Plant, Cell & Environment

Plant, Cell and Environment (2016) **39**, 2570–2579 doi: 10.1111/pce.12812

Original Article

The genetic architecture of freezing tolerance varies across the range of *Arabidopsis thaliana*

Matthew W. Horton[1,2†], Glenda Willems[3†], Eriko Sasaki[1], Maarten Koornneef[4] & Magnus Nordborg[1,3]

[1]*Gregor Mendel Institute, Vienna Biocenter (VBC), Austrian Academy of Sciences, Vienna 1030, Austria,* [2]*Department of Plant and Microbial Biology, University of Zurich, CH-8008 Zurich, Switzerland,* [3]*Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089, USA and* [4]*Max Planck Institute for Plant Breeding Research, D–50892 Cologne, Germany*

ABSTRACT

The capacity to tolerate freezing temperatures limits the geographical distribution of many plants, including several species of agricultural importance. However, the genes involved in freezing tolerance remain largely unknown. Here, we describe the variation in constitutive freezing tolerance that occurs among worldwide accessions of *Arabidopsis thaliana*. We found that although plants from high latitudes tend to be more freezing tolerant than plants from low latitudes, the environmental factors that shape cold adaptation differ across the species range. Consistent with this, we found that the genetic architecture of freezing tolerance also differs across its range. Conventional genome-wide association studies helped identify a priori and other promising candidate genes. However, simultaneously modelling climate variables and freezing tolerance together pinpointed other excellent a priori candidate genes. This suggests that if the selective factor underlying phenotypic variation is known, multi-trait mixed models may aid in identifying the genes that underlie adaptation.

Key-words: climate adaptation; genome-wide association studies; natural variation.

INTRODUCTION

The plant genetic model *Arabidopsis thaliana* is native to Eurasia and, in the last few hundred years, has spread across much of the world. Its vast species range, combined with the availability of inbred lines collected across its distribution (Platt *et al.* 2010), makes it well-suited for examining the ecological factors that drive adaptation to the environment.

In temperate areas and mountainous regions, a major component of climate adaptation is the ability to withstand freezing temperatures. Molecular studies have demonstrated that freezing tolerance is shaped in part by the C-repeat binding factor loci (*CBF1*, *CBF2* and *CBF3*). These three transcription factors are induced after exposure to low non-freezing temperatures (Gilmour *et al.* 1998) and regulate the expression of approximately 100 genes, some of which are known to be cryoprotective (Sieg *et al.* 1996; Steponkus *et al.* 1998; Hughes *et al.* 2013). Overexpressing the *CBF* loci results in increased freezing tolerance without cold acclimation (Jaglo-Ottosen *et al.* 1998; Gilmour *et al.* 2004), which further demonstrates their role in the development of cold hardiness.

The overall contribution of *CBF* genes to freezing tolerance variation, however, is unclear. Some studies have suggested that the *CBF* transcription factors are major determinants of natural variation in freezing tolerance (Alonso-Blanco *et al.* 2005; Oakley *et al.* 2014; Gehan *et al.* 2015), while other studies have found either a weak relationship or no relationship at all (McKhann *et al.* 2008; Zuther *et al.* 2012; Meissner *et al.* 2013). Notably, although *CBF* expression levels are negatively correlated with temperature (Gilmour *et al.* 1998), the freezing tolerance of cold-acclimated and non-acclimated accessions is strongly correlated (Zuther *et al.* 2012), which suggests that much of the among line variation in freezing tolerance is independent of cold acclimation and thus governed by a *CBF*-independent pathway.

One of the major limitations of published research on freezing tolerance is a reliance on linkage studies, which, compared with genome-wide association studies (GWAS), tend to identify candidate regions that span several megabases (Mb) and hundreds of candidate genes. Here, we describe the pattern of variation in freezing tolerance within *A. thaliana*, which we characterized by using three quantitative trait locus (QTL) families and nearly 500 accessions collected from around the world. Although our main aim was to investigate the genetic architecture of freezing tolerance at a global scale, we also fine-mapped major genetic variants by using regional mapping populations, in order to determine if the genetic basis of freezing tolerance varies across the range of *A. thaliana*.

MATERIALS AND METHODS

Plant material/phenotyping

We grew 24 replicate seeds of 499 worldwide accessions (Table S1) and 94 lines derived from each of three F3 families (Col-0 × Löv-1, Col-0 × Vår-26 and Col-0 × Edi-0). All seeds were sown in 98-cell trays (14 × 7), using a randomized block design, before being placed in Conviron climate chambers.

After stratifying the seeds for 5 d (4 °C), the chamber was warmed to 18 °C to facilitate germination; seeds that failed to

Correspondence: M. W. Horton; e-mail: matthew.horton@botinst.uzh.ch
[†]These authors contributed equally to the paper.

germinate within 7 d were excluded from further analyses. The plants were grown for 2 weeks at 18 °C (15 h of light; relative humidity 75%), followed by an additional 7 d at 10 °C (12 h of light, 75%). At the end of the third week, the temperature in the chamber was gradually decreased to −8 °C ($\Delta T = -2\,°C\,h^{-1}$) and held there for 3 h before returning to the starting conditions (10 °C; $\Delta T = 2\,°C\,h^{-1}$). Ten days later, the samples were scored for survivorship. Because of the moderate temperatures used, this approach is expected to mimic a sudden cold event in the fall, when plants are non-acclimated or weakly acclimated. A similar situation can occur in the spring, after plants deharden.

Correlations with geography and climate

Logistic regression was used to assess the relationship between survivorship and latitude or, in separate analyses, 19 variables representing temperature and precipitation (Hijmans *et al.* 2005). Each variable was modelled separately, and Nagelkerke's pseudo-R^2 (Nagelkerke 1991) was used to evaluate the goodness of fit of each model.

Linkage (QTL) analyses

The R/qtl package (Broman *et al.* 2003) was used in all analyses. To identify genomic regions associated with survivorship, we phenotyped 94F3 lines generated from three crosses: Col-0 × Edi-0, Col-0 × Löv-1 and Col-0 × Vår-26. Genotypes from the parental populations were used to create linkage maps for each mapping population, and QTLs were identified by using multiple QTL mapping (MQM). The phenotypes used in MQM were the best linear unbiased predictions from a mixed model that adjusts for the technical covariates batch and tray. Default parameters were used during marker augmentation (using the command, *mqmaugment*), co-factor selection (*mqmautocofactors*) and the QTL scan (*mqmscan*). To evaluate the significance of QTL, 1000 permutations were performed for each trait (*mqmpermutation*).

