

Supporting Online Material for Adaptation to Climate Across the *Arabidopsis thaliana* Genome

Angela M. Hancock, Benjamin Brachi, Nathalie Faure, Matthew W. Horton, Lucien B. Jarymowycz, F. Gianluca Sperone, Chris Toomajian, Fabrice Roux, Joy Bergelson*

*To whom correspondence should be addressed. E-mail: jbergels@uchicago.edu

Published 7 October 2011, *Science* **334**, 83 (2011)
DOI: 10.1126/science.1209244

This PDF file includes:

Materials and Methods
Figs. S1 to S11
Tables S1 and S2
References

Other Supporting Online Material for this manuscript includes the following: (available at <http://bergelson.uchicago.edu/regmap-data/climate-genome-scan/>)

Climate data for the 948 accessions used in the analyses, result files for the correlation analyses, and a browser that allows for viewing the results in their genomic context.

MATERIALS AND METHODS

Arabidopsis thaliana accessions and genotype data

We chose a set of accessions collected throughout the native Eurasian range of *A. thaliana* from the complete RegMap Project set (<http://regmap.uchicago.edu>) to be used in this analysis (Figure S5). We excluded accessions that were likely to be contaminants (23), leaving us with a set of 948 accessions.

Climate data

We obtained information about climate for each accession included in the analysis. Data were collected for 40 different climate variables. Details about each dataset are shown in Table S1. In summary, data for 23 bioclimatic variables that summarize information about extremes and variability in temperature and precipitation were obtained from the WorldClim project (2if4). We obtained relative humidity data from the NCEP-NCAR climate reanalysis project (25), and data for growing season lengths was obtained from the FAO GeoNetwork

(<http://www.fao.org/geonetwork/srv/en/main.home>). In addition, we used individual station data to calculate two additional variables related to growing season length: the number of consecutive cold and frost-free days. The lengths of the cold and frost-free periods were calculated using data collected from individual stations. Specifically, the numbers of consecutive days when the temperature was below 4 degrees Celsius (the vernalization temperature for *A. thaliana*) and the consecutive number of days when the temperature was above 0 degrees Celsius were calculated for each of 4189 stations across the Northern hemisphere for each of the years from 2005 to 2009. These years were selected because there were very few missing data points for them compared to the previous years for which data were available. For each station, we calculated an average over the 5-year period, and then we interpolated the results using the kriging function in ArcGIS to arrive at an estimate for each *A. thaliana* accession.

Next, we selected a representative subset of the total variables to include in the climate correlation analyses. The total set of 41 variables (latitude and 40 climate variables) was pruned based on the pairwise Pearson correlations of the variables (Figure S10) so that no two variables had an r^2 greater than 0.8. In cases where variables were strongly correlated with one another, the variable with the most obvious link to the ecology *A. thaliana* was selected. The variables used in the analyses were: aridity, number of consecutive cold days (below 4 degrees C), number of consecutive frost-free days, daylength in the spring, growing season length, maximum temperature in the warmest month, minimum temperature in the coldest month, temperature seasonality, photosynthetically active radiation, precipitation in the wettest month, precipitation in the driest month, precipitation seasonality, and relative humidity in the spring. Figures showing the global distribution of each variable that was used for the climate correlation analyses are available at <http://bergelson.uchicago.edu/regmap-data/climate-genome-scan/>.

Calculating correlations with climate

Across the distribution of a species, genetic variants or phenotypes may be strongly correlated with climate due simply to demographic history. Therefore, we used a

method that allowed us to control for population history when we calculated correlations with climate. A second challenge for this type of analysis is that linear model methods are heavily influenced by outliers, resulting in very strong correlations that are driven by one or a few observations. In fact, initial testing showed that linear model methods tended to often identify variants that were driven mainly or exclusively by climatic outliers.

Previous approaches to genome-wide association mapping using a similar scheme and similar statistical methods have dealt with this problem by removing low frequency variants (minor allele frequency < 10%) from the analysis (e.g., (7)). Here, we elected to include these variants in the analysis, but to use a non-parametric method to assess the strength of correlation between each environmental variable and genetic variant (or phenotype) while controlling for population structure. Specifically, we used a partial Mantel test (26, 27) to calculate the Spearman correlation between a given SNP and environmental variable while controlling for population structure using a kinship matrix based on genome-wide genetic variation data. The dependent variable in the model was either a distance matrix of the phenotypes (for analyses that lead to figs S1-S4) or a distance matrix of an individual genetic variant (for the genome-wide scan to identify variants that were strongly correlated with climate). The predictor variable was the pairwise distance matrix of the climate variable and the covariate was a distance matrix based on the pairwise kinship matrix based on the total set of SNPs. Partial Mantel tests were conducted using the ecodist package (28) in R (29) as the method to assess evidence of a correlation.

Because relationships among accessions in *A. thaliana* appear to follow a strong isolation by distance model rather than a model in which there are many discrete populations, we treated each individual separately in the analysis rather than grouping them into populations. This approach has the added benefit that the numbers of samples from any given region are taken into account in the model. The genetic distance matrix was computed based on the variance/covariance (similarity) matrix from the emma.kinship function in the R package EMMA (10).

In addition to calculating correlations with climate using the partial Mantel test, for comparison we also used a non-parametric Wilcoxon rank sum test to compare the distributions of climate between the two alleles of each SNP. This method does not control for demographic history. It should be noted that use of a kinship matrix to control for population structure results in false negatives when the climate variable itself is correlated with kinship. However, the difference between our enrichment results for the partial Mantel tests compared to the Wilcoxon rank sum tests (Fig. 1 and Fig. S6) indicate that we gain significant power by controlling for similarity among accessions.

Phenotype data used in the climate correlation analysis

Phenotypes used in this analysis were collected as part of a previously published project (7). Phenotypic variation represents genetic variation among accessions since, for each phenotype classification, all accessions were planted together in the same growth chamber or in a common garden and in replicates. We calculated correlations between the 13 climate variables and the 107 phenotypes included in Atwell et al. (7) using a partial Mantel test as described above.

Assessing significance for climate correlations

The classic method for assessing significance for the results of partial Mantel tests is to use permutations of the dependent variable (27). We could not use this approach for two reasons. First, it was computationally not practical to run enough permutations to obtain a p-value with precision necessary to find p-values low enough to equate to the level of genome-wide significance at the 0.05 level using a Bonferroni correction (i.e. 2.3×10^{-7}).

Second, with matrices of only 948x948 cells, we would not have enough information to run such a large number of independent permutations. Instead, we ran 1000 permutations for each SNP and each variable and found a very large excess of low p-values compared to that expected if p-values were uniformly distributed (as expected under the null). For the lowest p-value possible with 1000 permutations (0.001), we found between 33,388 and 78,795 SNPs compared to 214 expected. The abundance of very low p-values from the permutations suggests that including the kinship matrix may not completely control for the effects of population history even when permutations are used to assess significance. This is not unexpected and is analogous to situations observed for related methods that use a kinship matrix to control for population history (30).

Some of the observed excess of low p-values may be due to positive (or negative) selection and linkage disequilibrium (LD) between selected sites and surrounding neutral variation. We asked whether removing linked variation reduces the observed excesses of low p-values. To do this, we used the prune function in PLINK (31) to select representative sets of SNPs based on pairwise r^2 (with a window size of 1000 SNPs). Then, for a range of r^2 cutoffs, we calculated the proportion of SNPs with permutation p-values less than 0.001. These results show that the proportion of low p-values tends to decrease substantially as more SNPs are filtered from the dataset (Fig. S11), revealing an over-representation of SNPs in fairly strong LD among SNPs with permutation p-values less than 0.001. Given that we observed other evidence of sweeps (as described in the main text), this could imply a genetic draft scenario, in which a fairly large number of sweeps have carried linked genetic variation to high frequency (16), or widespread background selection in the genome of *A. thaliana* (15).

Given the uncertainty of the permutation p-values, we used the following alternative methods to ask whether the tail of the distributions of climate correlations were likely to be enriched for true signals of selection.

Comparison of climate correlation signals across annotation categories of variants

To determine whether SNPs likely to be functional were over-represented in the tail of the distribution with climate, we asked whether the proportions of variants in three categories (intergenic, synonymous (S) and nonsynonymous (NS)) relative to the proportions of the same variants overall differed from unity. First, to ask whether these categories were over-represented with the climate variables as a group, we calculated enrichment in the 1% tail of a composite variable, the ranked minimum rank across all climate variables. More specifically, we ranked the results from the individual climate variables to create a statistic that is sometimes referred to as an ‘empirical p-value’. Then, we found the minimum across the set of 13 climate variables and created a new rank statistic based on this vector of minima. This variable allowed us to test climate overall for evidence of enrichment in a single test. To assess significance for observed

enrichments, we conducted 10,000 permutations and compared the observed enrichment to that found in the permuted datasets. For each permutation, a new set of SNPs was chosen by shifting the locations of the SNPs in the functional category under test by a randomly chosen number between 1 and one less than the total number of genotyped SNPs. This scheme results in permutation sets that resemble the original set with respect to linkage disequilibrium. We used the same methodology described above for testing for enrichment for individual climate variables, a minimum rank statistic across variables from the Wilcoxon rank sum test results as well as for the tests of individual variables. Finally, we tested for enrichments in the proportion of NS relative to S SNPs and assessed significance using the permutation methodology described above.

Assessing whether alleles at strongly correlated variants predicts fitness in a particular environment

To ask whether the set of alleles with the strongest climate correlations could predict variation in fitness for a set of accessions planted in the same environment, we used fitness data collected for a set of 179 accessions originating from a wide geographic range but grown together in a common garden in Lille, France. The experimental design and growth conditions of the common garden experiment have been fully described elsewhere and correspond to the 2008-2009 experiment (9). Briefly, the experiment was organized in a three blocks design, each block being an independent randomization of two replicates per natural accession. For plants of two randomly chosen blocks, fitness was approximated by the total silique length (32), which is strongly correlated with seed count (21). The number of accessions used in our analysis after removing 15 likely contaminants (23) and 17 from outside of Eurasia was 147.

