Functional variants of DOG1 control seed chilling responses and variation in seasonal life-history strategies in Arabidopsis thaliana

Alejandra Martinez-Berdeja1, Michelle C. Stitzer2,3, Mark A. Taylor2, Miki Okada2, Exequiel Ezcurra4, Daniel E. Runcie5, and Johanna Schmitt2,3,1

1Department of Evolution and Ecology, University of California, Davis, CA 95616; 2Center for Population Biology, University of California, Davis, CA 95616; 3Department of Botany and Plant Sciences, University of California, Riverside, CA 92521; and 4Department of Plant Sciences, University of California, Davis, CA 95616

Contributed by Johanna Schmitt, November 22, 2019 (sent for review July 19, 2019; reviewed by Leonie Bentsink, Xavier Picó, and Peter Tiffin)

The seasonal timing of seed germination determines a plant’s realized environmental niche, and is important for adaptation to climate. The timing of seasonal germination depends on patterns of seed dormancy release or induction by cold and interacts with flowering-time variation to construct different seasonal life histories. To characterize the genetic basis and climatic associations of natural variation in seed chilling responses and associated life-history syndromes, we selected 559 fully sequenced accessions of the model annual species Arabidopsis thaliana from across a wide climate range and scored each for seed germination across a range of 13 cold stratification treatments, as well as the timing of flowering and senescence. Germination strategies varied continuously along 2 major axes: 1) Overall germination fraction and 2) induction vs. release of dormancy by cold. Natural variation in seed responses to chilling was correlated with flowering time and senescence to create a range of seasonal life-history syndromes. Genome-wide association identified several loci associated with natural variation in seed chilling responses, including a known functional polymorphism in the self-binding domain of the candidate gene DOG1. A phylogeny of DOG1 haplotypes revealed ancient divergence of these functional variants associated with periods of Pleistocene climate change, and Gradient Forest analysis showed that allele turnover of candidate SNPs was significantly associated with climate gradients. These results provide evidence that A. thaliana’s germination niche and correlated life-history syndromes are shaped by past climate cycles, as well as local adaptation to contemporary climate.

In annual plants, variation in seasonal germination timing creates alternative life-history strategies (SI Appendix, Fig. S1). Winter annuals germinate in autumn, overwinter, and then flower and disperse seed in spring, whereas spring or summer annuals overwinter as seeds and germinate, flower, and disperse seed in spring or summer (11, 17, 21, 23). Mixtures of fall and spring germination cohorts are also observed within populations (7, 10, 11, 17, 21), a form of within-year bet hedging (25). Whether and when seeds germinate in a given seasonal environment depends upon the germination niche, the range of conditions under which germination is possible (7, 11, 21, 26). Primary dormancy of freshly dispersed seeds varies among genotypes and seed maturation environments (22, 23, 27–33). Release from primary dormancy may occur through after-ripening at warm ambient temperatures, or through short exposure to chilling (21, 22, 34, 35). Both of these mechanisms allow germination of winter annual seedlings in fall. In many winter annuals, seeds remaining in the soil seed bank in fall enter secondary dormancy when exposed to prolonged winter chilling, preventing winter and spring germination (21, 22, 26). In

Significance

The seasonal timing of seed germination is critical for plant fitness in different climates. To germinate at the right time of year, seeds respond to seasonal environmental cues, such as cold temperatures. We characterized genetic variation in seed dormancy responses to cold across the geographic range of a widespread annual plant. Induction of secondary seed dormancy during winter conditions (which restricts germination to autumn) was positively correlated with flowering time, constructing winter and spring seasonal life-history strategies. Variation in seed chilling responses was strongly associated with functional variants of a known dormancy gene. These variants showed evidence of ancient diversification associated with Pleistocene glacial cycles, and were associated with climate gradients across the species’ geographical range.

Author contributions: A.M.-B. and J.S. designed research; A.M.-B. and M.O. performed research; A.M.-B., M.C.S., M.A.T., E.E., D.E.R., and J.S. analyzed data; and A.M.-B., M.C.S., and J.S. wrote the paper.

Reviewers: L.B., Wageningen University & Research; X.P., Estación Biológica de Doñana (Consejo Superior de Investigaciones Científicas); and P.T., University of Minnesota.