Genome-wide association analyses

As in the QTL analyses (in the preceding texts), the phenotypes used in GWAS were the best linear unbiased predictions from a mixed model that adjusts for technical covariates. For marginal single nucleotide polymorphism (SNP)-by-SNP tests, we used both a simple linear model and, separately, a model that accounts for population structure (Kang *et al.* 2010). The SNPs were described previously (The 1001 Genomes Consortium 2016).

As an attempt to account for allelic heterogeneity, which may undermine marginal tests, we estimated the sample relatedness at each gene model (±5000 bp surrounding DNA). This 'local' estimate of kinship was compared with the sample's remaining genome-wide relatedness (a 'global' estimate of kinship) by using a likelihood ratio test. For the multi-trait mixed model (MTMM), genetic variants that simultaneously contribute to freezing tolerance and the top climate variables for each population subset (Table 2) were mapped using the software LIMIX (Lippert *et al.* 2014). Climate scores for each accession were treated as the phenotypes and, together with survivorship, transformed to z-scores. The common effect was estimated by using the statistical model described earlier (Sasaki *et al.* 2015).

Gene-set enrichment analyses

To identify the mechanisms that are implicated in freezing tolerance, we asked whether any gene-ontology (GO) terms (ftp://ftp.arabidopsis.org/home/tair/Ontologies/Gene_Ontology/) are over-represented in the results from GWAS. The method was described earlier (Horton *et al.* 2014). Briefly, we split the results from GWAS into 10 kb windows and examined the distribution of individual GO terms to determine whether any of these were significantly over-represented in the top 1% of the GWAS results. Storey's approach (Storey & Tibshirani 2003) was used to account for multiple testing.

Data

The raw data and results from GWAS have been deposited in the Dryad Digital Repository (http://doi.org/10.5061/dryad.72m15).

RESULTS

The environmental factors that shape freezing tolerance

We phenotyped 499 accessions collected from around the world (Fig. 1 and Table S1) and found substantial variation in basal freezing tolerance. As an example, a widely used lab strain from the Cape Verde archipelago (Cvi-0) exhibits low freezing tolerance. Likewise, accessions from the Canary Islands and Italy (Can-0 and Mr-0) are also freezing intolerant.

Figure 1. The samples phenotyped for freezing tolerance were collected across the northern hemisphere; the green dots indicate the sites of origin.

In contrast, the most freezing tolerant accessions were collected in Germany and Sweden. Indeed, samples from low latitudes displayed lower freezing tolerance than accessions from more northerly areas (Fig. 2).

To better explore the relationship between an accession's freezing tolerance and its latitude of origin, we used binary logistic regression (logit) mixed models. The results from logistic regression were consistent with the results from earlier studies (Hannah *et al.* 2006; Zhen & Ungerer 2008; Zuther *et al.* 2012) and confirmed that the odds that an accession survived the freezing treatment increase with its latitude of origin (Table 1). Nevertheless, further analysis revealed that the relationship between latitude and survivorship differs among regional samples. For example, we found no relationship between latitude and freezing tolerance in North America, where *A. thaliana* is believed to have been recently introduced (O'Kane & Al-Shehbaz 1997). Perhaps more surprising is that the association between latitude and freezing tolerance becomes weaker at high latitudes in the native species range. In particular, the trend is stronger for accessions collected from mainland Europe than for accessions collected in Sweden. Given the strong negative correlation between latitude and temperature, this may indicate that there is a minimum temperature beyond which diminishing returns for additional freezing tolerance occur (i.e. a non-linear fitness function). However, it is also possible that other environmental factors are involved.

To examine this further, we considered the relationship between freezing tolerance and 19 climate variables related to seasonal and annual trends in precipitation and temperature (Hijmans *et al.* 2005). In agreement with the analysis of latitude, the climate variable that best predicts the freezing tolerance of accessions from continental Europe is the mean temperature of the coldest quarter (Fig. 3), with the odds of survivorship lower (Table 2) for accessions native to warm regions ($P = 8.13 \times 10^{-24}$). For Swedish accessions, freezing tolerance is not associated with variables related to the temperature at the site of origin. Instead, we found modest evidence that the odds of survival in Sweden increase with the amount of precipitation in the original habitat. Although snow is widely assumed to protect plants from sub-freezing air temperatures, it also maintains the ambient temperature near the melting point of water, which increases the likelihood of cellular damage due to ice recrystallization (Griffith & Yaish 2004; Gudleifsson 2010).

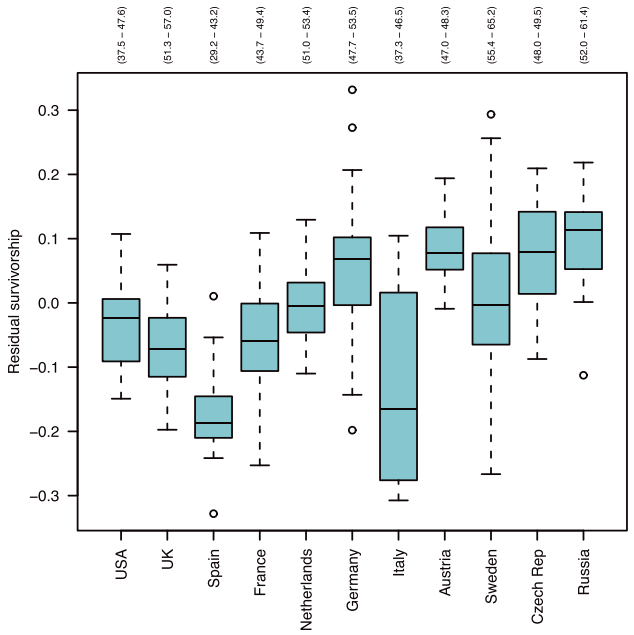

Figure 2. The distribution of freezing tolerance within *Arabidopsis thaliana*. Overall, freezing tolerance increases with the latitude of origin of *A. thaliana* accessions. However, the relationship between latitude and freezing tolerance varies across the species range. Each sample subset has a minimum of five accessions; the latitudinal range of the phenotyped accessions is listed in the margin.

