37 to P37 change within conserved signature sequence CMIII-3 (Fig. S4), which is essential for Arabidopsis CBF1 transcriptional activity\textsuperscript{86}. Whether any of the amino acid variations identified for CBFIVa-2.2, CBFIIId-6, or CBFII-5 affected CBF function in rye remains to be demonstrated.

### Bi-plot analysis supports WFS and associated traits are controlled by several common factors.

The four traits analyzed in the study were analyzed by PCA, which revealed relatively high amount of trait variation explained by PC1 (50.9 to 57.8%) and PC2 (9.7 to 11.3%) in the four trait bi-plots (Fig. 3). In all bi-plots, genotypes with very high and high WFS were mostly found in the first and fourth quadrants, whereas genotypes with low WFS were located in the second and third quadrants. Seven vectors representing SNP markers for candidate genes IRI1, COR413-PM, FRL4-like, JAR1, ICE1, BIPP2C1, and CHUP1 were visualized for each of the LTT, FLN, and PGH bi-plots, which all showed similar patterns (Fig. 3). In these three bi-plots, the first six vectors related relatively well with each other and contributed mainly to PC1, whereas CHUP1 vector showed a different direction and was mainly associated with PC2. The WFS bi-plot was also represented by IRI1, COR413-PM, FRL4-like, ICE1, and JAR1 vectors, but in contrast to the other bi-plots, the JAR1 vector was oriented closer to the CHUP1 vector. WFS was also represented by two additional vectors (markers Xos527421, Xos615123; Table S3) corresponding to gibberellin 2-beta-dioxygenase D11 (GA2OX-D11; 7R:105.61 Mb) and expansin-like A1 (EXLA1; 5R:823.29 Mb) genes, respectively (Fig. 4). Both WFS and PGH were associated with GA2OX-D11 (Xos527421_WFS; Xos527421_PGH; Table S3), which encodes an enzyme deactivating gibberellins and stimulating of compact growth\textsuperscript{87}, which is generally displayed by winter-hardy cereals. Expansins have a role in cell wall loosening causing changes to cell wall elasticity and are generally associated with stress responses affecting plant growth; however, the exact function for EXLA1 is not known\textsuperscript{88}.

### Discussion

**Rye, a model for study of winter hardiness in cereals.** The LD decay rate is the major factor limiting mapping resolution in GWAS and it is largely determined by the mating system (self or out-crossing), recombination rate, distribution of recombination hot spots on the genome, and population structure\textsuperscript{89}. In this respect, rye with a low average intra-genic LD of about 4.3 kb (Table S2) or less\textsuperscript{90} is well suited for association mapping, which provides a considerable higher mapping resolution than bi-parental mapping\textsuperscript{91}. In this study, the mapping precision was further improved by using a reduced-representation GBS library targeting low-copy hypomethylated regions for SNP identification\textsuperscript{92}. Successful accomplishment of this strategy was demonstrated by 90 out of 189 (47.6%) significant SNP markers mapped to annotated genes (Table S3).

The PCA and STRUCTURE analyses grouped the rye population of 96 genotypes into three main subpopulations (Figs. 1, S1, S2; Table S1). Many of the winter-hardy Canadian rye cultivars were grouped together with genotypes with relatively low winter-hardiness levels (Table S1). Overall, the clustering of rye genotypes did not correlate well to their geographical origin or WFS values (Table S1), but a slight preference around the Carsten genepool was indicated for genotypes with very high WFS (Fig. 1; Table S1). Although breeding has reduced the genetic pool of rye\textsuperscript{93}, it has still been possible to select/develop very cold hardy rye varieties with winter hardiness far exceeding that of wheat and barley cultivars, which have undergone loss of diversity for genes conferring high WFS. Thus, the key genetic factors required to improve winter hardness in tender winter cereals are present in certain rye genotypes.

