Multiple Genetic Trajectories to Extreme Abiotic Stress Adaptation in Arctic Brassicaceae

Siri Birkeland,*[1] A. Lovisa S. Gustafsson,[1] Anne K. Brysting,[2] Christian Brochmann,[1] and Michael D. Nowak[1]

[1]Natural History Museum, University of Oslo, Oslo, Norway
[2]Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Oslo, Norway

*Corresponding author: E-mail: siri.birkeland@nhm.uio.no.

Associate editor: Michael Purugganan

Abstract

Extreme environments offer powerful opportunities to study how different organisms have adapted to similar selection pressures at the molecular level. Arctic plants have adapted to some of the coldest and driest biomes on Earth and typically possess suites of similar morphological and physiological adaptations to extremes in light and temperature. Here, we compare patterns of molecular evolution in three Brassicaceae species that have independently colonized the Arctic and present some of the first genetic evidence for plant adaptations to the Arctic environment. By testing for positive selection and identifying convergent substitutions in orthologous gene alignments for a total of 15 Brassicaceae species, we find that positive selection has been acting on different genes, but similar functional pathways in the three Arctic lineages. The positively selected gene sets identified in the three Arctic species showed convergent functional profiles associated with extreme abiotic stress characteristic of the Arctic. However, there was little evidence for independently fixed mutations at the same sites and for positive selection acting on the same genes. The three species appear to have evolved similar suites of adaptations by modifying different components in similar stress response pathways, implying that there could be many genetic trajectories for adaptation to the Arctic environment. By identifying candidate genes and functional pathways potentially involved in Arctic adaptation, our results provide a framework for future studies aimed at testing for the existence of a functional syndrome of Arctic adaptation in the Brassicaceae and perhaps flowering plants in general.

Key words: molecular convergence, adaptation, positive selection, extremophiles.

Introduction

Strong selection pressures imposed by extreme environments and specialized niches can sometimes lead to independent evolution of similar morphological and/or physiological traits in different lineages (i.e., convergent evolution; Sackton and Clark 2019). Some striking examples are the evolution of antifreeze glycoproteins in Arctic cod and Antarctic notothenioid fishes (Chen et al. 1997), increased hemoglobin-oxygen affinity in high-altitude dwelling birds (Natarajan et al. 2016), and photosynthetic pathways that minimize photorespiration in hot and dry environments (Heyduk et al. 2019). Organisms adapted to extreme environments are, therefore, especially interesting for studying one of the fundamental questions in evolutionary biology: To what degree does adaptation follow predictable genetic trajectories? Convergent traits do not necessarily result from similar genetic changes, and the prevalence of molecular convergence is still heavily debated (Losos 2011; Conte et al. 2012; Hendry 2013; Bolnick et al. 2018; Sackton and Clark 2019). Increasingly, molecular convergence is viewed as a quantitative continuum rather than a presence–absence pattern (Bolnick et al. 2018). It may manifest at different levels of similarity, that is, 1) independently fixed nonsynonymous mutations at the same site,

2) positive selection acting on the same genes, or 3) positive selection acting on different genes that are part of the same functional pathway (Sackton and Clark 2019). By disentangling the impacts of these different levels of molecular convergence, we gain a more comprehensive perspective on the genomic landscape of adaptation.

In this study, we aim to identify sets of candidate genes for Arctic adaptations and search for signatures of codon, gene, and pathway convergence in plants that have adapted to live in some of the harshest terrestrial environments on Earth. The Arctic is characterized by very low year-round temperatures (average temperature of the warmest month $\leq 10\,^{\circ}$C) as well as major annual and minor diurnal fluctuations in light regime (Billings and Mooney 1968; Nilsen 1985; Elvebakk 1999). Arctic plants must therefore withstand a number of stresses directly or indirectly linked to highly fluctuating light and temperature conditions. Although there is considerable climatic variation within the Arctic, the combination of these abiotic stresses is overall more extreme relative to lower latitude environments from which most Arctic lineages are derived (see, e.g., Abbott and Brochmann [2003] on the origin of the Arctic flora). Here, we classify major Arctic stresses into five categories that we will focus on throughout the study: 1) protein misfolding and membrane stress, 2) physiological cold

Open Access

stress, 3) ice and snow stress, 4) light stress, and 5) oxidative stress. Low temperatures can affect nearly all aspects of plant cell biochemistry and metabolism, including the physical properties of cell membranes and proteins (categories 1 and 2; Heino and Palva 2004). Snowfall and ice formation can occur year-round, and Arctic soils are characterized by permafrost, with all soil 20–100 cm below the surface permanently frozen (Billings and Mooney 1968). The effects of such ice and snow stress (category 3) can manifest in many ways. It may, for example, cause mechanical stress and cellular rupture, and ice formation in extracellular spaces will draw water out of cells and increase solute concentration in the cytoplasm, leading to cellular dehydration stress (Wisniewski and Fuller 1999; Körner 2003; Wisniewski et al. 2004). The Arctic light environment is enriched in far-red and blue light, and at the highest latitudes in the Arctic, the period of midnight sun lasts the entire growing season (Salisbury 1985). The actual period available for reproduction, growth, and development can also be extremely short (e.g., ∼1 month in Arctic polar deserts; Jónsdóttir 2005). The Arctic *light stress* (category 4) might therefore be especially challenging for plants, as photoperiod and spectral composition are important cues for, for example, development, reproduction, and the onset of dormancy (Chory et al. 1996; Jackson 2009; Piskurewicz et al. 2009). Finally, the combined effects of high light intensity and low temperature can lead to increased *oxidative stress* (category 5; Heino and Palva 2004; Lütz 2010). This comes in addition to the increased accumulation of reactive oxygen species (ROS) within the cells due to many of the abovementioned abiotic stresses (accumulation of ROS is a common stress response in plants; Kilian et al. 2007). These five categories of stresses will serve as a framework for discussing molecular convergence in plant adaptation to the Arctic environment.

Arctic plants exhibit a range of similar morphological and physiological adaptations to cope with these five stress categories (Chapin 1983; Lütz 2010; Jónsdóttir 2011). Many Arctic plants are small, herbaceous perennials that cling to the ground, often forming cushions or tussocks that minimize exposure to abrasive winds and low temperatures (Billings and Mooney 1968). The leaves are often tough and leathery, and/or hairy (Hultén 1968; Alsos et al. 2019), providing protection from desiccation, mechanical stress from ice and snow, and strong winds. Arctic plants have adapted their metabolism, growth, and reproduction to the low temperatures, and have for instance optimum photosynthesis rates at lower temperatures than other plants (Chapin 1983). It is also likely that the Arctic photoperiod and spectral composition (such as the enrichment of far-red light and blue light) play important roles in plant growth and timing of development (Nilsen 1985; Salisbury 1985). It has been shown that some Arctic plants require a long critical day length for induction of flowering (Wehrmeister and Bonde 1977; Heide 2005) and that changes in the red to far-red light ratio might be important for the onset of dormancy when there still is midnight sun in late summer (Nilsen 1985). We can also hypothesize that the plasma membrane plays an important role in plant adaptation to Arctic cold and freezing stress (cf.,

Thomashow 1999; Knight and Knight 2012). Many temperate plants will, for example, acclimate to low temperatures by changing the lipid composition of the plasma membrane to avoid membrane rigidification and rupture during cold stress (Thomashow 1999; Wisniewski and Fuller 1999; Körner 2003). Response to freezing stress involves many of the same molecular components as response to dehydration and high salinity (Ryu et al. 1995; Bartels and Souer 2004; Verslues et al. 2006), and we could therefore also expect Arctic plants to be resistant against these stresses. Detailed studies of genes involved in plant cold response pathways have been conducted in temperate plants such as *Arabidopsis thaliana* (Jaglo-Ottosen et al. 1998; Kreps et al. 2002; Kilian et al. 2007; Thomashow 2010), but little is known for Arctic plants. By comparing molecular evolution in plant lineages independently adapted to the Arctic, we can learn more about the general traits and genes that are important for plant survival at high latitudes.

The focus of our study is three plant species in the Brassicaceae that have their main distribution above the Arctic Circle. The Brassicaceae is an optimal study system with its many publicly available molecular resources (TAIR 2019), and the family comprises several lineages that have independently colonized the Arctic (Carlsen et al. 2009; Jordon-Thaden et al. 2010; Koch 2012). The three selected species are *Cardamine bellidifolia*, *Cochlearia groenlandica*, and *Draba nivalis*, which all are diploid and represent three divergent clades that shared a common ancestor ∼30 Ma (Elven et al. 2011; Huang et al. 2016; Guo et al. 2017; fig. 1). Their distributions cover the entire Arctic (even the coldest bioclimatic subzone called Arctic polar desert; Elvebakk 1999; Elven et al. 2011), and they are therefore suitable candidates for studying adaptations to stresses associated with limited heat budget and low sun angle. We generated new transcriptome data for each species and conducted comparative analyses with previously published transcriptome/genome data from several temperate relatives. To identify candidate genes for Arctic adaptations, we performed genome-wide tests of positive selection and searched for positively selected genes (PSGs) associated with the five categories of Arctic stresses outlined above. To identify patterns of molecular convergence, we identified convergent amino acid substitutions and compared PSG sets in terms of orthology and functional annotations. This enabled us to estimate molecular convergence among Arctic lineages at the level of codons (the number of genes with convergent substitutions), genes (the number of orthologous genes under positive selection), and functional pathways (the degree of similarity between functional profiles of the PSGs). However, when estimating convergent substitutions, it has proven important to account for random accumulation of convergent substitutions with increasing divergence (i.e., the expected background convergence; Thomas and Hahn 2015; Xu et al. 2017). By estimating the relationship between convergent and divergent substitutions using species pairs in a phylogeny, we therefore tested whether Arctic lineage pairs exhibit an excess of convergence relative to neutral expectations (Castoe et al. 2009; Thomas and Hahn 2015). The aims of this study are to
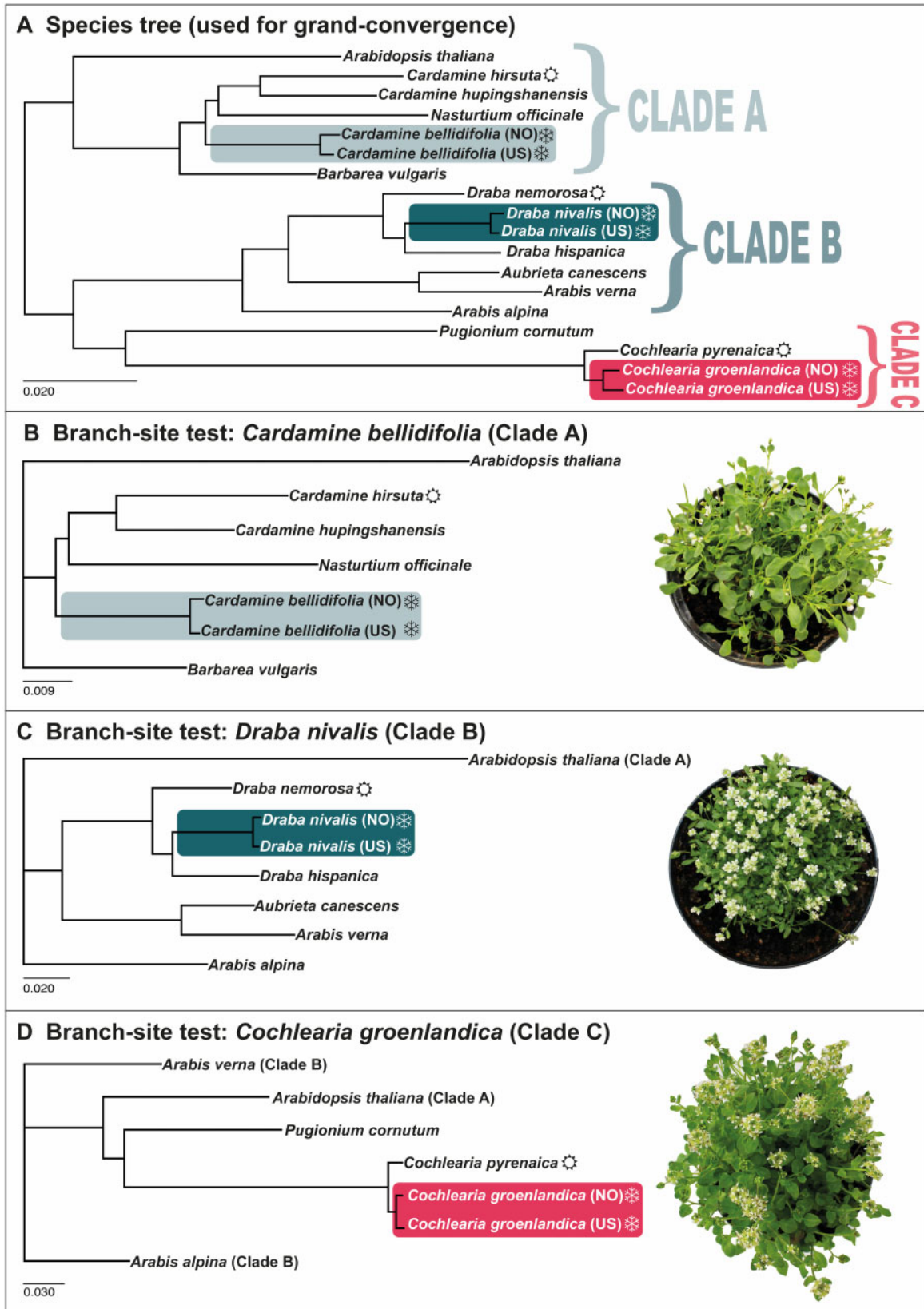
FIG. 1. Species trees used in the estimation of convergent and divergent substitutions (*A*), and in the branch-site tests of positive selection (*B*–*D*). Arctic focal species are indicated with a snow crystal, and non-Arctic comparisons are indicated with a sun. NO, Arctic genotype from Norway; US, Arctic genotype from the United States. Scale bars = substitutions/site. The topologies are in accordance with Huang et al. (2016).