Table 1. Freezing tolerance increases with the latitude of origin of *A. thaliana* accessions. Nevertheless, the relationship between latitude and freezing tolerance varies across the species range. All estimates were generated using logistic regression (logit) mixed models

Population subset	Odds ratio	R^2 (Nagelkerke)	*P*-value
All accessions	1.03	0.22	1.14×10^{-29}
Native range	1.04	0.28	3.86×10^{-36}
Continental Europe	1.08	0.65	2.14×10^{-52}
Americas	0.95	0.11	0.093
Sweden	0.99	0.003	0.37
UK	0.998	0.0001	0.97

The genetic factors that shape freezing tolerance

We found that constitutive freezing tolerance is highly heritable in *A. thaliana* and that it differs among populations (Table 3). To identify the genetic variants that underlie freezing tolerance,

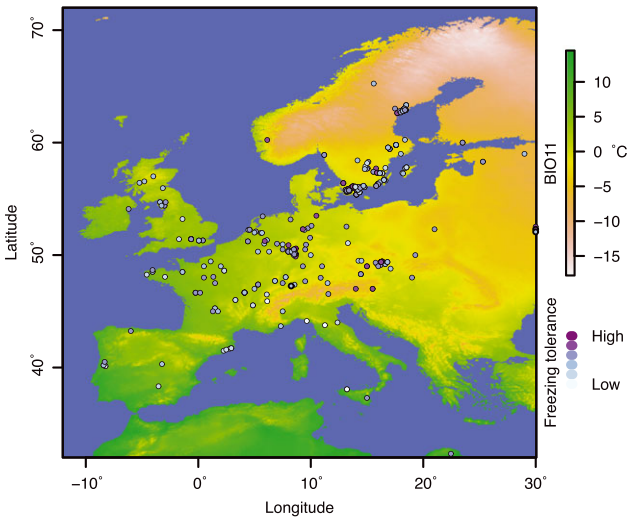

Figure 3. The distribution of the climatic variable that represents the mean temperature in the coldest quarter (BIO11), the variable that best predicts the freezing tolerance of continental European accessions. White dots represent accessions with low freezing tolerance, while purple dots represent accessions with high freezing tolerance.

Table 2. The climate variables that predict freezing tolerance differ across the species range. For example, with respect to accessions from mainland Europe, the odds of survivorship decrease with increasing (mean) temperature and accessions from warm areas are less likely than accessions from cold areas to survive freezing. In contrast, variables related to precipitation are associated with survivorship in Sweden

Region	Climate variable	Odds ratio	R^2 (Nagelkerke)	*P*-value
Continental Europe	Mean temperature of the coldest quarter (BIO11)	0.70	0.47	8.13×10^{-24}
Continental Europe	Annual mean temperature (BIO1)	0.71	0.46	4.20×10^{-23}
Continental Europe	Minimum temperature of the coldest month (BIO6)	0.70	0.46	6.35×10^{-23}
Sweden	Annual precipitation (BIO12)	1.11	0.06	2.39×10^{-04}
Sweden	Precipitation of the coldest quarter (BIO19)	1.11	0.06	3.43×10^{-04}
Sweden	Precipitation of the wettest month (BIO13)	1.11	0.05	7.11×10^{-04}

we used QTL mapping and GWAS. QTL analyses were conducted by using three separate linkage mapping populations (Materials and Methods section), while GWAS were performed on a global and two local mapping populations. Mixed models were used to correct for confounding due to technical artifacts and, in the case of GWAS, population structure (Materials and Methods section).

Using conventional (SNP-by-SNP) GWAS, we found several promising candidate genes in the results from both global and regional analyses (Figs S1–S3). As an example, in the results from global GWAS, the strongest association falls on chromosome 3 (~2.209 Mb) and peaks within a gene (*AT3G07000*) that contains a cysteine/histidine-rich C1 domain (Fig. 4). Cysteine-rich antifreeze proteins have been described in fishes and insects (Davies & Hew 1990; Tyshenko *et al.* 1997; Liou *et al.* 2000), but the limited studies that have been performed in plants have implicated defence-related loci that are believed to defend against cold-adapted pathogens (e.g. snow mold) while also inhibiting ice recrystallization (Griffith & Yaish 2004). Further complicating analyses, three other cysteine-rich loci fall within (+/−) 5 kb of *AT3G07000*. Among these, the low-molecular-weight cysteine-rich proteins 33 and 34 (LCR33 and LCR34) are putative defensins.

We identified several QTLs in the results from linkage studies, suggesting that the observed variation in freezing tolerance is due to variation at many loci. Across mapping populations, these QTLs are largely non-overlapping, which implies that the genetic basis of freezing tolerance may differ across the range of *A. thaliana*. Therefore, we separated accessions from continental Europe and Sweden into two mapping populations and used GWAS to fine-map locally common variants that may contribute to regional differences in freezing tolerance.

For accessions from continental Europe, the strongest association with freezing tolerance falls at a position of 17 793 902 on chromosome 5 (Fig. S2). Golden 2-like (*GLK2*), the closest candidate gene, is key to chloroplast development and

Table 3. Broad-sense heritability of freezing tolerance across populations

Sample subset	H^2	*N*	Number of accessions
All	0.76	11 819	499
Continental Europe	0.83	4322	174
Sweden	0.64	5477	246

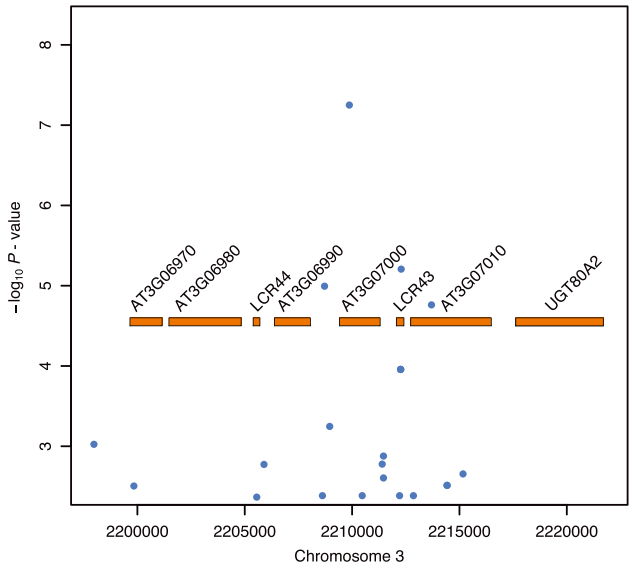

Figure 4. The strongest association in the results from global genome-wide association studies (marginal test) falls within the cysteine-rich gene *AT3G07000*.

regulated by *CBF* homologs in *Brassica napus* (Savitch *et al.* 2005); however, the two genes that lie next to *GLK2*, nuclear cap binding protein 20 and ethylene response factor 9, are also promising candidate genes (Jager *et al.* 2011; Licausi *et al.* 2013). For Swedish accessions, the top candidate genes include pleiotropic drug resistance 11, an ABC transporter whose substrate is unknown, and blade of petiole 2, which contributes to variability in the leaf abscission layer (McKim *et al.* 2008). Both leaf abscission (Jackson & Osborne 1970) and freezing tolerance (Shi *et al.* 2012) are regulated by ethylene, and leaf abscission can occur as a result of freezing stress (Young 1971). *AT3G07000*, or a neighbouring gene, also appears to affect variability in freezing tolerance in Sweden, but, as is the case with blade of petiole 2, SNPs within the region failed to reach genome-wide significance (Fig. S3).

