

Century-long timelines of herbarium genomes predict plant stomatal response to climate change

Received: 6 December 2022

Accepted: 21 June 2024

Published online: 8 August 2024

 Check for updates

Patricia L. M. Lang^{1,2,11}✉, Joel M. Erberich¹, Lua Lopez^{3,4}, Clemens L. Weiß⁵, Gabriel Amador⁶, Hannah F. Fung¹, Sergio M. Latorre^{7,8}, Jesse R. Lasky⁴, Hernán A. Burbano^{7,8}, Moisés Expósito-Alonso^{1,9,10,12,13} & Dominique C. Bergmann^{1,2}

Dissecting plant responses to the environment is key to understanding whether and how plants adapt to anthropogenic climate change. Stomata, plants' pores for gas exchange, are expected to decrease in density following increased CO₂ concentrations, a trend already observed in multiple plant species. However, it is unclear whether such responses are based on genetic changes and evolutionary adaptation. Here we make use of extensive knowledge of 43 genes in the stomatal development pathway and newly generated genome information of 191 *Arabidopsis thaliana* historical herbarium specimens collected over 193 years to directly link genetic variation with climate change. While we find that the essential transcription factors SPCH, MUTE and FAMA, central to stomatal development, are under strong evolutionary constraints, several regulators of stomatal development show signs of local adaptation in contemporary samples from different geographic regions. We then develop a functional score based on known effects of gene knock-out on stomatal development that recovers a classic pattern of stomatal density decrease over the past centuries, suggesting a genetic component contributing to this change. This approach combining historical genomics with functional experimental knowledge could allow further investigations of how different, even in historical samples unmeasurable, cellular plant phenotypes may have already responded to climate change through adaptive evolution.

Ongoing drastic increases in atmospheric CO₂ concentrations (subsequently [CO₂]) and the resulting changes in global temperatures and drought severity are altering our environment¹. Dissecting whether and how plants respond to this will be key to understanding plants' potential for adaptation to climate change, and to developing strategies to increase their chances of survival. Several phenotypic trends observed in large numbers of plant species and across continents are being reported, including the acceleration of flowering^{2,3} and other

life history events⁴, the increase in photosynthesis⁵ and the decrease in the number of stomatal pores in plant leaves⁶. However, it has thus far been difficult to resolve whether these trends reflect plastic phenotypic changes or result from evolutionary genetic change⁷. Stomata are one of the plant structures that are most directly relevant to multifactorial climatic changes, as these surface pores, essential for survival and productivity, are major contributors to plants' water-use efficiency (WUE): the ratio between CO₂ uptake for photosynthesis (*A*) and the

A full list of affiliations appears at the end of the paper. ✉ e-mail: patricialleng@gmail.com

release of O₂ and loss of water vapour with transpiration (E ; reviewed in ref. 8). Optimizing WUE to environmental conditions can improve plant fitness and yield when plant growth and cooling through transpiration are balanced with minimal water loss. WUE is partially fine-tuned through variation in stomatal size and density (that is, the amount of stomata per plant surface area^{9–11}). This stomatal variation can represent temporary plastic responses or result from evolutionary genetic change over generations that produces local adaptation^{12–14}.

A powerful way to differentiate between plastic and evolutionary plant responses to climate is to study different populations of a single species collected across geographic climatic gradients in combination with genetic analyses and common gardens¹⁵. Both stomatal size and density of natural populations of *Arabidopsis thaliana* measured in controlled environments correlate with climate variables of the populations' geographic origin, indicating a potential genetic basis for their climate responses¹⁶. In *Arabidopsis*, lower stomatal density typically follows increasing [CO₂] and temperature, while higher stomatal densities and drought-adjusted higher WUE result from decreased humidity. Stomatal size is generally anti-correlated with stomatal density^{16–20}. Given this connection of stomatal variation with climate gradients within a species, we would expect the past ~200 years of anthropogenic global change to have impacted such plant stomatal variation in complex unknown ways. Collections of pressed and dried specimens of different species are witnesses of plant responses to the [CO₂] increase from ~280 parts per million (p.p.m.) in 1750 to 419 p.p.m. in 2022 (August measurement; see trends at www.climate.gov and climate.nasa.gov), or the current global maximum temperature anomaly of approximately +1 °C (climate.nasa.gov), and, with that, we can track adaptation to climate change as it happens (reviewed in ref. 21).

Plant responses themselves can help infer historical climate trends. Measurements of stomatal densities in fossils indicate [CO₂] changes over geological time^{22–25}, and over the recent anthropogenic climate change, decreases in stomatal densities preserved in herbaria already reflect the industrialization-related increases in [CO₂] (ref. 6). Such analyses have thus far never extended beyond phenotypic quantification to also assess genetic changes underlying these potentially adaptive responses, mainly because fossil records lack quantifiable DNA. Now, sequencing of herbarium specimens can address this gap by directly exploring joint timelines of phenotypic and genotypic responses to climate change (for example, refs. 26–30).

Stomata are a unique system to use such timelines to understand plants' adaptive potential and its mechanisms, as their genetic pathway is dissected in minute detail in *A. thaliana* (for example, reviewed in refs. 8, 31, 32): From the sequential interactions of indispensable transcription factors that regulate cell production, fate and patterning to external regulators that fine-tune stomatal development in response to environmental and physiological stimuli. This knowledge provides a crucial advantage when investigating a genetic basis of potential

stomata involvement in climate adaptation. It allows selective study of genetic variants in already-validated causal genes and can complement genome-wide association 'discovery' approaches, which so far have yielded highly polygenic signals elsewhere in the genome that explain part of the observed phenotypic variation but are not easily connected to specific effects on phenotypes or functions^{16,33}. We can start accounting for this polygenic complexity by integrating genetic information and a functional understanding of the entire genetic developmental pathway and directly ask whether and how known stomatal development genes may promote adaptation.

In this Article, we use historical specimens as 'witnesses' of the ongoing climate change together with molecular genetics knowledge to ask: Can herbaria reveal climate change adaptation in the genetic pathways of essential plant features such as stomata? Can combination of historical and modern genomes with genes' known phenotypic effects circumvent currently lacking historical phenotyping to predict plant change?

Results and discussion

Stomata genes show a purifying natural selection signal

To investigate how stomata, or stomatal development, has responded to climatic change, we created a new temporal dataset^{34,35} of 191 broadly geographically distributed historical samples, covering the time period from 1817 to 2010 (Supplementary Fig. 1a and Supplementary Table 1), and paired it with the contemporary *A. thaliana* 1,001 genomes resource³⁶. So-called ancient DNA, retrieved from the historical herbarium specimens, was authenticated following the field's standards (for example, ref. 28). Samples show the expected patterns of age-related DNA fragmentation (merged fragments' median size of 98 bp; Supplementary Fig. 1b,c) and damage, that is, cytosine deamination (reflected as C-to-T substitutions in sequencing data; 0.7–4%; Supplementary Fig. 1d), accumulated particularly at DNA molecules' termini. As previously shown, deamination of a fragment's first base is highly correlated with the year of sample collection (one-sided Pearson's correlation test, correlation coefficient $r = -0.589$, $P = 1.693 \times 10^{-15}$; Supplementary Fig. 1e), a pattern that reflects post-collection ageing of the specimens as a primary cause of DNA damage but whose strength depends on multiple additional factors such as plant specimen collection and storage conditions^{37,38}.

Because stomatal density differences have been observed across geographic regions of *A. thaliana* and these changes are partially heritable^{16,33} we aimed to depict genetic variation in known genes of the stomatal pathway, which could constitutively alter stomatal densities. We surveyed the literature and defined a set of 43 genes experimentally validated to be involved in stomata development (Supplementary Table 2) that focuses on genes mediating the cell divisions and fate transitions central to stomatal development (Fig. 1a). Aiming to specifically inquire whether the central developmental pathway itself creates