First, we selected the set of 254 SNPs that were in the 0.01% tail of the 13 climate correlation distributions. Next, we pruned SNPs based on LD using PLINK ((31)). In this step, we selected SNPs with r^2 less than 50% from each 100 kb region where there was a climate-correlated SNP, and we preferentially chose the SNP with the strongest correlation. Then, we determined which variant was likely to be favorable in Lille by identifying the set of accessions that share a similar climate from the complete set of 948 accessions, i.e., those within 0.1 standard deviations from the Lille climate based on the distribution of climate for all 948 accessions. Next, we asked which allele had a higher frequency in this set relative to the frequencies outside of Lille. Finally, we counted the number of these alleles (i.e., those expected to be favorable in Lille) for each of the 147 accessions and asked whether this count of ‘favorable’ alleles predicted relative fitness (approximated by total silique length) among the accessions. We assessed evidence of correlation between the number of favorable alleles and fitness using the `cor.test()` function in R (29). To assess significance for the correlation coefficient, we compared the result from the actual data to a null distribution of coefficients from 1000 permutation sets. For each permutation, a new set of SNPs was chosen by shifting the positions of the actual climate correlated SNPs by a randomly chosen number between one and one less than the number of genotyped SNPs. For each of these sets of SNPs, we used the same analysis methodology as described above to prune based on linkage disequilibrium to determine which allele should be favorable in Lille.

The permutation scheme should correct for variation in fitness in Lille due to other factors related to geographic or genetic distance from this site so that the p-values

from this analysis should be conservative. However, the correlation coefficient may be overestimated when distance is not included in the model. Therefore, we also conducted the analyses using models that include either genetic and/or geographic distance from Lille (based on all 214,435 SNPs). We modeled the effect of the number of favorable alleles on fitness while controlling for either geographic or genetic distance from Lille. We found that the coefficient drops only slightly to 0.3516 ($p=0.009$) and 0.3442 ($p=0.005$) from 0.4202 ($p=0.008$) when geographic and genetic distance are included as covariates in the model, relative to a linear model with no covariates. Geographic distance was computed simply by calculating the Euclidean distance and genetic distance was approximated by the vector of genetic distances from the accession that was collected at the closest point to Lille from a distance matrix calculated using a kinship matrix produced in the R package EMMA (33).

Clarifying which biological functions were important for adaptation to the environment

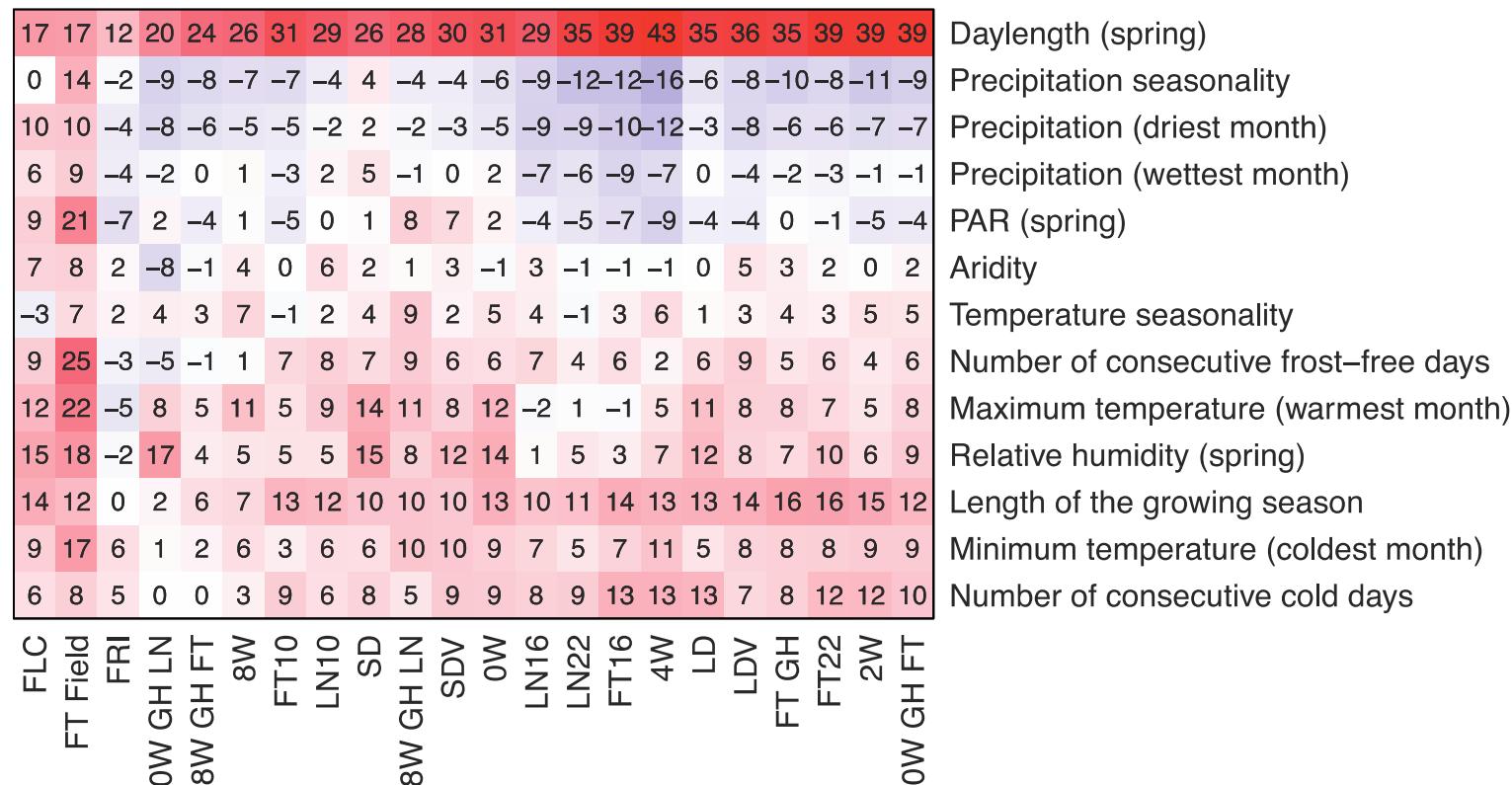
To determine which biological processes were important during adaptation to climate, we tested for an over-representation of SNPs from each of 732 Gene Ontology Biological Processes from the GOslim set (34) (obtained from (ftp://ftp.arabidopsis.org/home/tair/Ontologies/Gene_Ontology/) in the tails of the climate correlation distributions. Specifically, we asked whether each biological process was enriched in the 1% tail of the distribution of correlation coefficients for each climate variable relative to other genic SNPs. 10,000 permutations were run to assess significance using the same methodology described for annotation categories.

Assessing the distributions of geographic extents of SNPs

We calculated the geographic extent of each SNP using by finding its complex hull using the spatial analysis package for ArcGIS 10.1. Then we ranked SNPs based on where they fell in the overall distributions of complex hulls. For plots showing the distribution complex hulls in the 1% tail of the climate correlation distributions we clumped SNPs using the clump function in PLINK to remove redundancy in the data. The parameters used for clumping were the same as in the fitness analysis.

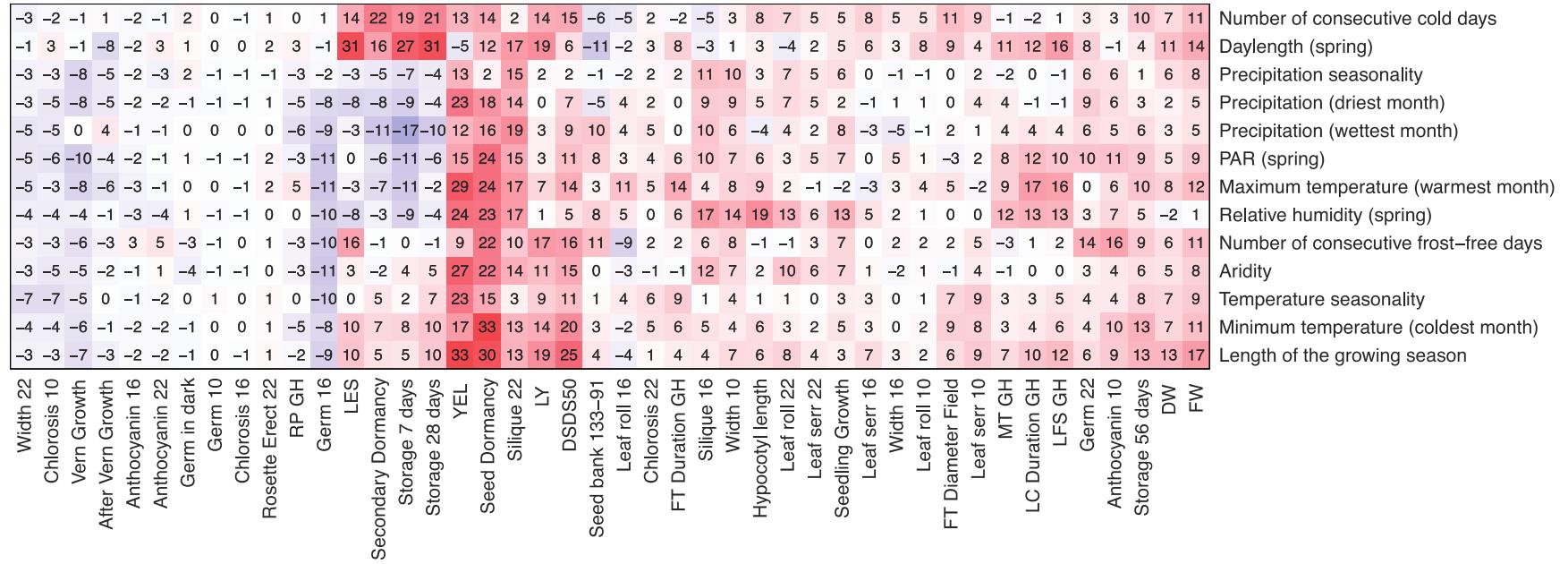
SUPPLEMENTARY FIGURES

Fig. S1.