One of the major challenges in GWAS is allelic heterogeneity, in which causal loci that maintain multiple ($n > 2$) alleles are difficult to map using marginal tests, especially analyses that rely on bi-allelic SNPs. Therefore, we re-examined freezing tolerance using a 'local' versus 'global' approach (Sasaki *et al.* 2015). That is, rather than estimating the (marginal) contribution of each SNP to freezing tolerance, we modelled the joint

effect of the SNPs at each gene (Materials and Methods section). Then, using a likelihood ratio test, we compared the estimate of this variance component to one based on genome-wide relatedness. In agreement with marginal SNP GWAS, our approach implicates *AT3G07000*, or a nearby gene, in freezing tolerance. However, we were also able to identify (Table S2) two a priori candidate genes (Catala *et al.* 2011; Hu *et al.* 2013) that were not identified by marginal analyses: elongated hypocotyl 5 and coronatine insensitive 1.

If loci that contribute to freezing tolerance have also diverged among populations, steps to correct for population structure may preclude their discovery using GWAS (Zhao *et al.* 2007). Therefore, after correcting for technical artifacts, we repeated these GWAS by using a simple linear model. As expected, the global analysis resulted in massively inflated *P*-values (Fig. S4). Nevertheless, we still found compelling candidate loci, and the strongest association implicates AT5G24316 in freezing tolerance. AT5G24316 is a proline-rich protein, others of which have been associated with the freezing tolerance of various plant species (Gothandam *et al.* 2010; Huang *et al.* 2011; Peng *et al.* 2015).

Analysing Swedish accessions with a simple linear model did not result in highly skewed *P*-values (Fig. S5). The strongest association (Chr 1: 29 993 180; $P = 1.38 \times 10^{-09}$) falls within the gene wrinkled 4 (*WRI4*). Like the *CBF* loci, *WRI4* is a member of the *APETALA2/ethylene* response factors. Moreover, *WRI4* has a demonstrated role in fatty-acid synthesis (To *et al.* 2012); future experiments will determine whether *WRI4* improves freezing tolerance by changes in membrane fluidity. Using a simple linear model to analyse continental European accessions pinpointed the membrane-tethered transcription factor *NAC089* (Fig. S6), a key component of the unfolded protein response (Yang *et al.* 2014) that occurs in response to stress at the endoplasmic reticulum (Hollien 2013).

A multi-trait mixed model to improve power in GWAS

A major challenge in evolutionary biology is to understand the genetic basis of adaptation. The standard approaches for identifying adaptive loci include (1) genome-wide tests of selection (i.e. 'selection scans'); (2) evaluating the association between genome-wide SNPs and environmental variables believed to underlie adaptation (Hancock *et al.* 2011); and (3) the genetic analysis of ecological traits (Fournier-Level *et al.* 2011). It is possible to combine the latter two approaches with MTMMs, which have been shown to improve power in the analysis of correlated phenotypes (Korte *et al.* 2012; van Heerwaarden *et al.* 2015). Therefore, having found evidence that the freezing tolerance of *A. thaliana* is driven by adaptation to the climate (Table 2), we used MTMMs to identify the polymorphisms that are associated with both freezing tolerance and the climate variable that best predicts freezing tolerance in each sample subset.

The joint (common test) analysis of BIO11 (minimum temperature in the coldest month) and the freezing tolerance of continental European accessions identified three peaks of association (Fig. 5). The most promising of these falls on chromosome 3 (0.874–0.891 Mb) and spans six genes; cold acclimation is known to alter the expression of two of these (Gehan *et al.* 2015), cysteine synthase 26 and Beta-glucosidase 25; the latter is a member of the so-called *CBF* regulon. The second strongest association (Chr 2: 1.905 Mb) falls between two DNAJ heat-shock proteins (AT2G05230 and AT2G05250) and acts as a reminder that gene names are often misleading. Studies of several model systems have revealed that heat shock proteins are either upregulated during cold acclimation or confer freezing tolerance (Porankiewicz & Clarke 1997; Chow & Tung 1998; Teigen *et al.* 2015). The third strongest association (Chr 5: 24.406 Mb) does not identify an obvious candidate gene; however, the expression of *AT5G60710* seems to be regulated (Tao *et al.* 2012), in part, by suppressor of overexpression of constans 1. Suppressor of overexpression of constans 1 also regulates the expression of the *CBF*s (Seo *et al.* 2009).

Despite using more accessions in the analysis of Swedish accessions and local annual precipitation (BIO12), we found only one significant association in the common SNP test (Chr 3: 15.9 Mb; Fig. S7). The association peaks within a pseudogene (*AT3G44175*), and it is unclear which, if any, of the surrounding genes are causal. The list of likely candidates, however, includes the nearby (<6 kb) gene never in mitosis A-related kinase 6. The expression of never in mitosis A-related kinase 6 is altered by ethylene signalling and salt stress (Zhang *et al.* 2011), and the salt response pathway shares genetic overlap with the cold response pathway (Zeller *et al.* 2009). For the interaction test, the top scoring SNP (Chr 1: 29.761 Mb) falls next to BOI-related gene 2. Boi-related genes are frequently implicated in resistance against the fungal pathogen *Botrytis cineria* and, curiously, are regulated by the same transcription factor (SPEECHLESS) that regulates inducer of *CBF* expression 1.

The biological processes that shape freezing tolerance

Finally, to understand the biological processes that affect freezing tolerance, we identified the GO terms over-represented (FDR < 0.01) in the results from GWAS. An analysis of all accessions emphasized phenomena that vary at large scales, including processes related to the circadian rhythm and the red or far-red light-signalling pathway (Table 4). Key components of the clock are involved in cold acclimation (Seo *et al.* 2012), and the circadian expression of genes is shaped through cold acclimation (Bieniawska *et al.* 2008). Furthermore, the circadian clock regulates the diurnal production of reactive oxygen species, and the response to reactive oxygen species was significantly enriched for both the global and Swedish population sets. For Sweden, the most strongly enriched category relates to the synthesis of pentacyclic triterpenes, a diverse family of metabolites that are largely uncharacterized. Many triterpenes, however, are known to accumulate in the wax layer on the surfaces of plants, which may indicate that triterpenes alter the leaf water interface.