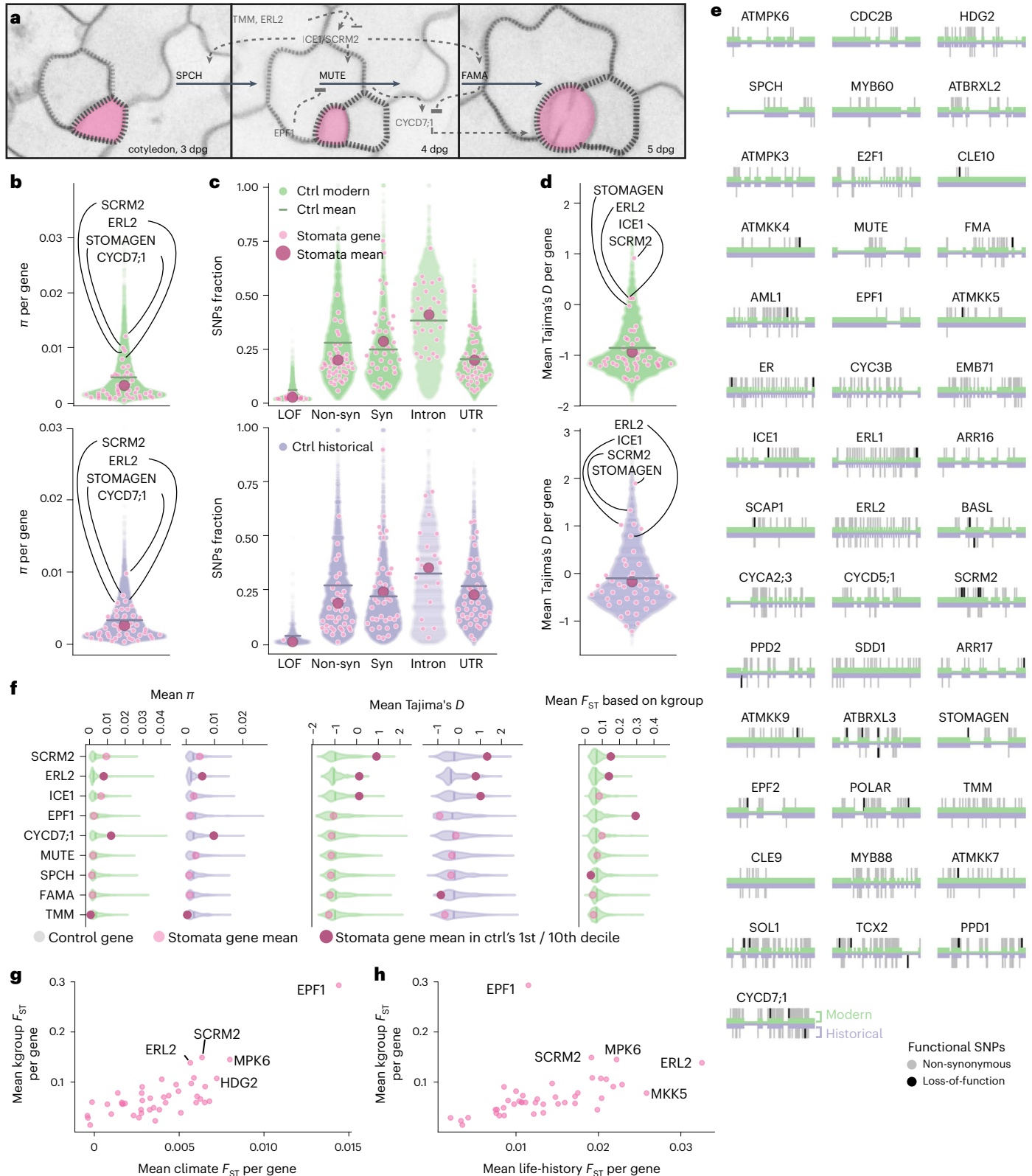
Fig. 1 | Conserved core stomata genes and regulatory genes show local adaptation signals. **a**, Stomatal development in *A. thaliana* (simplified, example regulators in grey, central (core) transcription factors in black. Stomata false-coloured in magenta. Cotyledons imaged at 3, 4 and 5 days post germination. **b**, Genetic diversity in stomatal genes is significantly lower than in length-matched control genes (gene names mark outliers; nucleotide diversity π per gene, empirical $P_{\text{mod}} = 0.004$, $P_{\text{hist}} = 0.046$). **c**, Significantly fewer SNPs in stomatal genes are putative LOF or non-synonymous than in the control genes (empirical $P_{\text{mod}}^{\text{non-syn}} = 0$, $P_{\text{hist}}^{\text{non-syn}} = 0.002$, $P_{\text{mod}}^{\text{LOF}} = 0$, $P_{\text{hist}}^{\text{LOF}} = 0.047$; Supplementary Table 3). Ctrl, control; Non-syn, non-synonymous. **d**, Mean per-gene Tajima's D indicates selection signals. Stomatal gene group is not significantly different ($P_{\text{mod}}, P_{\text{hist}} > 0.1$) from the control, but several genes are outliers (labelled). Significance tests for panels **b–d** asked whether means of the group of stomatal genes were outliers compared with the means of 1,000 control gene groups. Purple for historical, green for modern datasets, horizontal line indicates full control dataset's mean. Magenta circles for stomatal genes, large dark circles

for stomatal gene group mean. **e**, LOF (black) and non-synonymous (grey) SNPs in the 43 focus genes in historical (bottom, purple) and modern (top, green) dataset. **f**, Mean per-gene values for nucleotide diversity π , Tajima's D and $F_{\text{ST}}^{\text{kgroup}}$ values for outlier genes, compared with conserved stomatal factors SPCH, MUTE and FAMA (for all 43 genes, see Supplementary Figs. 2 and 3 and Supplementary Table 2). $F_{\text{ST}}^{\text{kgroup}}$ is calculated for populations defined by whole-genome genetic variation (from ref. 49). Gene values are displayed as transparent pink circles on violin plots representing distribution of values for the respective length-matched control genes, vertical line indicating distribution's 0.5 quantile. Solid pink circles indicate that the gene mean value lies within the 1st/10th decile of the control distribution. **g,h**, Stomatal gene differentiation as mean F_{ST} per gene, with $F_{\text{ST}}^{\text{kgroup}}$ (y-axis), compared with F_{ST} for populations clustered by climate of origin (precipitation, temperature, BIO4 and BIO15, Bioclim dataset; ref. 50) (**g**) and life-history traits (data; ref. 51) (**h**). Genes with the highest F_{ST} values across the three analyses are labelled.

constitutive stomatal density differences that can undergo positive selection, we excluded genes that are predominantly characterized as environmental sensors affecting stomatal stress responses.

We hypothesized that if there was an increased genetic diversity in these 43 stomatal genes, it might reflect broad variation in stomatal size and density (the number of stomata per plant surface area) that may have allowed local adaptation to the different environments

encountered by the species across its geographic range. To investigate this, we examined our modern and historical genomes as ‘snapshots’ of the species’ current and past ~200 years of global genetic diversity (Fig. 1b,c). Overall, expectedly, we found fewer single-nucleotide polymorphisms (SNPs) in the 191 historical than the 1,135 modern samples. This is consistent with the historical dataset’s much smaller size and stringent SNP-calling that has to account for characteristics



and sequencing challenges intrinsic to historical DNA (Supplementary Fig. 1b–d; for example, ref. 28). In addition, a modern increase in SNP numbers is also consistent with a recent human-linked population expansion of the species^{39,40}. To compare genetic diversity in these genes in the context of the broader genome, we generated 1,000 sets of 43 randomly drawn control genes, with a gene length distribution matched to that of the 43 stomatal genes. In comparison to this control, stomatal genes harbour drastically reduced genetic diversity, as estimated by Watterson's θ (Supplementary Fig. 2; ref. 41; $P_{\text{mod}} = 0.006$, $P_{\text{hist}} = 0.069$) and lower nucleotide diversity π (Fig. 1b and Supplementary Fig. 2; ref. 42, $P_{\text{mod}} = 0.004$, $P_{\text{hist}} = 0.046$). This would be expected if purifying selection is purging genetic variation in the developmentally important stomatal genes.

Analyses of non-synonymous and synonymous SNPs in stomata and control genes further show that in stomata genes there are significantly fewer variants annotated as potentially affecting protein function, which we expect to be generally detrimental (partial or full loss of function (LOF) from missense, frameshift or gain of stop codon), compared with likely neutral variants (located in introns, untranslated regions (UTRs) or degenerate codons; annotation with SnpEff⁴³; $P_{\text{mod}}^{\text{non-syn}} = 0.011$, $P_{\text{hist}}^{\text{non-syn}} = 0.003$, $P_{\text{mod}}^{\text{LOF}} = 0.003$, $P_{\text{hist}}^{\text{LOF}} = 0.049$; Fig. 1c and Supplementary Table 3). Despite the overall low variation, in the modern data all but one stomatal gene do harbour non-synonymous variation (from 4 (*CLE10*, AT1G69320) to 124 (*CYCD7;1*, AT5G02110)), and some even contain LOF variants at low frequency (from 1 to 4, in 23 out of 43 genes; Fig. 1e and Supplementary Table 2). We hereafter refer to non-synonymous and LOF variants jointly as putatively functional variation. As expected, the transcription factors that are essential for stomatal development are among the genes with the lowest genetic variation (*SPCH* = 0.006 SNPs per bp, *MUTE* = 0.008 SNPs per bp, *FAMA* = 0.008 SNPs per bp; Fig. 1e, genes ordered by modern data-based SNPs per bp, from left to right, top to bottom). We concluded that stomata genes generally seem to be under purifying selection—especially master transcription factor regulators that lack variation and are unlikely involved in local adaptation—although some genes still harbour non-synonymous variants at low frequency that could have strong phenotypic effects.

This led us to think of the stomatal pathway as composed of two main groups of genes. One group represents highly conserved essential core genes, where the functional loss of any single gene drastically affects stomatal morphology and density or is even lethal^{44–46}. Here, examples are the above-mentioned non-redundant master transcription factors *SPEECHLESS* (AT5G53210), *MUTE* (AT5G53210) and *FAMA* (AT3G24140). The second group consists of their direct and indirect regulators that fine-tune the core pathway's activity and outcome, including duplicated, redundant genes in the pathway that as pairs—but not individually—are similarly essential (Fig. 1a). As loss or functional changes of single 'regulator' genes have on average less impact on overall plant development, they are under less evolutionary constraint and thus may be more likely to harbour genetic variation available for positive selection.

Local adaptation signals in stomata regulator genes

If any of the functional natural variation in stomatal genes is adaptive, population genetics statistics might detect potential signals of selection and population differentiation. We used a battery of Tajima's D and F_{ST} statistics among populations grouped by population history, phenotypes or climates. Tajima's D , which aims to distinguish between neutrally evolving polymorphic sites and those that may be under positive or balancing selection⁴⁷, does not significantly differ between the groups of stomatal and control genes (Fig. 1d; $P_{\text{mod}}, P_{\text{hist}} > 0.1$). However, when comparing each individual stomatal gene with its specific control distribution, we saw a negative Tajima's D value for the genes encoding essential core transcription factors *SPCH*, *MUTE* and *FAMA* (Fig. 1f, Supplementary Fig. 2 and Supplementary Table 2). This fits with their key roles as necessary and sufficient drivers of stomatal

development^{44–46,48}. By contrast, we identified several outliers among the regulatory genes with both high π , high positive Tajima's D and high population differentiation measured by Wright's F_{ST} between geographically separated *A. thaliana* populations ($F_{\text{ST}}^{\text{kgroup}}$, based on $k = 11$ groups⁴⁹). The combination of high species-wide Tajima's D and high cross-population F_{ST} may reflect above-average differentiation of alleles between populations relative to within-population differentiation, and maintenance of multiple alleles of the same gene that may be involved in local adaptation. We find indications of this in several genes, all among the 10% highest values of the control distribution for either one or both Tajima's D and $F_{\text{ST}}^{\text{kgroup}}$ (Fig. 1f and Supplementary Fig. 2): *EPFL1* (AT2G20875), *ERL2* (AT5G07180)—which also has the second highest π after *CYCD7;1*, *ICE1* (AT3G26744), *SCRM2* (AT1G12860) and *STOMAGEN/EPFL9* (AT4G12970).