### Adjustment of photosynthesis during cold acclimation.

Allele differences for CHUP1 and WEB1 encoding proteins involved in chloroplast movement\textsuperscript{94,95} were associated with WFS levels within the rye population (Table S3). CHUP1 allele variation was also suggested by the PCA bi-plot analysis to affect PGH, FLN, and LTT trait values (Fig. 3), which are determined during the cold acclimation process. Chloroplast movement to
the periclinal cell wall is suggested to enhance photosynthesis and biomass production under low light conditions, whereas movement to anticlinal cell wall was initially postulated to protect the photosynthetic apparatus from high light. Later studies suggest the avoidance reaction fine-tunes light signals or regulates signaling pathways. Chloroplasts have a very important role during cold acclimation in winter cereals, as they adjust the photosynthesis efficiency upon response to environmental temperature and light signals. This allows certain cold-hardy winter cereals to increase their photosynthesis efficiency during cold acclimation when compared to less winter-hardy genotypes. A study in Arabidopsis supports the photosynthetic acclimation depends on chloroplast relocation which coordinates photosynthesis with down-stream carbohydrate metabolism. Retrograde and anterograde signaling during cold stress may also be affected by chloroplast repositioning as nuclei attach...
Figure 4. Location of significant MTAs identified on rye physical map. Locations refers to the Lo7 v1 pseudomolecules 2021. Most significant markers ($p < 1.49e^{-04}$; Table 2) are indicated in bold. Location of FR-R2, VRN1, and discussed candidate genes are shown.
and co-migrate with chloroplasts to new destinations on the plasma membrane. Although no amino acid difference was associated with SNP variations for CHUP1 or WEB1, it is possible these genes are differentially regulated by light and/or temperature signals in rye genotypes with varying degree of WFS. Further analysis of rye CHUP1 and WEB1 genes and expression levels during cold acclimation are needed to confirm their role during cold acclimation in rye.

**Protection of membrane integrity during cold acclimation.** Like the organelle membranes, the plasma membrane is prone to freezing injury and must be protected to withstand cell dehydration caused by apoplastic freezing and membrane fracturing by growing ice crystals. The protection mechanisms generally include membrane rigidification, production of proteins with membrane stabilization properties, and prevention of ice crystal formation and growth. One of the cold-induced plasma membrane proteins is the multi-spanning COR413-PM, for which sequence variation was associated with WFS and FLN (Table 2). Overexpression of PsCOR413PM2 isolated from Phlox subulata in Arabidopsis causes an increased Ca\(^{2+}\) flux in roots and higher expression of CBF1, CBF3 and certain COR genes not induced by CBFs. Thus, COR413 acts early in the cold-induction pathway and both COR413-PM1 and related chloroplast COR413-IM in Arabidopsis are included in a set of 56 plant core environmental stress genes triggering systemic stress responses; however, the systemic signal inducing this response is not known. Although the precise COR413-PM function has not been determined, it has been suggested the membrane-bound protein maintains membrane fluidity at low temperature, acts as a G protein-coupled receptor involved in signaling, or stabilizes the membrane lipid structure.

Like COR413-PM, IRIPs are important for membrane protection. In this study, WFS, LTT, and PGH for the rye population were strongly associated with allelic variation for one of many IRIP genes encoded from a 7R locus (Table 2). IRIPs contribution to frost-hardiness is supported by an increased IRIP expression during cold acclimation in wheat and improved freezing tolerance in transgenic Arabidopsis expressing LpIRI-a, LpIRI-b, LpIRI2, or LpIRI3 from Lolium perenne. A reduced freezing tolerance caused by down-regulation of IRIPs in Brachypodium distachyon further confirms ice-binding proteins contribute to development of frost tolerance in cereals. Prevention of ice crystal growth in the apoplast is particularly critical during freeze/thaw cycles, which may occur during winter or early spring for winter cereals.

Unexpectedly, our study suggested IRIP sequence variation affected PGH (Table 2), which is caused by curvature of shoots emerging from the crown during cold acclimation. The phenotype is likely due to altered gravitropic responses affecting auxin distribution at the shoot base. The shoot is one of first plant tissues that freezes, despite the lower part of the plant is warmer than the upper part. Thus, very cold hardy cereals have evolved a more efficient frost protection for SAM than less cold-hardy genotypes to assure survival of the crown and certain higher expression of COR genes not induced by CBFs98. Thus, IRIPs contribution to frost-hardiness is supported by an increased IRIP expression during cold acclimation in wheat and improved freezing tolerance in transgenic Arabidopsis expressing LpIRI-a, LpIRI-b, LpIRI2, or LpIRI3 from Lolium perenne. One of the cold-induced plasma membrane proteins is the multi-spanning COR413-PM, for which sequence variation was associated with WFS and FLN (Table 2). Overexpression of PsCOR413PM2 isolated from Phlox subulata in Arabidopsis causes an increased Ca\(^{2+}\) flux in roots and higher expression of CBF1, CBF3 and certain COR genes not induced by CBFs. Thus, COR413 acts early in the cold-induction pathway and both COR413-PM1 and related chloroplast COR413-IM in Arabidopsis are included in a set of 56 plant core environmental stress genes triggering systemic stress responses; however, the systemic signal inducing this response is not known. Although the precise COR413-PM function has not been determined, it has been suggested the membrane-bound protein maintains membrane fluidity at low temperature, acts as a G protein-coupled receptor involved in signaling, or stabilizes the membrane lipid structure.