Table 1. Number of PSGs Found in the Branch-Site Tests, as well as Corresponding topGO Results.

Species	Lineage Tested	Alignments Retained for the Branch-Site Test	PSGs $P < 0.05$	Enriched GO-terms[a]BP, CC, MF
Cardamine bellidifolia	Arctic	11,874	201	274 (159), 38 (25), 57 (50)
Cochlearia groenlandica	Arctic	13,306	159	198 (92), 28 (18), 34 (23)
Draba nivalis	Arctic	13,891	360	359 (256), 54 (36), 87 (73)
Cardamine hirsuta	Non-Arctic	11,874	395	268 (168), 105 (48), 72 (67)
Cochlearia pyrenaica	Non-Arctic	13,305	350	217 (168), 52 (49), 77 (69)
Draba nemorosa	Non-Arctic	13,881	380	298 (179), 22 (18), 67 (66)

[a]Numbers of significant GO-terms ($P < 0.05$) when applying the classic algorithm. Numbers of enriched GO-terms when applying the elim algorithm are given in brackets.

1) identify sets of candidate genes for Arctic adaptations in *C. bellidifolia*, *D. nivalis* and *Co. groenlandica*, 2) search for evidence of convergent molecular evolution across codons, genes, and pathways, and 3) test for a signature of convergent evolution that exceeds neutral expectations between Arctic species pairs. We show that there is little evidence for convergence at the codon and gene level, but that all positively selected and convergent gene sets were enriched for genes associated with typical Arctic stresses. Thus, our results suggest that there are many genetic trajectories to Arctic adaptation, perhaps due to the multifaceted nature of stress response pathways in plants.

Results

Final Data Set

We acquired genome-wide coding sequences from 15 species representing clades A, B, and C in the Brassicaceae (fig. 1, supplementary tables S1–S3, Supplementary Material online). These included the three selected Arctic species, for which we generated RNA sequences for this study. In addition to already published genomes, the data set consisted of de novo transcriptome assemblies yielding 18,830–34,226 transcripts per assembly and containing 67.0–89.8% universal single-copy orthologs (complete/fragmented BUSCOs; supplementary table S2, Supplementary Material online). Orthologous gene alignments were generated separately for the selection tests and for the estimation of convergent substitutions (supplementary fig. S1, Supplementary Material online) and were based on gene subsets from 8,572 (clade A), 9,941 (clade B), and 8,705 (clade C) orthogroups (∼gene families) in the tests of positive selection and 6,325 orthogroups in the estimation of convergent substitutions.

Tests of Positive Selection

To detect genes that evolved under positive selection in the branches leading to the Arctic species, we ran branch-site tests separately for clades A, B, and C (fig. 1B–D; ∼12–13,000 orthologous gene alignments/clade; supplementary fig. S1, Supplementary Material online). We identified 360 PSGs in *D. nivalis*, 201 in *C. bellidifolia*, and 159 in *Co. groenlandica* (table 1, supplementary tables S4–S6 and S7–S9, Supplementary Material online). To identify genes that might be under positive selection in both Arctic and non-Arctic lineages (e.g., immunity genes) and to evaluate whether gene functional annotations exclusively were associated with the Arctic, we performed complementary analyses in three non-Arctic reference species (fig. 1B–D), resulting in

380 PSGs in *D. nemorosa*, 395 in *C. hirsuta*, and 350 in *Co. pyrenaica* (table 1, supplementary tables S10–S12 and S13–S15, Supplementary Material online).

To identify candidate genes for Arctic adaptations, we manually reviewed the PSG annotations for documented associations with at least one of the five categories of Arctic stresses (i.e., gene ontology [GO] terms or publications indicating a mechanistic association with some physiological or biochemical aspect of the stress categories). We identified a total of 138 candidate genes spanning the five categories, of which 40 were found in *C. bellidifolia*, 31 in *Co. groenlandica*, and 67 in *D. nivalis* (supplementary table S16, Supplementary Material online). In *C. bellidifolia* and *D. nivalis*, there were many candidate genes associated with *light stress*, for example, blue and far-red light response (like *FAR1*-related sequences found in both PSG sets, and *ZTL* found in *D. nivalis*), whereas many candidate genes in *D. nivalis* and *Co. groenlandica* were associated with *physiological cold stress* and *ice and snow stress* (e.g., *COR15b* and *CSDP1* in *D. nivalis*, and *SFR2* and *RAV1* in *Co. groenlandica*). All Arctic PSG sets contained one or more late embryogenesis abundant (*LEA*) genes, putatively associated with *physiological cold stress* and *ice and snow stress*. Heat shock protein-related genes (e.g., heat shock chaperonin-binding or containing heat shock protein domains), which are putative candidates for adaptation to *protein misfolding and membrane stress*, were especially common in the PSG sets of *D. nivalis* and *C. bellidifolia*. Examples of candidate genes from the *oxidative stress* category were *CAT1* and *CAT2*, important for controlling ROS (both found in the *D. nivalis* PSG set). The three Arctic sets contained many MYB and MYC type transcription factors, involved in abiotic stress-activated signal transduction pathways. We also generated a heatmap showing the proportion of PSGs associated with the five categories, demonstrating that all Arctic PSG sets included considerably more genes associated with *protein misfolding and membrane stress* and *ice and snow stress* than the non-Arctic reference species (supplementary fig. S2, Supplementary Material online).

To assess whether positive selection is acting on the same genes in the three Arctic species, we compared the PSG sets with respect to orthology. We did not find any orthologous genes under positive selection in all three species. Thirteen PSGs were orthologous in two of the Arctic species (fig. 2) and four of these were identified as PSGs in at least one non-Arctic reference species (supplementary table S17, Supplementary Material online). Significance testing (the supertest; Wang et al. 2015) indicated that the PSG intersections observed
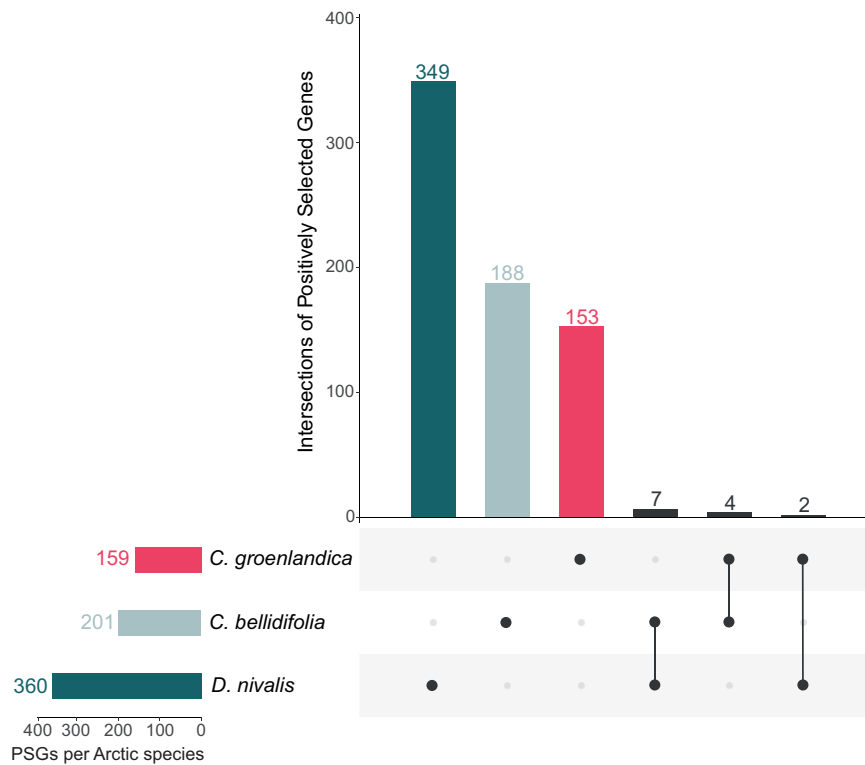
Fig. 2. UpSet plot of orthologous, PSGs identified in Arctic species. The plot in the left corner shows total number of PSGs, and the main plot shows the number of unique PSGs, followed by PSG intersections/overlaps between species (connected dots).

between any pair of Arctic species were not greater than expected by chance (supplementary table S18, Supplementary Material online). Among the 13 PSGs orthologous between Arctic species pairs, we found seven candidates possibly associated with three categories of Arctic stresses: *protein misfolding and membrane stress* (three genes involved in endoplasmic reticulum morphology or protein folding; *ERMO3*, DNAJ heat shock N-terminal domain-containing protein, and prefoldin 2), *light stress* (two genes involved in photoautotrophic growth or auxin-regulated plant growth and development; *IAA16* and *PAC*), and *oxidative stress* (*PCR2* and *OGOX1*; supplementary table S17, Supplementary Material online).