If these outlier genes are indeed involved in local adaptation, the distribution of their genetic variation might follow environmental (climate) gradients such as variability in temperature and humidity^{16,19}. We tested this by grouping samples into populations either based on their collection locations' precipitation and temperature seasonality (BIOCLIM 4 and BIOCLIM 15; ref. 50) or based on plant life history traits typically aligned with climate, such as the timing of germination or flowering for optimal survival and reproductive success in a given environment (Fig. 1g,h and Supplementary Fig. 3; ref. 51). Using F_{ST} statistics, we then assessed genetic differentiation between the resulting populations. Despite variation in the absolute F_{ST} values, there is clear overlap in the genes that are most differentiated between populations, independent of using population ancestry groups to compute a traditional F_{ST} , or climate or life history to delineate said populations, with *SCRM2* and *EPFL1* among the top four climate differentiators (Fig. 1g and Supplementary Fig. 3) and *SCRM2* and *ERL2* in the top four life-history genes (Fig. 1h and Supplementary Fig. 3; see also Supplementary Text 1 and Supplementary Fig. 6b–d for stomatal gene differentiation over time). A single LOF of the top four adaptation candidates *SCRM2*, *ERL2*, *EPFL1* and *ICE1* tends to minimally affect plant development beyond the stomatal context^{52–54}, with the exception of *ICE1*'s role in endosperm breakdown and embryo development⁵⁵. Interestingly, the basic Helix-Loop-Helix transcription factors and paralogs *ICE1* (*SCRM*) and *SCRM2*, both involved in cold tolerance, act redundantly as direct interaction partners and expression regulators of *SPCH*, *MUTE* and *FAMA*⁵⁴. This fits the expectation of strong purifying selection acting in indispensable master transcription factors and local adaptation through contribution of flexible regulators.

Functional score recovers phenotype changes across geography

Although the stomatal development pathway at first glance may appear relatively simple, leaf stomatal density is a complex trait (Supplementary Text 2). It is likely affected by a combination of (environmental) factors and complex fine-tuning of the aforementioned core and regulatory genes and likely many others. Genome-wide associations of stomatal density thus far yielded moderate heritabilities across many genomic regions, explaining only small fractions of the variation with high uncertainty on specific causal variants¹⁶. As numerous previous studies provide causal evidence that well-known stomatal genes affect the trait, we study these genes' SNPs located in coding regions where we infer protein amino-acid composition change, that is, non-synonymous changes leading to loss or gain of function. If such putatively functional SNPs are then under climate-driven natural selection, we should expect them to follow geographic and historical climate gradients and that the non-reference functional variants in multiple genes appear in concerted fashion⁵⁶. Indeed, when visualizing the distribution of such variants in six genes with the most putatively functional SNPs, they appear geographically segregated (Fig. 2a).

In our set of 43 stomatal development genes, 24 have confirmed phenotypes in increasing or decreasing abaxial leaf stomatal density

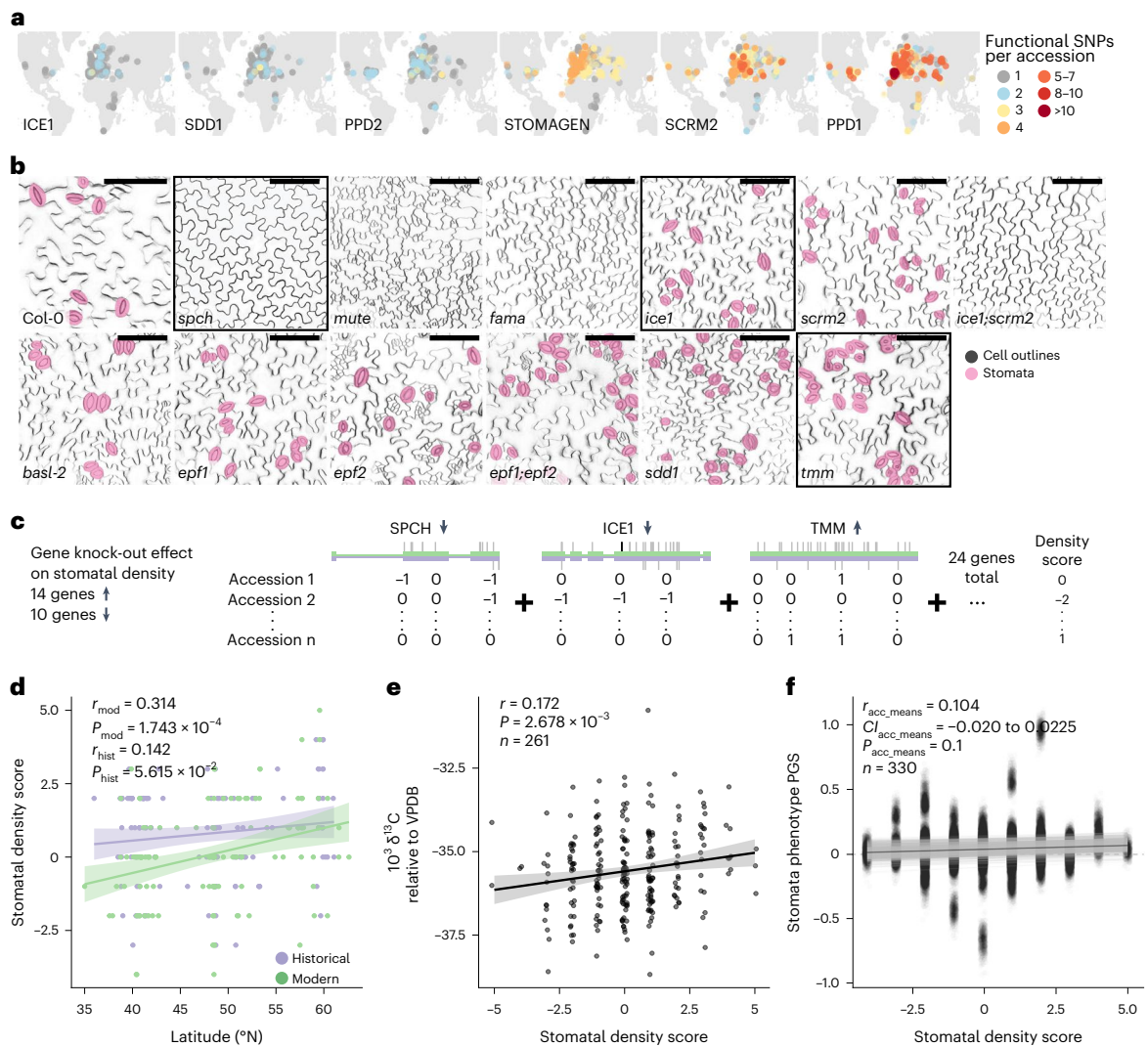


Fig. 2 | Functional score predicts stomatal density patterns. **a**, Geographic distribution of functional (non-synonymous, LOF) SNP accumulation in historical and modern samples for six genes with the overall highest amounts of putatively functional SNPs, overlaid on a continental map. Colour gradient from grey to dark red indicates samples with 1 to over 10 functional SNPs. **b**, Gradient of stomatal density differences resulting from loss of major stomatal development genes visualized with confocal microscopy. Stomata are false-coloured in magenta; scale bars, 100 μm . Black frames around three example genes used in **c**. **c**, Schematic overview detailing the generation of the experimentally informed stomatal density score. Of 24 genes with known effect on stomatal density, loss of 14 increases and loss of 10 decreases stomatal density. Putatively functional SNPs are assigned a ‘-1’ when located in genes whose loss decreases density and a ‘+1’ in genes whose loss increases density. Density scores for each historical and modern sample are calculated as the sum of these values across the 24 genes, counting a single functional SNP per gene. **d**, Linear regression ($\pm\text{s.d.}$)

of the stomatal density score with paired samples’ latitude of origin, separated into historical (purple) and modern (green) samples (for each $n = 126$, one-sided Pearson’s correlation test $P_{\text{mod}} = 1.743 \times 10^{-4}$, $P_{\text{hist}} = 5.615 \times 10^{-2}$, correlation coefficient $r_{\text{mod}} = 0.314$, $r_{\text{hist}} = 0.142$). Analyses exclude samples from North America and the African continent. **e**, Correlation ($\pm\text{s.d.}$) of the density score with the $\delta^{13}\text{C}$ measurement, a proxy of WUE, in 261 *A. thaliana* accessions. This is based on isotope amount ratios of stable carbon isotopes $^{13}\text{C}/^{12}\text{C}$, expressed as ‰ against the Vienna Pee Dee Belemnite (VPDB; ref. 111) standard (one-sided Pearson’s correlation test $P = 2.678 \times 10^{-3}$, correlation coefficient $r = 0.172$; $\delta^{13}\text{C}$ (stable carbon isotope ratio) data from ref. 16). **f**, Correlation ($\pm\text{s.d.}$) of the stomatal density score with genome-wide association-based traditional PGS for stomatal density (one-sided Pearson’s correlation test, positive correlation 997/1,000 re-trainings, 141/1,000 significant one-sided Pearson’s correlation tests with $P < 0.05$, correlation coefficient $r_{\text{median}} = 0.092$; stomatal density data from ref. 16).