**Candidate genes belonging to the ICE-CBF-COR regulon.** Like other Triticeae members, rye Fr-R2 carries a cluster of CBF genes, for which haplotype variation is suggested to underlie differences in freezing tolerance among rye accessions. However, the significant associations between WFS and three different CBF genes (CBFIIId-D19, CBFV-a-2.2, and CBFIIId-a-6) at Fr-R2 were not among the strongest MTAs (Table S3). Transcript analysis of rye CBF genes indicates existence of three CBFV-a-2.2 and two CBFIIId-a-6 genes. This could possibly explain the low significance for individual CBF alleles as complementary functions may be exerted by additional copies of near-identical CBFs; thus, SNP variations for multi-copy CBF genes contribute little to phenotype differences within the population. Lack of strong association between SNP variations for individual CBF genes at Fr-A2 and frost resistance is also seen in a panel of 1739 European winter wheat genotypes studied by association mapping. However, most mapping studies of bi-parental populations in wheat, barley, and rye show Fr-2 has a major role for frost tolerance. Due to the limited number of recombination events and narrow genetic variation within the populations studied, the bi-parental confidence intervals for QTL regions are large, and for the Fr loci on group five chromosomes, defined by a composite effect from several CBF genes, their expression levels, interactions, and copy number effects.

Although variation for individual CBF genes appears to have a minor role for WFS, several candidate genes belonging to the ICE-CBF-COR regulon were identified. One strong candidate gene was ICE1 (Table 2), for which amino acid variation at the end of the DNA binding bHLH region and/or start of zipper region was associated with WFS, LTT, FLN and PGH (Fig. S4). In a previous study, allele variation within an intron for ScICE2, probably allelic with ICE1 identified in this study, was associated with variation for winter hardiness and freezing tolerance among 201 European rye genotypes. Thus, specific ICE alleles may be important for winter hardiness in rye. Since genotypes with low or very low WFS generally encoded ICE1 containing the rare H283 modification, these ICE1 proteins may display reduced affinity for the MYC recognition elements (CANNTG) present in promoters of CBF genes and several other cold-regulated genes. Alternatively, the H283 ICE1 variation, which is also located at the start of the ZIP domain, could affect dimer formation. Through heterodimers formed with various bHLH factors, both ICE1 and ICE2 have a role in the establishment and differentiation of stomata in dicot plants and stomata lineage formation in monocot plants. Through similar heterodimer formations, additional developmental roles for rye ICE1 may exist at SAM as ICE1 allele variation in this study was associated with both FLN and PGH besides WFS and LTT (Table 2).
Study suggests jasmonic acid influences cold acclimation and floral transition. Allele variations for JAR1 was associated with WFS and FLN in our study (Table S3). In support of JAR1 causing variation for WFS in rye are previous studies demonstrating JA has a role in the development of cold stress resistance in Arabidopsis. These studies show cold-induced JA-Ile production leads to activation of the ICE-CBF-COR regulon, which could explain the association between JAR1 and WFS observed in our study (Table 2). The association between JAR1 and FLN (Table 2) further suggested JA-Ile influences timing of vegetative to reproductive transition at SAM during cold acclimation. Such a role for JAR1 is supported by studies in maize and rice associating JAR1 activity with delayed juvenile-to-adult transition at SAM in both species. During vernalization in wheat, the MeJA levels are shown to increase, but they decline post-vernalization when plants return to normal temperature. An increased expression of VRN1 and VRN3 and delayed flowering is demonstrated by exogenous application of MeJA to spring wheat, which suggests JA regulates vernalization requirement and modulates flowering time in wheat. Thus, it is possible the association between JAR1, WFS and FLN observed in our study involves JAR1-mediated regulation of VRN1 with effects on timing of floral transition.

Another JA-induced effect during cold acclimation in Arabidopsis is biosynthesis of cold-protective secondary metabolites, such as polyamines, glutathione, and anthocyanins. The production of glycosylated cyanins in particular is associated with enhanced winter hardiness in rye population analyzed in this study. Thus, allele variations for JAR1 and two of the genes encoding early enzymes in the phenylpropanoid pathway, PAL8, CHS2 (Table 2), may have a role in increasing anthocyanin levels during cold acclimation in several of the most winter-hardy genotypes. JA can also affect cold acclimation independent of the CBF pathway whereby growth is modulated flowering time in wheat. Thus, JAR1 can have many different roles during and after cold acclimation, which may be reflected in the different positions of the JAR1 vector in the WFS bi-plot versus the LTT, FLN, PGH bi-plots (Fig. 3).