The authors declare no competing interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: Germination, phenology and DOG1 haplotype data are available on Dryad repository, https://doi.org/10.5061/dryad.br6v4. DOG1 haplotype data and code are also available on GitHub, https://github.com/mstitzer/martinez-berdeja_dog1.

1To whom correspondence may be addressed. Email: jschmitt@ucdavis.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1912451117/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1912451117 PNAS Latest Articles | 1 of 9
contrast, summer annual seeds lose dormancy with chilling over winter, allowing germination in spring (21, 23). Seeds that lose dormancy in fall and do not enter secondary dormancy over the winter may germinate in either fall or spring resulting in a mixture of fall and spring annual life histories. Natural variation in dormancy responses to seasonal chilling cues thus determines the germination niche and the expression of seasonal life-history strategies within and among populations in different environments.

Germination timing also determines the selective environment for subsequent life-history traits, such as flowering time (1, 15, 16, 18). Seedlings that germinate in fall may be selected to delay flowering until winter is past, which may favor late flowering and strong vernalization (winter chilling) requirements. However, strong vernalization requirements and late flowering may be maladaptive for summer annuals, especially if the growing season is short. Consequently, natural selection on germination and later life-history traits may be correlated, and adaptation to climate across a species' range may therefore involve coordinated evolution of suites of life-history traits (11, 12, 19, 20, 36, 37).

To understand the adaptive evolution of the germination niche across different climates, it is important to elucidate its genetic basis. It is also of interest to examine the genetic basis of correlated traits that may form adaptive life-history syndromes, such as winter vs. summer annual life histories. If different traits share common genetic mechanisms, this pleiotropy may constrain or facilitate adaptive evolution of multivariate syndromes (38). On the other hand, if the loci contributing to variation in different traits do not overlap, then adaptive divergence of life-history syndromes could only occur through correlated selection response at different sets of loci across a climate gradient. Once loci underlying life-history variation are identified, we can test for a geographic signature of selection by climate. If allelic variants at those loci are involved in adaptation to climate, we expect them to be significantly associated with relevant climate variables across the landscape (39–41).

The model plant A. thaliana is an ideal system for dissecting the genetic and environmental determinants of climate adaptation in the germination niche and correlated life-history traits. Extensive sequence data are available for this annual species, facilitating genome-wide association (GWA) (42, 43). A. thaliana inhabits a wide native climate range across Eurasia and Africa, and has shifted across the landscape with Pleistocene climate changes (44–47). Across this range, the species exhibits substantial life-history variation (11, 13, 48). The pathways involved in seed dormancy and flowering time have been characterized through functional studies (24, 49). Natural variation in primary dormancy and after-ripening is also well documented (11, 12, 19, 20, 50), and underlying allelic variants have been identified (13, 27, 50–54). However, less is known about the genetic basis of natural variation in seed chilling responses in this species, despite their importance for fine-tuning dormancy cycles to shape seasonal germination timing (19, 22).

DELAY OF GERMINATION 1 (DOG1), a key regulator of seed dormancy and a member of a small gene family of unknown molecular function (27), is a particularly important candidate gene for natural variation in germination niche traits. This gene controls primary seed dormancy through multiple mechanisms (55–62). Allelic variants of DOG1 are associated with natural variation in primary dormancy, after-ripening, and germination timing in the field (5, 13, 50, 52, 63). DOG1 expression is associated with dormancy variation measured as after-ripening period (27, 53) and exhibits clinal variation; southern ecotypes have higher DOG1 expression associated with longer after-ripening (12, 51, 63). DOG1 variation is also associated with natural variation in flowering time (42, 43, 64) and may have pleiotropic effects both indirectly through cascading effects of germination timing (6) and directly through its effects on levels of associated micro RNAs (60). However, little is known about the contribution of DOG1 functional polymorphisms to natural variation in germination responses to chilling.

Here we combine experimental phenotypic data from a geographically diverse set of accessions, whole-genome polymorphism data (43), and geographic and climate information to address the following questions:

1) What is the range of natural variation of germination responses to chilling? How does natural variation in germination responses to chilling covary with other life-history traits to shape winter and spring annual life-history syndromes across the landscape?