(mostly shown by knock-out mutant analysis; see exemplary phenotypes in Fig. 2b; refs. 44–46,48,54,57–69). We combined this information on functional variation in the 24 stomatal density genes in our historical and modern samples with the genes’ putative effects on increasing or decreasing stomatal density. With this genetic and phenotypic information, we devised a simple cumulative ‘stomatal density score’ to indicate increasing (positive) or decreasing (negative score) stomatal density compared with the baseline accession Col-0 (in which functional variation is typically characterized). This functional score is conceptually similar to polygenic scores but is based on summing over

putatively functional variation of genes with known increase/decrease effects on a trait (for distinction from genome-wide association (GWA), see also Supplementary Text 3). Putatively functional SNPs are assigned a ‘-1’ in genes whose knock-out decreases stomatal density (10 genes) and a ‘+1’ in genes whose knock-out increases stomatal density (14 genes; Supplementary Table 4), where genes with multiple SNPs are only counted once (Fig. 2c). This recovered a gradient of ‘stomatal density scores’ (Fig. 2d–f).

We found that this density score significantly correlates with modern samples’ latitude of origin, which in turn dictates life cycle

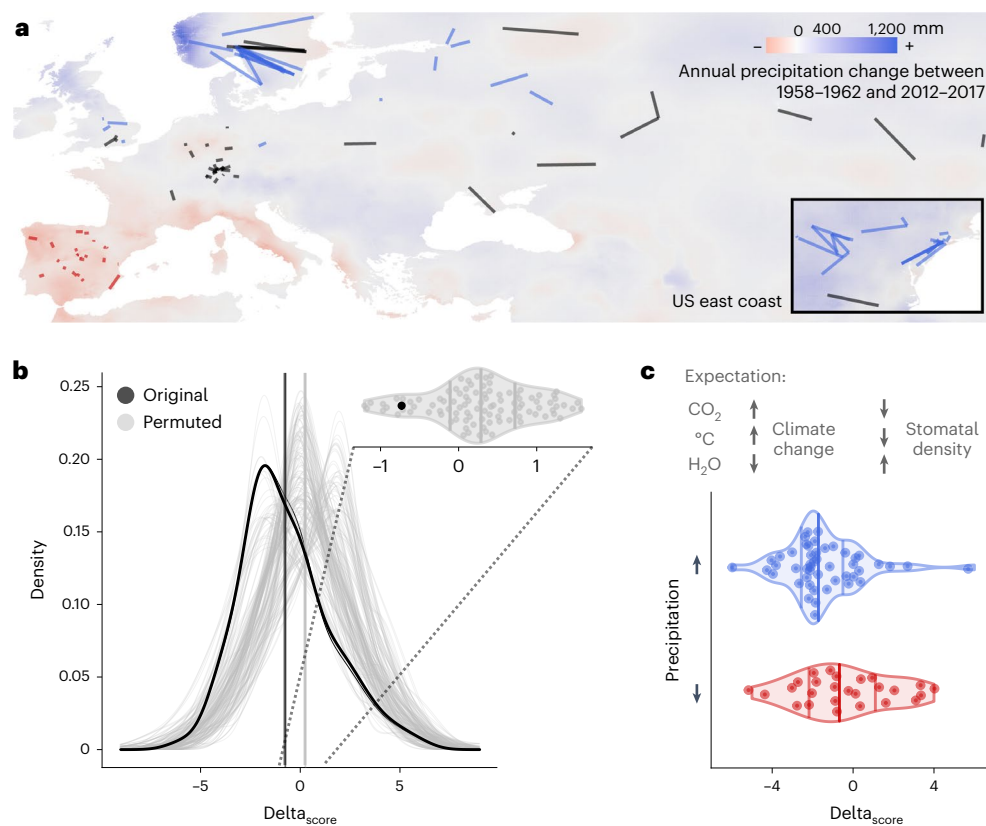


Fig. 3 | Stomatal density decrease over time fits climate change expectations.

a, Map with historical and modern sample pairs as used for stomatal density change analyses (in **b** and **c**). Connecting lines between sample pairs are coloured by the precipitation change in the sample locations; red indicates a significant decrease in precipitation from 1958 to 2017 in both the historical and modern sample location, blue indicates a significant increase in precipitation, and black no significant change or different changes in paired locations. Background colour gradient indicates change in the mean annual precipitation between 1958–1962 and 2012–2017, with colours as above. Inlay shows sample pairs located on the North American East Coast. **b**, Distribution of per sample-pair calculated difference in stomatal density scores ($\text{delta}_{\text{score}} = \text{modern} - \text{historical}$) for original data (black) and 100 permutations (grey) between genes and their assigned effect (decrease/increase) on stomatal density, with genetic variation itself remaining un-permuted; mean $\text{delta}_{\text{score}} = -0.730$, Wilcoxon signed rank test, $P < 2.2 \times 10^{-16}$. Density distribution means are marked by solid black and

grey vertical lines. Violin plot of the distribution means, with the non-permuted mean $\text{delta}_{\text{score}}$ lower than 92/100 permutations. **c**, Expected effects of climate-change-related shifts in $[\text{CO}_2]$, temperature and water availability on stomatal density (based on published experiments, for example, refs. 17–20). Change in the stomatal density score ($\text{delta}_{\text{score}} = \text{modern} - \text{historical}$) in sample pairs with significantly increased (blue, $n = 48$) or decreased (red, $n = 26$) precipitation in geographic locations of origin (excluding sample pairs where precipitation did not change significantly or where a pair's locations did not change in the same way). Horizontal lines indicate 0.25, 0.5 and 0.75 distribution quantiles. Increased, but not decreased, precipitation is significantly associated with decreased $\text{delta}_{\text{score}}$ (linear regression, $\text{delta}_{\text{score}} \approx \text{precipitation}_{\text{directionality}}$, $P_{\text{incr.precipitation}} = 0.024$, $P_{\text{decr.precipitation}} = 0.889$, $P_{\text{model}} = 0.045$; Supplementary Table 9). Analyses include samples from North America and exclude samples from the African continent as well as pairs between island and mainland samples.

length (Fig. 2d; one-sided Pearson's correlation test $P_{\text{mod}} = 1.743 \times 10^{-4}$, $P_{\text{hist}} = 5.615 \times 10^{-2}$, correlation coefficient $r_{\text{mod}} = 0.314$, $r_{\text{hist}} = 0.142$; correlation consistently significant upon population structure correction; Supplementary Text 4). In fact, experimentally measured stomatal density also follows a latitudinal trend attributed to *A. thaliana*'s ecological and life-history adaptations to latitudinal climate gradients (Fig. 2d and Supplementary Fig. 4a,b; refs. 16,51,70). Despite the score being derived from simple functional single-gene knock-out experiments, we found that the score correlates positively with a measure of integrated life-time WUE based on CO_2 exchange and water loss (Fig. 2e; $\delta^{13}\text{C}$, difference in the ratio of ^{13}C to ^{12}C compared to a standard; one-sided Pearson's correlation test $P = 2.678 \times 10^{-3}$, correlation coefficient $r = 0.172$). A positive trend was found with stomatal density measured in modern *A. thaliana* populations (Supplementary Fig. 4c; $\delta^{13}\text{C}$ and stomatal density data from ref. 16). Permuting the functional (positive or negative) effects of the 24 genes and recomputing the density score removed all significant relationships above (Supplementary Fig. 4d,e), indicating that the density score is not a result of internal dataset biases or population structure. Finally, we compared our functionally derived density score with a traditional GWA-based polygenic score (PGS; refs. 71,72) based

on published stomatal density data¹⁶. PGS trained on 80% of the data explain on average 9.2% of the variance in the reserved 20% of phenotype data (positive correlation in 1,000 out of 1,000 re-trainings, 745 out of 1,000 significant one-sided Pearson's correlation tests with $P < 0.05$, correlation coefficient $r_{\text{min-max}} = 0.010\text{--}0.561$, $r_{\text{median}} = 0.304$; Supplementary Fig. 5a,b). They also correlated with the functionally derived density score (on average 0.9% of the density score's variance explained; positive correlation in 997 out of 1,000 re-trainings (genome-wide), 141 out of 1,000 significant one-sided Pearson's correlation tests, $P < 0.05$, correlation coefficient $r_{\text{min-max}} = -0.020\text{--}0.204$, $r_{\text{median}} = 0.092$; Fig. 2f and Supplementary Fig. 5c (only stomata genes); correlations with experimentally measured density and density score lost upon randomization of the phenotype–genotype associations in the training dataset; Supplementary Fig. 5d,e and Supplementary Table 5).

Functional score predicts reduction of stomatal density

Ultimately, we aim to understand how stomatal variation in *A. thaliana* may have changed over the past centuries of climate change. For this, we calculated our stomatal density score on historical genomes. To avoid geographical biases past and present, we made comparisons within a

subset of historical and modern samples, paired based on geographic proximity (126 pairs, minimum–maximum distance = 0.5–495 km, mean = 139 km; median = 61 km; Pearson's correlation of sample pairs' latitudes of origin, $r = 0.990$, $P < 2.2 \times 10^{-16}$; Fig. 3a, Supplementary Fig. 6a and Supplementary Table 6). The historical set for these comparisons ranged between 1817 and 2002 (mean collection year 1923) and the modern set between 1992 and 2012 (mean 2001), with a mean age difference of 79 years in historical–modern sample pairs (from 1 to 182 years). Analyses included only shared SNPs across historical and modern data, to avoid biases related to dataset-specific 'private' SNPs.