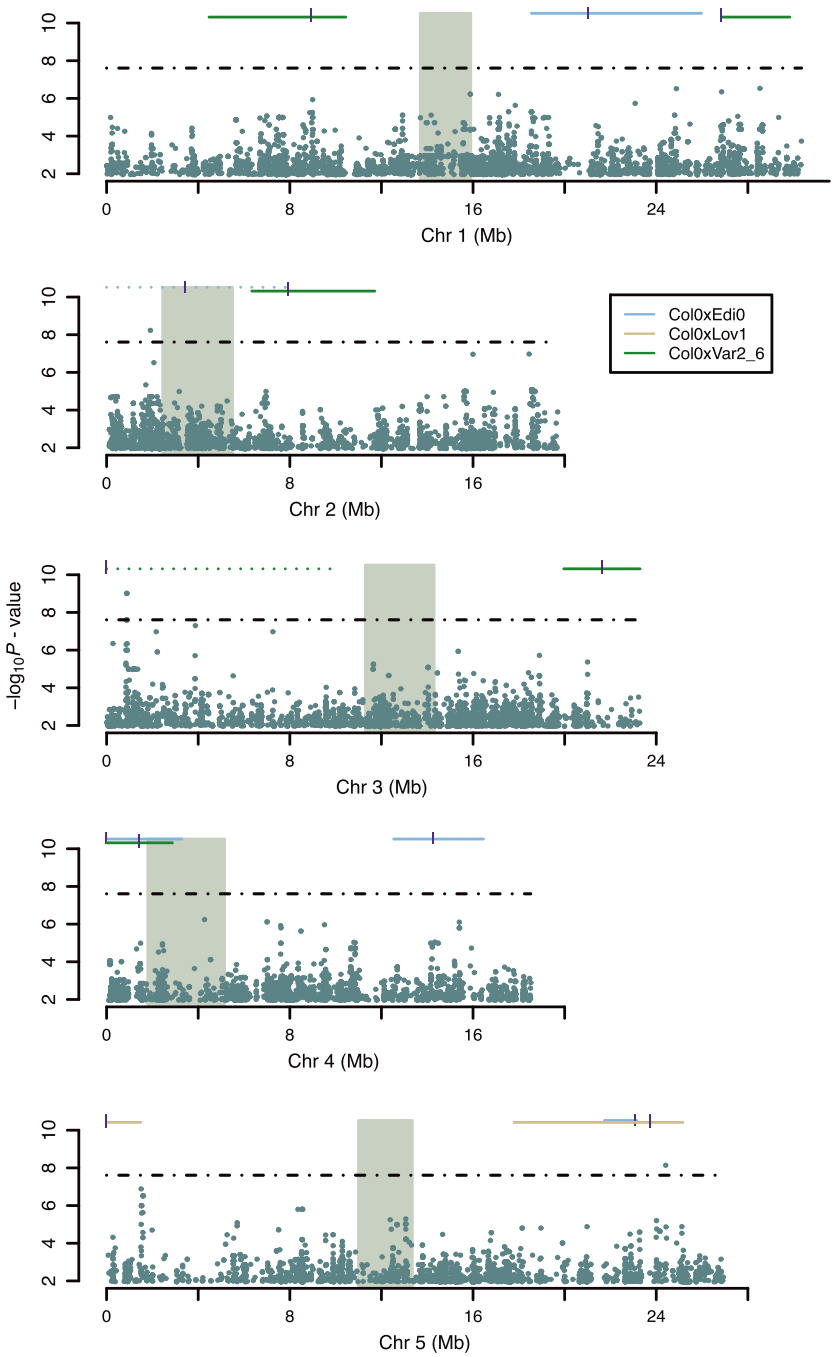

Figure 5. Simultaneously modelling the freezing tolerance of continental European accessions with the climate variable that best predicts survivorship after freezing (BIO11) aids in identifying the a priori candidate gene Beta-glucosidase 25 (Chr 3: ~0.88 Mb). Beta-glucosidase 25 is a member of the C-repeat binding factor regulon. Other strong associations include two DNAJ heat shock proteins on chromosome 2 (1.906 Mb). The QTLs identified using three separate mapping populations are shown above each chromosome. The significance levels were estimated using permutations and are illustrated using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line. For the multi-trait mixed model, the Bonferroni threshold is shown. The centromere is shaded.

DISCUSSION

In *A. thaliana*, freezing tolerance has been extensively studied by using mutants (Gilmour *et al.* 1998; Jaglo-Ottosen *et al.* 1998; Xin & Browse 1998; Gilmour *et al.* 2004) and the progeny of interaccession crosses (Alonso-Blanco *et al.* 2005; Meissner *et al.* 2013; Oakley *et al.* 2014; Gehan *et al.* 2015). However, the genes that contribute to freezing tolerance variation remain largely unknown. We used 499 accessions of *A. thaliana* to explore the spatial distribution of constitutive freezing tolerance. We found that variation in freezing tolerance within *A. thaliana* is extensive (Fig. 2), and that the environmental factors that select for freezing tolerance differ across the species range. With

Table 4. The biological processes enriched in the 1% tail from GWAS of freezing tolerance

Population subset	Biological process	Enrichment	FDR $q < 0.01$
All	Transferase activity	7.8	0.00024
All	Response to reactive oxygen species	13.9	0.00187
All	Red or far-red light signalling pathway	12.2	0.00205
All	Circadian rhythm	6.1	0.00249
All	Two-component response regulator activity	8.5	0.00501
Continental Europe	Positive regulation of transcription, DNA-dependent	5.0	0.00396
Sweden	Pentacyclic triterpenoid biosynthetic process	27.8	0.00035
Sweden	Response to reactive oxygen species	13.9	0.00312
Sweden	Translational initiation	6.8	0.00993

Storey's procedure was used to correct for multiple testing.

respect to mainland Europe, the most tolerant accessions originate from colder areas than the least tolerant accessions (Table 2), which supports the notion that differences in climate contribute to genetic differentiation among populations (Ågren & Schemske 2012). In contrast, variability in temperature does not seem to be associated with variation in freezing tolerance within Sweden. Although this implies that there is a minimum temperature beyond which selection for additional freezing tolerance is ineffective, we cannot rule out the possibility that other environmental factors also affect freezing tolerance. In particular, the most freezing tolerant Swedish accessions come from areas that have historically experienced higher amounts of precipitation than intolerant accessions. This is consistent with reports that, in cold areas, plant mortality increases with increasing precipitation (Gudleifsson 2010).