To assess whether positive selection has acted on different genes, but on similar functional pathways, we performed GO enrichment tests within the Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) domains for each of the three Arctic PSG sets (table 1, supplementary tables S19–S21 and fig. S3A–C, Supplementary Material online). We compared the sets of enriched GO-terms and used the results from the non-Arctic reference species to filter out terms unlikely to be associated with Arctic adaptation (supplementary tables S22–S24, Supplementary Material online). All Arctic PSG sets were enriched for genes associated with "response to stress" (BP, fig. 3A) and two sets for "response to abiotic stimulus" (BP, fig. 3B). The enriched "children" of these two terms (i.e., more specific terms nested under "response to stress" and "response to abiotic stimulus" in the GO hierarchy) varied somewhat among species (fig. 4), but included several potential associations with the five stress categories.

All three Arctic PSG sets were significantly enriched for "response to salt stress," and the PSG sets of *D. nivalis* and *Co. groenlandica* were significantly enriched for "response to temperature stimulus," "response to cold," and "response to osmotic stress." *Draba nivalis* and *C. bellidifolia* shared an enrichment for "MAPK-cascade" and "response to oxidative stress." *Draba nivalis* stood out in being significantly enriched for a range of GO-terms associated with all five categories, including "response to light stimulus," "photoperiodism," and "response to misfolded protein." Notably, "plasma membrane" (CC domain) was significantly enriched in all Arctic PSG sets, but in none of the non-Arctic reference sets (supplementary fig. S4, Supplementary Material online), and the related BP term "endomembrane system organization" was also significantly enriched in *D. nivalis* and *C. bellidifolia*. Thus, we found evidence both for positive selection on genes associated with abiotic stress in all three Arctic species and for considerable convergence in the functional pathways in which these genes might be involved.

Convergent Amino Acid Substitutions

To estimate molecular convergence at the level of codons, we used Grand Convergence (Qian and de Koning 2018; Qian et al. 2018) to calculate posterior expected numbers of convergent and divergent substitutions for 7,765 gene alignments between pairs of Arctic and non-Arctic branches (fig. 1A). We identified 217 convergent genes between Arctic branch pairs (defined as containing at least one substitution with >0.8 posterior probability (PP) of convergence; table 2, supplementary tables S25–S27, Supplementary Material online). The
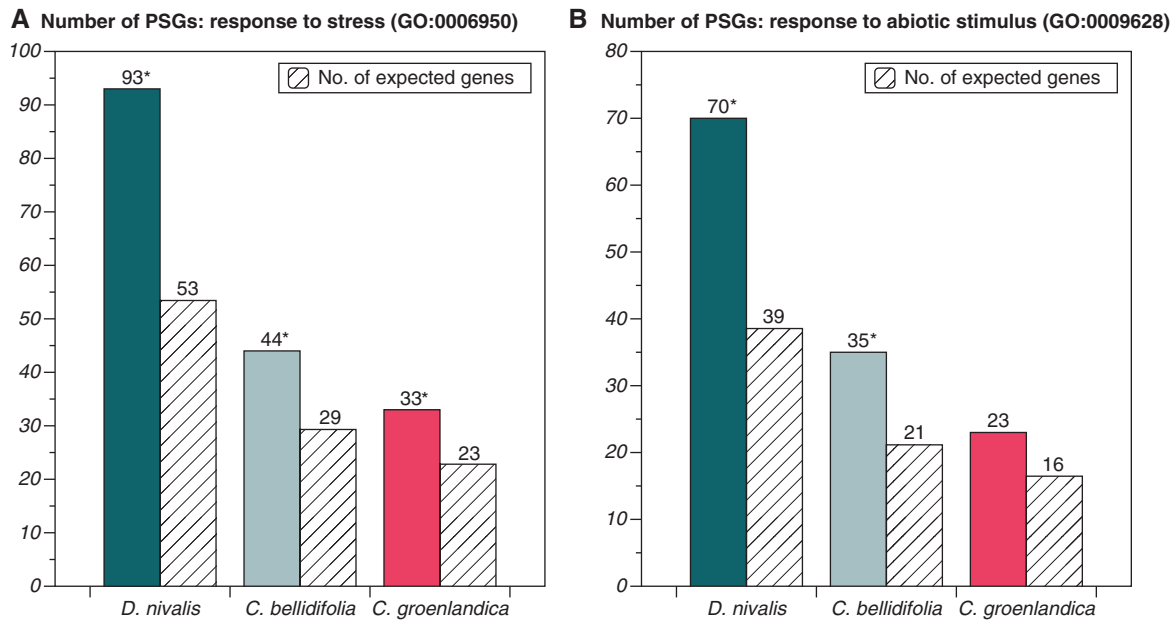
Fig. 3. Number of PSGs annotated with (A) "response to stress" and (B) "response to abiotic stimulus" in the Arctic species. Expected numbers of genes (if randomly drawn) are indicated as hatched bars. Stars (*) indicate that the GO-term was significantly enriched in the PSG set of this species when applying the classic algorithm in topGO ($P < 0.05$).

highest number of convergent genes (126) was observed between *D. nivalis* and *C. bellidifolia*, followed by *Co. groenlandica* and *C. bellidifolia* (58), and *D. nivalis* and *Co. groenlandica* (33, fig. 5, table 2). For comparison, a total of 362 convergent genes were identified between pairs of the three non-Arctic reference species (table 2, supplementary tables S28–S30, Supplementary Material online). Only a single gene was identified as convergent in all three Arctic pairwise comparisons (fig. 5): *EMB2742*, involved in, for example, embryo development ending in seed dormancy (supplementary table S31, Supplementary Material online). Three convergent genes were identified in both the *D. nivalis*–*C. bellidifolia* and the *C. bellidifolia*–*Co. groenlandica* branch pairs (*ORRM6*, *SK4*, and *MAPKKK14*), but only one of these had an apparent putative association with the five categories of stresses: *MAPKKK14*, belonging to a group of protein kinases involved in stress-induced signaling pathways, potentially involved in far-red light induced seedling deetiolation (supplementary table S31, Supplementary Material online).

To test for an excess of convergent substitutions in Arctic species pairs, we summed the probabilities for divergent and convergent substitutions over all alignments in 50 branch pairs from figure 1A, using one (arbitrarily chosen) alignment per orthogroup. These included the three Arctic branch pairs, the three non-Arctic reference branch pairs, and 44 random nonsister branch pairs from the phylogeny (supplementary table S32, Supplementary Material online). The number of convergent substitutions (y) could be predicted from the number of divergent substitutions (x) with the linear model $y = (−18.9642) + 0.5807x$ with $R^2 = 0.84$, indicating a strong linear relationship (fig. 6). Removing three outliers resulted in $y = (−32.566) + 0.59324x$ with $R^2 = 0.91$ (supplementary fig. S5, Supplementary Material online; outliers defined as falling

outside 1.5σ, where σ is the standard deviation of the residual). None of the Arctic branch pairs could be categorized as outliers, and thus they did not show any sign of excess convergence.

Next, we identified convergent genes that had functional associations with the five categories of Arctic stresses and compared the functional profiles of the convergent gene sets. In addition to the genes that had convergent substitutions in more than two comparisons (*EMB2742*, *MAPKKK14*, *ORRM6*, and *SK4*; cf., above), we found candidates for Arctic adaptation in the individual convergent gene sets (supplementary table S31, Supplementary Material online). Examples of the latter were *CRY2* (a blue light receptor mediating blue-light cotyledon expansion and flowering time), the cold and light responsive *KCS1* (encoding a condensing enzyme involved in the critical fatty acid elongation in wax biosynthesis), the cold and drought stress-activated *CAT2* (also found to be under positive selection in *D. nivalis* and possibly associated with *oxidative stress*, *physiological cold stress*, and *light stress*), *LSH4* (encoding a light-dependent short hypocotyls-like protein), and several genes possibly associated with *oxidative stress*, such as *GPX7* and *GCTF9*.

GO-term enrichment tests (table 2, supplementary tables S33–S35, Supplementary Material online) showed that "CTP synthase activity" (MF) was significantly enriched in all three Arctic comparisons. This GO-term was only associated with *EMB2742* (see above), and there were several cases where only a few genes were associated with a significantly enriched GO-term (perhaps a statistical artifact for GO-terms with few annotations in the background). Several GO-terms possibly associated with Arctic stresses were significantly enriched in the convergent genes shared by two Arctic comparisons, and none of the non-Arctic comparisons (supplementary tables
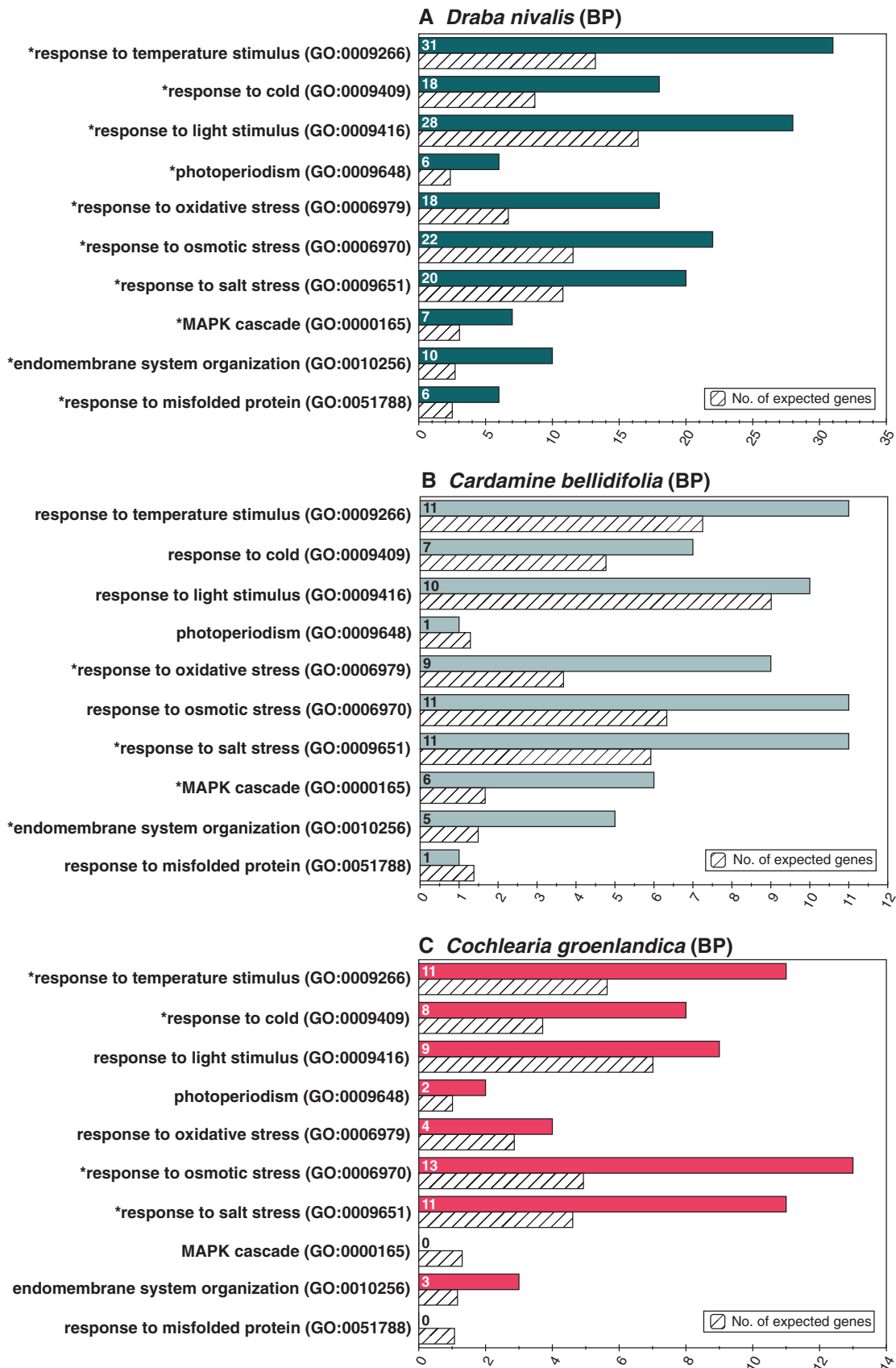
FIG. 4. Number of PSGs annotated with ten selected GO-terms in the Biological Process domain (BP). (A) *Draba nivalis*, (B) *Cardamine bellidifolia*, and (C) *Cochlearia groenlandica*. Expected numbers of genes (if randomly drawn) are indicated as hatched bars. An asterisk (*) indicates that the GO-term was significantly enriched in the PSG set of this species when applying the classic algorithm in topGO (P < 0.05).