2) What is the genetic basis of natural variation in germination responses to chilling and associated life-history traits? Does allelic variation in DOG1 contribute to variation in seed chilling responses to shape the seasonal germination niche?

3) Do these genetic variants exhibit a geographic signature of adaptation to climate?

Results and Discussion

Natural Variation in Seed Chilling Responses Creates Diverse Seasonal Life Histories. We grew 559 Arabidopsis accessions to reproductive maturity at 14 °C with a 12/12-h photoperiod following 6 wk of vernalization at 4 °C. Fresh seeds were harvested from each plant when siliques matured, and were tested for germination at 22 °C immediately (i.e., base germination), or after 4, 8, 11, 15, 22, or 32 d of dark stratification at 4 °C and 10 °C. We performed a principal component analysis (PCA) on all germination phenotypes. The first 2 principal components (PCs) of germination timing and dormancy level explained 91% of the germination variation (PC1germ = 83%, PC2germ = 7%) (Fig. 1 and SI Appendix, Fig. S2). We also measured days to flowering (DTF), and days to senescence (DTS) of the maternal plants. Additionally, we assayed germination of dry-stored seeds at 6-wk intervals to quantify after-ripening requirements (days of seed storage required to reach 50% germination, DSDS50) (Table 1).

All germination variables measured were positively associated to PC1germ (SI Appendix, Table S1), and ecotypes with a positive score along this axis had overall high germination. Accessions with high overall germination (PC1germ) were late-flowering and late-senescing and had short after-ripening times, whereas accessions with low germination were early-flowering and early-senescing and had long after-ripening times (Table 1). Variation in overall germination fraction reflects variation in primary dormancy, consistent with previous observations of natural variation in dormancy measured as after-ripening requirement (11, 12, 50). Accessions with high germination were distributed in central and northern Europe (SI Appendix, Fig. S3A). Low dormancy allows immediate germination after dispersal, which is favored by selection in northern climates (13) and may result in multiple generations per year in midlatitudes with wet summers (7, 65). In contrast, primary dormancy that cannot be broken by chilling exposure, correlated with strong after-ripening requirements, would maintain a large population in the seed bank (7), a potential bet-hedging strategy (66). Accessions with low germination were distributed in south and central Spain and southern Europe (SI Appendix, Fig. S3A).

PC2germ represents the germination response to cold. Base germination and germination percent of all stratification times at 10 °C were positively associated with PC2germ, while germination percent at 11 to 32 d at 4 °C were negatively associated to this axis (SI Appendix, Table S1). Accessions with a positive score on PC2germ had high base germination and secondary dormancy induced by prolonged cold exposure at 4 °C (Fig. 1 and SI Appendix, Fig. S2 and Table S1), flowered and senesced later, and had longer after-ripening periods (Fig. 2 and Table 1). These accessions would behave as winter annuals, with germination
restricted to late summer or fall, and flowering and seed dispersal the following summer. Many of these accessions occur in Scandinavia and the Iberian Peninsula (SI Appendix, Fig. S3).

In contrast, accessions with a negative score on PC2germ that break primary dormancy with brief chilling exposure and do not enter secondary dormancy with prolonged chilling also displayed fast phenological transitions and had long after-ripening times, preventing summer germination. Dormancy release by cold would allow germination in both fall and spring, and coupled with rapid reproduction, would facilitate rapid cycling or spring annual life histories. Seeds that do not germinate in fall would overwinter in the soil under cold temperatures and would be ready to germinate the next spring (26). This seasonally flexible germination may be favored in disturbed, ruderal landscapes.

These accessions were common at midlatitudes in Europe and England (SI Appendix, Fig. S3).

The timing of germination defines the temporal environment an annual plant experiences, and interacts with later life-cycle traits to generate different life histories. Our results are consistent with previous observations of covariation in seed dormancy (measured as after-ripening requirements) and flowering-time traits across climatic gradients in Arabidopsis thaliana (12, 19, 20, 37). Trait syndromes allow local adaptation, as covariation in life-history traits, growth rate, and stress responses influence fitness (37). However, our observations of germination responses to chilling add a new dimension to this picture of multivariate life histories, showing that cold-induced secondary dormancy is correlated with flowering and senescence to construct a wide range of seasonal life-history strategies.