Expansion of stress related alleles in winter-hardy rye. Rye genotypes with high and very high WFS showed an overall higher frequency of allele variants for candidate genes than genotypes with low or very low cold hardiness (Fig. 3; Table S4). This tendency was particularly obvious for IRIP1, COR413-PM, FRL4-like, BIPP2C1, CHUP1, CHS2 and PAL8. The trend suggested the most winter-hardy genotypes have evolved higher fitness to winter conditions through acquisition of additional alleles for some of the important cold-responsive genes. A wide-spread duplications of rye genes as compared to other cereals is supported by a recent near full-length genome assembly of Chinese rye line Weining with a low heterozygosity rate of 0.26%. Similar to the expansion of CBF and IRIP gene families that occurred when Pooidae lineages adapted to extreme temperatures during the Eocene–Oligocene transition, the extra alleles for some of the cold-regulated candidate genes identified in this study may have evolved during periods of cold stress in the history of rye. Gene expansion is not unique to cereals, but has occurred throughout the evolutionary history of plants and is predominantly observed for stress-responsive genes. Whether the additional gene copy number for cold-responsive genes identified by our study provides that extra winter-hardiness observed for some of the Canadian rye cultivars when compared to other rye genotypes needs to be explored further.

Concluding remarks

The phenotypic relationships reported earlier for the rye population studied here showed that sub-traits like LTT, FLN, and PGH with high heritability estimates ( 0.45 to 0.81), can be used to dissect a complex trait like WFS. With the abundance of SNP markers provided by GS combined with a draft genome sequence for rye, the association mapping conducted on the rye population identified in total 90 annotated candidate genes (p<0.01). The accuracy of the mapping is strengthened by finding seven of the nine strongest candidate genes have roles in plant adaptation during cold acclimation and contribute to enhanced winter hardiness in cereals and other plants. For two strong candidate genes, BIPP2C and FRL4-like, there are no obvious associations to WFS in Triticeae species. As similar genes in other monocot species are implicated in floral transition, studies of the rye BIPP2C and FRL4-like is likely to advance our knowledge of the cold tolerance process in winter cereals. Thus, future studies are planned with a diverse panel of rye genotypes with varying LTT and/or WFS to validate all the candidate genes identified in this study, many of which seem promising but were not discussed here. The long-term objective is to verify gene variants that enhance winter hardiness in rye and use this information to increase WFS in temperate winter cereals of interest for the Northern hemisphere. Identification of genetic factors contributing to WFS in winter cereals is of particular importance considering the immediate challenges of global warming increasing the demand for crops with higher resilience to variable weather patterns.

Data availability

All data used in the study is included in the manuscript and in the supplementary data files.

Received: 20 September 2021; Accepted: 25 March 2022 Published online: 06 April 2022