Table 2. Number of Convergent Genes (one result per orthogroup), as well as Corresponding topGO Results.

Branches Compared [7,765 alignments]	Lineages Tested	No. of Convergent Genes (at least one site PP $\geq$ 0.80)[a]	Enriched GO-terms[b] (BP, CC, MF)
Draba nivalis, Cardamine bellidifolia	Arctic, Arctic	126	258 (161), 39 (21), 90 (59)
Draba nivalis, Cochlearia groenlandica	Arctic, Arctic	33	149 (80), 19 (17), 26 (19)
Cardamine bellidifolia, Cochlearia groenlandica	Arctic, Arctic	58	144 (81), 38 (22), 41 (31)
Draba nemorosa, Cardamine hirsuta	Non-Arctic, Non-Arctic	181	432 (253), 59 (41), 78 (64)
Draba nemorosa, Cochlearia pyrenaica	Non-Arctic, Non-Arctic	109	191 (127), 57 (37), 74 (51)
Cardamine hirsuta, Cochlearia pyrenaica	Non-Arctic, Non-Arctic	72	180 (105), 65 (36), 69 (51)

[a]Reporting only one result per orthogroup.
[b]Numbers of significant GO-terms ($P < 0.05$) when applying the classic algorithm. Numbers of enriched GO-terms when applying the elim algorithm are given in brackets.
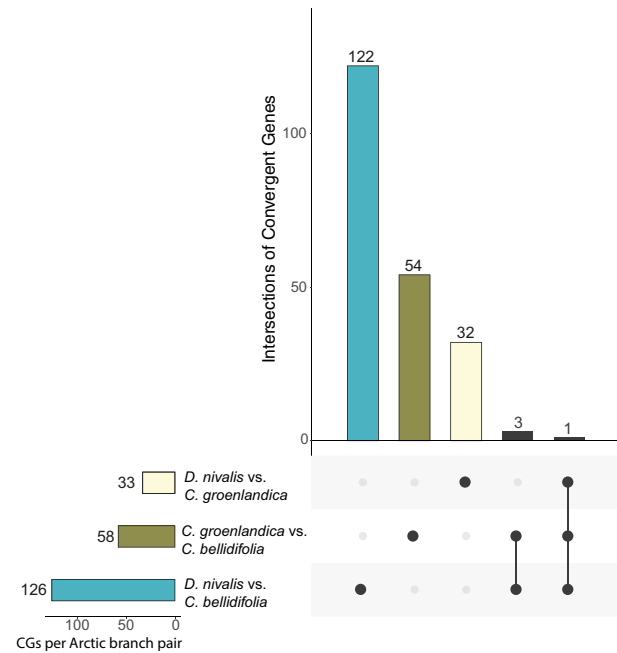

Fig. 5. UpSet plot of convergent genes (CGs) identified between Arctic branch pairs (containing at least one substitution with > 0.8 PP of convergence). The plot on the left shows total number of CGs, and the plot on the right shows the number of unique CGs, followed by CG intersections/overlaps between branch pairs (connected dots).

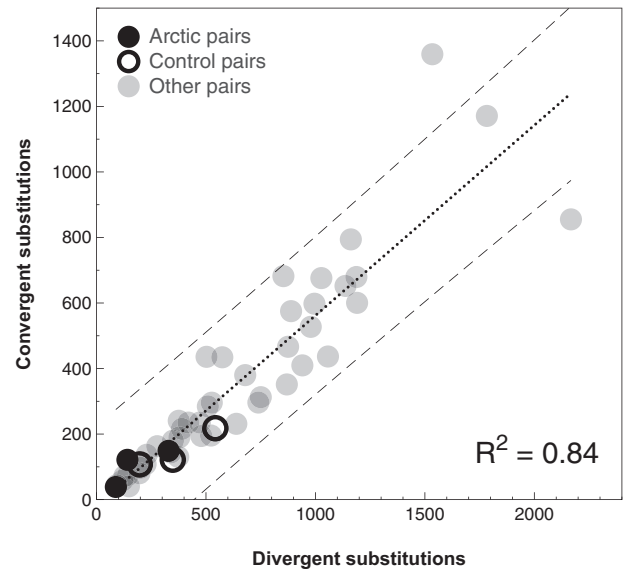**Fig. 6.** Posterior expected numbers of convergent and divergent substitutions between 50 branch pairs from the full species phylogeny (fig. 1A). Numbers are sums over all alignments. Dotted line, the regression line; dashed line, 95% prediction intervals; filled black circles, the three Arctic branch pair comparisons; open circles, the branch pairs of the non-Arctic reference species; filled gray circles, remaining branch pairs.

S36–S38, Supplementary Material online). Examples are terms associated with general abiotic stress responses such as "positive regulation of abscisic acid-activated signaling pathway" (BP), "regulation of stomatal movement" (BP), and "calcium channel activity" (MF), and terms associated with protection of plant cells from oxidative stress (i.e., "glutathione peroxidase activity"; MF). We found several putative connections to the five Arctic stress categories in the individual convergent gene sets (i.e., GO-terms enriched in only one set); and interestingly, some of these terms also occurred in the functional profiles of the PSGs. Examples include "response to temperature stimulus," "cellular response to osmotic stress," "cellular response to salt stress," and GO-terms associated with light stress (e.g., "circadian regulation of calcium ion oscillation," "response to far red light," "response to blue light," "blue light photoreceptor activity," "cellular response to UV-B," "response to high light intensity," and

"regulation of shade avoidance"). Thus, whereas we found limited evidence for convergence at the codon level in Arctic lineages, the convergent gene sets contained several putative candidate genes for Arctic adaptations.

Limited Overlap between PSG Sets and Convergent Gene Sets

Because genes that have evolved under positive selection and contain convergent substitutions provide strong evidence for adaptive convergence at the level of codons, we investigated intersections of PSG sets and convergent gene sets. We found that none of the genes carrying convergent substitutions in Arctic branch pairs had evolved under positive selection in both species, and only 8–12% of the convergent genes likely experienced positive selection in one of the Arctic species (supplementary tables S25–S27, Supplementary Material online). We also found that 9–16% of the convergent genes
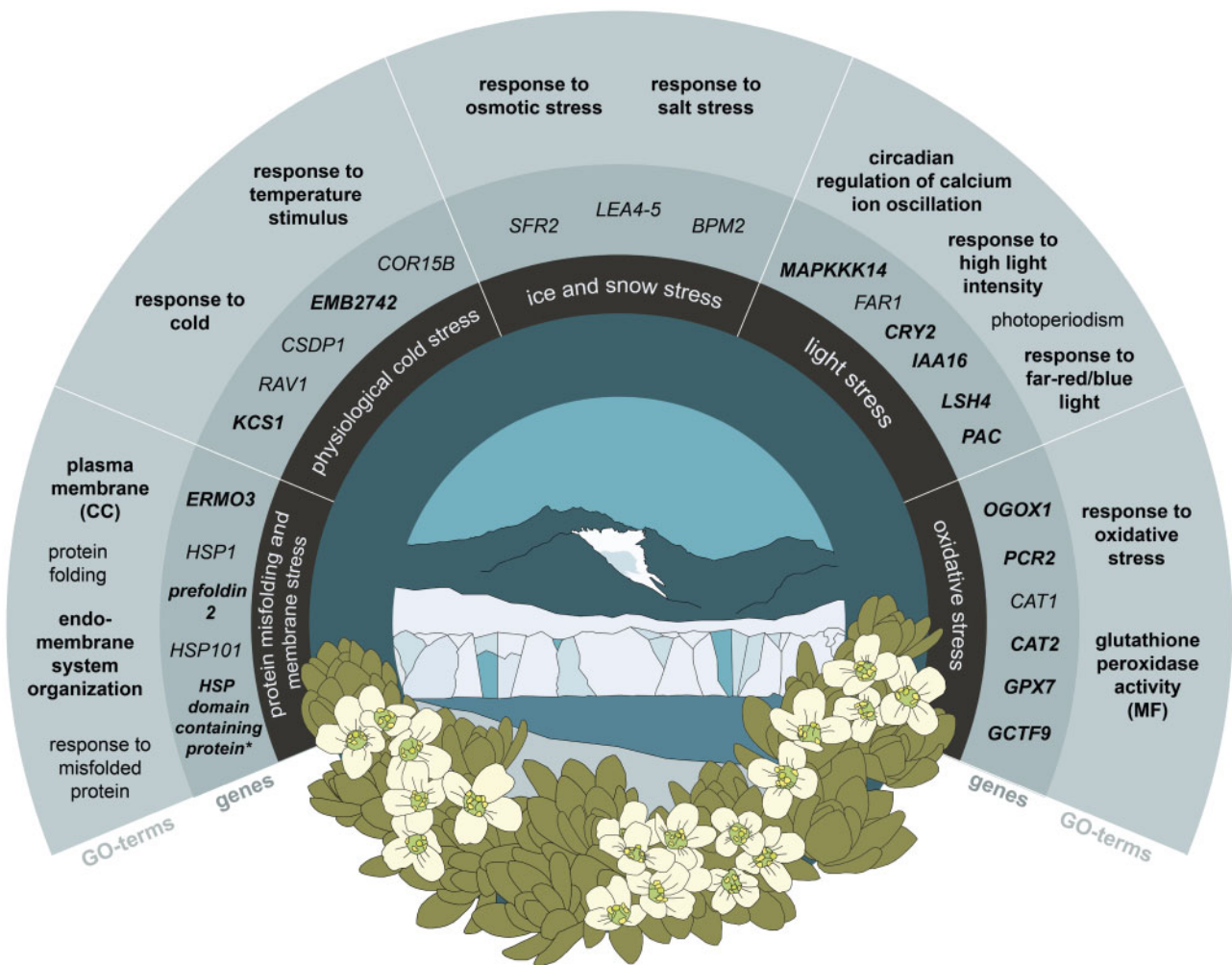
Fig. 7. Summary of putative molecular adaptations to five categories of Arctic stresses (inner black circle). The outer gray circle shows examples of enriched GO-terms (within the BP domain, unless indicated with CC or MF). The inner gray circle shows examples of candidate genes. Bold text indicates that the finding was made in more than one Arctic lineage. The results are based on PSG sets and convergent gene sets. *DNAJ heat shock N-terminal domain-containing protein.

probably evolved under positive selection in non-Arctic species, possibly indicating that these represent rapidly evolving genes. Thus, we found limited evidence for positive selection being responsible for fixing convergent substitutions in Arctic lineages per se.