Genetic Architecture of Seed Responses to Chilling: A Major Role for DOG1.

To understand the genetic architecture of the germination niche, we performed GWA analyses on the germination PCs using 498 fully sequenced A. thaliana accessions and 3,483,598 single nucleotide polymorphisms (SNPs; 1001 Genomes Consortium). A region on chromosome (Chr.) 5 was the most highly associated to both the first and second PCs of germination. This region included 1 SNP significantly associated with PC1germ and 42 SNPs significantly associated with PC2germ. Several of these SNPs likely tag the same functional variant, as they are organized into only 4 linkage disequilibrium (LD) blocks spanning multiple candidate genes.

Different SNPs were most highly associated to PC1germ and PC2germ, although the closest gene to both was DOG1 (Fig. 3 and SI Appendix, Tables S2 and S3). The SNP significantly associated to PC1germ and PC2germ (Chr. 5: 18,592,365) (SI Appendix, Table S2) had previously been shown to tag the promoter region of DOG1 and to be associated with after-ripening time variation (50) (SI Appendix, Fig. S4). However, in our experiments, DSDS50 was associated with a different set of SNPs in several LD blocks, which did not include DOG1 (SI Appendix, Fig. S5 and Table S2). Using more permissive thresholds, we found that the top 1,000 SNPs

---

**Table 1. Correlations between the first 2 PCs from the germination variables PCA with phenology variables of the mother plants**

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>PC1germ r</th>
<th>PC1germ P value</th>
<th>PC2germ r</th>
<th>PC2germ P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTF</td>
<td>523</td>
<td>-0.07</td>
<td>0.09</td>
<td>0.43</td>
<td>&lt;2.2×10⁻¹⁶</td>
</tr>
<tr>
<td>DTS</td>
<td>523</td>
<td>0.18</td>
<td>5.19×10⁻⁵</td>
<td>0.50</td>
<td>&lt;2.2×10⁻¹⁶</td>
</tr>
<tr>
<td>DSDS50</td>
<td>517</td>
<td>-0.49</td>
<td>&lt;2.2×10⁻¹⁶</td>
<td>-0.43</td>
<td>&lt;2.2×10⁻¹⁶</td>
</tr>
</tbody>
</table>

Fig. 1. (A) PC scores of germination variables (base, and cold treatments: 4, 8, 11, 15, 22, and 32 d at 4 and 10 °C). Percent variance explained by each axis: PC1germ = 83%, PC2germ = 7%. (B) Mean percent germination in each treatment for the lines with the 10% highest and lowest scores on PC1germ and PC2germ. Colors indicate the amino acid sequence at the DOG1 self-binding domain (53), which we use to define haplotypes, and grey color indicates accessions for which we were not able to assign a haplotype.
associated to each trait were enriched for seed-specific gene ontologies (SI Appendix, Tables S4 and S5).

As SNPs near DOG1 were associated in both analyses, we further characterized genetic variation at DOG1. Known indel polymorphism within the self-binding domain of DOG1 (amino acid positions 13 to 16) (53) is not fully captured by published polymorphism within the self-binding domain of DOG1 (amino acid variants present). We identified the ancestral self-binding haplotypes. Haplotypes in the self-binding domain have been defined by the amino acid variants present. We identified the ancestral self-binding DOG1 E-haplotypes (ECCY) and 2 loss-of-function DOG1 D-haplotypes (D-RY, D-SY) previously reported (53), as well as 2 additional rare derived haplotypes without an amino acid deletion in the self-binding domain (ECFSY and ECFSY) found in less than 5% of individuals and restricted to Iberia (SI Appendix, Figs. S6 and S7). DOG1 haplotype identities are significantly associated with PC2germ in the D-haplotype accessions, this SNP has a similar allele frequency among the D- and E-haplotypes (SI Appendix, Table S8). Thus, this association arises not because the causative variant is absent in 1 haplotypic background, but rather because different SNPs have different effect sizes in different backgrounds. The epistatic interaction between the DOG1 E-haplotype and the associated allele in this region of Chr. 5 give additional insight into the complex gene interactions involved in the regulation of germination responses to chilling. Although not measured in our study, it is likely that DOG1 expression levels are involved in the chilling response as well (12, 53, 57, 59, 62, 63, 69–73). Other mechanisms regulating DOG1 expression might also be involved in the chilling germination responses: For example, the antisense transcript asDOG1, which negatively regulates the expression of DOG1 (74, 75).