References

1. Schlegel, R. H. J. RYE. Genetics, Breeding, and Cultivation (CRC Press, Boca Raton, 2014).
2. Martis, M. M. et al. Reticulate evolution of the rye genome. Plant Cell 25, 3685–3698 (2013).
3. Bauer, E. et al. Towards a whole-genome sequence for rye (Secale cereale L.). Plant J. 89, 853–869 (2017).
4. Rabanus-Wallace, M. T. et al. Chromosome-scale genome assembly provides insights into rye biology, evolution and agronomic potential. Nat. Genet. 53, 564–573 (2021).
5. Li, Y. et al. High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (Secale cereale L.) genes involved in frost response. BMC Plant Biol. 11, 6 (2011).
6. Targonska, M., Bobbok-Bregoezowska, H. & Rakoczy-Trojanowska, M. Assessment of genetic diversity in Secale cereale based on SSR markers. Plant Mol. Biol. Rep. 34, 37–51 (2016).
7. Fowler, D. B., Byrna, B. M. & Greer, K. J. Overwinter low-temperature responses of cereals: Analyses and simulation. Crop Sci. 54, 2395–2405 (2014).
8. Thomashow, M. F. Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50, 571–599 (1999).
9. Trevaskis, B., Hemming, M. N., Dennis, E. S. & Peacock, W. J. The molecular basis of vernalization-induced flowering in cereals. Trends Plant Sci. 12, 352–357 (2007).
10. Franklin, K. A. & Whitelam, G. C. Light-quality regulation of freezing tolerance in Arabidopsis thaliana. Nat. Genet. 39, 1410–1413 (2007).
11. Kurepin, L. V. et al. Role of CBFs as integrators of chloroplast redox, phytochrome and plant hormone signaling during cold acclimation. Int. J. Mol. Sci. 14, 12729–12763 (2013).
12. Fowler, D. B. & Limin, A. E. Interactions among factors regulating phenological development and acclimation rate determine low-temperature tolerance in wheat. Ann. Bot. 94, 717–724 (2004).
13. Livingston III, D. P. The second phase of cold hardening: Freezing tolerance and fructan isomer changes in winter cereal crowns. Crop Sci. 36, 1568–1573 (1996).
14. Le, M. Q., Pagter, M. & Hincha, D. K. Global changes in gene expression, assayed by microarray hybridization and quantitative RT-PCR, during acclimation of three Arabidopsis thaliana accessions to sub-zero temperatures after cold acclimation. Plant Mol. Biol. 87, 1–15 (2015).
15. Hüner, N. P. A., Dahal, K., Bode, R., Kurepin, L. V. & Ivanov, A. G. Photosynthetic acclimation, vernalization, crop productivity and the grand design of photosynthesis. J. Plant Physiol. 203, 29–43 (2016).
16. Limin, A. E. & Fowler, D. B. Morphological and cytological characters associated with low-temperature tolerance in wheat (Triticum aestivum L. em Thell.). Can. J. Plant Sci. 80, 687–692 (2000).
17. Willick, I. R., Takahashi, D., Fowler, D. B., Uemura, M. & Tanino, K. K. Tissue-specific changes in apoplastic proteins and cell wall structure during cold acclimation of winter wheat crowns. J. Exp. Bot. 69, 1221–1234 (2018).
18. Bohn, M., Lütke, S., Sperling, P., Heinz, E. & Dorfling, K. Plasma membrane lipid alterations induced by cold acclimation and abscisic acid treatment of winter wheat seedlings differing in frost resistance. J. Plant Physiol. 164, 146–156 (2007).
19. Griffith, M. & Vaish, M. W. F. Antifreeze proteins in overwintering plants: A tale of two activities. Trends Plant Sci. 9, 399–405 (2004).
20. Guy, C., Kaplan, F., Kopka, J., Selbig, J. & Hincha, D. K. Metabolomics of temperature stress. Plant Physiol. 132, 220–235 (2008).
21. Dahal, K. et al. The effects of phenotypic plasticity on photosynthetic performance in winter rye, winter wheat and Brassica napus. Plant Physiol. 144, 169–188 (2012).
22. Oliver, S. N., Finnegan, E. J., Dennis, E. S., Peacock, W. J. & Trevaskis, B. Vernalization-induced flowering in cereals is associated with changes in histone methylation at the VERNALIZATION1 gene. Proc. Natl. Acad. Sci. U.S.A. 106, 8386–8391 (2009).
23. Luo, X. & He, Y. Experiencing winter for spring flowering: A molecular epigenetic perspective on vernalization. J. Integr. Plant Biol. 62, 104–117 (2020).