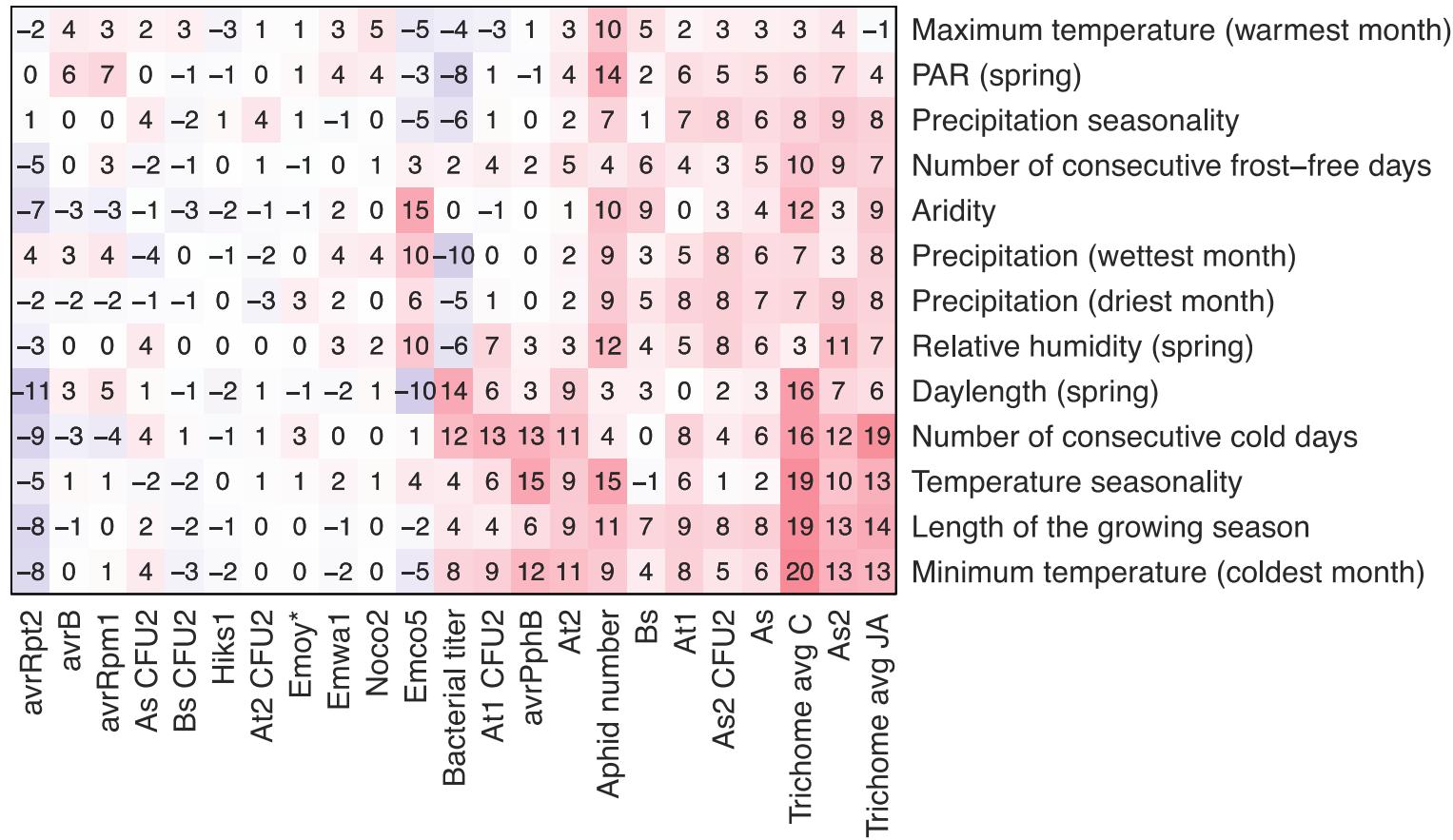
Correlation matrix of partial mantel correlations between climate variables and flowering time phenotypes, where kinship is included as a covariate in this model. Numbers shown in the heatmap are Spearman partial correlation coefficients multiplied by 100. Detailed definitions for each phenotype can be accessed at <https://cynin.gmi.oeaw.ac.at/home/resources/atpolydb>.

Fig. S2



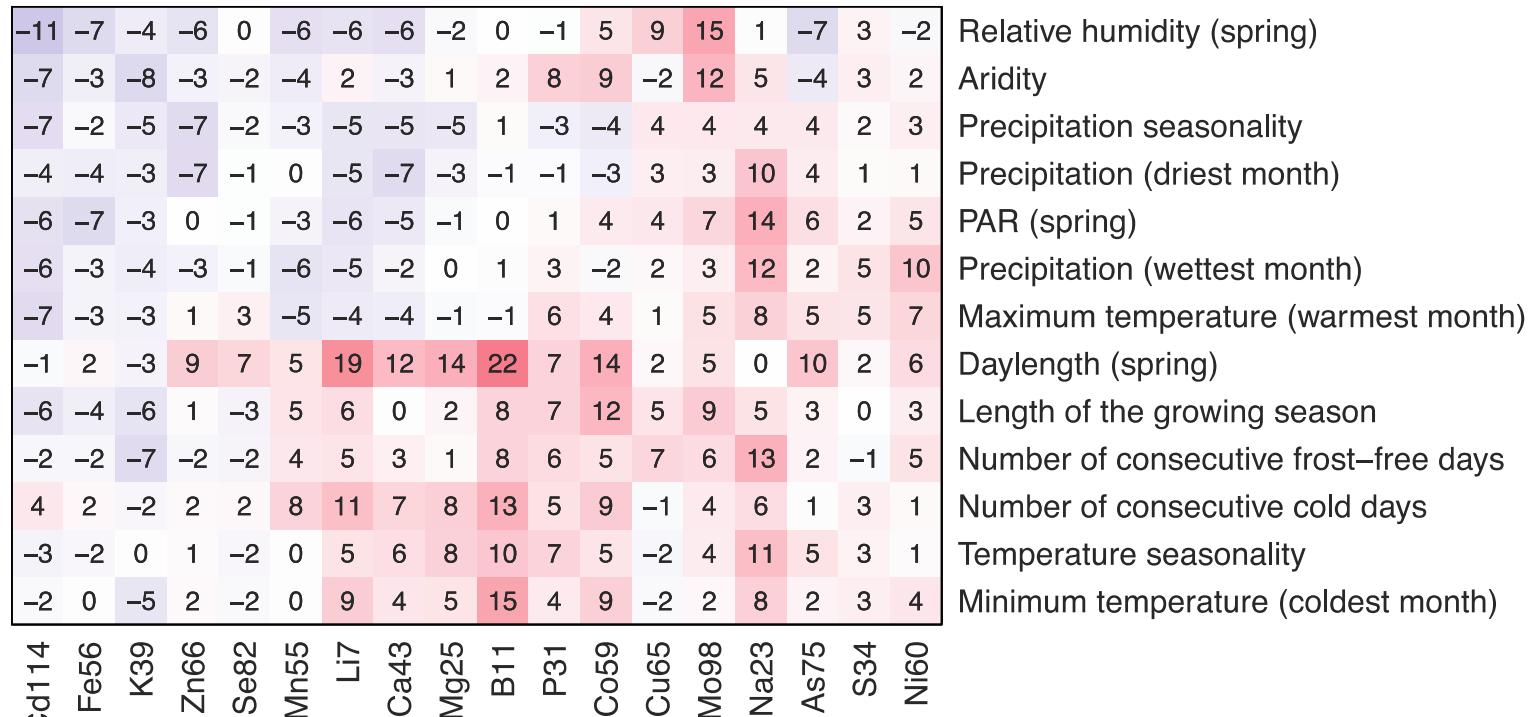
Correlation matrix of partial mantel correlations between climate variables and developmental phenotypes, where kinship is included as a covariate in this model. Numbers shown in the heatmap are Spearman partial correlation coefficients multiplied by 100. Detailed definitions for each phenotype can be accessed at <https://cynin.gmi.oeaw.ac.at/home/resources/atpolydb>.

Fig. S3



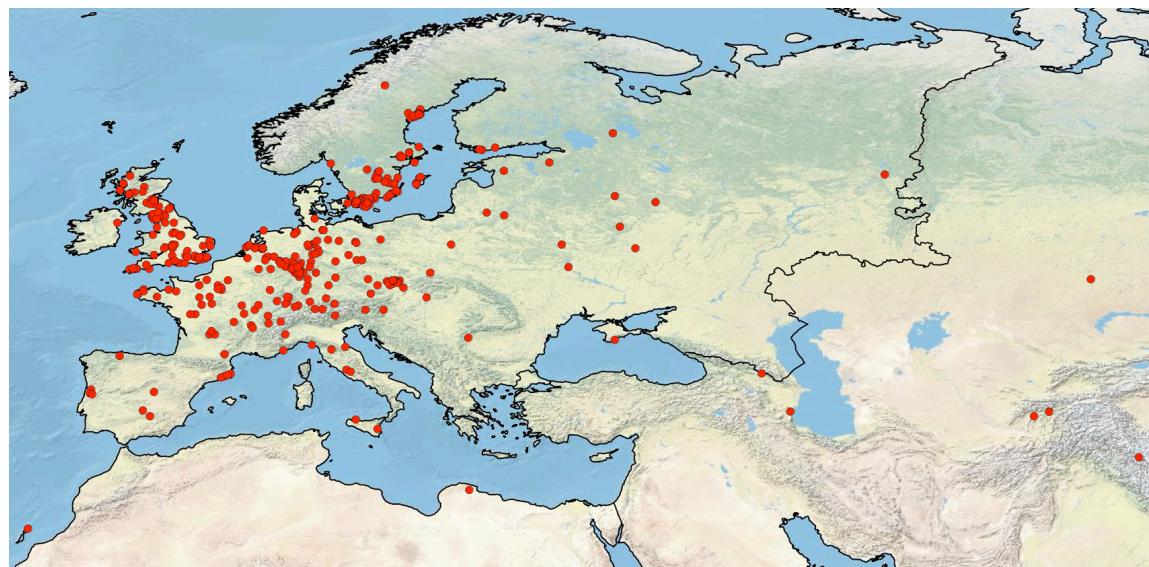
Correlation matrix of partial mantel correlations between climate variables and defense phenotypes, where kinship is included as a covariate in this model. Numbers shown in the heatmap are Spearman partial correlation coefficients multiplied by 100. Detailed definitions for each phenotype can be accessed at <https://cynin.gmi.oeaw.ac.at/home/resources/atpolydb>.