Before calculating temporal trends, we established that the stomatal density score correlation with latitude remains consistent also in the historical dataset, suggesting patterns of local adaptation have long been established in the species (see historical trends in Supplementary Fig. 4a,b,d,e). We then calculated the difference in density score (Δ_{score}) for each individual historical–modern geographic sample pair (Fig. 3a). The resulting distribution of score changes corroborates previous results that suggest a decrease in stomatal density in modern samples (black distribution, mean $\Delta_{\text{score}} = -0.730$, Wilcoxon signed rank test, $P < 2.2 \times 10^{-16}$; Fig. 3b and Supplementary Fig. 7d). This decrease was robust to reanalyses with population group subsampling by removal of samples along a latitudinal gradient or of modern samples with uncertain collection dates (mean $\Delta_{\text{score}}^{\text{noUS}} = -0.433$, $\Delta_{\text{score}}^{\text{noUS,noSpain}} = -0.423$, $\Delta_{\text{score}}^{\text{noScandinavia}} = -0.745$, $\Delta_{\text{score}}^{\text{noUncertainty}} = -0.717$, Wilcoxon signed rank test, all $P < 1.25 \times 10^{-15}$; Supplementary Fig. 7a–c, Supplementary Table 7 and 8 and Supplementary Text 5). Recalculated temporal trends using permuted positive and negative stomatal density gene effects did not recover the originally predicted decrease in stomatal density. This suggests that the trend is not due to confounding structure of the data or population history but rather results from the coordinated gain or loss of non-synonymous nucleotide variation in positive and negative stomatal density regulator genes (Fig. 3b and Supplementary Fig. 7a–d, grey distributions).

Such a putative decrease in stomatal density over the past century would fit the expectation based on experimental evidence showing that increases in $[\text{CO}_2]$ and temperatures lead to leaves with lower stomatal density^{6,17,20,23,25,73,74}. However, while $[\text{CO}_2]$ and temperature have mostly increased, climate change has also altered weather patterns and heterogeneously increased or decreased precipitation depending on the region on Earth¹. Because experiments simulating decreased water availability show directly opposite stomatal responses (that is, density increases with less water availability; refs. 18,75), we wondered whether we could use the differential changes in precipitation as a 'natural counterfactual' experiment to validate our genetic prediction of stomatal changes over time. We thus conducted a sensitivity analysis, grouping sample pairs based on their temperature and precipitation trajectories and magnitudes of change over the past 60 years (Fig. 3a,c). Sample pairs from locations with increased precipitation are significantly more likely to have decreased stomatal density scores over time (Fig. 3c, Supplementary Table 9 and Supplementary Fig. 7e; almost 2.5-fold odds of stomatal decrease with Fisher's exact test, $\text{odds}_{\text{ppt,high}} = 2.291$ (ppt = precipitation), $P_{\text{ppt,high}} = 0.05$, and consistent results for the same analysis with less stringent filters in Supplementary Fig. 7f; ref. 76; Supplementary Table 10 and Supplementary Text 5). Locations with decreased precipitation, a change counteracting the expected effects of increased $[\text{CO}_2]$ and temperature, showed a (non-significant) trend of stomatal density increase (Fisher's exact test, $\text{odds}_{\text{ppt,low}} = 0.589$, $P_{\text{ppt,low}} = 0.283$, Supplementary Table 10; for example, refs. 18,75). This observation is consistent with the hypothesis that *A. thaliana* under drought would favour more and smaller stomata, as these open and close more rapidly.

Despite the spatial sample pairing and climate splits in our analyses, temporal trends in stomatal density score changes might still contain residual biases such as population structure. We therefore conducted a series of analyses including fitting the first three main

genomic principal components (PCs) to capture population structure and describe the consistent stability of various estimates (Supplementary Text 5). The signal of lowering average stomatal density remains after population structure correction ($P_{\text{model2}} = 0.001$; Supplementary Text 5 and Supplementary Table 9), and also the more pronounced decrease in stomatal density in regions with increased precipitation is mostly consistent after corrections for each genomic PC axis ($P = 0.034$, 0.081, 0.019, after PC1, PC2, PC3 corrections, while no variable is significant with a full model; Supplementary Text 5 and Supplementary Table 9). Taken together, despite the noise inherent to naturally evolving structured populations and the counteracting effects of $[\text{CO}_2]$ and temperature versus precipitation, there is a consistent signal-to-noise ratio in the identified trends of decreasing stomatal density. These trends correspond well with both experimental and historical observations of stomatal density responses to the climate variation connected with global change.

Conclusions and outlook

Global change has led to rapid and drastic changes of multiple climate parameters—atmospheric CO_2 concentration, temperature, precipitation—that have a strong impact on plant development. Here we studied responses of *A. thaliana* leaf stomatal development to anthropogenic climate change, using historical herbarium genomes as witnesses of this multi-factorial 'global change experiment' (for example, refs. 6,77). Despite the overall high conservation of stomatal development genes, our integrative approach allowed us to identify evolutionary signals in several of these genes that are consistent with local adaptation across *A. thaliana*'s geographic range, suggesting that the well-described stomatal development pathway itself could also evolve to climate change conditions. We developed a novel functional score based on functional and experimental molecular knowledge of stomatal development genes that agrees with a historical trend of stomatal density decrease in *A. thaliana*, classically observed across species but of unknown genetic basis^{6,24}. This trend is consistent with experiment-based expectations and climate counterfactuals but may certainly be influenced by additional factors beyond climate change itself. While evolutionary processes can be fast (for example, refs. 40,78,79), our analyses show that even with hundreds of historical genomes, the described trends of putative adaptive evolution are significant yet close to the detection limit above the noise of genetic drift, phenotypic plasticity, counteracting climatic factors and methodological idiosyncrasies, as seen with the re-analyses of different geographical subsets of the data. Our discovery will stimulate follow-up investigations such as studying the molecular mechanisms of how these genes promote adaptation—now enabled, for instance, by single base-editing with CRISPR (clustered regularly interspaced short palindromic repeats) to recreate historical variants^{80,81}—characterizing cellular phenotypes directly from historical specimen tissues using customized microscopy techniques (for example, ref. 82) and generating denser timelines of historical genomes possible with high-throughput ancient DNA technologies^{29,83}. The functional historical genomics approach presented here adds an exciting avenue to leverage the power of genomics to reconstruct phenotypic impacts of climate change on species, even for those phenotypes that are cellular or sub-cellular. Ultimately, our approach could help uncover complex responses involving WUE, photosynthetic capacity or drought resistance that are not preserved or measurable in historical collections. Disentangling this will be key to understanding plants' past and future adaptive potential and design targets for engineering plants for the future.

Methods

Sequence data

Contemporary data. Published⁴³ *A. thaliana* SNPs³⁶ were annotated (TAIR10_GFF3_genes_transposons.gff; ref. 84). We extracted 'genic' SNPs using bash, BCFtools v1.10.2⁸⁵, PLINK v1.90b6.16 64-bit^{71,86} and R⁸⁷

within RStudio v1.2.1335 (ref. 88). Samples lacking geographic origin coordinates or collection year were excluded as indicated.

Herbarium data. Sample processing and sequencing. African ($n = 9$), North American ($n = 33$), German ($n = 34$) and broadly geographically distributed ($n = 204$) *A. thaliana* historical whole-genome sequences were downloaded^{34,35,40,89}.

Historical sequencing was processed as described^{28,90}. We mapped the merged reads (Adapterremoval v2.3.1; ref. 91; BWA v0.7.15-r1140; ref. 92) to TAIR10 (ref. 84), filtered for mapping quality ≥ 20 (Samtools v1.9; ref. 85), removed PCR duplicates (DeDup v0.12.8; ref. 93) and confirmed the authenticity of historical samples with fragmentation (Samtools v1.9; ref. 85; median merged fragment length of 98 bp, insert sizes for sheared, unmerged fragments of 234 bp) and deamination patterns (MapDamage v2.2.1; ref. 94; Supplementary Fig. 1). For double-processed (sheared and unsheared) broadly geographically distributed samples³⁵, we concatenated the merged fraction of unsheared samples with the merged and unmerged fractions of sheared samples.

Based on sequencing and mapping statistics from Samtools (stats), DeDup and MapDamage, we defined quality thresholds. We retained samples with $>1,000,000$ bp or $>10,000$ reads sequenced, $>95\%$ of reads mapped, a duplication rate <0.3 and an error rate <0.02 for de novo SNP calling and excluded samples with >0.5 missing genotypes, resulting in a total of 191 samples with an $\sim 1.5\text{--}39\times$ genome coverage (mean $7.1\times$, median $5.9\times$). We did not differentiate between ‘regular’^{34,35} and UDG-treated libraries^{40,89}. Samples lacking geographic origin or collection year were excluded as indicated.

De novo SNP calling. For genome-wide de novo calling of SNPs in the historical dataset, we created and indexed a reference dictionary with TAIR10⁸⁴ using Picard’s CreateSequenceDictionary (Picard v2.18.29-0; ‘Picard Toolkit.’ 2019. Broad Institute, GitHub Repository. <https://broadinstitute.github.io/picard/>; Broad Institute) and samtools faidx (Samtools v1.9,⁸⁵). We called variants across our herbarium samples by calling haplotypes for individual samples (GATK HaplotypeCaller), combining .gvcf files into a single input file (GenomicDBImport) and calling variants de novo (not based on an existing SNPs set, GenotypeGVCF) following GATK best practice recommendations (gatk4-4.2.0.0-0; ref. 95). Keeping only SNPs (vcftools --remove-indels, VCFtools/0.1.16; ref. 96), we either included only or excluded all sites present in the published 1001G vcf file (vcftools --positions³⁶). To determine parameters for quality-based SNP filtering, we compared quality parameters in the two subsets and the full dataset (vcftools --gzvcf <infile.vcf.gz> --get-INFO QD --get-INFO FS --get-INFO ReadPosRankSum --get-INFO MQRankSum --get-INFO BaseQRankSum --out <outfile>) using R. Assuming that the distributions of quality parameter values of the 1,001-only dataset are representative for high-quality SNPs, we defined cut-offs for SNP filtering with vcflib (vcflib -f DP > 22 & FS $< .2$ & ReadPosRankSum $> (0-2)$ & ReadPosRankSum < 2 ; vcflib/20161123-git,⁹⁷). Briefly, DP refers to the combined depth of a site across all samples; FS indicates strand bias as estimated by Fisher’s exact test; and ReadPosRankSum compares the positioning of reference versus alternative alleles within a read, allowing to filter alleles biased to read ends. Subsequently, we excluded samples with missing call frequencies $>50\%$ and filtered for biallelic positions with PLINK (plink --mind .5; PLINK v1.90b6.16 64-bit^{71,86}). We used PLINK to remove variants where <3 individuals carried the alternative allele (--mac 3) and samples with site missingness $>15\%$ (--geno .15; Supplementary Fig. 8). We used variant-based PCAs to visualize filters’ effects on samples’ genetic diversity and identify outliers (PLINK --pca <sample_number>). The resulting vcf files were annotated using SnpEff v5.0e⁴³ and the TAIR10 *A. thaliana* reference genome⁸⁴ and further annotated and subset as described above for contemporary data.