Life-History Syndromes Resulted from Correlated Selection on Multiple Loci. Accessions with loss-of-function DOG1 D-haplotypes showed strong primary dormancy release and no secondary dormancy induced by cold, and also flowered and senesced early. In contrast, individuals with self-binding DOG1 E-haplotypes showed secondary dormancy induced by prolonged cold, and flowered and senesced later on average (Fig. 2). Besides PC2germ, DOG1 haplotypes explained variation in DTF and DTS (SI Appendix, Fig. S10 and Table S6A). Similar D- and E-haplotype germination patterns were observed for geographically restricted samples within Spain and Sweden (SI Appendix, Figs. S11 and S12 and Tables S6 B and C). Pleiotropic effects of DOG1 on flowering and germination mediated by miR156 and miR172 may be a mechanism behind these trait correlations (60, 76).

Our results also showed that GWAs for DTF were polygenic (39 SNPs organized in 11 LD blocks) (Fig. 3 and SI Appendix, Table S2) and SNPs tagging DOG1 were not significantly associated with DTF. DTS showed no associated SNPs above a permutation threshold for significance (SI Appendix, Fig. S5 and Table S2). However, the 5 SNPs associated to PC2germ tagging DOG1 have been previously associated to different flowering-time phenotypes under various conditions (Fig. 3 and SI Appendix, Fig. S13 and Table S2) (42, 43, 64, 77). The lack of association of DTF with DOG1 in our study could partly be due to the accessions sampled, as we performed an additional GWA analysis using only our accessions with subsampled 1001 Genomes Consortium flowering phenotypes (43) and did not find evidence of association of DTF with SNPs tagging DOG1 (SI Appendix, Table S8).

Fig. 2. Correlations between PC2germ and DTF (A) and DTS (B). Colors indicate the amino acid sequence at the self-binding domain for different DOG1 haplotypes.
in Iberia with hot, dry summers and are associated with relict and Spanish accessions (Fig. 4 and SI Appendix, Figs. S6, S7, and S15). ECSY and EFSY likely originated in the Iberian refugium (46, 79–81), and may remain endemic in that region because they require hot, dry Mediterranean summers to persist.

D-SY and D-RY haplotypes are sister to one another, arising from an ECCY ancestor (Fig. 4A). Pairwise diversity within all D-clade individuals suggests this split occurred about 365 Kya (ranging from 281 to 450 Kya using differing assumptions about generation time) (Materials and Methods) during a period of climate cooling following an interglacial period (Fig. 4B and SI Appendix, Table S9), possibly during contraction into different refugia. However, we do not observe further diversification within the extant D-haplotypes until much later, potentially due to cycles of expansion and contraction out of refugia followed by local extinction. The extant, predominately Asian and east Eurasia D-SY clade diversified 115 Kya (88 to 141 Kya with varying generation times) during a cool period of the last interglacial period (Fig. 4B and SI Appendix, Table S9). Pairwise diversity among extant D-RY haplotypes suggests this west Eurasian and North American clade expanded recently 16 Kya (ranging 12 to 19 Kya), during rapid warming after the last glacial maximum (Fig. 4B and SI Appendix, Table S9), but the gene tree suggests an earlier origin (Fig. 4A).