Fig. S4



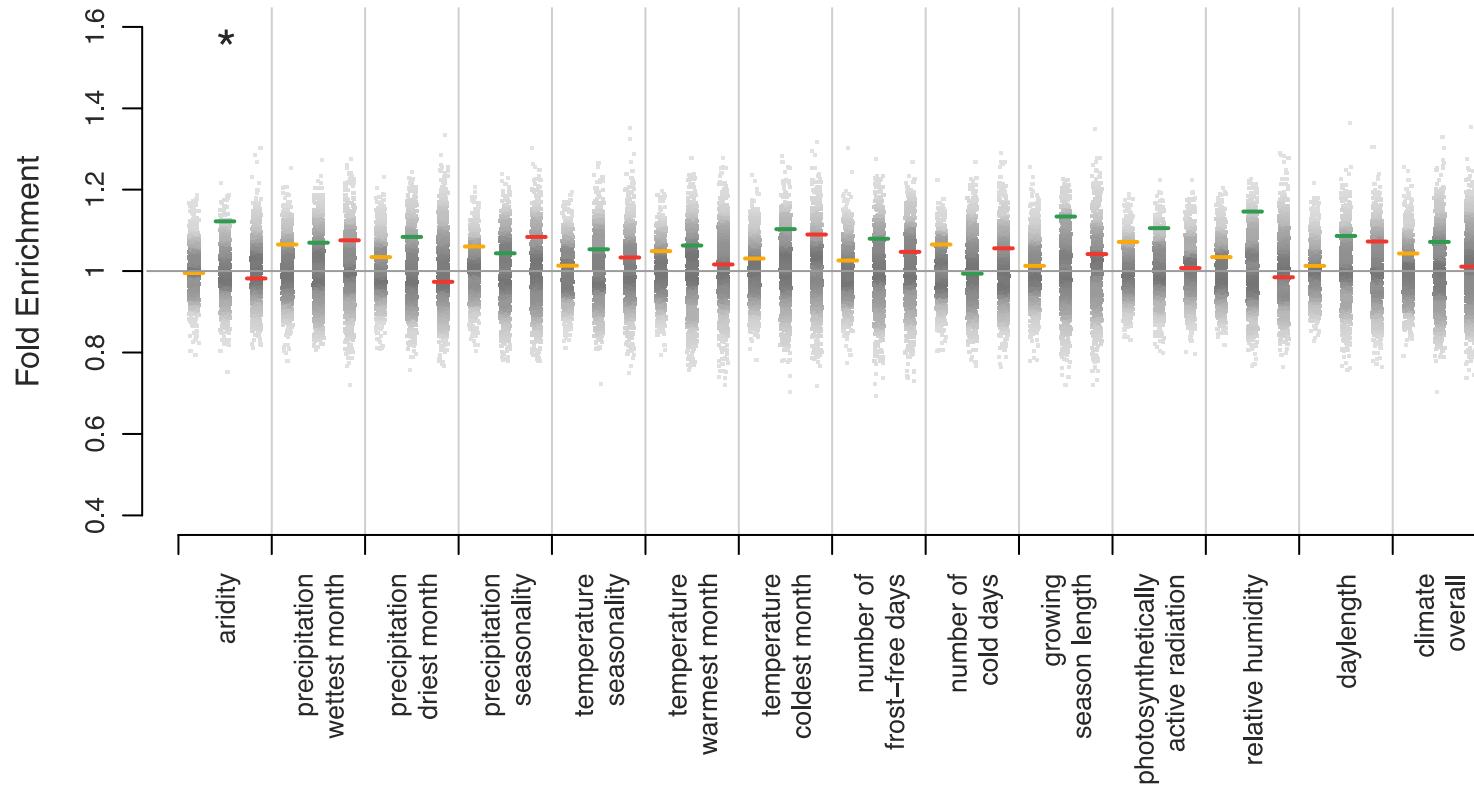
Correlation matrix of partial mantel correlations between climate variables and ionomic phenotypes. Kinship is included as a covariate in this model. Numbers shown in the heatmap are Spearman partial correlation coefficients multiplied by 100. Detailed definitions for each phenotype can be accessed at <https://cynin.gmi.oeaw.ac.at/home/resources/atpolydb>.

Fig. S5



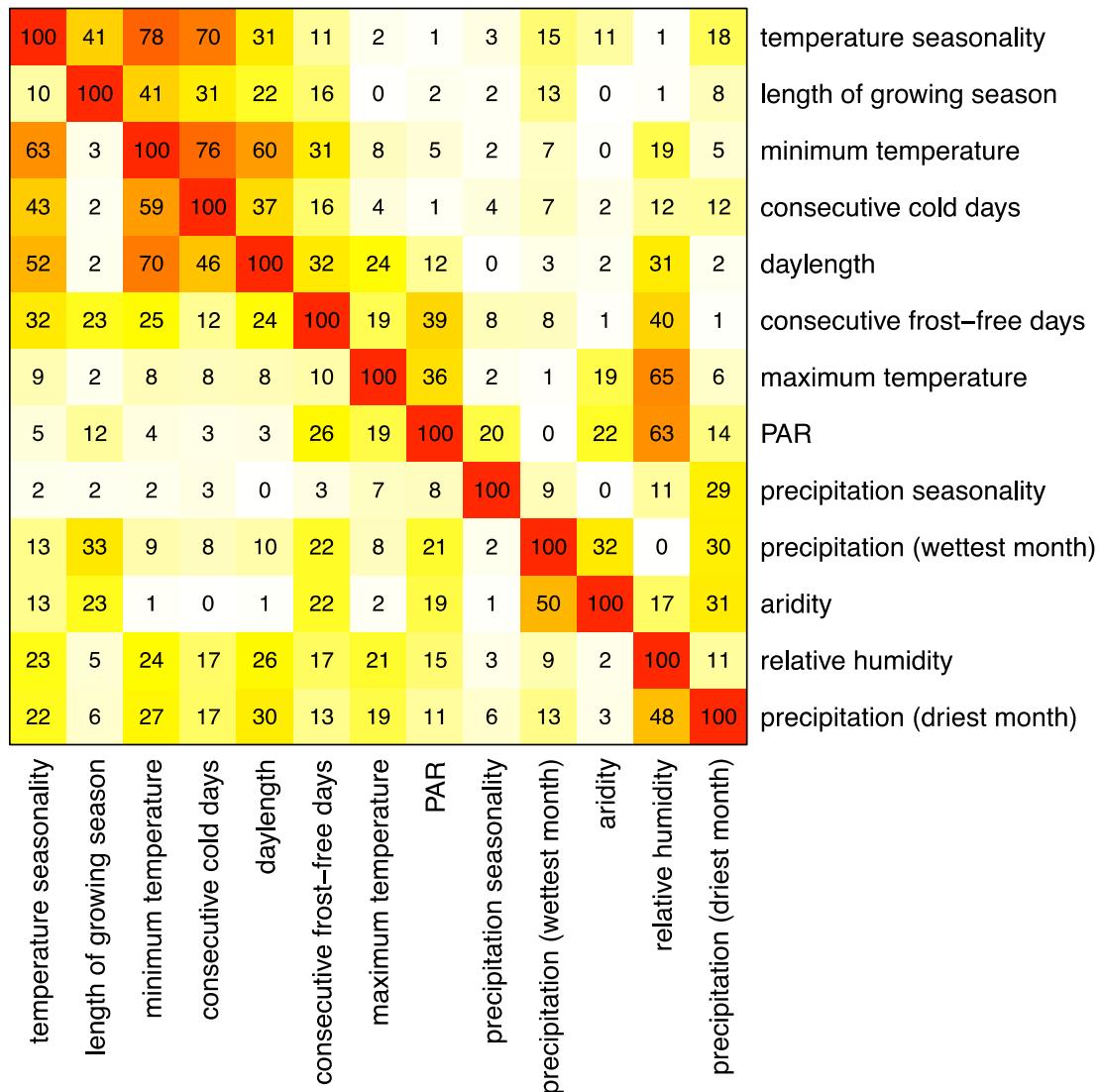
Total set of 948 accessions used in this analysis

Fig. S6



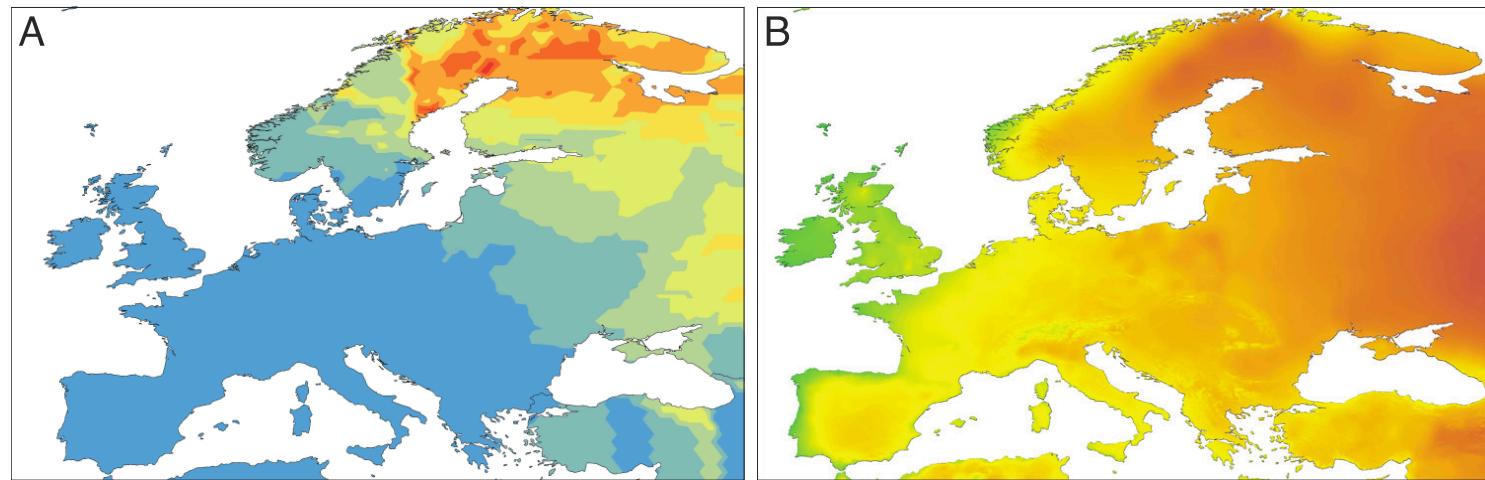
Enrichment of 3 SNP classes with the Wilcoxon signed rank test scan, which does not control for population history. The figure shows enrichment of amino acid changing SNPs (red), synonymous SNPs (green), and intergenic SNPs (yellow) in the 1% tail of the distributions for each individual climate variable as well as for climate overall (using a rank statistic based on the minimum rank across climate variables). Enrichments shown are relative to the proportion of each class of SNPs in the genome overall. Gray dots show the distribution of results of 1000 permutations. The gray line shows the expected enrichment under the null hypothesis of no enrichment. Only one category was significantly enriched ($p < 0.05$); it is denoted by an asterisk.