The main goal of our study was to map the genetic basis of constitutive freezing tolerance, specifically, to a sudden cold snap in the fall. To do so, we used a QTL mapping approach in combination with GWAS. As noted in the preceding texts, we found several promising candidate genes, including loci regulated by *CBF*s. Somewhat surprisingly, however, GWAS did not pinpoint the *CBF* loci themselves. Presumably because of the different parents used, linkage studies, including our own (e.g. Fig. S1), have provided conflicting results; some studies (Alonso-Blanco *et al.* 2005; Oakley *et al.* 2014) have identified a QTL that overlaps the *CBF* loci (Chr 4: ~13.02 Mb), while others have not (Meissner *et al.* 2013). Although the mapping resolution provided by these crosses is coarse and unidentified loci may contribute to these QTL, molecular studies have confirmed that the *CBF* loci play a role in cold acclimation and freezing tolerance (Jaglo-Ottosen *et al.* 1998; Gilmour *et al.* 2004). Moreover, genetic differences at the *CBF* loci are known to contribute to variation in cold acclimation requirements among accessions (Gehan *et al.* 2015). Although it is possible that the *CBF* genes do not affect constitutive freezing tolerance, one might expect genetic and allelic heterogeneity to pose problems for GWAS of freezing tolerance. Indeed, several independent losses of function *CBF* alleles have been discovered in southern accessions (Alonso-Blanco *et al.* 2005; Kang *et al.* 2013; Gehan *et al.* 2015), perhaps due to relaxed selection or the costs of freezing tolerance in warm areas. The problem of allelic heterogeneity is frequently overlooked but likely undermines other studies, including the analysis of flowering time (Li *et al.* 2014; Sasaki *et al.* 2015).

The top result from the global GWAS of freezing tolerance highlights *AT3G07000*, a cysteine-rich locus that is surrounded by three other cysteine-rich loci. If the candidate gene is confirmed in functional analyses, plants will join fishes and insects in being known to use cysteine-rich antifreeze proteins. Although the peak that highlights *AT3G07000* is not significant in GWAS of freezing tolerance for either continental European or Swedish accessions, it is nonetheless evident (Figs S2 and S3); whether it would prompt functional confirmation after analysing these regional samples would likely depend on research priorities.

Because the climatic factors that shape freezing tolerance differ across the species range (Table 2), and because the QTL identified with three separate mapping populations are largely non-overlapping (e.g. Fig. S1), we separated our global mapping population into two large regional mapping populations. By doing so, we found that the peak that implicates pleiotropic drug resistance 11 (or a nearby gene) in the freezing tolerance of Swedish accessions (Fig. S2) does not (seem to) underlie the variation in freezing tolerance found on the continent (Fig. S3). In contrast, an analysis of continental European accessions implicates *GLK2* in freezing tolerance, a gene whose expression is regulated by *CBF* homologs in *B. napus* (Savitch *et al.* 2005); this peak does not appear in the analysis of Swedish accessions. Taken together, we interpret this as evidence that the genetic architecture of freezing tolerance differs across the species range of *Arabidopsis*.

Although conventional GWAS are a powerful tool in genetics, there are, of course, opportunities for improvement. We used a combined model (MTMM) to identify loci that simultaneously contribute to freezing tolerance and climate adaptation. The top candidates identified by this approach, when applied to continental European accessions, include a member of the *CBF* regulon (Beta-glucosidase 25) and two DNAJ heat shock proteins. For Swedish accessions, the joint analysis of freezing tolerance and annual precipitation highlights (Chr 3: ~15.87–15.9 Mb) yet another DNAJ homolog (*ATJ3*), a gene that is also implicated in flowering time (Shen *et al.* 2011). A handful of heat shock proteins are known to affect freezing tolerance, and further studies will examine whether these specific loci contribute to differences in freezing

tolerance among populations of *A. thaliana*. Overall, these results suggest that MTMMs may help in identifying genes that underlie both phenotypic variation and adaptation.

ACKNOWLEDGMENTS

This work was supported by multiple sources, including NIH P50HG002790 (MN), the University of Zurich Research Priority Program 'Evolution in Action' (MH), core funding at MPIPZ (MK) and a Marie Curie International Outgoing Fellowship, number 220833 (GW). Seeds for the three linkage mapping populations were kindly provided by Caroline Dean.

REFERENCES

Ågren J. & Schemske D.W. (2012) Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. *The New Phytologist* **194**, 1112–1122.

Alonso-Blanco C., Gomez-Mena C., Llorente F., Koornneef M., Salinas J. & Martinez-Zapater J.M. (2005) Genetic and molecular analyses of natural variation indicate CBF2 as a candidate gene for underlying a freezing tolerance quantitative trait locus in *Arabidopsis*. *Plant Physiology* **139**, 1304–1312.

Bieniawska Z., Espinoza C., Schlereth A., Sulpice R., Hincha D.K. & Hannah M. A. (2008) Disruption of the *Arabidopsis* circadian clock is responsible for extensive variation in the cold-responsive transcriptome. *Plant Physiology* **147**, 263–279.

Broman K.W., Wu H., Sen S. & Churchill G.A. (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890.

Catala R., Medina J. & Salinas J. (2011) Integration of low temperature and light signaling during cold acclimation response in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16475–16480.

Chow K.C. & Tung W.L. (1998) Overexpression of dnaK/dnaJ and groEL confers freeze tolerance to *Escherichia coli*. *Biochemical and Biophysical Research Communications* **253**, 502–505.

Davies P.L. & Hew C.L. (1990) Biochemistry of fish antifreeze proteins. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology* **4**, 2460–2468.

Fournier-Level A., Korte A., Cooper M.D., Nordborg M., Schmitt J. & Wilczek A.M. (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86–89.

Gehan M.A., Park S., Gilmour S.J., An C., Lee C.M. & Thomashow M.F. (2015) Natural variation in the C-repeat binding factor cold response pathway correlates with local adaptation of *Arabidopsis* ecotypes. *The Plant Journal* **84**, 682–693.

Gilmour S.J., Fowler S.G. & Thomashow M.F. (2004) *Arabidopsis* transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. *Plant Molecular Biology* **54**, 767–781.

Gilmour S.J., Zarka D.G., Stockinger E.J., Salazar M.P., Houghton J.M. & Thomashow M.F. (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *The Plant Journal* **16**, 433–442.

Gothandam K.M., Nalini E., Karthikeyan S. & Shin J.S. (2010) OsPRP3, a flower specific proline-rich protein of rice, determines extracellular matrix structure of floral organs and its overexpression confers cold-tolerance. *Plant Molecular Biology* **72**, 125–135.

Griffith M. & Yaish M.W.F. (2004) Antifreeze proteins in overwintering plants: a tale of two activities. *Trends in Plant Science* **9**, 399–405.

Gudleifsson B.E. (2010) Ice tolerance and metabolite accumulation of herbage crops in Iceland and impact of climate change. *Icelandic Agricultural Sciences* **23**, 111–122.