Joined historical and modern data. Before merging, the historical and modern datasets were individually filtered with PLINK for alleles

with a minimum minor allele count of 3 and for samples with $<15\%$ site missingness (‘De novo SNP call’). Datasets were then intersected using BCFtools (bcftools isec -n = 2, v1.10.2⁸⁵) and filtered for shared SNPs (PLINK --extract <shared_SNPs>) to avoid dataset-specific biases. Gene-specific and control subsets were generated as described above.

For direct comparisons of historical and modern samples, we subset the datasets to 126 geographically matched sample pairs (‘Geographic-distance based sample pairing’) and re-filtered (PLINK --keep-geno .15 --mac 3).

Gene-specific subsets. Stomatal gene list. From the literature^{31,98,99} and lab-internal experience, we generated a list of 43 genes that are central to stomatal growth, development and the stomatal lineage (Supplementary Table 2).

Control gene lists. The majority of our analyses focus on per-gene statistics and are sensitive to differences in the number of SNPs, which is strongly correlated with gene length. Control genes were thus selected to match the gene length distribution of the original dataset. For each stomatal focus gene, we subsampled all genes of the same length ($\pm 2.5\%$), randomly picked one gene and generated a list of 43 control genes for each stomatal focus gene (up to 1,000 re-samplings per analysis).

Subsetting. SNPs were extracted from vcf files based on stomata-specific and control gene lists using bash and PLINK v1.90b6.16 64 bit. We filtered annotations for the first and second splice isoforms of the focus genes. SNP types (synonymous, non-synonymous, LOF and so on, based on TAIR/SnpEff annotation) were assessed based on the first splice isoform, grouping several SNP-types together: ‘loss-of-function’ (disruptive_inframe_deletion, disruptive_inframe_insertion, inframe_deletion, inframe_insertion, frameshift, start_lost, stop_lost, stop_gained), ‘non-synonymous’ (missense), ‘synonymous’ (synonymous), ‘UTR’ (5_prime_UTR_premature_start_codon_gain, 3_prime_UTR, 5_prime_UTR), ‘intron’ (intron) and ‘other’ (none, splice_region, splice_donor, splice_acceptor, intron, stop_retained, non_coding_transcript_exon, upstream_gene, downstream_gene).

Population genetics statistics

Population genetics analyses were conducted per focus group and per gene to assess whether stomatal development genes individually or as a group are outliers. To compare stomatal developmental genes (focus genes) as a group against the control genes (‘Control gene lists’), we calculated mean values for each of the $1,000 \times 43$ control groups and assessed whether the stomatal gene group value was higher or lower than the majority of control group means. For per-gene assessments, we calculated mean values for each focus gene and each of 1,000 gene-specific control genes. For comparing, we calculated whether the focus gene lies outside of the 0.1 or 0.9 quantile of the control value distribution. All summary calculations, statistical analysis and plotting were done in R/RStudio.

Genetic diversity— π and SNPs per gene. We summed SNPs for each gene and calculated raw SNPs/bp (that is, per gene length) and Watterson’s θ . To estimate nucleotide diversity π per polymorphic site, we used VCFtools default settings (<--site-pi >; VCFtools/0.1.16,⁹⁶) and extracted the maximum per-site π value per gene. We also calculated $\frac{\pi}{\text{bp}}$ by dividing the sum of all π values for a single gene by the gene length (bp), assuming that positions lacking a π value are invariant.

Tajima’s D . To identify signals of selection in our genes of interest, we calculated Tajima’s D ⁴⁷ on the full 1001G vcf file (--TajimaD 100; VCFtools/0.1.16; ref. 96). We then associated these values to our biallelic SNPs of interest and summarized them using R and RStudio. We calculated the mean Tajima’s D value per gene for both length-matched control and focus gene sets.

For comparison of Tajima's D between historical and modern samples, we calculated Tajima's D for the independent subsets present in the geographically paired dataset ('Joined historical and modern data').

Population fixation index F_{ST} . Population differentiation (fixation index F_{ST}) was estimated with Plink (plink --fst --within <groups>; PLINK v1.90b6.16 64 bit; refs. 71,86). Population subdivisions for the ('modern') 1001G dataset were based on either published, genetics-based global *A. thaliana* population structure (11 groups⁴⁹), on accessions' life strategies⁵¹ or on experienced (micro-)climates⁵⁰. For the former, admixed individuals as defined in ref. 36 were excluded. For the latter two, accessions were grouped by temperature and precipitation seasonality of their geographic origins (WorldClim.org) or by their environment-independent germination rate and cold-induced dormancy (see refs. 51,100, excluding accessions with imputed phenotypes). Based on accessions' silhouette scores (cluster R-library, silhouette, <https://CRAN.R-project.org/package=cluster>; ref. 101; R package version 2.1.0.) for 2 through 15 clusters, we identified and used the number of clusters with the highest score (6 and 4, respectively) in a single k -means clustering. Environmental and life history-based F_{ST} was calculated as described for stomatal genes and 1,000 control re-samplings.

Using F_{ST} to assess the genetic differentiation between populations not across (geographic) space but over time, we also calculated F_{ST} for the paired dataset ('Joined historical and modern data'), with the samples' identity as 'modern' or 'historical' defining the two populations.

Matching historical and modern samples

Geographic-distance-based sample pairing. Historical samples were matched to the geographically closest modern sample from the 1001G *A. thaliana* dataset³⁶. For each historical sample, we calculated pairwise distances to the geographic origin of all modern samples (geosphere R-library, distVincentyEllipsoid; ref. 102; R package version 1.5-10; <https://CRAN.R-project.org/package=geosphere>). We selected the closest modern sample and removed it from subsequent pairwise distance calculations. Historical-modern pairs >500 km apart were removed, generating a final dataset of 126 pairs. Historical samples from Africa and pairs between islands and mainland or matched across bodies of water or mountain ranges were excluded (Supplementary Table 6). For some analyses, samples from North America (originating west of longitude -25°) were removed, retaining 116 sample pairs. Only sample pairs with the historical collection date predating the modern one were considered.

Climate change trajectories. Despite this close pairing, sample pairs' geographic origins may still be sufficiently far apart to have experienced diverging climatic changes between historical and modern sampling. This may translate into differing selection pressures and genetic makeups already in the past, which complicates attributing genetic differences between historical and modern samples to temporal (adaptive) processes alone. We reduced these confounders by modelling the directionality of climate change between 1958 and 2017 (TerraClimate dataset, resolution ~ 4 km; ref. 76). For the geographic location of each sample, we extracted precipitation and the maximum monthly temperatures (raster R package; ref. 103; R package version 3.4-10; <https://CRAN.R-project.org/package=raster>). Temperatures were then subset to one value per year, the month with the highest (t_{max}) recorded temperature. To extract monthly precipitation trends, we transformed records into a time series and decomposed it to separate the gradual trend over time from the periodic seasonal precipitation variation (R/Rstudio, stats-package). With the annual values for t_{max} and precipitation per sample location, we assessed directionality (increase or decrease) of climate change by extracting the slope of a linear regression of the climate parameter over time. Using Spearman correlation, we calculated the P values of these climate change trends,

corrected for multiple testing (Benjamini-Hochberg), and assigned trends of $P_{BH} < 0.01$ as significant. Sample pairs were then classified as matching (same significant increase or decrease in temperature or precipitation), not showing any significant change, or not matching (opposing directionalities of change, or only one sample showing a significant change in either direction).

Stomatal density

Experimentally informed genetics-based stomatal density proxy.

To reconstruct phenotypes from genotypes, we generated a proxy for stomatal density by summing over variants of involved genes with known functional effect. We defined the effects of mutation in the 43 stomatal genes: higher stomatal density (14 genes), lower stomatal density (10 genes) or none (18 genes; Supplementary Table 4). We then used as protein function affecting (non-synonymous, putative LOF; referred to as putatively 'functional') assigned SNPs ('Gene-specific subsets') to calculate a stomatal density proxy for each historical and modern sample. We refer to this as a 'functional score' to distinguish it from a traditional GWA-based polygenic risk score ('Traditional polygenic score model').

Accessions that carry the reference allele for a functional SNP were assigned a value of '0' (no stomatal density effect). The alternative allele translates into '+1' in a gene whose mutation increases, and '-1' in a gene whose mutation decreases stomatal density. To calculate a density score per sample, we summed these values across all density-affecting genes, taking one functional SNP per gene into account.

The functional score correlated with published experimentally measured stomatal densities and $\delta^{13}C$ (ref. 16). Assessment of the score's correlation with latitude was performed on the full set of 191 historical samples. To account for geographical sampling bias, the modern 1,135 accessions were subsampled 100 times without replacement to mirror the size of the historical dataset. Linear regression was calculated for all sample sets. To avoid extrapolation beyond the geographical space covered by samples, we re-calculated the intercept using the median latitude of the historical and modern sample set, respectively. To assess the contribution of *A. thaliana*'s global population structure to the score, we permuted the association of density phenotypes (increase/decrease) with stomatal genes 100 times. This aimed to validate the stomatal density score's latitudinal trajectory and the density shift over time. For the latter, we subtracted the historical sample's score from its modern match within each geographic distance-based historical-modern sample pair ('Matching historical and modern samples'). This aims to reduce population structure differences between historical and modern samples, increasing the probability that identified differences result from change over time (Supplementary Text 5).