Fig. S7



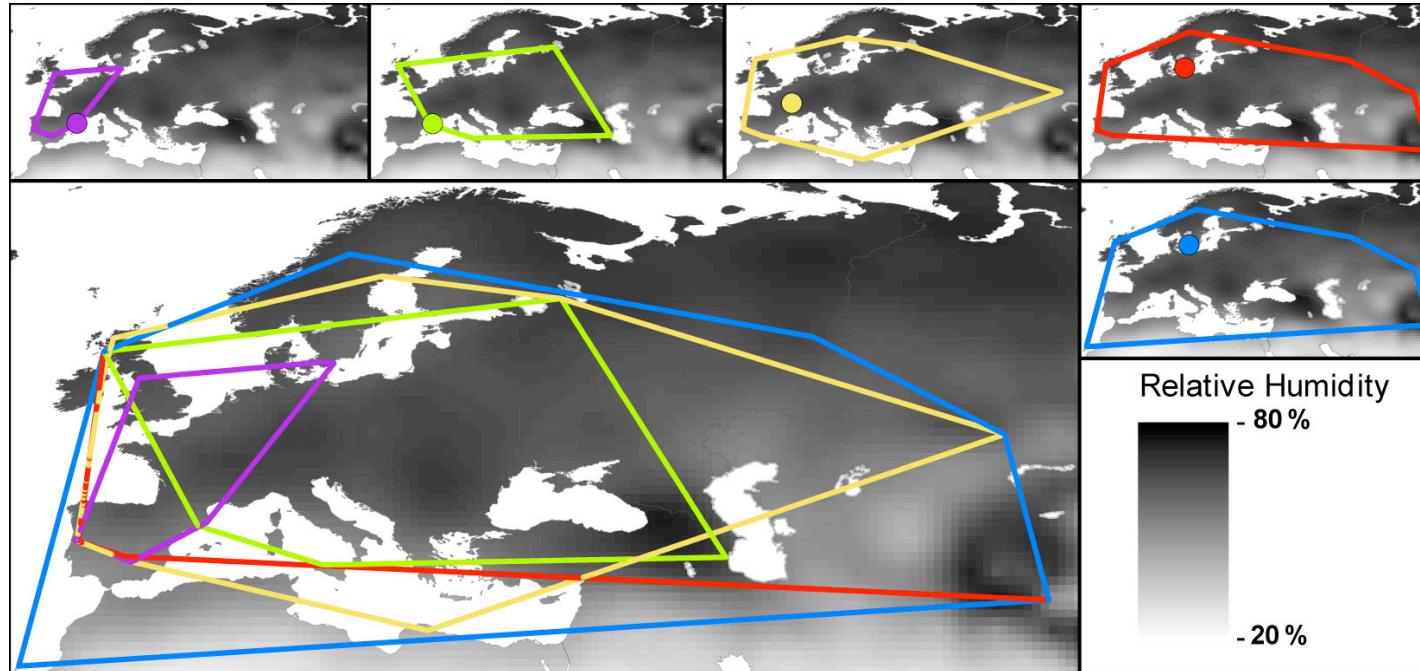
Pairwise matrix of similarity between climate variables. The numbers in the top-right half are the Pearson's r^2 between pairs of variables multiplied by 100. The numbers in the bottom-left are the proportion of SNPs that overlap between the two variables compared.

Fig. S8



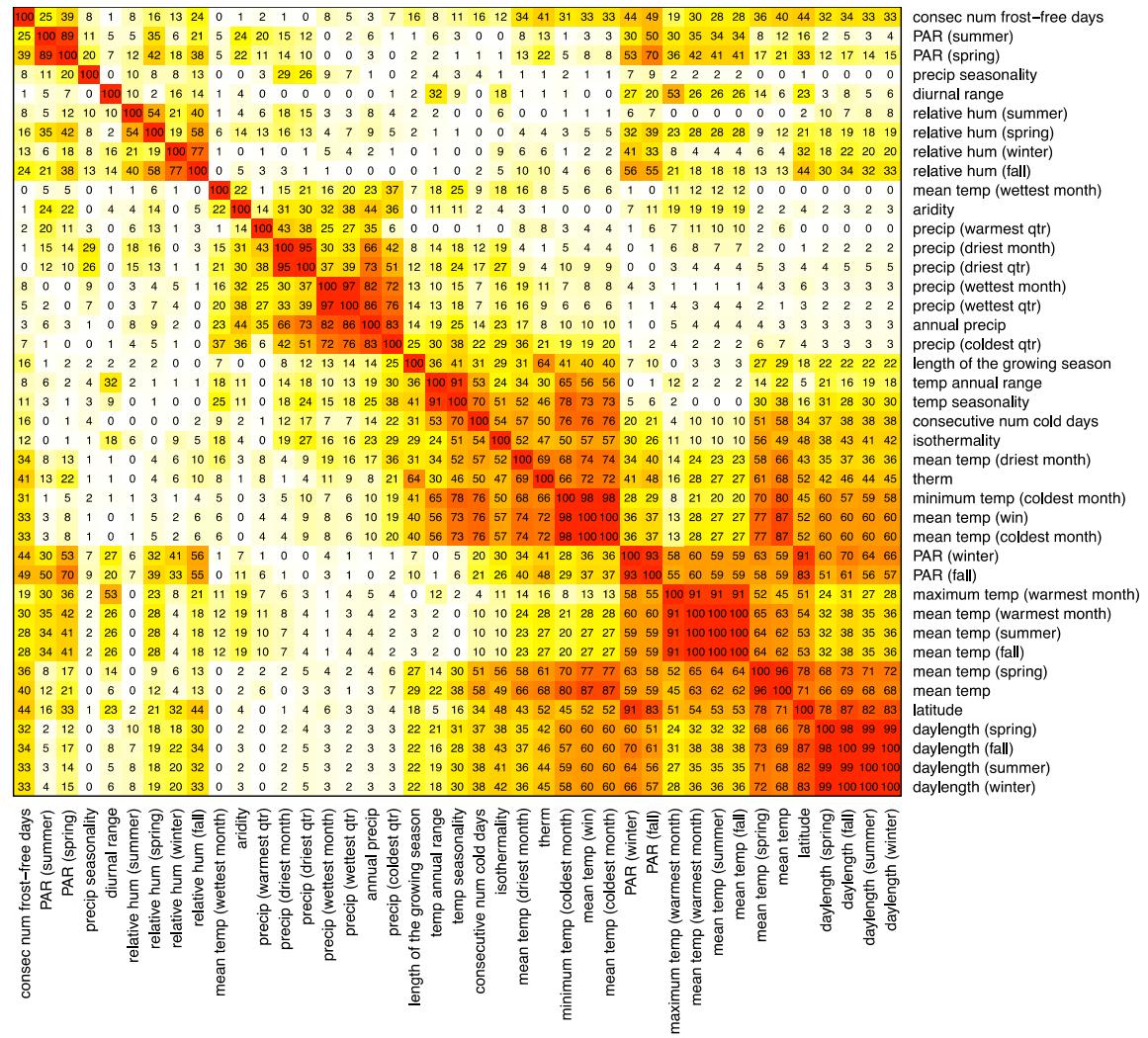
An example of an allele that may have pleiotropic phenotypic effects and thus be correlated with multiple climate variables is tagged by a SNP in *TTG1* (transparent test glabra). This SNP has the strongest correlations in scans with temperature seasonality and minimum temperature and is also strongly correlated with daylength (rank = 50/214435). The gene is involved in purple anthocyanin production, trichome patterning and epidermal cell fate specification (35), three phenotypes that are likely to play roles in both leaf temperature regulation and UV damage. (A) Interpolated distribution of the *TTG1* SNP with the strongest correlation with temperature seasonality and minimum temperature (frequencies range from 0 (blue) to 0.8 (red)). (B) Distribution of temperature seasonality (values range from 27 (green) to 112 standard deviations (orange)).

Fig. S9



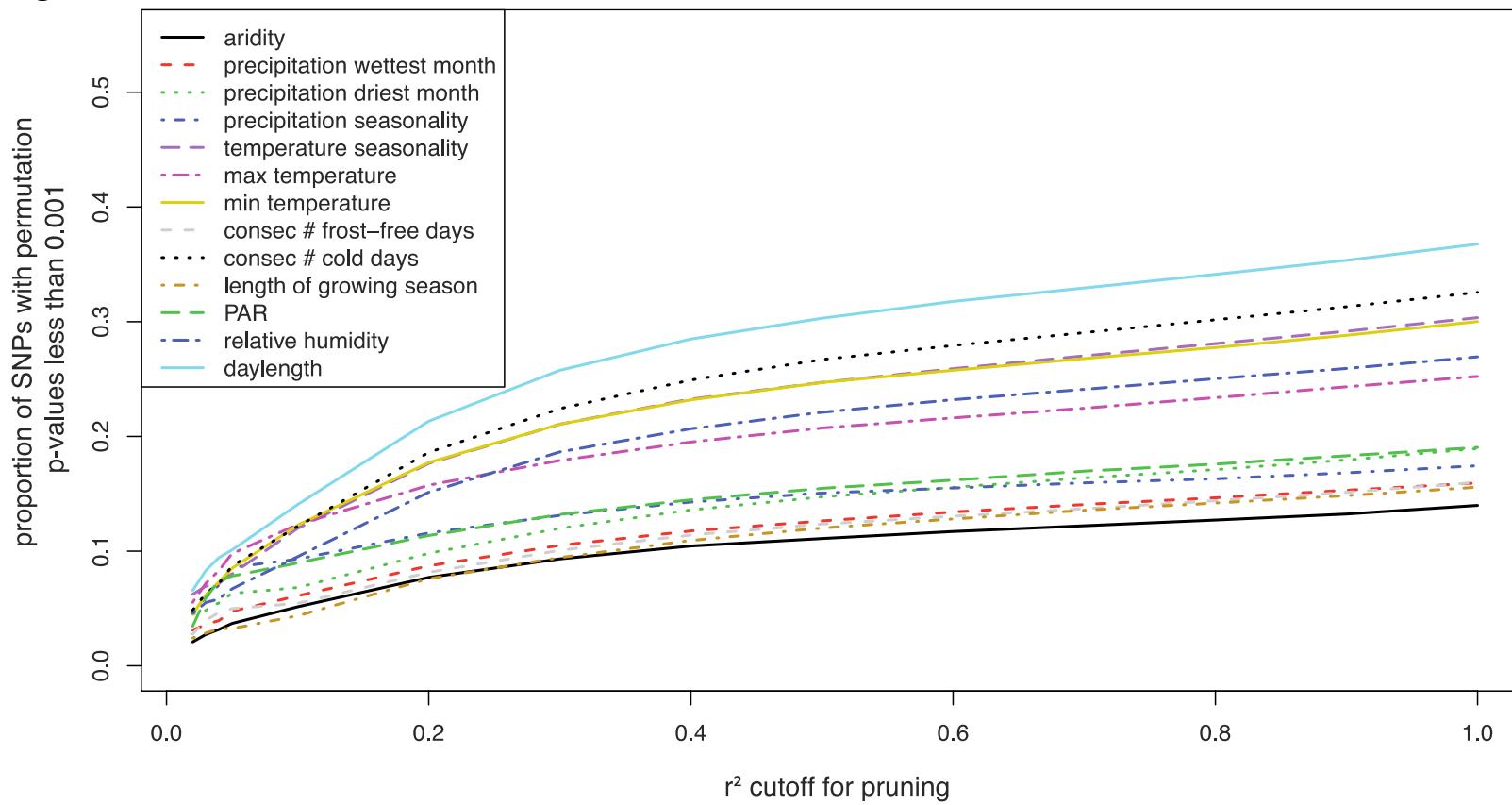
Distributions of the top 5 regions for relative humidity. For each SNP, the central sub-panel contains polygons showing the geographic extents of all 5 SNPs and sub-panels show the central feature and extent of each individual SNP.