Discussion

Our findings reveal that positive selection has been acting on different genes associated with similar functional pathways in the three Arctic species and that these pathways can be associated with abiotic stress types frequently encountered in the Arctic (fig. 7). However, there was little evidence for independently fixed mutations at the same sites and/or positive selection acting on the same genes. The three species might therefore have evolved similar suites of adaptations through distinct evolutionary pathways, implying that there can be

many genetic trajectories for adaptation to the Arctic environment.

Limited Evidence for Molecular Convergence in Arctic Brassicaceae: Comparisons and Causes

Considering the growing body of literature showing that the genomic landscape of adaptation may involve a high degree of repeatability also at the level of substitutions and genes (Conte et al. 2012; Stern 2013), an important question arises: Why do we primarily find molecular convergence at the level of pathways in Arctic Brassicaceae? The amount of molecular convergence reported here is considerably lower than that found in other studies (Foote et al. 2015; Hu et al. 2017; Xu et al. 2017; Davies et al. 2018). Notably, we found no orthologous genes under positive selection in all three Arctic species, and only 2% of all PSGs were found to be under positive selection in Arctic species pairs. None of these genes

contained convergent substitutions, which implies that gene function could potentially have diverged rather than converged in the two lineages (Conte et al. 2012). The numbers are lower than those found in a study reporting limited evidence for parallel molecular adaptation in four lineages of subterranean mammals (Davies et al. 2018). In their study, Davies et al. (2018) found that 11% of all PSGs were under selection in more than one lineage, <1% were under selection in three lineages, and none in all four lineages. Our methodology is similar to that of Davies et al. (2018), but the results may not be directly comparable as they represent quite different organismal groups and divergence times (~90 Ma for subterranean mammals vs. ~30 Ma for Arctic Brassicaceae; Foley et al. 2016; Huang et al. 2016). In plants, a comparable study has been conducted in three species of mangroves (Xu et al. 2017). Xu et al. (2017) calculated the amount of nonneutral molecular convergence using only conserved sites and found that as many as ~400 genes contained nonneutral convergent substitutions in at least two of three mangrove species (this was interpreted as a genome-wide pattern of convergent evolution; Xu et al. 2017). In comparison, we found half as many genes with convergent substitutions in Arctic species pairs (~200 genes) without distinguishing between nonneutral and neutral substitutions. Xu et al. (2017) may have identified more evidence for convergent evolution because they used whole genome data rather than transcriptomes, which are more likely to contain the complete set of genes for a given species, but they also limited their analyses to a subset of the sites and included fewer alignments. It is also possible that specific features of the Arctic environment make molecular convergence less likely, or that our conservative study design made it difficult to detect molecular convergence.

One reason why molecular convergence may be less likely in the Arctic is the vast opportunity space for adaptation. There are numerous potential evolutionary changes that could lead to increased fitness in the Arctic, because temperature and light affect nearly all aspects of plant cell biochemistry and metabolism (Salisbury 1985; Heino and Palva 2004). This implies that adaptive molecular convergence might be uncommon in Arctic plants solely due to the high number of possible evolutionary trajectories leading to better performance under low temperatures and specific light conditions. Furthermore, plants living in the Arctic need to cope with a wide range of abiotic stresses, and research on plant stress responses has shown considerable overlap in gene expression patterns between various stress stimuli (Kilian et al. 2007). This could imply that there is considerable evolutionary constraint on these pathways and/or that adaptation to the Arctic environment might have occurred by modifying different elements in a variety of general stress response pathways. Arctic plants may thus represent an interesting example of how general pathways may lead to the evolution of different solutions to similar environmental challenges. It is also possible that the low levels of molecular convergence we document in Arctic Brassicaceae could be due to different levels of preadaptation to cold and stressful environments in non-Arctic ancestral lineages. The major clades of the Brassicaceae radiated around the Eocene–Oligocene boundary (33.9 Ma), possibly in response to a cooler and drier environment (Huang et al. 2016), and parts of the Arctic flora are likely derived from montane ancestors that migrated northwards as global temperatures dropped in the mid-Miocene and onwards (Abbott and Brochmann 2003). The Arctic also spans a wide range of growing season lengths and summer temperatures (Elvebakk 1999) and may be quite environmentally heterogeneous (see, e.g., Jónsdóttir 2005 on the heterogeneity of a High Arctic Archipelago). Although there are broad similarities in the selection pressures encountered by Arctic plants, the specific combination and strength of these pressures likely vary among different Arctic habitats. Such variation may lead to different evolutionary trade-offs and a reduced probability of molecular convergence. The three Arctic species examined here occur in the same range of bioclimatic subzones (ranging from extreme Arctic polar desert to Arctic shrub tundra; Elvebakk 1999; Elven et al. 2011), but they are often found in species-specific habitats. Cochlearia groenlandica grows, for example, along sea shores and on lush bird cliffs, C. bellidifolia on moist mineral soils and in snowbeds, and D. nivalis in well-drained habitats such as rocky outcrops, crevices, and rock ledges (Alsos et al. 2019). Nonetheless, our results provide evidence for considerable molecular convergence at the level of pathways, which means that the three lineages likely have experienced similar types of selection pressure. It is also possible that there may be other molecular mechanisms of Arctic adaptation that we were unable to detect. For example, convergent evolution in gene regulatory regions has been shown to be more prevalent than convergent evolution in protein-coding regions in the independent loss of flight in several lineages of paleognathous birds (Sackton et al. 2019). Future studies on molecular convergence in Arctic Brassicaceae should therefore include data on sequence evolution in regulatory regions and patterns of differential gene expression.

Conservative Estimates of Molecular Convergence

The choice of methods and study design makes our estimates of molecular convergence in Arctic Brassicaceae relatively conservative in four ways: First, the branch-site test of positive selection has proven to be relatively powerful and result in few false positives (Zhang et al. 2005; Yang and dos Reis 2011). Second, since the branch-site test can be sensitive to sequence and alignment errors (Yang and dos Reis 2011), we applied a stringent sequence and alignment filtering scheme that led to the removal of many alignments. This means that we may to some degree have underestimated the number of PSGs, as well as genes with convergent substitutions. Third, we included several species representatives from each of the three relevant clades in the Brassicaceae, because inappropriate taxonomic sampling can lead to overestimation of molecular convergence (Thomas et al. 2017). We also included two genotypes from different areas in the Arctic for each of our focal species, which narrowed our focus to evolutionary changes that were likely fixed at species level. Finally, because this study largely is based on transcriptome assemblies, the whole gene space was not recovered, and there might be

genes that are only expressed in certain tissues or that require specific environmental cues to initiate expression which are not included in our data set. This could also contribute to relatively conservative estimates of molecular convergence. Once whole-genome sequences are available for these Arctic species and temperate relatives, it is possible that more convergent genes could be identified. Nonetheless, we expect signatures of adaptation to be present in genes that are constitutively expressed in these species, because Arctic plants experience relatively low temperatures and extreme light regimes year-round, and this was confirmed by the strong signal of Arctic adaptation in our analyses. It is therefore likely that our results provide a good, but conservative estimate of the amount of molecular convergence in protein-coding regions in these three lineages of Arctic Brassicaceae. Furthermore, the relationship between convergent and divergent substitutions should be relatively unaffected by the stringent filtering scheme, and we found no indication of excess convergence, consistent with our low estimates of molecular convergence.

A Functional Syndrome of Arctic Adaptation: A Hypothesis for Experimental Verification

In this study, we have presented some of the first genetic evidence for putative plant adaptations to the Arctic environment. We found signatures of positive selection and convergent substitutions in genes associated with general abiotic stress responses, such as the "MAPK-cascade," "calcium channel activity," "positive regulation of abscisic acid-activated pathway," and "regulation of stomatal movement" (figs. 4 and 7). Most abiotic stress responses are initiated with an increase in cytoplasmic calcium, and numerous abiotic stresses also activate MAPK-cascades (Knight and Knight 2001). The plant hormone abscisic acid similarly plays important roles in stress signaling (de Zelicourt et al. 2016), and abscisic acid-level increase in response to both high light and during cold acclimation (Heino and Palva 2004; Tuteja 2007; Galvez-Valdivieso et al. 2009). Finally, stomatal regulation plays an important role in plant responses to environmental changes, and plants will, for example, close their stomata in response to drought or cold stress (Wilkinson et al. 2001). It is likely that extreme abiotic stress directly and/or indirectly linked to the dramatic light and temperature conditions of the Arctic may have influenced the evolution of such general stress response pathways.

Our results also revealed more specific pathways and genes related to the five categories of Arctic stresses. Relevant to *protein misfolding and membrane stress* (category 1), we found strong support for the plasma membrane playing a vital role in adaptation to the Arctic environment (see fig. 7 and supplementary fig. S4, Supplementary Material online). In cold-adapted organisms, the plasma membrane will often act as a semipermeable barrier that allows water to flow out of the cell in the presence of ice, while also preventing the spread of ice crystals within the cell (Uemura and Steponkus 1999). The Arctic species also contained many heat shock-related proteins (Hsps/chaperones) under positive selection, which are involved in membrane stabilization and protein folding under normal and stressful cellular conditions (Wang et al. 2004). Three genes under positive selection in more than one Arctic species were involved in protein folding and/or traffic, as well as endoplasmic reticulum morphology. These results highlight the important role that protein misfolding and the accumulation of misfolded proteins in the endoplasmic reticulum can play in cellular responses to the abiotic stresses, particularly cold stress, imposed by Arctic environments (Zhu 2016).

Of relevance to *physiological cold stress* (category 2), we found that many cold-regulated genes are under positive selection and carry convergent substitutions in at least two Arctic species. Convergent substitutions were identified in *EMB2742* (Embryo defective 2742/CTP synthase family protein) in all Arctic branch pairs. This gene has been found to be upregulated during cold acclimation in *A. thaliana* (Oono et al. 2006). In addition, all three Arctic species had one or more *LEA* genes under positive selection. These genes may be associated with *physiological cold stress*, as most *LEA* genes in *A. thaliana* contain abscisic acid response elements and/or low-temperature response elements in their promoters (Hundertmark and Hincha 2008). One of the *LEA* genes under positive selection in *D. nivalis* is the well-known cold-responsive *COR15b* (Cold Regulated 15b), which functions in stabilizing chloroplast membranes, and whose overexpression yields increased freezing tolerance in *A. thaliana* (Navarro-Retamal et al. 2016). Our identification of *COR15b* as a candidate gene for Arctic adaptation is consistent with a previous study showing that *Draba* contains a lineage-specific *COR15* duplication and that the *COR15b* duplicate appears to be evolving under positive selection (Zhou et al. 2009). Other important cold-regulated genes include *CSDP1* (Cold Shock Domain Protein 1; *D. nivalis*) and *RAV1* (related to ABI2/VP1; *Co. groenlandica*). Cold shock domains are conserved nucleic acid-binding domains found in a variety of organisms, and overexpression of *CSDP1* in *A. thaliana* has been shown to be associated with delayed seed germination in stressful conditions (Sasaki et al. 2013). The low-temperature induced transcription factor *RAV1* is involved in the regulation of plant growth under cold stress in *A. thaliana* and is known to exhibit circadian regulation (Chinnusamy et al. 2007; TAIR 2019).