The extant distribution of D-RY in Iberia and Western Europe (SI Appendix, Figs. S6 and S7) suggests that this haplotype may have arisen and spread from the Iberian refugium. These observations suggest that changing Pleistocene climate may have favored the rise and spread of the D-haplotypes, possibly through selection for associated traits promoting spring annual or rapid cycling life histories. The loss of cold-induced secondary dormancy in D-haplotypes, correlated with shorter life cycles, may have provided life-history flexibility facilitating persistence in changing climates and the invasion of new habitats. The ability to germinate in spring is advantageous in montane regions (17), such as Central Asia, and rapid cycling life histories may be favored in disturbed habitats with ample summer rainfall. Both D-SY and D-RY currently occupy climates with wetter summers than the E-haplotypes (SI Appendix, Fig. S15B). The wide Eurasian distribution of D-SY haplotypes suggests that their expanded germination niche may have facilitated postglacial expansion and spread of “nonrelict” genotypes out of central Asia into Europe and China (47, 82). The recent expansion of the D-RY clade is consistent with invasion of new ruderal habitats made available by the spread of agriculture through Europe (43, 82, 83). Our results show that derived loss-of-function D-haplotypes are more common in midlatitudes and England, and could have resulted from the nonrelict east–west expansion following human-mediated habitat disturbance. Relict populations in the Iberian Peninsula and Scandinavia might have preserved the ancestral DOG1 E-haplotype that promotes a winter annual cycle, advantageous in more undisturbed habitats (81, 82). Moreover, the D-RY haplotype of DOG1 is enriched in the invaded range of North America (SI Appendix, Fig. S6), where A. thaliana has arrived as a human commensal in the last 300 y (84). Although this could be due to a founder effect, phenological traits associated to this haplotype might also have played a role in invading disturbed habitats suitable for rapid cycling.

We caution that selection, geographic sampling, and demography could affect these estimates of age. But, with the exception of a more recent D-RY origin, these clade ages match the relative branching pattern on the gene tree. The origin of DOG1 haplotype variants suggests that the germination niche of A. thaliana has been shaped by climate cycling throughout the Pleistocene, as the species range repeatedly expanded out of and contracted into glacial refugia (46, 79, 80) and the changing climate and physical environment filtered which individuals were able to persist (85). Novel germination strategies, measured via

**Appendix, Fig. S14**. Our sampling includes mostly European accessions while the 1001 Genomes phenotypes also include Asian accessions (78). Moreover, the lack of association could also be due to the different vernalization and growth conditions used in each experiment.

We also tested for epistatic effects of DOG1 haplotypes on flowering and senescence and found that SNPs significantly associated with DTF for the E-haplotype were located at different positions from those identified with the whole dataset, and the most significantly associated SNP on Chr. 1 was found closest to AT1G33440, a major facilitator superfamily protein (Fig. 3 and SI Appendix, Fig. S9B and Tables S2, S7, and S8).

Taken together, our results support that correlations between seed chilling responses and reproductive timing may result in winter and summer annual life histories. DOG1 variation is critical for cold induction of secondary dormancy that determines seasonal germination timing, and may also have pleiotropic effects on flowering time. However, a number of other loci are important to flowering-time variation, suggesting that the multivariate life-history syndromes we observe may be shaped by correlated selection at multiple loci in addition to DOG1.

**Arabidopsis Germination Niche Today Is Explained by Current and Past Climate**. To understand the evolution of DOG1, we constructed a gene tree of the first exon of DOG1, which includes the self-binding domain (Fig. 4A). The earliest DOG1 diversification represented in extant individuals is that of the ancestral and widely distributed ECCY haplotype, associated with traits that promote a winter annual life history. The tree topology showed a deep divergence within the ancestral ECCY clade, especially for relict, African, and Swedish accessions and a few anciently diverged haplotypes in China (Fig. 4A and SI Appendix, Figs. S6 and S7, and Table S9).