Fig. S10



Correlation matrix of all climate variables

Fig. S11



Proportions of SNPs with permutation p-values less than 0.001 across SNP sets sampled to include varying levels of redundancy due to LD. An r^2 cutoff of 1 represents the case where all SNPs were included in the analysis and r^2 cutoffs near zero represent cases in which each SNP represents a nearly independent region of the genome.

Table S1.

Summary of 41 climate variables gathered for this analysis and sources of information.

variable	source	grid resolution	url	reference
Length of the growing season (based on temperature and potential evaporation relative to precipitation by month)	FAO GeoNetwork	0.5°x 0.5° grid (~55 km)	http://www.fao.org/geonetwork/srv/en/main.home	NA
Length of the growing season (based on temperature alone; by month)		5 arc-Minutes (~9 km)		
Mean Temperature (winter)	WorldClim	30 arc-seconds (~1 km)	http://www.worldclim.org/	Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas.
Mean Temperature (spring)				International Journal of Climatology 25: 1965-1978.
Mean Temperature (summer)				
Mean Temperature (fall)				
Annual Mean Temperature				
Mean Diurnal Range (Mean of monthly (max temp - min temp))				
Isothermality (P2/P7) (* 100)				
Temperature Seasonality (standard deviation *100)				
Max Temperature of Warmest Month				
Min Temperature of Coldest Month				
Temperature Annual Range (P5-P6)				
Mean Temperature of Wettest Quarter				
Mean Temperature of Driest Quarter				
Mean Temperature of Warmest Quarter				
Mean Temperature of Coldest Quarter				
Annual Precipitation				
Precipitation of Wettest Month				
Precipitation of Driest Month				
Precipitation Seasonality (Coefficient of Variation)				
Precipitation of Wettest Quarter				
Precipitation of Driest Quarter				
Precipitation of Warmest Quarter				
Precipitation of Coldest Quarter				
Photosynthetically active radiation (fall)	NASA/GEWE X Surface Radiation Budget (SRB)	1°x 1° grid (~111 km)	http://eosweb.larc.nasa.gov/PRODOCS/srb/table_srb.html	These data were obtained from the NASA Langley Research Center Atmospheric Science Data Center.
Photosynthetically active radiation (winter)				
Photosynthetically active radiation (spring)				
Photosynthetically active radiation (summer)				
Number of consecutive cold days (below 4 degrees C)	NCDC	0.37°x 0.37° grid (~40 km)	http://www.ncdc.noaa.gov/cgi-bin/res40.pl?page=gsod.html	The NCDC Climate Services Branch (CSB)
Number of consecutive frost-free days (above 0 degrees C)				
Relative humidity (winter)	NCAR/NCEP	5 arc-Minutes (~9 km)	http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-.NCAR/.CDAS-1/.MONTHLY/.Intrinsic/.Pressure_Level/.rhum/	Kalnay, E., M. Kanamitsu, R. Kistler, et all. The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society, March, 1996
Relative humidity (summer)				
Relative humidity (spring)				
Relative humidity (fall)				
Daylength (winter)		NA		Meeus, Jean. (1991) Astronomical algorithms.
Daylength (spring)				Richmond, Va.: Willmann-Bell.
Daylength (summer)				ISBN 0943396352
Daylength (fall)				

Table S2.

Enrichment of biological processes in the 1% tail (with $p \leq 0.01$) for any of the 13 climate variables.

Climate	Biological Process	Enrichment	p-value
aridity	pyridine nucleotide biosynthetic process	17.39	0.001
aridity	pyridine nucleotide biosynthetic process	17.39	0.00097
aridity	root hair initiation	15.38	0.008
aridity	base-excision repair	13.59	0.00014
aridity	two-component signal transduction system (phosphorelay)	12.79	0.00013
aridity	NAD(P)H dehydrogenase complex (plastoquinone)	12.50	0.003
aridity	glutamate decarboxylation to succinate	11.76	0.006
aridity	DNA replication initiation	11.40	0.0006
aridity	DNA unwinding during replication	11.11	0.001
aridity	starch metabolic process	11.11	0.001
aridity	DNA unwinding during replication	11.11	0.00097
aridity	starch metabolic process	11.11	0.00106
aridity	seed dormancy	10.52	0.001
aridity	N-glycan processing	10.00	0.008
aridity	protein import into chloroplast stroma	9.75	0.007
aridity	salicylic acid mediated signaling pathway	9.09	0.002
aridity	sugar mediated signaling pathway	8.97	0.00036
aridity	asparagine biosynthetic process	8.33	0.008
aridity	defense response, incompatible interaction	7.69	0.007
aridity	floral organ abscission	6.97	0.006
aridity	DNA topological change	6.84	0.005
aridity	sucrose biosynthetic process	5.67	0.007
aridity	response to virus	5.52	0.004
aridity	photosynthesis	5.50	0.00106
aridity	microtubule-based movement	5.31	0.002
aridity	microtubule-based movement	5.31	0.00026
aridity	positive regulation of flower development	4.63	0.005
aridity	glycolysis	4.11	0.003
aridity	regulation of cell cycle	3.87	0.003
aridity	vegetative to reproductive phase transition of meristem	3.84	0.008
aridity	response to stress	3.07	0.005
aridity	DNA repair	2.82	0.004
aridity	regulation of transcription	1.63	0.008
consec cold days	cytidine deamination	26.66	0.003
consec cold days	cytidine metabolic process	26.66	0.005
consec cold days	protein ubiquitination during ubiquitin-dependent protein catabolic process	24.99	0.005
consec cold days	regulation of actin filament polymerization	19.99	0.005
consec cold days	methylation-dependent chromatin silencing	16.66	0.005
consec cold days	Golgi vesicle transport	15.68	0.005
consec cold days	cellular aromatic compound metabolic process	15.38	0.005
consec cold days	plant-type cell wall organization	14.75	0.001
consec cold days	epidermal cell fate specification	14.28	0.004
consec cold days	mitochondrial electron transport, succinate to ubiquinone	13.79	0.005
consec cold days	protein import into nucleus	13.33	0.001
consec cold days	protein import into nucleus	13.33	0.00014
consec cold days	peptidyl-cysteine S-nitrosylation	12.76	0.002
consec cold days	post-translational protein modification	12.50	0.001
consec cold days	post-translational protein modification	12.50	0.00056
consec cold days	cell wall biogenesis	12.30	0.004
consec cold days	transcription initiation from RNA polymerase II promoter	11.47	0.004
consec cold days	regulation of protein metabolic process	10.97	0.001
consec cold days	cysteine biosynthetic process	9.23	0.007
consec cold days	D-ribose metabolic process	8.82	0.003
consec cold days	trichome differentiation	8.33	0.006
consec cold days	hyperosmotic response	7.35	0.007