24. Smékalová, V., Doskočilová, A., Komis, G. & Šamaj, J. Crosstalk between secondary messengers, hormones and MAPK modules during abiotic stress signalling in plants. Biotechnol. Adv. 32, 2–11 (2014).
25. Eremina, M., Rozhon, W. & Poppenberger, B. Hormonal control of cold stress responses in plants. Cell. Mol. Life Sci. 73, 797–810 (2016).
26. Hu, Y. et al. Jasmonate regulates leaf senescence and tolerance to cold stress: Crosstalk with other phytohormones. J. Exp. Bot. 68, 1361–1369 (2017).
27. Crosatti, C., Rizza, F., Badeck, F. W., Mazzucotelli, E. & Cattivelli, L. Harden the chloroplastic to protect the plant. Physiol. Plant. 147, 55–63 (2013).
28. Cattivelli, L. et al. Chromosome regions and stress-related sequences involved in resistance to abiotic stress in Triticum. Plant Mol. Biol. 48, 649–665 (2002).
29. Zhu, J. et al. Copy number and haplotype variation at the VRN-A1 and central FR-A2 loci are associated with frost tolerance in hexaploid wheat. Theor. Appl. Genet. 127, 1183–1197 (2014).
30. Bäga, M., Fowler, D. B. & Chibbar, R. N. Identification of genomic regions determining the phenological development leading to floral transition in wheat (Triticum aestivum L.). J. Exp. Bot. 60, 3575–3585 (2009).
31. Dhillon, T. et al. Regulation of freezing tolerance and flowering in temperate cereals: The VRN-I connection. Plant Physiol. 153, 1846–1858 (2010).
32. Pearce, S. et al. Large deletions in the CBP gene cluster at the Fr-B2 locus are associated with reduced frost tolerance in wheat. Theor. Appl. Genet. 126, 2683–2697 (2013).
33. Würschum, T., Longin, C. H. F., Hahn, V., Tucker, M. R. & Leiser, W. L. Copy number variations of CBP genes at the Fr-A2 locus are essential components of winter hardiness in wheat. Plant J. 89, 764–773 (2017).
34. Francia, E. et al. Fine mapping of a HvCBF gene cluster at the frost resistance locus Fr-H2 in barley. Theor. Appl. Genet. 115, 1083–1091 (2007).
35. Erath, W. et al. Exploring new alleles for frost tolerance in winter rye. Theor. Appl. Genet. 130, 2151–2164 (2017).
36. Li, Y. et al. Association analysis of frost tolerance in rye using candidate genes and phenotypic data from controlled, semi-controlled, and field phenotyping platforms. BMC Plant Biol. 11, 146 (2011).
37. Stockinger, E. J., Gilmour, S. J. & Thomashow, M. F. Arabidopsis thaliana CBF1 encodes an AP2-domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc. Natl. Acad. Sci. U.S.A. 94, 1035–1040 (1997).
38. Fowler, S. & Thomashow, M. F. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14, 1675–1690 (2002).
39. Jia, Y. et al. The cbf3 triple mutants reveal the essential functions of CBF3 in cold acclimation and allow the definition of CBF regulations in Arabidopsis, New Phytol. 212, 345–353 (2016).
40. Bahmani, H. et al. The relationships between plant developmental traits and winter field survival in rye (Secale cereale L.). Plants 10, 1–17 (2021).
41. Alqudah, A. M., Sallam, A., Baenziger, P. S. & Börner, A. GWAS: Fast-forwarding gene identification and characterization in temperate cereals: Lessons from barley—a review. J. Adv. Res. 22, 119–135 (2020).
42. Nguyen, K. L., Grondin, A., Courtois, B. & Gantet, P. Next-generation sequencing accelerates crop gene discovery. Trends Plant Sci. 24, 263–274 (2019).
43. Piepho, H. P., Möhring, J., Melching, A. E. & Büchse, A. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161, 209–228 (2008).
44. Poland, J. A., Brown, P. J., Sorrells, M. E. & Jannink, J.-L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 8, e67925 (2013).
45. Mascher, M., Wu, S., St. Amand, P., Stein, N. & Poland, J. Application of genotyping-by-sequencing on semiconductor sequencing platforms: A comparison of genetic and reference-based marker ordering in barley. PLoS One 8, 76925 (2013).
Tremblay, K., Ouellet, F., Fournier, J., Danyluk, J. & Sarhan, F. Molecular characterization and origin of novel bipartite cold-tolerant maize genes based on comparative genomics. *BMC Genomics* 21, 357–419 (2019).