We further subset the sample pairs by their experienced climate change ('Climate change trajectories'). Only locations where the change from 1968 to 2017 is significant were included. With Fisher's exact test for count data (stats::fisher.test(), R) we calculated the odds of stomatal density decrease in modern samples under certain climate conditions, identifying density score changes for each sample pair individually.

Traditional polygenic score model. For each of 1,000 iterations, phenotypes¹⁶ were randomly split 4:1 into training and test sets (scikit-learn sampling¹⁰⁴). With GEMMA (v0.89.1; ref. 105), we calculated genome-wide associations and associations with the 43 stomata genes as described above, using a univariate linear mixed model on the training phenotypes and the joined historical and modern genotypic datasets.

We then generated an additive polygenic score (PGS, Plink (v1.9)) model on the genome-wide/stomata gene associations with the stomata phenotypes, using a P -score threshold < 0.05 to select the most predictive SNPs. Plink and R scripts for PGS analysis followed published methods⁷². Phenotypes predicted by the PGS model were compared with the test set or with the functional stomatal density score

(‘Experimentally informed genetics-based stomatal density proxy’) using one-sided Pearson’s correlation tests. To test whether the PGS models capture the phenotypes’ genetic signatures, we ran 1,000 iterations where the phenotypic data were permuted and reassigned to random genotypes in the training set and subsequently performed GWA and Plink modelling as above. For all comparisons from the 1,000 iterations, we calculated the mean predicted PGS phenotype per accession and tested these phenotypes’ correlation with measured stomatal density or the stomatal density score with one-sided Pearson’s tests.

Plant growth and conditions. *A. thaliana* mutant line seeds were sterilized with chlorine gas (50 ml bleach with 2.5 ml 37% hydrochloric acid) and stratified on MS plates (1/2 MS (Caisson Labs), 1% agar (*w/v*), pH 5.7) for 2 days at 4 °C before transfer to a 22 °C chamber with 16 h light/8 h dark cycles (110 $\mu\text{mol m}^{-2} \text{s}^{-1}$).

The following mutants and transgenic lines were reported previously: *basl-2* (ref. 106), *epf1-1* (ref. 53), *epf2-1* (ref. 107), *ice1-D* (*scrm-D*; ref. 54), *ice1-2* (ref. 54), *mute* (ref. 45), *scrm2-1* (ref. 54), *sdd1-1* (ref. 58), *spch-3* (ref. 44), *tmm-1* (ref. 108), *fama* (ref. 46), *epf1-1;epf2-1* (ref. 107) and *ice1-2;scrm2-1* (ref. 54). Natural *A. thaliana* accessions were published previously (Col-0; for example, ref. 36).

Microscopy and image analysis. To visualize cell outlines, we stained 9 days post germination seedlings of *A. thaliana* mutant lines with FM4-64 (N-(3-triethylammoniumpropyl)-4-(6-(4-(diethylamino)phenyl)hexatrienyl)pyridinium dibromide, ThermoFisher catalogue number T13320). The *spch-3* mutant (ref. 44) contains a plasma membrane marker, *pATML1::mCherry-RC12A*, and was not FM4-64 stained. Cotyledons were imaged on a Leica SP5 confocal microscope with HyD detectors using a 40 \times NA1.1 water objective at a resolution of 1,024 \times 1,024 pixels. Images were post-processed (contrast enhancement and noise reduction) using Fiji (V2.1.0/1.53c; ref. 109) and Adobe Illustrator V26.3.1.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

A. thaliana 1001 variant data were published in ref. 36 and respective kgroups in ref. 49. Historical sequencing data are publicly available (North American accessions⁴⁰, African accessions⁸⁹, German accessions³⁴, broadly geographically distributed accessions³⁵). Supplementary tables are available online via figshare at <https://doi.org/10.6084/m9.figshare.25996414> (ref. 110).

References

- IPCC, 2022: *Climate Change 2022: Impacts, Adaptation, and Vulnerability*. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge Univ. Press, 2022).
- Panchen, Z. A., Primack, R. B., Anisko, T. & Lyons, R. E. Herbarium specimens, photographs, and field observations show Philadelphia area plants are responding to climate change. *Am. J. Bot.* **99**, 751–756 (2012).
- Primack, D., Imbres, C., Primack, R. B., Miller-Rushing, A. J. & Del Tredici, P. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *Am. J. Bot.* **91**, 1260–1264 (2004).
- Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37–42 (2003).
- Walker, A. P. et al. Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric CO₂. *New Phytol.* **229**, 2413–2445 (2021).
- Woodward, F. I. Stomatal numbers are sensitive to increases in CO₂ from pre-industrial levels. *Nature* **327**, 617–618 (1987).
- Parmesan, C. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637–669 (2006).
- Han, S.-K., Kwak, J. M. & Qi, X. Stomatal lineage control by developmental program and environmental cues. *Front. Plant Sci.* **12**, 751852 (2021).
- Bertolino, L. T., Caine, R. S. & Gray, J. E. Impact of stomatal density and morphology on water-use efficiency in a changing world. *Front. Plant Sci.* **10**, 225 (2019).
- Faralli, M., Matthews, J. & Lawson, T. Exploiting natural variation and genetic manipulation of stomatal conductance for crop improvement. *Curr. Opin. Plant Biol.* **49**, 1–7 (2019).
- Berry, J. A., Beerling, D. J. & Franks, P. J. Stomata: key players in the earth system, past and present. *Curr. Opin. Plant Biol.* **13**, 233–240 (2010).
- Vinton, A. C., Gascoigne, S. J. L., Sepil, I. & Salguero-Gómez, R. Plasticity’s role in adaptive evolution depends on environmental change components. *Trends Ecol. Evol.* <https://doi.org/10.1016/j.tree.2022.08.008> (2022).
- Bay, R. A. et al. Predicting responses to contemporary environmental change using evolutionary response architectures. *Am. Nat.* **189**, 463–473 (2017).
- Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A. & Merilä, J. Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* **17**, 167–178 (2008).
- Clausen, J., Keck, D. D. & Hiesey, W. M. Regional differentiation in plant species. *Am. Nat.* **75**, 231–250 (1941).
- Dittberner, H. et al. Natural variation in stomata size contributes to the local adaptation of water-use efficiency in *Arabidopsis thaliana*. *Mol. Ecol.* <https://doi.org/10.1111/mec.14838> (2018).
- Crawford, A. J., McLachlan, D. H., Hetherington, A. M. & Franklin, K. A. High temperature exposure increases plant cooling capacity. *Curr. Biol.* **22**, R396–R397 (2012).
- Vile, D. et al. *Arabidopsis* growth under prolonged high temperature and water deficit: independent or interactive effects? *Plant Cell Environ.* **35**, 702–718 (2012).
- Yan, W., Zhong, Y. & Shangguan, Z. Contrasting responses of leaf stomatal characteristics to climate change: a considerable challenge to predict carbon and water cycles. *Glob. Chang. Biol.* **23**, 3781–3793 (2017).
- Lau, O. S. et al. Direct control of SPEECHLESS by PIF4 in the high-temperature response of stomatal development. *Curr. Biol.* **28**, 1273–1280.e3 (2018).
- Lang, P. L. M., Willems, F. M., Scheepens, J. F., Burbano, H. A. & Bossdorf, O. Using herbaria to study global environmental change. *New Phytol.* <https://doi.org/10.1111/nph.15401> (2018).
- Beerling, D. J. & Chaloner, W. G. Stomatal density as an indicator of atmospheric CO₂ concentration. *Holocene* **2**, 71–78 (1992).
- Van Der Burgh, J., Visscher, H., Dilcher, D. L. & Kürschner, W. M. Paleatmospheric signatures in Neogene fossil leaves. *Science* **260**, 1788–1790 (1993).
- Beerling, D. J. & Chaloner, W. G. Evolutionary responses of stomatal density to global CO₂ change. *Biol. J. Linn. Soc. Lond.* **48**, 343–353 (1993).
- McElwain, J. C. & Chaloner, W. G. Stomatal density and index of fossil plants track atmospheric carbon dioxide in the Palaeozoic. *Ann. Bot.* **76**, 389–395 (1995).
- Yoshida, K. et al. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. *Elife* **2**, e00731 (2013).
- Lang, P. L. M. et al. Hybridization ddRAD-sequencing for population genomics of non-model plants using highly degraded historical specimen DNA. *Mol. Ecol. Resour.* <https://doi.org/10.1111/1755-0998.13168> (2020).