Arctic plants are always at risk of experiencing *ice and snow stress* (category 3), and we found one candidate gene associated with "response to freezing" in *Co. groenlandica* (*SFR2*, sensitive to freezing 2) that plays a role in the protection of the chloroplast envelope against freezing injuries (Fourrier et al. 2008). *Ice and snow stress* can also lead to loss of cellular water and increase solute concentration in the cytoplasm (Wisniewski and Fuller 1999; Körner 2003). In this way, the stress response pathways associated with freezing tolerance are partially homologous with those of dehydration and salt tolerance (Bartels and Souer 2004), and thus the >50 candidate genes functionally associated with salt stress and osmotic stress may also be connected to this stress category. Some of the previously mentioned LEA proteins are also known to be involved in protection from freezing damage

(Hundertmark and Hincha 2008; Cuevas-Velazquez et al. 2014). That Arctic Brassicaceae may be well adapted to ice stress is consistent with a recent study demonstrating high tolerance to ice encasement in other Arctic plants (Bjerke et al. 2018).

Our results also highlight several putative molecular adaptations to the Arctic light regime (category 4; *light stress*). Many genes annotated with far-red or blue light response were found to be under positive selection or containing convergent substitutions in the Arctic species. Specifically, we identified convergent substitutions in Arctic branch pairs in the genes *MAPKKK14* (involved in far-red light induced change to photomorphogenesis in seedlings; Khanna et al. 2006) and *CRY2* (a blue light receptor involved in cotyledon expansion and flowering time; TAIR 2019). We also found that *IAA16* and *PAC* (both involved in light-controlled growth and development; Rinaldi et al. 2012; TAIR 2019) are under positive selection in two different Arctic species. The evolution of these genes may be connected to changes in day length and/or the enrichment of far-red and blue light with increasing latitude (cf., Nilsen 1985; Salisbury 1985). These genes can be interesting starting points for future experiments aimed at understanding plant adaptations to polar light climates.

Consistent with the fact that high irradiance in combination with low temperatures leads to excessive production of ROS (*oxidative stress*, category 5; Inzé and Montagu 1995; Heino and Palva 2004), we found that genes associated with oxidative stress were enriched in the PSG sets of *D. nivalis* and *C. bellidifolia*. Genes associated with "glutathione peroxidase activity" were also overrepresented among convergent genes in Arctic branch pairs, and these enzymes are involved in both cellular redox homeostasis and H_2O_2 detoxification. Their activity can be induced by a range of environmental stresses, including, for example, cold stress, drought, and UV-B light (Bela et al. 2015). Some interesting candidate genes for adaptation to oxidative stress in Arctic environments include *PCR2* and *OGOX1* (both under positive selection in *C. bellidifolia* and *Co. groenlandica*) and two catalase genes under positive selection in *D. nivalis* (*CAT1* and *CAT2*). *CAT1* and *CAT2* are involved in controlling cellular redox homeostasis (Du et al. 2008). *CAT2* was also found to contain convergent substitutions in *Co. groenlandica* and *D. nivalis*. Intriguingly, *CAT2* is both cold and drought stress activated and involved in preventing oxidative stress under photorespiration, while also being conditioned by photoperiod (Queval et al. 2007; TAIR 2019).

To summarize, we found putative molecular adaptations to cold stress, freezing stress, oxidative stress, and the unique light regime of polar environments in all three Arctic species. We also found ample evidence that Arctic Brassicaceae possess adaptations associated with the plasma membrane and protein folding, a finding which is highly relevant for coping with typical Arctic stresses. Our results thus suggest that Arctic plants may have evolved similar ecological preferences and phenotypes, but through different genetic trajectories. This potential "functional syndrome" of Arctic adaptation

will need experimental verification, and the hypothesized molecular adaptations suggested here can serve as a starting point for future studies.

Materials and Methods

Taxon Sampling, Transcriptome Assemblies, and Final Data Set

We acquired genome-wide coding sequences from 15 species belonging to clades A, B, and C in the Brassicaceae, including the three Arctic focal species *C. bellidifolia*, *Co. groenlandica*, and *D. nivalis* (fig. 1A). RNA sequences for the three Arctic species were generated de novo for this study (Supplementary Material online, GenBank SRA ID: SRP240404 and BioProject ID: PRJNA599963), whereas data for the other 12 species were derived from published genomes or transcriptomes (supplementary table S3, Supplementary Material online). Because we were mainly interested in substitutions fixed at the species level, we included two genotypes from each focal species, representing two distant Arctic regions (NO/US, supplementary table S1, Supplementary Material online). Seeds were harvested from field-collected plants and germinated in a growth chamber at the University of Oslo (cultivation conditions as specified by Brochmann et al. [1992]).

We downloaded RNA sequences from eight of the non-Arctic species (supplementary tables S2 and S3, Supplementary Material online). Short-read data were selected and downloaded from NCBI (https://www.ncbi.nlm.nih.gov/genbank/) based on the following criteria: 1) The species should belong to the same genus or clade as one of the Arctic focal species in the phylogenies of Huang et al. (2016) and Guo et al. (2017), 2) the data should have been generated on an Illumina machine, and 3) the files should contain more than 11 M fragments in total. The raw RNA-seq reads were quality-checked with FastQC v.0.11.4 (Andrews 2010) and assembled with Trinity v.2.4.0 (Grabherr et al. 2011) using the integrated Trimmomatic option and a minimum assembled contig length of 300 bp. The transcript abundance in each assembly was estimated with kallisto v.0.43.0 (Bray et al. 2016), and only the highest expressed isoform for each "trinity gene" was retained. We also acquired an early version of the now published transcriptome of the non-Arctic species *Co. pyrenaica* (for assembly details, see Lopez et al. 2017), which was filtered based on the longest isoform as we did not have access to the raw reads at the time. The following assembly processing was performed on all transcriptome assemblies: Chimeric transcripts, likely resulting from misassemblies, were filtered out according to the procedure of Yang and Smith (2013), and coding regions within the isoform-filtered transcripts were predicted with TransDecoder v.3.0.0 (Haas et al. 2013) using default settings. The final transcriptome quality was evaluated based on several metrics (see Supplementary Material online), including the number of complete and fragmented BUSCOs (BUSCO v.3.0.1, Simão et al. 2015). The transcriptomic data were then combined with coding sequences derived from the already published genomes of three additional non-Arctic species (*A. thaliana*,

Arabis alpina, and *C. hirsuta*; supplementary table S3, Supplementary Material online). Sequences containing missing data or ambiguity characters were excluded, and the sequence sets of *Arabis alpina* and *C. hirsuta* were run through TransDecoder to ensure that they, for example, did not contain stop codons. The filtered Arctic transcriptome assemblies are available at https://doi.org/10.5061/dryad.v41ns1rs0.

Orthogroup Inference and Alignment Generation

Alignments were generated separately for the positive selection tests and for the calculation of posterior expected numbers of convergent and divergent substitutions. For the positive selection tests, we ran OrthoFinder v.1.1.8 (Emms and Kelly 2015) separately for each of the three clades/groups illustrated in figure 1B–D. This was done in order to identify sets of genes descended from a single gene in the last common ancestor of the species (i.e., orthogroups; Emms and Kelly 2015). These orthogroups were used as a basis for subsequent alignment construction. Due to the low number of single-copy orthogroups (not uncommon in plants), orthogroups that contained multiple gene copies/transcripts per species were divided into subsets based on smallest genetic distance. All orthogroups were aligned based on protein sequence using MAFFT v.7.300 (Katoh et al. 2005), genetic distances between the gene copies/transcripts were calculated as Kimura protein distances (Kimura 1983), and one gene copy/transcript per species or accession was extracted based on the smallest genetic distance to the transcript of the Alaskan accession of either *C. bellidifolia*, *D. nivalis* or *Co. groenlandica*, depending on the clade in question. If there were more than one transcript from the Arctic accessions, a new subset was made for each such transcript. The resulting orthogroup subsets contained only one transcript per species/accession and were realigned with PRANK (Löytynoja and Goldman 2005) using GUIDANCE v2.02 (Sela et al. 2015) with ten bootstraps. GUIDANCE allows assessment of alignment confidence and subsequent removal of unreliable aligned columns and sequences. Previous studies have shown that a phylogeny aware aligner such as PRANK in combination with alignment filtering with GUIDANCE improves positive selection inference on simulated data (Jordan and Goldman 2012; Privman et al. 2012). All alignments with a sequence score <0.6, as well as all columns scoring <0.80, were removed from the data set. The three final alignment sets had an average length of 1,544–1,586 bp and contained on average 17.0–20.7% missing data, but sites containing alignment gaps were deleted within codeml (cleandata = 1).

For estimation of posterior expected numbers of convergent and divergent substitutions, we ran OrthoFinder v.2.3.1 (Emms and Kelly 2019) for the total data set (fig. 1A). The resulting number of single-copy orthogroups was extremely small, and we therefore extracted subsets from multi-copy orthogroups as described above for the positive selection tests. For simplicity, we extracted subsets based on only one of the Arctic focal species, that is, we extracted one transcript per species/accession based on the smallest genetic distance to the transcripts of *D. nivalis* (arbitrarily chosen). The alignments were then generated as explained above (final

average alignment length was 1,524 bp with an average of 25.8% missing data). Note that all alignments for both the positive selection tests and the calculation of posterior expected numbers of convergent and divergent substitutions contained all relevant species (cf., fig. 1).

The Branch-Site Test of Positive Selection

The branch-site test (model A; Zhang et al. 2005) implemented in codeml (PAML v.4.9d, Yang 1997) was used to test for positive selection on individual codons along the lineage leading to each of the Arctic species. In this test, an alternative model allowing positive selection on the foreground lineage (e.g., one of the Arctic lineages) is contrasted to a null-model that does not allow such positive selection using a likelihood-ratio test (Zhang et al. 2005). To limit the influence of missing data (a challenge when using transcriptomic data), we ran the tests separately for clades A, B, and C (Huang et al. 2016; fig. 1B–D). However, the model species *A. thaliana* was included in all runs, and due to the low availability of high-quality transcriptomic data from clade C, we also included two species from clade B in the *Co. groenlandica* data set (fig. 1D). Codeml was run from four to six times per model with different initial parameter values, and the run with the highest likelihood score was used in subsequent analyses (Wong et al. 2004). The likelihood-ratio tests were performed in R v.3.4.4 (R Core Team 2019) using extRemes (Gilleland and Katz 2016). The significance level was set to 0.05 with 1 degree of freedom. Phylogenetic trees (representing the three clades) used in the analyses were constructed using RAxML based on codon alignments of ~2,000–4,000 single-copy genes (Stamatakis 2014; see fig. 1B–D, Supplementary Material online). We also ran the branch-site test for each topology in figure 1B–D with a close non-Arctic relative (*C. hirsuta*, *Co. pyrenaica*, and *D. nemorosa*) as the foreground branch. This was done in order to filter out genes that always seem to be under positive selection and to gain a better understanding of which genes that actually could be tied to Arctic adaptation. The resulting sets of PSGs were compared between species in order to identify genes with convergent selection patterns. Genes were considered as being under convergent selection if the same *A. thaliana* ortholog was found to be under positive selection in two or more species. The significance of the gene intersections was then assessed with the supertest function in SuperExactTest v.1.0.7 (Wang et al. 2015) and visualized with UpSetR 1.3.3 (Conway et al. 2017).