Sandve, S. R., Rudi, H., Asp, T. & Rognli, O. A. Tracking the evolution of a cold stress associated gene family in cold tolerant barley (*Hordeum vulgare*). *Theor. Appl. Genet.* 120, 291–299 (2010).

Xue, T. et al. Genome-wide expression analysis of protein phosphatase 2C in rice and Arabidopsis. *BMC Genomics* 9, 550 (2008).

Singh, A., Giri, I., Kapoor, S., Tyagi, A. K. & Pandey, G. K. Protein phosphatase complement in rice: Genome-wide identification and transcriptional analysis under abiotic stress conditions and reproductive development. *BMC Genomics* 11, 435 (2010).

Yu, X. et al. Genome-wide identification and homologous expression analysis of PP2C genes in wheat (*Triticum aestivum L.*). *Front. Genet.* 10, 561 (2019).

Hu, X., Song, F. & Zheng, Z. Molecular characterization and expression analysis of a rice protein phosphatase 2C gene, OsBPP2C1, and overexpression in transgenic tobacco conferred enhanced disease resistance and abiotic tolerance. *Physiol. Plant.* 127, 225–236 (2006).

Ni, L. et al. Abscissic acid inhibits rice protein phosphatase PP45 via H2O2 and relieves repression of the Ca2+/CaM-dependent protein kinase DM13. *Plant Cell* 31, 128–152 (2019).

Tremblay, K., Oueltet, F., Fournier, J., Danyluk, J. & Sarhan, F. Molecular characterization and origin of novel bipartite cold-regulated ice recrystallization inhibition proteins from cereals. *BMC Genomics* 21, 357–419 (2019).

Breton, G., Danyluk, J., Frenette Charron, J.-B. & Sarhan, F. Expression profiling and bioinformatic analyses of a novel stress-regulated gene 1 operator (G1O) in maize (*Zea mays*). *Plant Physiol.* 179, 2117–2129 (2008).

Cosgrove, D. J. Plant expansins: Diversity and interactions with plant cell walls. *Curr. Opin. Plant Biol.* 25, 162–172 (2015).

Flint-Garcia, S. A., Thornberry, J. M. & Buckler IV, E. S. Structure of linkage disequilibrium in plants. *Ann. Rev. Plant Biol.* 54, 357–374 (2003).
90. Wilson, S. & Ruban, A. V. Rethinking the influence of chloroplast movements on non-photosynthetic quenching and photo-protection. *Plant Physiol.* **183**, 1213–1223 (2020).
91. Wada, M. Chloroplast movement. *Plant Sci.* **210**, 177–182 (2013).
92. Húner, N. P. A. et al. Chloroplast redox imbalance governs phenotypic plasticity: the ‘grand design of photosynthesis’ revisited. *Front. Plant Sci.* **3**, 255 (2012).
93. Kitashova, A. et al. Impaired chloroplast positioning affects photosynthetic capacity and regulation of the central carbohydrate metabolism during cold acclimation. *Photophys. Res.* **147**, 49–60 (2021).
94. Higa, T.; Suetsumu, N.; Kong, S.-G. & Wada, M. Actin-dependent plastid movement is required for motive force generation in directional nuclear movement in plants. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4327–4331 (2014).
95. Örvar, B. L.; Sangwan, V.; Omann, F. & Dhindsa, R. S. Early steps in cold sensing by plant cells: The role of actin cytoskeleton and membrane fluidity. *Plant J.* **23**, 785–794 (2000).
96. Uemura, M. Chloroplast movement. *Front. Plant Sci.* **2**, 2153 (2017).
97. Zhou, A. et al. PsCor43pm2, a plasma membrane-localized, cold-regulated protein from *Phlox subulata*, confers low temperature tolerance in *Arabidopsis*. *Int. J. Mol. Sci.* **19**, 2379 (2018).
98. Hahn, A. et al. Plant core environmental stress response genes are systemically coordinately during abiotic stresses. *Int. J. Mol. Sci.* **14**, 7617–7641 (2013).
99. Su, C. et al. Proteomics analysis of the function of a novel cold-regulated multispansing transmembrane protein COR413-PM1 in *Arabidopsis*. *Int. J. Mol. Sci.* **19**, 2572 (2018).
100. Houde, M. et al. Wheat EST resources for functional genomics of abiotic stress. *BMC Genomics* **7**, 149 (2006).
101. Zhang, C.; Fei, S.; Arora, R.; & Hannapel, D. J. Ice recrystalization inhibition proteins of perennial ryegrass enhance freezing tolerance. *Plants* **232**, 155–164 (2010).
102. Dong, Z. et al. Maize LAZY1 mediates shoot gravitropism and inflorescence development through regulating auxin transport, auxin signaling, and light response. *Plant Physiol.* **163**, 1306–1322 (2013).
103. Campoli, C.; Matus-Cádiz, M. A.; Pozniak, C. J.; Cattivelli, L. & Fowler, D. B. Comparative expression of Chl genes in the *Triticaceae* under different acclimation induction temperatures. *Mol. Genet. Genomics* **282**, 141–152 (2009).
104. Jaworski, J. & Seo, Y. W. Identification of novel C-repeat binding factor (CBF) genes in *rye (Secale cereale)* and expression studies. *Gene* **564**, 82–94 (2019).
105. Zhao, Y. et al. Dissecting the genetic architecture of frost tolerance in Central European winter wheat. *J. Exp. Bot.* **64**, 4453–4460 (2013).
106. Bäga, M. et al. Identification of quantitative trait loci and associated candidate genes for low-temperature tolerance in cold-hardy winter wheat. *Genome* **57**, 53–68 (2007).
107. Francis, E. et al. Two loci on chromosome 5H determine low-temperature tolerance in a ‘Nure’ (winter) × ‘Tremois’ (spring) barley map. *Theor. Appl. Genet.* **108**, 670–680 (2004).
108. Kanaoka, M. M. et al. SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* **20**, 1775–1785 (2008).
109. Raisig, M. T.; Abrash, E.; Bettadapur, A.; Vogel, J. P.; & Bergmann, D. C. Grasses use an alternatively wired bHLH transcription factor network to establish stomatal identity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8326–8331 (2016).
110. Hu, Y.; Jiang, L.; Wang, F. & Yu, D. Jasmonate regulates the inducer of cbf expression–c-repeat binding factor/DRE binding factor1 cascade and freezing tolerance in *Arabidopsis*. *Plant Cell* **25**, 2907–2924 (2013).
111. Híbara, K.-I. et al. Jasmonate regulates juvenile-to-adult phase transition in rice. *Development* **143**, 3407–3416 (2016).
112. Osadchuk, K.; Cheng, C.-L.; & Irish, E. E. Jasmonic acid levels decline in advance of the transition to the adult phase in maize. *Plant Direct* **3**, 1–13 (2019).
113. Dimario, A. O. et al. Transcriptome analysis of an mvp mutant reveals important changes in global gene expression and a role for methyl jasmonate in vernalization and flowering in wheat. *J. Exp. Bot.* **65**, 2271–2286 (2014).
114. Han, X.; Zhang, Y.; Peng, W.; Wang, Z.; & Xie, D. Molecular mechanisms for jasmonate-induced of anthocyanin accumulation in *rice*. *Plant Direct* **5**, 1–13 (2019).
115. Zhang, H. et al. Proteomic analysis of the function of a novel cold-regulated multispansing transmembrane protein COR413-PM1 in *Arabidopsis*. *Int. J. Mol. Sci.* **19**, 2572 (2018).
116. Bahrami, H. et al. Preferential accumulation of glycosylated cyanidins in winter-hardy *rye* (*Secale cereale*) genotypes during cold acclimation. *Environ. Exp. Bot.* **164**, 203–219 (2021).
117. Yang, D.-L. et al. Plant hormone jasmonate prioritizes defense over growth by interfering with gibberellin signaling cascade. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1192–1200 (2012).
118. Li, G. et al. A high-quality genome assembly highlights *rye* genomic characteristics and agronomically important genes. *Nat. Genet.* **53**, 574–584 (2021).
119. Sandve, S. R. & Fjellheim, S. Did gene family expansions during the Eocene–Oligocene boundary climate cooling play a role in Pooideae adaptation to cool climates? *Mol. Ecol.* **19**, 2075–2088 (2010).
120. Hanada, K.; Zou, C.; Lehti-Shiu, M. D.; Shinozaki, K.; & Shiu, S. H. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiol.* **148**, 993–1003 (2008).