The rare ECSY haplotype arose within the ECCY clade, followed by EFYSY, during periods of cooler climate after the last interglacial period (Fig. 4A and SI Appendix, Table S9); both of these derived haplotypes are geographically restricted to regions...
SNPs Associated with Chilling Responses and DOG1 Functional Haplotypes Show a Signature of Climate Adaptation. To understand the climatic associations between genetic variants associated to various phenotypes, we used Gradient Forest, a tree-based machine-learning regression approach, to describe nonlinear turnover functions of allele frequencies along environmental gradients (40, 86, 87). We performed Gradient Forest regression analyses between environmental gradients and LD-pruned index SNPs from the 1,000 most highly associated SNPs from our GWA results for PC1_germ (23 SNPs), PC2_germ (9 SNPs), and the DOG1 haplotypes. After accounting for spatial autocorrelation by including Moran’s eigenvector map (MEM) variables, mean temperature of the wettest quarter (Bio8) explained the highest amount of turnover in allele frequency of SNPs associated to PC2_germ and was among the top 3 predictors for PC1_germ as well (Fig. 5 and SI Appendix, Table S10). Additionally, altitude, isothermality (Bio3), and temperature mean diurnal range (Bio2) predicted allele frequencies of both germination PCs (SI Appendix, Fig. S16 and Table S10). These environmental gradients structure turnover in index SNPs more than a set of 150 random SNPs (Fig. 5 and SI Appendix, Fig. S16), evidence of local adaptation to climate at these loci. Given that the E- and D-haplotypes are very old, our data suggest that ecotypes carrying both of these haplotypes could have arrived at locations throughout the range, but that environmental filtering and selection likely give rise to the geographic patterns that we see today. As a predominately selfing annual plant, A. thaliana populations are structured across the landscape, giving rise to geographic differences in ecologically relevant alleles (88); however, our results support local adaptation to climate, as allele turnover from index SNPs is more strongly structured by temperature and precipitation gradients than alleles from random SNPs.

The response-to-cold PC2_germ cumulative importance function for mean temperature of the wettest quarter (Bio8) showed a threshold turnover in allele frequencies around 14 to 15 °C (Fig. 5), and alleles from SNPs associated to DOG1 were among the ones with the highest importance values for this function (SI Appendix, Table S11). Accessions with the reference allele, with high dormancy release under chilling and no secondary dormancy, are found in these southern regions, while accessions with the alternative DOG1 allele have strong secondary dormancy. This temperature threshold divides the winter growing season in southern regions (<14 to 15 °C) from a warm wet summer (>14 to 15 °C) growing season in northern latitudes, thus the shape of the allele turnover function suggests that germination traits might have been selected by these climatic gradients. Environmental predictors’ importance varied at explaining the distribution of different DOG1 haplotypes (SI Appendix, Figs. S17 and Tables S12 and S13). Allele distributions of DOG1 might have resulted from its effect on germination responses to chilling and on other life-history traits as well. For example, 14 °C acts both as a temperature threshold for the induction of strong maternal effects on seed dormancy (11, 22) and for vernalization disruption in flowering time (89).

Conclusions

The seasonal germination niche shapes phenology and life history (1, 2, 5–8), and may be an essential component of adaptation to climate (11–13). Natural variation in dormancy and germination responses to seasonal environmental cues is often observed (21), and the genetic basis of variation in primary dormancy and after-ripening has been well studied in A. thaliana. However, much less is known about the genetic basis of natural variation in seed responses to seasonal chilling. Our data reveal a new axis of natural variation in germination and dormancy responses to cold that may drive the expression of winter annual vs. spring annual life histories. Moreover, this variation in response
Materials and Methods

Seed Bulking. Seeds from 559 A. thaliana fully sequenced accessions from the 1001 Genomes Project (ABRC stocks) were stratified at 0.15% agar at 4 °C for 7 d. Seeds were sown in soil and allowed to germinate for 10 d; seedlings were vernalized for 6 wk at 4 °C with a 12/12-h photoperiod and were grown at 14 °C with a 12/12-h photoperiod in a walk-in growth chamber (Conviron E7/2 Controlled Environment Chamber). Two individuals of each accession were grown in the same chamber (a total of 1,118 plants), planted 2 wk apart in 2 temporal replicate blocks. Plants were watered twice a week with nutrient water until they showed 50 to 60% ripe siliques and seeds were individually harvested from each plant when 70 to 80% of the siliques were ripe. Fresh harvested seeds from each plant were immediately used for 2 different experiments: 1) Cold stratification germination experiments and 2) dry seed storage germination experiments to test for seed after-ripening times. Phenology variables were recorded on the maternal plants including DTF and DTS. DTF

Dry Seed Storage Germination Experiments. Additional germination experiments were done to test for seed after-ripening. Fresh seeds were stored under dry laboratory conditions and were tested for total germination every 6 wk until 75% germination was recorded in 2 consecutive tests, which we considered as after-ripened seeds. Some accessions were after-ripened after 6 wk while some others showed little evidence of after-ripening even after dry storage in the laboratory for up to 788 d. Germination induction conditions were the same as for the cold stratification experiment. Seed after-ripening was assessed by calculating the number of days to 50% germination by fitting a polynomial regression to the dry-stored seed germination data for each accession (DSDS50) (92).