Posterior Expected Numbers of Convergent and Divergent Substitutions

We used Grand Convergence (Qian and de Koning 2018; Qian et al. 2018) to calculate the posterior expected numbers of convergent and divergent substitutions across all branch pairs in a super phylogeny generated with OrthoFinder v.2.3.1 (Emms and Kelly 2019; fig. 1A). This was done to identify sets of convergent genes (i.e., genes containing at least one convergent substitution in Arctic species pairs) and to examine deviations from the expected relationship between convergent and divergent substitutions. We also estimated the

number of convergent genes in the non-Arctic reference species. In Grand Convergence, a convergent substitution is defined as changes resulting in the same amino acid at the same site in two lineages, and divergent substitutions as changes resulting in different amino acids at the same site in two lineages (silent substitutions are not considered; Castoe et al. 2009). The program was run with branch length estimations from OrthoFinder, the Le Gascuel amino acid substitution model (Le and Gascuel 2008), and all gaps and ambiguity characters excluded. Genes containing at least one convergent site with PP >0.80 were considered as convergent, but since a gene/transcript potentially could be included in more than one alignment following our orthogroup subsetting procedure (see above), we only included unique TAIR IDs (i.e., *A. thaliana* orthologs) in our list of convergent genes. The resulting sets of convergent genes were compared with each other in the same way as the PSGs in order to identify genes that were shared between more than one Arctic branch pair.

To investigate the relationship between divergent and convergent substitutions in our data set, we chose one alignment per orthogroup and summed all pairwise convergent substitutions and all pairwise divergent substitutions over all alignments. Using the branch totals from 50 branch pairs, we performed a simple linear regression to investigate whether our data fitted a linear model as expected from other studies (Castoe et al. 2009; Thomas and Hahn 2015). The 50 branch pairs were mostly arbitrarily chosen, except that all pairwise combinations between the Arctic species and between the non-Arctic reference species were included (see supplementary table S32, Supplementary Material online, for a complete list of the 50 branch pairs).

Gene Annotations and GO Enrichment Tests

The final sets of PSGs and convergent genes were annotated with DAVID v.6.8 (Huang et al. 2009) based on the *A. thaliana* ortholog. We also tested for overrepresented GO-terms within the three domains: BP, CC and MF using a Fisher's exact test in combination with the "classic," "elim," and "weight" algorithms implemented in topGO v.2.32 of Bioconductor (Gentleman et al. 2004; Alexa et al. 2006). The "classic" algorithm processes each term independently without taking the GO-hierarchy into account, the "elim" algorithm traverses the GO-graph bottom-up, discarding genes that have already been mapped to significant child terms, and the "weight" algorithm decides which GO-terms best describe the gene based on a weighting scheme that takes the enrichment scores of the neighboring GO-terms into account (Alexa et al. 2006). Simulations have shown that the "elim" and "weight" algorithms tend to be more conservative than the "classic" algorithm, and the creators of topGO recommend using "elim" for identifying important areas in the graph due to its simplicity (Alexa et al. 2006). The gene sets were annotated based on homology to *A. thaliana* using annotations from org.At.tair.db v.3.6 (Carlson 2018), and *A. thaliana* was also used as the background gene universe in all gene set enrichment analyses. Although it could be considered less than ideal to apply the *A. thaliana*

background in this type of analysis, it is likely to be more consistent than creating new backgrounds based on the annotation of de novo transcriptome assemblies for each species, because transcriptome assemblies represent a less complete sampling of gene space and are prone to varying degrees of sampling bias. We did not correct for multiple testing as we wanted to avoid being overly conservative, and because such corrections cannot directly be applied to the results of the "elim" and "weight" algorithms due to non-independent testing (the topGO manual; Alexa and Rahnenfuhrer 2018). Enriched GO-terms that might be tied to Arctic adaptation were assessed based on literature, and the results from the non-Arctic species were used to filter out more general GO-terms that were enriched in all species (e.g., GO-terms related to immunity). Finally, we also generated a heatmap of the proportions of PSGs associated with the five categories of Arctic stresses, by using the number of genes annotated with a selection of these GO-terms. The annotations were identical to those used in the topGO analyses, and the heatmap was generated with gplots package v.3.0.1.1 in R (Warnes et al. 2019). (For a flow chart of the analysis pipeline, see supplementary fig. S1, Supplementary Material online.)

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

References

Abbott RJ, Brochmann C. 2003. History and evolution of the Arctic flora: in the footsteps of Eric Hultén. *Mol Ecol*. 12(2):299–313.

Alexa A, Rahnenfuhrer J. 2018. topGO: enrichment analysis for Gene Ontology. R package version 2.34.0. Available from: http://bioconductor.org/packages/release/bioc/html/topGO.html.

Alexa A, Rahnenführer J, Lengauer T. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22(13):1600–1607.

Alsos I, Arnesen G, Elven R, Sandbakk BE. 2019. The flora of Svalbard. Tromsø (Norway): The Arctic University of Norway. [cited 2019 Sep 27]. Available from: https://svalbardflora.no/

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bartels D, Souer E. 2004. Molecular responses of higher plants to dehydration. In: Hirt H, Shinozaki K, editors. Plant responses to abiotic stress. Berlin (Germany): Springer. p. 9–38.

Bela K, Horváth E, Gallé Á, Szabados L, Tari I, Csiszár J. 2015. Plant glutathione peroxidases: emerging role of the antioxidant enzymes in plant development and stress responses. *J Plant Physiol*. 176:192–201.

Billings WD, Mooney HA. 1968. The ecology of Arctic and alpine plants. *Biol Rev*. 43(4):481–529.

Bjerke JW, Elverland E, Jaakola L, Lund L, Bochenek Z, Kłos A, Tømmervik H. 2018. High tolerance of a high-arctic willow and graminoid to simulated ice encasement. *Boreal Environ Res*. 23:329–338.

Bolnick DI, Barrett RDH, Oke KB, Rennison DJ, Stuart YE. 2018. (Non)parallel evolution. *Annu Rev Ecol Evol Syst*. 49(1):303–330.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 34(5):525–527.

Brochmann C, Soltis PS, Soltis DE. 1992. Multiple origins of the octoploid Scandinavian endemic *Draba cacuminum*: electrophoretic and morphological evidence. *Nord J Bot*. 12(3):257–272.

Carlsen T, Bleeker W, Hurka H, Elven R, Brochmann C. 2009. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Ann Mo Bot Gard*. 96(2):215–236.

Carlson M. 2018. org.At.tair.db: Genome wide annotation for *Arabidopsis*. R package version 3.7.0. Available from: http://bioconductor.org/packages/release/data/annotation/html/org.At.tair.db.html.

Castoe TA, de Koning APJ, Kim H-M, Gu W, Noonan BP, Naylor G, Jiang ZJ, Parkinson CL, Pollock DD. 2009. Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc Natl Acad Sci U S A*. 106(22):8986–8991.

Chapin F. 1983. Direct and indirect effect of temperature on Arctic plants. *Polar Biol*. 2(1):47–52.

Chen L, DeVries AL, Cheng C-HC. 1997. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. *Proc Natl Acad Sci U S A*. 94(8):3817–3822.

Chinnusamy V, Zhu J, Zhu JK. 2007. Cold stress regulation of gene expression in plants. *Trends Plant Sci*. 12(10):444–451.

Chory J, Chatterjee M, Cook RK, Elich T, Fankhauser C, Li J, Nagpal P, Neff M, Pepper A, Poole D, et al. 1996. From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc Natl Acad Sci U S A*. 93(22):12066–12071.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc Biol Sci*. 279(1749):5039–5047.

Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33(18):2938–2940.

Cuevas-Velazquez CL, Rendón-Luna DF, Covarrubias AA. 2014. Dissecting the cryoprotection mechanisms for dehydrins. *Front Plant Sci*. 5:1–6.

Davies KTJ, Bennett NC, Faulkes CG, Rossiter SJ. 2018. Limited evidence for parallel molecular adaptations associated with the subterranean niche in mammals: a comparative study of three superorders. *Mol Biol Evol*. 35(10):2544–2559.

Du YY, Wang PC, Chen J, Song CP. 2008. Comprehensive functional analysis of the catalase gene family in *Arabidopsis thaliana*. *J Integr Plant Biol*. 50(10):1318–1326.

Elvebakk A. 1999. Bioclimatic delimitation and subdivision of the Arctic. In: Nordal I, Razzhivin VY, editors. The species concept in the high north: a panarctic flora initiative. Oslo (Norway): The Norwegian Academy of Science and Letters. p. 81–112.

Elven R, Murray DF, Razzhivin VY, Yurtsev BA. 2011. Annotated checklist of the Panarctic Flora (PAF). Vascular plants. Fairbanks (USA): University of Alaska. [cited 2019 Sep 27]. Available from: http://panarcticflora.org/

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol*. 16(1):157.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 20(1):1–14.

Foley NM, Springer MS, Teeling EC. 2016. Mammal madness: is the mammal tree of life not yet resolved? *Philos Trans R Soc Lond B Biol Sci*. 371(1699): pii: 20150140.

Foote AD, Liu Y, Thomas GWC, Vinař T, Alföldi J, Deng J, Dugan S, van Elk, CE, Hunter ME, Joshi V, et al. 2015. Convergent evolution of the genomes of marine mammals. *Nat Genet*. 47(3):272–275.

Fourrier N, Bédard J, Lopez-juez E, Barbrook A, Bowyer J, Jarvis P, Warren G, Thorlby G. 2008. A role for SENSITIVE TO FREEZING2 in protecting chloroplasts against freeze-induced damage in *Arabidopsis*. *Plant J*. 55(5):734–745.

Galvez-Valdivieso G, Fryer MJ, Lawson T, Slattery K, Truman W, Smirnoff N, Asami T, Davies WJ, Jones AM, Baker NR, et al. 2009. The high light response in *Arabidopsis* involves ABA signaling between vascular and bundle sheath cells. *Plant Cell* 21(7):2143–2162.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 5(10):R80.

Gilleland E, Katz RW. 2016. extRemes 2.0: an extreme value analysis package in R. *J Stat Softw*. 72(8):1–39.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat Biotechnol*. 29(7):644–652.

Guo X, Liu J, Hao G, Zhang L, Mao K, Wang X, Zhang D, Ma T, Hu Q, Al-Shehbaz IA, et al. 2017. Plastome phylogeny and early diversification of Brassicaceae. *BMC Genomics* 18(1):176.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Philip D, Bowden J, Couger MB, Eccles D, Li B, Macmanes MD, et al. 2013. *De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nat Protoc*. 8(8):1494–1443.

Heide OM. 2005. Ecotypic variation among European arctic and alpine populations of *Oxyria digyna*. *Arct Antarct Alp Res*. 37(2):233–238.

Heino P, Palva ET. 2004. Signal transduction in plant cold acclimation. In: Hirt H, Shinozaki K, editors. Plant responses to abiotic stress. Berlin (Germany): Springer. p. 151–186.

Hendry AP. 2013. Key questions in the genetics and genomics of eco-evolutionary dynamics. *Heredity* 111(6):456–466.

Heyduk K, Moreno-Villena JJ, Gilman IS, Christin P-A, Edwards EJ. 2019. The genetics of convergent evolution: insights from plant photosynthesis. *Nat Rev Genet*. 20(8):485–493.

Hu Y, Wu Q, Ma S, Ma T, Shan L, Wang X, Nie Y, Ning Z, Yan L, Xiu Y, et al. 2017. Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc Natl Acad Sci U S A*. 114(5):1081–1086.