consec cold days	copper ion transport	7.02	0.008
consec cold days	aging	6.28	0.001
consec cold days	mRNA processing	6.18	0.009
consec cold days	multidrug transport	4.94	0.001
consec cold days	multidrug transport	4.94	0.0009
consec cold days	carbohydrate metabolic process	1.91	0.009
consec frost-free days	maintenance of root meristem identity	25.71	0.00047
consec frost-free days	indoleacetic acid biosynthetic process	23.68	0.001
consec frost-free days	threonine biosynthetic process	18.74	0.001
consec frost-free days	gynoecium development	18.36	0.001
consec frost-free days	cellular response to water deprivation	18.18	0.002
consec frost-free days	mitochondrial electron transport, NADH to ubiquinone	17.85	0.002
consec frost-free days	NAD(P)H dehydrogenase complex (plastoquinone)	16.66	0.001
consec frost-free days	epidermal cell differentiation	16.39	0.002
consec frost-free days	cotyledon vascular tissue pattern formation	15.51	0.00083
consec frost-free days	proteasomal ubiquitin-dependent protein catabolic process	14.28	0.003
consec frost-free days	positive gravitropism	13.09	0.00031
consec frost-free days	root epidermal cell differentiation	12.50	0.003
consec frost-free days	epidermal cell fate specification	11.90	0.005
consec frost-free days	regulation of chlorophyll biosynthetic process	11.62	0.002
consec frost-free days	two-component signal transduction system (phosphorelay)	11.62	0.00038
consec frost-free days	starch biosynthetic process	8.74	0.003
consec frost-free days	cotyledon development	8.62	0.001
consec frost-free days	cotyledon development	8.62	0.00103
consec frost-free days	tubulin complex assembly	8.57	0.007
consec frost-free days	cell morphogenesis	7.69	0.009
consec frost-free days	phloem or xylem histogenesis	7.44	0.002
consec frost-free days	mRNA processing	7.30	0.001
consec frost-free days	salicylic acid mediated signaling pathway	7.27	0.003
consec frost-free days	trichome differentiation	7.14	0.009
consec frost-free days	nucleotide-excision repair	6.97	0.007
consec frost-free days	protein targeting to vacuole	6.89	0.004
consec frost-free days	cellular amino acid biosynthetic process	6.66	0.009
consec frost-free days	microtubule-based movement	6.58	0.001
consec frost-free days	microtubule-based movement	6.58	0.00005
consec frost-free days	photorespiration	6.29	0.002
consec frost-free days	double fertilization forming a zygote and endosperm	5.64	0.006
consec frost-free days	cytokinin mediated signaling pathway	4.22	0.002
consec frost-free days	flower development	3.65	0.002
consec frost-free days	DNA repair	3.20	0.005
daylength	regulation of signal transduction	26.66	0.00024
daylength	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	22.22	0.002
daylength	cytidine deamination	19.99	0.004
daylength	cytidine metabolic process	19.99	0.004
daylength	methylation-dependent chromatin silencing	16.66	0.002
daylength	jasmonic acid and ethylene-dependent systemic resistance	16.00	0.004
daylength	epidermal cell fate specification	14.28	0.003
daylength	amylopectin biosynthetic process	12.50	0.002
daylength	histone modification	10.00	0.006
daylength	trichome differentiation	9.52	0.001
daylength	sodium ion transport	8.94	0.003
daylength	sodium ion transport	8.94	0.00079
daylength	post-translational protein modification	8.33	0.003
daylength	maintenance of floral meristem identity	7.69	0.008
daylength	cysteine biosynthetic process	7.69	0.005
daylength	hyperosmotic response	7.35	0.007
daylength	regulation of protein metabolic process	7.31	0.006
daylength	methionine biosynthetic process	6.78	0.007
daylength	response to virus	5.52	0.007
daylength	photomorphogenesis	4.08	0.002
daylength	response to salicylic acid stimulus	3.13	0.003
daylength	carbohydrate metabolic process	2.25	0.00293
growing season length	mitochondrial electron transport, NADH to ubiquinone	24.99	0.001

growing season length	mitochondrial electron transport, NADH to ubiquinone	24.99	0.0005
growing season length	asparagine biosynthetic process	16.66	0.003
growing season length	seed oilbody biogenesis	15.38	0.002
growing season length	tubulin complex assembly	14.28	0.002
growing season length	protein sumoylation	13.63	0.002
growing season length	base-excision repair	13.59	0.00008
growing season length	epidermal cell differentiation	13.11	0.0003
growing season length	photoinhibition	11.76	0.002
growing season length	peptidyl-cysteine S-nitrosylation	10.64	0.001
growing season length	mitochondrial electron transport, succinate to ubiquinone	10.34	0.006
growing season length	nucleotide-sugar transport	10.00	0.001
growing season length	N-glycan processing	10.00	0.007
growing season length	epidermal cell fate specification	9.52	0.008
growing season length	PSII associated light-harvesting complex II catabolic process	8.82	0.004
growing season length	vernization response	8.41	0.001
growing season length	regulation of meristem growth	6.82	0.008
growing season length	chloroplast RNA processing	6.25	0.009
growing season length	response to virus	6.21	0.005
growing season length	folic acid and derivative biosynthetic process	5.88	0.009
growing season length	negative regulation of flower development	4.76	0.001
growing season length	amino acid transport	4.04	0.003
growing season length	DNA repair	3.08	0.001
growing season length	embryonic development ending in seed dormancy	2.25	0.00022
growing season length	regulation of transcription	1.66	0.003
max T (warmest month)	L-phenylalanine biosynthetic process	18.60	0.00039
max T (warmest month)	mitochondrial electron transport, NADH to ubiquinone	17.85	0.008
max T (warmest month)	defense response signaling pathway, resistance gene-independent	14.28	0.00086
max T (warmest month)	translational termination	12.30	0.001
max T (warmest month)	regulation of nitrogen utilization	12.24	0.001
max T (warmest month)	galactolipid biosynthetic process	12.12	0.005
max T (warmest month)	mRNA cleavage involved in gene silencing by miRNA	11.53	0.009
max T (warmest month)	unsaturated fatty acid biosynthetic process	11.11	0.007
max T (warmest month)	plastid organization	8.75	0.005
max T (warmest month)	tRNA processing	8.57	0.001
max T (warmest month)	tRNA processing	8.57	0.00005
max T (warmest month)	polysaccharide biosynthetic process	7.84	0.003
max T (warmest month)	histone deacetylation	7.84	0.004
max T (warmest month)	tricarboxylic acid cycle	7.69	0.003
max T (warmest month)	nuclear mRNA splicing, via spliceosome	7.36	0.001
max T (warmest month)	nuclear mRNA splicing, via spliceosome	7.36	0.00049
max T (warmest month)	red, far-red light phototransduction	7.14	0.007
max T (warmest month)	reciprocal meiotic recombination	6.25	0.006
max T (warmest month)	response to cyclopentenone	5.80	0.003
max T (warmest month)	photoperiodism, flowering	4.54	0.008
max T (warmest month)	DNA mediated transformation	4.52	0.006
max T (warmest month)	trichome morphogenesis	4.23	0.007
max T (warmest month)	cell wall modification	3.81	0.003
max T (warmest month)	oxidation reduction	2.80	0.005
max T (warmest month)	response to cadmium ion	2.69	0.00031
max T (warmest month)	response to cold	2.52	0.003
min T (coldest month)	regulation of signal transduction	26.66	0.00032
min T (coldest month)	cytidine deamination	19.99	0.004
min T (coldest month)	regulation of actin filament polymerization	19.99	0.005
min T (coldest month)	cytidine metabolic process	19.99	0.008
min T (coldest month)	threonine biosynthetic process	18.74	0.003
min T (coldest month)	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	16.66	0.004
min T (coldest month)	epidermal cell fate specification	14.28	0.006
min T (coldest month)	glycerol-3-phosphate metabolic process	12.50	0.005
min T (coldest month)	jasmonic acid and ethylene-dependent systemic resistance	12.00	0.007
min T (coldest month)	amylopectin biosynthetic process	10.00	0.008
min T (coldest month)	post-translational protein modification	9.72	0.002
min T (coldest month)	trichome differentiation	9.52	0.001
min T (coldest month)	cysteine biosynthetic process	9.23	0.003

min T (coldest month)	copper ion transport	8.77	0.006
min T (coldest month)	regulation of protein metabolic process	8.53	0.009
min T (coldest month)	protein import into nucleus	8.33	0.006
min T (coldest month)	hyperosmotic response	7.35	0.005
min T (coldest month)	D-ribose metabolic process	7.35	0.008
min T (coldest month)	protein targeting to vacuole	6.89	0.004
min T (coldest month)	response to gibberellin stimulus	5.36	0.001
min T (coldest month)	response to gibberellin stimulus	5.36	0.0004
min T (coldest month)	pollen germination	4.40	0.005
min T (coldest month)	response to salicylic acid stimulus	3.97	0.002
min T (coldest month)	defense response to fungus	3.47	0.009
min T (coldest month)	response to jasmonic acid stimulus	2.97	0.004
min T (coldest month)	response to cadmium ion	2.63	0.00047
min T (coldest month)	response to salt stress	2.39	0.002
min T (coldest month)	carbohydrate metabolic process	1.96	0.005
PAR	maintenance of root meristem identity	31.42	0.00001
PAR	indoleacetic acid biosynthetic process	28.94	0.00001
PAR	cellular response to water deprivation	27.26	0.00006
PAR	regulation of defense response	24.24	0.00028
PAR	gynoecium development	22.44	0.00002
PAR	red light signaling pathway	21.62	0.00014
PAR	stomatal complex development	21.62	0.00011
PAR	cotyledon vascular tissue pattern formation	18.96	0.00005
PAR	mitochondrial electron transport, NADH to ubiquinone	17.85	0.001
PAR	NAD(P)H dehydrogenase complex (plastoquinone)	16.66	0.001
PAR	jasmonic acid and ethylene-dependent systemic resistance	16.00	0.002
PAR	positive gravitropism	15.47	0.00001
PAR	root hair initiation	15.38	0.009
PAR	methionine metabolic process	12.50	0.007
PAR	induced systemic resistance, jasmonic acid mediated signaling pathway	11.53	0.008
PAR	tyrosine biosynthetic process	11.11	0.002
PAR	protein deubiquitination	11.11	0.009
PAR	cotyledon development	10.34	0.001
PAR	cotyledon development	10.34	0.00021
PAR	leucine biosynthetic process	9.61	0.003
PAR	proanthocyanidin biosynthetic process	9.26	0.006
PAR	phloem or xylem histogenesis	9.09	0.00031
PAR	base-excision repair	8.74	0.004
PAR	tubulin complex assembly	8.57	0.009
PAR	gravitropism	8.33	0.001
PAR	regulation of seed germination	7.87	0.003
PAR	starch biosynthetic process	7.76	0.005
PAR	jasmonic acid mediated signaling pathway	7.57	0.001
PAR	phototropism	7.50	0.005
PAR	thylakoid membrane organization	6.06	0.005
PAR	lactose catabolic process, using glucoside 3-dehydrogenase	6.00	0.007
PAR	systemic acquired resistance	5.62	0.004
PAR	photosynthesis	4.59	0.001
PAR	photosynthesis	4.59	0.007
PAR	regulation of stomatal movement	4.52	0.006
PAR	positive regulation of transcription	4.37	0.004
PAR	vegetative to reproductive phase transition of meristem	4.27	0.008
PAR	seed development	3.74	0.009
PAR	flower development	3.16	0.003
PAR	response to cold	2.98	0.0069
PAR	regulation of transcription, DNA-dependent	1.89	0.001
precip (driest month)	maintenance of root meristem identity	22.85	0.00036
precip (driest month)	indoleacetic acid biosynthetic process	21.05	0.00013
precip (driest month)	mitochondrial electron transport, succinate to ubiquinone	20.68	0.001
precip (driest month)	cotyledon vascular tissue pattern formation	13.79	0.001
precip (driest month)	glucuronoxylan biosynthetic process	13.04	0.005
precip (driest month)	epidermal cell fate specification	11.90	0.002
precip (driest month)	peptidyl-cysteine S-nitrosylation	10.64	0.001