Data Analyses.

Germination strategies (PGerm). We excluded accessions that had misleading location data due to being misidentified according to ref. 93. We ran a PCA on mean percent germination from the 2 plantings for the base and cold treatment logit-transformed germination data (prcomp function in R). PCA scores and loadings were rescaled to the axes 2 SD.

Germination responses and phenological associations. We tested for correlations between the first 2 PCs of germination variables, after-ripening (DSDS50), and phenology variables of the mother plants (i.e., DTF and DTS; Pearson's product-moment correlation, cor.test function in R).

GWAs on germination and phenology traits. GWAs analyses were performed on the scores of the first 2 germination PCs, base germination, DSDS50, DTF, and DTS to identify the associated genetic variants using a linear mixed model (LMM) corrected for population structure, implemented in GEMMA (94). Genotypes from sampled individuals (n = 498) were obtained from the 1001 Genomes Project (3,483,598 SNPs, minor allele frequency [MAF] 0.01, imputed genotypes) (43). Permutation-based thresholds were calculated for each trait by running a GWA 100 times with phenotypes permuted over genotypes and getting the average P-value of the top fifth quantile from each analysis. We annotated SNPs to the closest TAIR10 gene (distanceToNearest function in GenomicRanges) with SNPs > 1-kb distance from a gene annotated as intergenic regions. We also used TAGGIT to annotate the 1,000 most highly associated SNPs with respect to seed-specific gene ontology categories (95). GWAs results were used to group SNPs into LD blocks. SNPs that were associated at a significance threshold of P ≤ 0.0001 and that were not included in other LD blocks were selected as index SNPs. LD blocks around index SNPs were defined by all other associated SNPs at a significant threshold of P < 0.01 that were in LD with the index SNP (r² = 0.50) within a physical distance of 250 kb (default settings from clump command, PLINK).

DOG1 functional haplotypes dated phylogeny. A 3-bp indel known to affect DOG1 function (53) is not genotyped in the available 1001 Genomes Project vcf files. In order to characterize our accessions at this known functional site, we mapped whole-genome resequencing reads from 1,135 A. thaliana accessions from the 1001 Genomes Consortium (43), 64 African A. thaliana accessions (45), and 118 Chinese A. thaliana accessions (46) to the primary cDNA of DOG1 (ATSG45830.1) using bwa-mem v0.7.12-r1039 (96). We retained mapped reads and their pair, and assembled these using phrap v1.090518 (http://www.phrap.org). We aligned assemblies to exons of DOG1 using MAFFT v7.27 (97), and extracted regions from the haplotype at amino acid positions 13 to 16, as defined by ref. 53. We filtered minor alleles with fewer than 2 individuals, which eliminated 3 rare haplotypes: C-CY, D-CY, and EYSY. Due to alignment and assembly issues, we recovered sequence for 972 individuals. DOG1 local reassemblies of the 393 bp of exon 1 were pruned for redundant sequences, resulting in 74 unique individual sequences (98, 99) used to generate a Bayesian likelihood phylogeny in BEAST 2.5.2 with a strict clock and constant-sized coalescent prior, using an Arabidopsis lyrata and Capsella rubella DOG1 sequence as outgroups. We ran BEAST for 10 million generations, confirmed Markov chain Monte Carlo convergence by eye, discarded 10% of trees as burn-in, and summarized trees with a maximum clade credibility tree. Attempts to run BEAST including
59. K. Nakabayashi et al., The time required for dormancy release in Arabidopsis is determined by DELAY OF GERMINATION1 protein levels in freshly harvested seeds. Plant Cell 28, 2822–2838 (2016).
63. G. C. Chiang et al., DOG1 expression is predicted by the seed-maturation environment and contributes to geographical variation in Arabidopsis thaliana. Mol. Ecol. 20, 3336–3349 (2011).