Huang CH, Sun R, Hu Y, Zeng L, Zhang N, Cai L, Zhang Q, Koch MA, Al-Shehbaz IA, Edger PP, et al. 2016. Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Mol Biol Evol*. 33(2):394–412.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 4(1):44–57.

Hultén E. 1968. Flora of Alaska and neighboring territories: a manual of the vascular plants. Stanford (CA): Stanford University Press.

Hundertmark M, Hincha DK. 2008. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9(1):118–122.

Inzé D, Montagu MV. 1995. Oxidative stress in plants. *Curr Opin Biotechnol*. 6(2):153–158.

Jackson SD. 2009. Plant responses to photoperiod. *New Phytol*. 181(3):517–531.

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF. 1998. *Arabidopsis CBF1* overexpression induces *COR* genes and enhances freezing tolerance. *Science* 280(5360):104–107.

Jónsdóttir IS. 2005. Terrestrial ecosystems on Svalbard: heterogeneity, complexity and fragility from an Arctic island perspective. *Proc R Ir Acad*. 105B(3):155–165.

Jónsdóttir IS. 2011. Diversity of plant life histories in the Arctic. *Preslia* 83:281–300.

Jordan G, Goldman N. 2012. The effects of alignment error and alignment filtering on the sitewise detection of positive selection. *Mol Biol Evol.* 29(4):1125–1139.

Jordon-Thaden I, Hase I, Al-Shehbaz I, Koch MA. 2010. Molecular phylogeny and systematics of the genus *Draba* (Brassicaceae) and identification of its most closely related genera. *Mol Phylogenet Evol.* 55(2):524–540.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33(2):511–518.

Khanna R, Shen Y, Toledo-ortiz G, Kikis EA, Johannesson H, Hwang Y, Quail PH. 2006. Functional profiling reveals that only a small number of phytochrome-regulated early-response genes in *Arabidopsis* are necessary for optimal deetiolation. *Plant Cell* 18(9):2157–2171.

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K. 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J.* 50(2):347–363.

Kimura M. 1983. The neutral theory of molecular evolution. Cambridge: Cambridge University Press.

Knight H, Knight MR. 2001. Abiotic stress signalling pathways: specificity and cross-talk. *Trends Plant Sci.* 6(6):262–267.

Knight MR, Knight H. 2012. Low-temperature perception leading to gene expression and cold tolerance in higher plants. *New Phytol.* 195(4):737–751.

Koch MA. 2012. Mid-Miocene divergence of *Ionopsidium* and *Cochlearia* and its impact on the systematics and biogeography of the tribe Cochlearieae (Brassicaceae). *Taxon* 61(1):76–92.

Körner C. 2003. Climatic stress. In: Körner C, editor. Alpine plant life. Functional plant ecology of high mountain ecosystems. 2nd ed. Berlin (Germany): Springer. p. 101–119.

Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper JF, Mesa T, Row M, Diego S, California JAK, et al. 2002. Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiol.* 130(4):2129–2141.

Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol Biol Evol.* 25(7):1307–1320.

Lopez L, Wolf EM, Pires JC, Edger PP, Koch MA. 2017. Molecular resources from transcriptomes in the Brassicaceae family. *Front Plant Sci.* 8:1–15.

Losos JB. 2011. Convergence, adaptation, and constraint. *Evolution* 65(7):1827–1840.

Löytynoja A, Goldman N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A.* 102(30):10557–10562.

Lütz C. 2010. Cell physiology of plants growing in cold environments. *Protoplasma* 244(1–4):53–73.

Natarajan C, Hoffmann FG, Weber RE, Fago A, Witt CC, Storz JF. 2016. Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science* 354(6310):336–340.

Navarro-Retamal C, Bremer A, Alzate-Morales J, Caballero J, Hincha DK, González W, Thalhammer A. 2016. Molecular dynamics simulations and CD spectroscopy reveal hydration-induced unfolding of the intrinsically disordered LEA proteins COR15A and COR15B from *Arabidopsis thaliana*. *Phys Chem Chem Phys.* 18(37):25806–25816.

Nilsen J. 1985. Light climate in northern areas. In: Kaurin Å, Junttila O, Nilsen J, editors. Plant production in the north: Proceedings from "Plant Adaptation Workshop"; 1983 Sep 4–9; Tromsø, Norway. 1st ed. Tromsø (Norway): Norwegian University Press. p. 62–82.

Oono Y, Seki M, Satou M, Iida K, Akiyama K, Sakurai T, Fujita M, Yamaguchi-Shinozaki K, Shinozaki K. 2006. Monitoring expression profiles of *Arabidopsis* genes during cold acclimation and deacclimation using DNA microarrays. *Funct Integr Genomics.* 6(3):212–234.

Piskurewicz U, Turečková V, Lacombe E, Lopez-Molina L. 2009. Far-red light inhibits germination through DELLA-dependent stimulation of ABA synthesis and ABI3 activity. *EMBO J.* 28(15):2259–2271.

Privman E, Penn O, Pupko T. 2012. Improving the performance of positive selection inference by filtering unreliable alignment regions. *Mol Biol Evol.* 29(1):1–5.

Qian C, Bryans N, Kruykov I, de Koning APJ. 2018. Visualization and analysis of statistical signatures of convergent evolution. Alberta (Canada): University of Calgary. [cited 2019 Sep 27]. Available from: https://github.com/dekoning-lab/grand-conv.

Qian C, de Koning A. 2018. Rapid discovery of convergent molecular evolution across entire phylogenies. Alberta (Canada): University of Calgary. [cited 2019 Sep 27]. Available from: https://github.com/dekoning-lab/grand-conv.

Queval G, Issakidis-Bourguet E, Hoeberichts FA, Vandorpe M, Gakière B, Vanacker H, Miginiac-Maslow M, Van Breusegem F, Noctor G. 2007. Conditional oxidative stress responses in the *Arabidopsis* photorespiratory mutant *cat2* demonstrate that redox state is a key modulator of daylength-dependent gene expression, and define photoperiod as a crucial factor in the regulation of H_2O_2-induced cell death. *Plant J.* 52(4):640–657.

R Core Team. 2019. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: https://www.R-project.org/.

Rinaldi MA, Liu J, Enders TA, Bartel B, Strader LC. 2012. A gain-of-function mutation in *IAA16* confers reduced responses to auxin and abscisic acid and impedes plant growth and fertility. *Plant Mol Biol.* 79(4–5):359–373.

Ryu SB, Costa A, Xin Z, Li PH. 1995. Induction of cold hardiness by salt stress involves synthesis of cold- and abscisic acid-responsive proteins in potato (*Solanum commersonii* Dun). *Plant Cell Physiol.* 36(7):1245–1251.

Sackton TB, Clark N. 2019. Convergent evolution in the genomics era: new insights and directions. *Philos Trans R Soc Lond B Biol Sci.* 374(1777):20190102.

Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, Gardner PP, Clarke JA, Baker AJ, Clamp M, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* 364(6435):74–78.

Salisbury FB. 1985. Plant adaptations to the light environment. In: Kaurin Å, Junttila O, Nilsen J, editors. Plant production in the north: Proceedings from "Plant Adaptation Workshop"; 1983 Sep 4–9; Tromsø, Norway. 1st ed. Tromsø (Norway): Norwegian University Press. p. 43–61.

Sasaki K, Kim M-H, Imai R. 2013. *Arabidopsis* COLD SHOCK DOMAIN PROTEIN 2 is a negative regulator of cold acclimation. *New Phytol.* 198(1):95–102.

Sela I, Ashkenazy H, Katoh K, Pupko T. 2015. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* 43(W1):W7–W14.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.

Stern DL. 2013. The genetic causes of convergent evolution. *Nat Rev Genet.* 14(11):751–764.

[TAIR] The Arabidopsis Information Resource. 2019. Fremont: Phoenix Bioinformatics Corporation. [cited 2019 Oct 23]. Available from: www.arabidopsis.org.

Thomas GWC, Hahn MW. 2015. Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. *Mol Biol Evol.* 32(5):1232–1236.

Thomas GWC, Hahn MW, Hahn Y. 2017. The effects of increasing the number of taxa on inferences of molecular convergence. *Genome Biol Evol.* 9(1):213–221.

Thomashow MF. 1999. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annu Rev Plant Physiol Plant Mol Biol.* 50(1):571–599.

Thomashow MF. 2010. Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. *Plant Physiol.* 154(2):571–577.

Tuteja N. 2007. Abscisic acid and abiotic stress signaling. *Plant Signal Behav.* 2(3):135–138.

Uemura M, Steponkus PL. 1999. Cold acclimation in plants: relationship between the lipid composition and the cryostability of the plasma membrane. *J Plant Res.* 112(2):245–254.

Verslues PE, Agarwal M, Katiyar-Agarwal S, Zhu J, Zhu JK. 2006. Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant J.* 45(4):523–539.

Wang M, Zhao Y, Zhang B. 2015. Efficient test and visualization of multi-set intersections. *Sci Rep.* 5:16923.

Wang W, Vinocur B, Shoseyov O, Altman A. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 9(5):13–15.

Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, et al. 2019. gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1.1. Available from: https://cran.r-project.org/web/packages/gplots/index.html.

Wehrmeister RR, Bonde EK. 1977. Comparative aspects of growth and reproductive biology in Arctic and Alpine populations of *Saxifraga cernua* L. *Arct Alp Res.* 9(4):401–406.

Wilkinson S, Clephan AL, Davies WJ. 2001. Rapid low temperature-induced stomatal closure occurs in cold-tolerant *Commelina communis* leaves but not in cold-sensitive Tobacco leaves, via a mechanism that involves apoplastic calcium but not abscisic acid. *Plant Physiol.* 126(4):1566–1578.

Wisniewski M, Fuller M. 1999. Ice nucleation and deep supercooling in plants: new insights using infrared thermography. In: Margesin R, Schinner F, editors. Cold-adapted organisms. Ecology, physiology, enzymology and molecular biology. Berlin (Germany): Springer. p. 105–118.

Wisniewski M, Fuller M, Palta J, Carter J, Arora R, Fuller M, Palta J, Carter J, Arora R. 2004. Ice nucleation, propagation, and deep supercooling in woody plants. *J Crop Improv.* 10(1–2):5–16.

Wong WSW, Yang Z, Goldman N, Nielsen R. 2004. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* 168(2):1041–1051.

Xu S, He Z, Guo Z, Zhang Z, Wyckoff GJ, Greenberg A, Wu CI, Shi S, Zhang J. 2017. Genome-wide convergence during evolution of mangroves from woody plants. *Mol Biol Evol.* 34(4):1008–1015.

Yang Y, Smith SA. 2013. Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. *BMC Genomics* 14(1):328–311.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci.* 13(5):555–556.

Yang Z, dos Reis M. 2011. Statistical properties of the branch-site test of positive selection. *Mol Biol Evol.* 28(3):1217–1228.

de Zelicourt A, Colcombet J, Hirt H. 2016. The role of MAPK modules and ABA during abiotic stress signaling. *Trends Plant Sci.* 21(8):677–685.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol.* 22(12):2472–2479.

Zhou D, Zhou J, Meng L, Wang Q, Xie H, Guan Y, Ma Z, Zhong Y, Chen F, Liu J. 2009. Duplication and adaptive evolution of the *COR15* genes within the highly cold-tolerant *Draba* lineage (Brassicaceae). *Gene* 441(1–2):36–44.

Zhu J. 2016. Abiotic stress signaling and responses in plants. *Cell* 167(2):313–324.