Hancock A.M., Brachi B., Faure N., Horton M.W., Jarymowycz L.B., Sperone F. G., … Bergelson J. (2011) Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* **334**, 83–86.

Hannah M.A., Wiese D., Freund S., Fiehn O., Heyer A.G. & Hincha D.K. (2006) Natural genetic variation of freezing tolerance in *Arabidopsis*. *Plant Physiology* **142**, 98–112.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. & Jarvis A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965–1978.

Hollien J. (2013) Evolution of the unfolded protein response. *Biochimica et Biophysica Acta* **1833**, 2458–2463.

Horton M.W., Bodenhausen N., Beilsmith K., Meng D., Muegge B.D., Subramanian S., … Bergelson J. (2014) Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nature Communications* **5**, 5320.

Hu Y., Jiang L., Wang F. & Yu D. (2013) Jasmonate regulates the inducer of CBF expression-C-repeat binding factor/DRE binding factor1 cascade and freezing tolerance in *Arabidopsis*. *The Plant Cell* **25**, 2907–2924.

Huang G., Gong S., Xu W., Li P., Zhang D., Qin L., … Li X. (2011) GhHyPRP4, a cotton gene encoding putative hybrid proline-rich protein, is preferentially expressed in leaves and involved in plant response to cold stress. *Acta Biochimica et Biophysica Sinica* **43**, 519–527.

Hughes S.L., Schart V., Malcolmson J., Hogarth K.A., Martynowicz D.M., Tralman-Baker E., … Graether S.P. (2013) The importance of size and disorder in the cryoprotective effects of dehydrins. *Plant Physiology* **163**, 1376–1386.

Jackson M.B. & Osborne D.J. (1970) Ethylene, the natural regulator of leaf abscission. *Nature* **225**, 1019–1022.

Jager K., Fabian A., Tompa G., Deak C., Hohn M., Olmedilla A., … Papp I. (2011) New phenotypes of the drought-tolerant cbp20 *Arabidopsis thaliana* mutant have changed epidermal morphology. *Plant Biology* **13**, 78–84.

Jaglo-Ottosen K.R., Gilmour S.J., Zarka D.G., Schabenberger O. & Thomashow M.F. (1998) *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**, 104–106.

Kang H.M., Sul J.H., Service S.K., Zaitlen N.A., Kong S.Y., Freimer N.B., … Eskin E. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* **42**, 348–354.

Kang J., Zhang H., Sun T., Shi Y., Wang J., Zhang B., … Gu H. (2013) Natural variation of C-repeat-binding factor (CBFs) genes is a major cause of divergence in freezing tolerance among a group of *Arabidopsis thaliana* populations along the Yangtze River in China. *New Phytologist* **199**, 1069–1080.

Korte A., Vilhjalmsson B.J., Segura V., Platt A., Long Q. & Nordborg M. (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics* **44**, 1066–1071.

Li P., Filiault D., Box M.S., Kerdaffrec E., van Oosterhout C., Wilczek A.M., … Dean C. (2014) Multiple FLC haplotypes defined by independent cis-regulatory variation underpin life history diversity in *Arabidopsis thaliana*. *Genes and Development* **28**, 1635–1640.

Licausi F., Ohme-Takagi M. & Perata P. (2013) APETALA2/ethylene responsive factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytologist* **199**, 639–649.

Liou Y.C., Tocilj A., Davies P.L. & Jia Z. (2000) Mimicry of ice structure by surface hydroxyls and water of a beta-helix antifreeze protein. *Nature* **406**, 322–324.

Lippert C., Casale F.P., Rakitsch B. & Stegle O. (2014) LIMIX: genetic analysis of multiple traits. *bioRxiv*. DOI:10.1101/003905

McKhann H.I., Gery C., Berard A., Leveque S., Zuther E., Hincha D.K., … Teoule E. (2008) Natural variation in CBF gene sequence, gene expression and freezing tolerance in the Versailles core collection of *Arabidopsis thaliana*. *BMC Plant Biology* **8**, 105.

McKim S.M., Stenvik G.E., Butenko M.A., Kristiansen W., Cho S.K., Hepworth S.R., … Haughn G.W. (2008) The BLADE-ON-PETIOLE genes are essential for abscission zone formation in *Arabidopsis*. *Development* **135**, 1537–1546.

Meissner M., Orsini E., Ruschhaupt M., Melchinger A.E., Hincha D.K. & Heyer A.G. (2013) Mapping quantitative trait loci for freezing tolerance in a recombinant inbred line population of *Arabidopsis thaliana* accessions Tenela and C24 reveals REVEILLE1 as negative regulator of cold acclimation. *Plant, Cell & Environment* **36**, 1256–1267.

Nagelkerke N.J.D. (1991) A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691–692.

O'Kane S.L. Jr. & Al-Shehbaz I.A. (1997) A synopsis of *Arabidopsis* (Brassicaceae). *Novon* **7**, 323–327.

Oakley C.G., Agren J., Atchison R.A. & Schemske D.W. (2014) QTL mapping of freezing tolerance: links to fitness and adaptive trade-offs. *Molecular Ecology* **23**, 4304–4315.

Peng T., Jia M.M. & Liu J.H. (2015) RNAi-based functional elucidation of PtrPRP, a gene encoding a hybrid proline rich protein, in cold tolerance of *Poncirus trifoliata*. *Frontiers in Plant Science* **6**, 808.

Platt A., Horton M., Huang Y.S., Li Y., Anastasio A.E., Mulyati N.W., … Borevitz J.O. (2010) The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics* **6**e1000843.