precip (driest month)	positive gravitropism	9.52	0.002
precip (driest month)	L-phenylalanine biosynthetic process	9.30	0.002
precip (driest month)	specification of floral organ identity	8.00	0.009
precip (driest month)	cell adhesion	7.84	0.006
precip (driest month)	cotyledon development	7.76	0.001
precip (driest month)	DNA mediated transformation	5.08	0.001
precip (driest month)	DNA replication	4.01	0.002
precip (driest month)	flower development	3.41	0.007
precip (driest month)	intracellular protein transport	2.33	0.006
precip (wettest month)	pyridine nucleotide biosynthetic process	17.39	0.00068
precip (wettest month)	mitochondrial electron transport, NADH to ubiquinone	14.28	0.005
precip (wettest month)	base-excision repair	13.59	0.00009
precip (wettest month)	hyperosmotic response	11.76	0.002
precip (wettest month)	glutamate decarboxylation to succinate	11.76	0.005
precip (wettest month)	root hair cell tip growth	11.68	0.001
precip (wettest month)	signal peptide processing	11.43	0.003
precip (wettest month)	stomatal complex morphogenesis	11.11	0.00074
precip (wettest month)	nicotianamine biosynthetic process	11.11	0.006
precip (wettest month)	seed dormancy	10.52	0.002
precip (wettest month)	amylopectin biosynthetic process	10.00	0.004
precip (wettest month)	N-glycan processing	10.00	0.006
precip (wettest month)	starch metabolic process	9.72	0.001
precip (wettest month)	starch metabolic process	9.72	0.002
precip (wettest month)	PSII associated light-harvesting complex II catabolic process	8.82	0.007
precip (wettest month)	photoinhibition	8.82	0.004
precip (wettest month)	post-embryonic root development	8.33	0.008
precip (wettest month)	salicylic acid mediated signaling pathway	7.27	0.005
precip (wettest month)	protein catabolic process	6.42	0.00029
precip (wettest month)	defense response, incompatible interaction	6.15	0.007
precip (wettest month)	response to abiotic stimulus	6.00	0.008
precip (wettest month)	cell division	5.75	0.001
precip (wettest month)	cell division	5.75	0.0003
precip (wettest month)	organ morphogenesis	5.26	0.005
precip (wettest month)	thylakoid membrane organization	4.54	0.006
precip (wettest month)	vegetative to reproductive phase transition of meristem	4.27	0.003
precip (wettest month)	cell wall modification	3.17	0.006
precip (wettest month)	translational initiation	3.13	0.007
precip (wettest month)	vesicle-mediated transport	2.56	0.004
precip seasonality	maintenance of root meristem identity	22.85	0.00115
precip seasonality	mitochondrial electron transport, NADH to ubiquinone	17.85	0.006
precip seasonality	gynoecium development	16.32	0.001
precip seasonality	protein import into nucleus	15.00	0.00021
precip seasonality	response to symbiotic fungus	14.28	0.003
precip seasonality	cotyledon vascular tissue pattern formation	13.79	0.002
precip seasonality	secondary shoot formation	12.50	0.006
precip seasonality	lipid storage	12.19	0.001
precip seasonality	coenzyme A biosynthetic process	10.81	0.006
precip seasonality	positive gravitropism	9.52	0.004
precip seasonality	very-long-chain fatty acid metabolic process	7.84	0.008
precip seasonality	cotyledon development	7.76	0.002
precip seasonality	cuticle development	7.54	0.005
precip seasonality	cell death	7.00	0.001
precip seasonality	phloem or xylem histogenesis	6.61	0.005
precip seasonality	toxin catabolic process	6.01	0.008
precip seasonality	mismatch repair	5.73	0.006
precip seasonality	negative regulation of flower development	5.08	0.004
precip seasonality	response to ethylene stimulus	3.54	0.004
precip seasonality	flower development	3.41	0.009
relative humidity	regulation of signal transduction	29.99	0.001
relative humidity	folic acid and derivative metabolic process	23.07	0.003
relative humidity	glucuronoxylan biosynthetic process	17.39	0.004
relative humidity	signal peptide processing	17.14	0.001
relative humidity	asparagine biosynthetic process	16.66	0.007
relative humidity	mitochondrial electron transport, NADH to ubiquinone	14.28	0.009

relative humidity	synapsis	12.30	0.0005
relative humidity	cellular component organization	11.43	0.005
relative humidity	meristem development	10.97	0.003
relative humidity	peptidyl-cysteine S-nitrosylation	10.64	0.002
relative humidity	sister chromatid cohesion	10.57	0.002
relative humidity	electron transport chain	10.00	0.009
relative humidity	epidermal cell fate specification	9.52	0.008
relative humidity	reciprocal meiotic recombination	8.93	0.001
relative humidity	oxygen and reactive oxygen species metabolic process	8.66	0.002
relative humidity	protein import into nucleus	8.33	0.001
relative humidity	specification of floral organ identity	8.00	0.009
relative humidity	microsporogenesis	6.33	0.004
relative humidity	actin cytoskeleton organization	5.88	0.006
relative humidity	intracellular protein transport	2.67	0.004
relative humidity	response to salt stress	2.27	0.004
temp seasonality	regulation of signal transduction	26.66	0.00052
temp seasonality	cytidine deamination	19.99	0.009
temp seasonality	cytidine metabolic process	19.99	0.007
temp seasonality	electron transport chain	15.00	0.003
temp seasonality	base-excision repair	12.62	0.00042
temp seasonality	trichome differentiation	9.52	0.006
temp seasonality	negative regulation of seed germination	9.30	0.004
temp seasonality	auxin biosynthetic process	9.26	0.007
temp seasonality	cysteine biosynthetic process	9.23	0.009
temp seasonality	protein targeting to vacuole	8.62	0.001
temp seasonality	two-component signal transduction system (phosphorelay)	8.14	0.003
temp seasonality	mRNA processing	7.86	0.001
temp seasonality	salicylic acid mediated signaling pathway	7.27	0.008
temp seasonality	gravitropism	6.66	0.009
temp seasonality	response to gibberellin stimulus	5.06	0.001
temp seasonality	pollen tube growth	4.69	0.002
temp seasonality	pollen germination	4.40	0.008
temp seasonality	response to salicylic acid stimulus	3.55	0.004
temp seasonality	microtubule-based movement	3.40	0.008
temp seasonality	DNA repair	3.20	0.004
temp seasonality	response to auxin stimulus	2.85	0.003
temp seasonality	response to salt stress	2.08	0.004
temp seasonality	protein amino acid phosphorylation	1.69	0.008

References and Notes

1. W. E. Bradshaw, C. M. Holzapfel, Genetic shift in photoperiodic response correlated with global warming. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14509 (2001).
2. S. J. Franks, S. Sim, A. E. Weis, Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1278 (2007).
3. D. B. Lobell, W. Schlenker, J. Costa-Roberts, Climate trends and global crop production since 1980. *Science* **333**, 616 (2011).
4. M. Lynch, R. Lande, in *Biotic Interactions and Global Change*, P. M. Kareiva, J. G. Kingsolver, R. B. Huey, Eds. (Sinauer Associates, Sunderland, MA, 1993), pp. 234–250.
5. P. A. Umina, A. R. Weeks, M. R. Kearney, S. W. McKechnie, A. A. Hoffmann, A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science* **308**, 691 (2005).
6. A. A. Hoffmann, C. M. Sgrò, Climate change and evolutionary adaptation. *Nature* **470**, 479 (2011).
7. S. Atwell *et al.*, Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627 (2010).
8. J. Bergelson, F. Roux, Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nat. Rev. Genet.* **11**, 867 (2010).
9. B. Brachi *et al.*, Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet.* **6**, e1000940 (2010).
10. C. Weinig *et al.*, Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics* **162**, 1875 (2002).
11. D. H. Kim, M. R. Doyle, S. Sung, R. M. Amasino, Vernalization: Winter and the timing of flowering in plants. *Annu. Rev. Cell Dev. Biol.* **25**, 277 (2009).
12. B. Rathcke, E. P. Lacey, Phenological patterns of terrestrial plants. *Annu. Rev. Ecol. Syst.* **16**, 179 (1985).
13. See supporting material in *Science* Online.
14. A. M. Hancock *et al.*, Adaptations to climate-mediated selective pressures in humans. *PLoS Genet.* **7**, e1001375 (2011).
15. B. Charlesworth, M. T. Morgan, D. Charlesworth, The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289 (1993).
16. J. H. Gillespie, Genetic drift in an infinite population. The pseudohitchhiking model. *Genetics* **155**, 909 (2000).
17. H.-u, -R. Athar, M. Ashraf, in *Handbook of Photosynthesis*, M. Pessarakli, Ed. (CRC Press, Boca Raton, FL, 2005), p. 928.
18. H. A. Orr, Adaptation and the cost of complexity. *Evolution* **54**, 13 (2000).

19. Z. Wang, B. Y. Liao, J. Zhang, Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18034 (2010).
20. G. P. Wagner, J. Zhang, The pleiotropic structure of the genotype-phenotype map: The evolvability of complex organisms. *Nat. Rev. Genet.* **12**, 204 (2011).
21. F. Roux, J. Gasquez, X. Reboud, The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. *Genetics* **166**, 449 (2004).
22. C. Toomajian *et al.*, A nonparametric test reveals selection for rapid flowering in the *Arabidopsis* genome. *PLoS Biol.* **4**, e137 (2006).
23. A. E. Anastasio *et al.*, Source verification of mis-identified *Arabidopsis thaliana* accessions. *Plant J.* **67**, 554 (2011).
24. R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965 (2005).
25. R. Kistler *et al.*, The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc.* **82**, 247 (2001).
26. N. Mantel, The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209 (1967).
27. P. E. Smouse, J. C. Long, R. R. Sokal, Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Biol.* **35**, 627 (1986).
28. S. C. Goslee, D. L. Urban, The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22**, 1 (2007).
29. R Development Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2009).
30. G. Coop, D. Witonsky, A. Di Rienzo, J. K. Pritchard, Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**, 1411 (2010).
31. S. Purcell *et al.*, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559 (2007).
32. F. Roux, L. Gao, J. Bergelson, Impact of initial pathogen density on resistance and tolerance in a polymorphic disease resistance gene system in *Arabidopsis thaliana*. *Genetics* **185**, 283 (2010).
33. H. M. Kang *et al.*, Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709 (2008).
34. Gene Ontology Consortium, The Gene Ontology project in 2008. *Nucleic Acids Res.* **36** (Database issue), D440 (2008).
35. A. R. Walker *et al.*, The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell* **11**, 1337 (1999).