Acknowledgements
Several research technicians and summer students over the last six years are acknowledged for help in the field and greenhouse work.

Author contributions
M.B.: Conceptualization, funding acquisition, experimental, data analysis, manuscript first draft preparation and editing. H.B.: Experimental, data analysis, manuscript writing J.L.: Conceptualization, funding acquisition, resources, manuscript review. B.H.: Data analysis, manuscript review and editing R.J.G.: Conceptualization, resources, manuscript review and editing. A.L.: Conceptualization, funding acquisition, data analysis, manuscript review and editing. R.N.C.: Conceptualization, funding acquisition, project management, data analysis, manuscript review and editing. All authors reviewed the manuscript and approved for publication.

Funding
This work was supported by the Canada Research Chairs (RNC), Natural Sciences and Engineering Research Council Canada (NSERC) Discovery grant (RNC), NSERC collaborative research project grant with Western Grains Research Foundation, Saskatchewan Winter Cereals Development Commission, Winter Cereals Manitoba Inc. and Alberta Wheat Commission, and a NSERC Alliance grant with Western Grains Research Foundation and Saskatchewan Winter Cereals Development Commission. HB is a grateful recipient of the Harvey scholarship and Developed scholarship by the Department of Plant Sciences. BH was funding by the European Union's
Horizon 2020 research and innovation programme under grant agreement No 771134 for the project RYE-SUS. RYE-SUS is carried out under the ERA-NET Cofund SusCrop (Grant No. 771134), being part of the Joint Programming Initiative on Agriculture, Food Security and Climate Change (FACCE-JPI).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-09582-2.

Correspondence and requests for materials should be addressed to R.N.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022