Porankiewicz J. & Clarke A.K. (1997) Induction of the heat shock protein ClpB affects cold acclimation in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Journal of Bacteriology* **179**, 5111–5117.
Sasaki E., Zhang P., Atwell S., Meng D. & Nordborg M. (2015) "Missing" G × E variation controls flowering time in *Arabidopsis thaliana*. *PLoS Genetics* **11**e1005597.
Savitch L.V., Allard G., Seki M., Robert L.S., Tinker N.A., Huner N.P., … Singh J. (2005) The effect of overexpression of two Brassica CBF/DREB1-like transcription factors on photosynthetic capacity and freezing tolerance in Brassica napus. *Plant & Cell Physiology* **46**, 1525–1539.
Seo E., Lee H., Jeon J., Park H., Kim J., Noh Y.S. & Lee I. (2009) Crosstalk between cold response and flowering in *Arabidopsis* is mediated through the flowering-time gene SOC1 and its upstream negative regulator FLC. *The Plant Cell* **21**, 3185–3197.
Seo P.J., Park M.J., Lim M.H., Kim S.G., Lee M., Baldwin I.T. & Park C.M. (2012) A self-regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian clock regulation of temperature responses in *Arabidopsis*. *The Plant Cell* **24**, 2427–2442.
Shen L., Kang Y.G., Liu L. & Yu H. (2011) The J-domain protein J3 mediates the integration of flowering signals in *Arabidopsis*. *The Plant Cell* **23**, 499–514.
Shi Y., Tian S., Hou L., Huang X., Zhang X., Guo H. & Yang S. (2012) Ethylene signaling negatively regulates freezing tolerance by repressing expression of CBF and type-A ARR genes in *Arabidopsis*. *The Plant Cell* **24**, 2578–2595.
Sieg F., Schroder W., Schmitt J.M. & Hincha D.K. (1996) Purification and characterization of a cryoprotective protein (cryoprotectin) from the leaves of cold-acclimated cabbage. *Plant Physiology* **111**, 215–221.
Steponkus P.L., Uemura M., Joseph R.A., Gilmour S.J. & Thomashow M.F. (1998) Mode of action of the COR15a gene on the freezing tolerance of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14570–14575.
Storey J.D. & Tibshirani R. (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9440–9445.
Tao Z., Shen L., Liu C., Liu L., Yan Y. & Yu H. (2012) Genome-wide identification of SOC1 and SVP targets during the floral transition in *Arabidopsis*. *The Plant Journal* **70**, 549–561.
Teigen L.E., Orczewska J.I., McLaughlin J. & O'Brien K.M. (2015) Cold acclimation increases levels of some heat shock protein and sirtuin isoforms in threespine stickleback. Comparative biochemistry and physiology. *Part A, Molecular & integrative physiology* **188**, 139–147.
The 1001 Genomes Consortium (2016) 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491.
To A., Joubes J., Barthole G., Lecureuil A., Scagnelli A., Jasinski S., … Baud S. (2012) WRINKLED transcription factors orchestrate tissue-specific regulation of fatty acid biosynthesis in *Arabidopsis*. *The Plant Cell* **24**, 5007–5023.
Tyshenko M.G., Doucet D., Davies P.L. & Walker V.K. (1997) The antifreeze potential of the spruce budworm thermal hysteresis protein. *Nature Biotechnology* **15**, 887–890.
van Heerwaarden J., van Zanten M. & Kruijer W. (2015) Genome-wide association analysis of adaptation using environmentally predicted traits. *PLoS Genetics* **11**e1005594.
Xin Z. & Browse J. (1998) Eskimo1 mutants of *Arabidopsis* are constitutively freezing-tolerant. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 7799–7804.
Yang Z.T., Wang M.J., Sun L., Lu S.J., Bi D.L., Song Z.T., … Liu J.X. (2014) The membrane-associated transcription factor NAC089 controls ER-stress-induced programmed cell death in plants. *PLoS Genetics* **10**e1004243.
Young R. (1971) Effect of exposure to subfreezing temperatures on ethylene evolution and leaf abscission in citrus. *Plant Physiology* **48**, 724–727.
Zeller G., Henz S.R., Widmer C.K., Sachsenberg T., Ratsch G., Weigel D. & Laubinger S. (2009) Stress-induced changes in the *Arabidopsis thaliana* transcriptome analyzed using whole-genome tiling arrays. *The Plant Journal* **58**, 1068–1082.
Zhang B., Chen H.W., Mu R.L., Zhang W.K., Zhao M.Y., Wei W., … Zhang J.S. (2011) NIMA-related kinase NEK6 affects plant growth and stress response in *Arabidopsis*. *The Plant Journal* **68**, 830–843.
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.
Zhen Y. & Ungerer M.C. (2008) Clinal variation in freezing tolerance among natural accessions of *Arabidopsis thaliana*. *New Phytologist* **177**, 419–427.
Zuther E., Schulz E., Childs L.H. & Hincha D.K. (2012) Clinal variation in the non-acclimated and cold-acclimated freezing tolerance of *Arabidopsis thaliana* accessions. *Plant, Cell & Environment* **35**, 1860–1878.

Received 6 June 2016; accepted for publication 28 July 2016

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. The accessions used in the study.
Table S2. The candidate genes identified by a local–global analysis of all sequenced accessions.
Figure S1. The genetic bases of freezing tolerance. Shown are the genome wide *P*-values, using all accessions. The dash-dotted line corresponds to a genome-wide significance threshold estimated by using a permutation approach that takes into account population structure; the centromere is shaded. The QTLs identified using 3 separate mapping populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line.
Figure S2. GWAS results, continental Europe. Shown are the genome wide *P*-values, using accessions from continental Europe. The dash-dotted line corresponds to a genome-wide significance threshold estimated by using a permutation approach that takes into account population structure; the centromere is shaded. The QTLs identified by using 3 separate mapping populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line.
Figure S3. GWAS results, Sweden. Shown are the genome wide *P*-values, using Swedish accessions. The dash-dotted corresponds to a genome-wide significance threshold estimated by using a permutation approach that takes into account population structure; the centromere is shaded. The QTLs identified by using 3 separate mapping populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line.
Figure S4. The genetic bases of freezing tolerance, without correcting for population structure. Shown are the genome wide *P*-values, using all accessions. The dash-dotted line corresponds to a genome-wide Bonferroni significance threshold (5%); the centromere is shaded. The QTLs identified by using 3 separate mapping populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line. Note the massive *P*-value inflation.
Figure S5. Uncorrected *P*-values, Sweden. Shown are the genome wide *P*-values, using Swedish accessions. The dash-dotted line corresponds to a genome-wide Bonferroni significance threshold (5%); the centromere is shaded. The QTLs identified by using 3 separate mapping populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line.
Figure S6. Uncorrected *P*-values, continental Europe. Shown are the genome wide *P*-values, using accessions from continental Europe. The dash-dotted line corresponds to a genome-wide Bonferroni significance threshold (5%); the centromere is shaded. The QTLs identified by using 3 separate mapping

populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line.

Figure S7. MTMM, Sweden + annual precipitation. Simultaneously modelling the freezing tolerance of Swedish accessions with the climate variable that best predicts survivorship after freezing (BIO12) pinpoints a region on chromosome 3 (15.87–15.9 Mb). The QTLs identified by using 3 separate mapping populations are shown above each chromosome. The significance levels for each QTL were estimated by using permutations and are illustrated by using either a solid ($P < 0.05$) or a dotted ($P < 0.1$) line. For the MTMM, the Bonferroni threshold is shown; the centromere is depicted as a vertical pillar.