28. Latorre, S. M., Lang, P. L. M., Burbano, H. A. & Gutaker, R. M. Isolation, library preparation, and bioinformatic analysis of historical and ancient plant DNA. *Curr. Protoc. Plant Biol.* **5**, e20121 (2020).
29. Kistler, L. et al. Ancient plant genomics in archaeology, herbaria, and the environment. *Annu. Rev. Plant Biol.* <https://doi.org/10.1146/annurev-arplant-081519-035837> (2020).
30. Burbano, H. A. & Gutaker, R. M. Ancient DNA genomics and the renaissance of herbaria. *Science* **382**, 59–63 (2023).
31. Lee, L. R. & Bergmann, D. C. The plant stomatal lineage at a glance. *J. Cell Sci.* **132**, jcs228551 (2019).
32. Kinoshita, T., Toh, S. & Torii, K. U. Chemical control of stomatal function and development. *Curr. Opin. Plant Biol.* **60**, 102010 (2021).
33. Delgado, D., Alonso-Blanco, C., Fenoll, C. & Mena, M. Natural variation in stomatal abundance of *Arabidopsis thaliana* includes cryptic diversity for different developmental processes. *Ann. Bot.* **107**, 1247–1258 (2011).
34. Latorre, S. M., Lang, P. L. M. & Burbano, H. A. Historical *Arabidopsis thaliana* genomes from Germany. *Zenodo* <https://doi.org/10.5281/zenodo.7156189> (2022).
35. Lopez, L., Marciniak, S., Perry, G. H. & Lasky, J. R. Historical *Arabidopsis thaliana* genomes across its native range. *Zenodo* <https://doi.org/10.5281/zenodo.7187528> (2022).
36. 1001 Genomes Consortium. 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491 (2016).
37. Weiß, C. L. et al. Temporal patterns of damage and decay kinetics of DNA retrieved from plant herbarium specimens. *R. Soc. Open Sci.* **3**, 160239 (2016).
38. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Pääbo, S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS ONE* **7**, e34131 (2012).
39. Lee, C.-R. et al. On the post-glacial spread of human commensal *Arabidopsis thaliana*. *Nat. Commun.* **8**, 14458 (2017).
40. Exposito-Alonso, M. et al. The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.* **14**, e1007155 (2018).
41. Watterson, G. A. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**, 256–276 (1975).
42. Hahn, M. W. *Molecular Population Genetics* (Oxford Univ. Press, 2018).
43. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
44. MacAlister, C. A., Ohashi-Ito, K. & Bergmann, D. C. Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature* **445**, 537–540 (2007).
45. Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L. & Torii, K. U. Termination of asymmetric cell division and differentiation of stomata. *Nature* **445**, 501–505 (2007).
46. Ohashi-Ito, K. & Bergmann, D. C. *Arabidopsis* FAMA controls the final proliferation/differentiation switch during stomatal development. *Plant Cell* **18**, 2493–2505 (2006).
47. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
48. Bergmann, D. C., Lukowitz, W. & Somerville, C. R. Stomatal development and pattern controlled by a MAPKK kinase. *Science* **304**, 1494–1497 (2004).
49. Exposito-Alonso, M. et al. Genomic basis and evolutionary potential for extreme drought adaptation in *Arabidopsis thaliana*. *Nat. Ecol. Evol.* **2**, 352–358 (2018).
50. Hijmans, R. J., Cameron, S. E. & Parra, J. L. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
51. Exposito-Alonso, M. Seasonal timing adaptation across the geographic range of *Arabidopsis thaliana*. *Proc. Natl Acad. Sci USA* **117**, 9665–9667 (2020).
52. Shpak, E. D., McAbee, J. M., Pillitteri, L. J. & Torii, K. U. Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science* **309**, 290–293 (2005).
53. Hara, K., Kajita, R., Torii, K. U., Bergmann, D. C. & Kakimoto, T. The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. *Genes Dev.* **21**, 1720–1725 (2007).
54. Kanaoka, M. M. et al. SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* **20**, 1775–1785 (2008).
55. Denay, G. et al. Endosperm breakdown in *Arabidopsis* requires heterodimers of the basic helix-loop-helix proteins ZHOUP1 and INDUCER OF CBP EXPRESSION 1. *Development* **141**, 1222–1227 (2014).
56. Berg, J. J. & Coop, G. A population genetic signal of polygenic adaptation. *PLoS Genet.* **10**, e1004412 (2014).
57. Yang, M. & Sack, F. D. The too many mouths and four lips mutations affect stomatal production in *Arabidopsis*. *Plant Cell* **7**, 2227–2239 (1995).
58. Berger, D. & Altmann, T. A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in *Arabidopsis thaliana*. *Genes Dev.* **14**, 1119–1131 (2000).
59. Nadeau, J. A. & Sack, F. D. Control of stomatal distribution on the *Arabidopsis* leaf surface. *Science* **296**, 1697–1700 (2002).
60. Boudolf, V. et al. B1-type cyclin-dependent kinases are essential for the formation of stomatal complexes in *Arabidopsis thaliana*. *Plant Cell* **16**, 945–955 (2004).
61. Hara, K. et al. Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in *Arabidopsis* leaves. *Plant Cell Physiol.* **50**, 1019–1031 (2009).
62. Sugano, S. S. et al. Stomagen positively regulates stomatal density in *Arabidopsis*. *Nature* **463**, 241–244 (2010).
63. Lampard, G. R., Wengier, D. L. & Bergmann, D. C. Manipulation of mitogen-activated protein kinase signaling in the *Arabidopsis* stomatal lineage reveals motifs that contribute to protein localization and signaling specificity. *Plant Cell* **26**, 3358–3371 (2014).
64. Gonzalez, N. et al. A repressor protein complex regulates leaf growth in *Arabidopsis*. *Plant Cell* **27**, 2273–2287 (2015).
65. Castorina, G., Fox, S., Tonelli, C., Galbiati, M. & Conti, L. A novel role for STOMATAL CARPENTER 1 in stomata patterning. *BMC Plant Biol.* **16**, 172 (2016).
66. Han, S.-K. et al. MUTE directly orchestrates cell-state switch and the single symmetric division to create stomata. *Dev. Cell* **45**, 303–315.e5 (2018).
67. Vatén, A., Soyars, C. L., Tarr, P. T., Nimchuk, Z. L. & Bergmann, D. C. Modulation of asymmetric division diversity through cytokinin and SPEECHLESS regulatory interactions in the *Arabidopsis* stomatal lineage. *Dev. Cell* **47**, 1–14 (2018).
68. Zoulias, N., Harrison, E. L., Casson, S. A. & Gray, J. E. Molecular control of stomatal development. *Biochem. J.* **475**, 441–454 (2018).
69. Rowe, M. H., Dong, J., Weimer, A. K. & Bergmann, D. C. A plant-specific polarity module establishes cell fate asymmetry in the *Arabidopsis* stomatal lineage. Preprint at *bioRxiv* <https://doi.org/10.1101/614636> (2019).
70. Debieu, M. et al. Co-variation between seed dormancy, growth rate and flowering time changes with latitude in *Arabidopsis thaliana*. *PLoS ONE* **8**, e61075 (2013).
71. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
72. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).

73. Li, Y., Xu, J., Haq, N. U., Zhang, H. & Zhu, X.-G. Was low CO₂ a driving force of C₄ evolution: *Arabidopsis* responses to long-term low CO₂ stress. *J. Exp. Bot.* **65**, 3657–3667 (2014).
74. Samakovli, D. et al. YODA-HSP90 module regulates phosphorylation-dependent inactivation of SPEECHLESS to control stomatal development under acute heat stress in *Arabidopsis*. *Mol. Plant* **13**, 612–633 (2020).
75. Doheny-Adams, T., Hunt, L., Franks, P. J., Beerling, D. J. & Gray, J. E. Genetic manipulation of stomatal density influences stomatal size, plant growth and tolerance to restricted water supply across a growth carbon dioxide gradient. *Phil. Trans. R. Soc. Lond. B* **367**, 547–555 (2012).
76. Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* **5**, 170191 (2018).
77. DeLeo, V. L., Menge, D. N. L., Hanks, E. M., Juenger, T. E. & Lasky, J. R. Effects of two centuries of global environmental variation on phenology and physiology of *Arabidopsis thaliana*. *Glob. Chang. Biol.* <https://doi.org/10.1111/gcb.14880> (2019).
78. Rudman, S. M. et al. Direct observation of adaptive tracking on ecological time scales in *Drosophila*. *Science* **375**, eabj7484 (2022).
79. Franks, S. J., Kane, N. C., O'Hara, N. B., Tittes, S. & Rest, J. S. Rapid genome-wide evolution in *Brassica rapa* populations following drought revealed by sequencing of ancestral and descendant gene pools. *Mol. Ecol.* **25**, 3622–3631 (2016).
80. Tan, J., Zhang, F., Karcher, D. & Bock, R. Expanding the genome-targeting scope and the site selectivity of high-precision base editors. *Nat. Commun.* **11**, 629 (2020).
81. Kang, B.-C. et al. Precision genome engineering through adenine base editing in plants. *Nat. Plants* **4**, 427–431 (2018).
82. Haus, M. J., Kelsch, R. D. & Jacobs, T. W. Application of optical topometry to analysis of the plant epidermis. *Plant Physiol.* **169**, 946–959 (2015).
83. Gutaker, R. M. & Burbano, H. A. Reinforcing plant evolutionary genomics using ancient DNA. *Curr. Opin. Plant Biol.* **36**, 38–45 (2017).
84. Berardini, T. Z. et al. The *Arabidopsis* information resource: making and mining the 'gold standard' annotated reference plant genome. *Genesis* **53**, 474–485 (2015).
85. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
86. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
87. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2019); <https://www.R-project.org/>
88. RStudio Team. *RStudio: Integrated Development for R* (RStudio, 2018); <http://www.rstudio.com/>
89. Durvasula, A. et al. African genomes illuminate the early history and transition to selfing in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **114**, 5213–5218 (2017).
90. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448 (2010).
91. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).
92. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <https://arxiv.org/abs/1303.3997> (2013).
93. Peltzer, A. et al. EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).
94. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
95. Van der Auwera, G. A. & O'Connor, B. D. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. (O'Reilly Media, 2020).
96. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
97. Garrison, E., Kronenberg, Z. N., Dawson, E. T., Pedersen, B. S. & Prins, P. A spectrum of free software tools for processing the VCF variant call format: vcfliib, bio-vcf, cyvcf2, hts-nim and slivar. *PLoS Comput. Biol.* **18**, e1009123 (2021).
98. Chater, C. C. C., Caine, R. S., Fleming, A. J. & Gray, J. E. Origins and evolution of stomatal development. *Plant Physiol.* **174**, 624–638 (2017).
99. Simmons, A. R. & Bergmann, D. C. Transcriptional control of cell fate in the stomatal lineage. *Curr. Opin. Plant Biol.* **29**, 1–8 (2016).
100. Martínez-Berdeja, A. et al. Functional variants of DOG1 control seed chilling responses and variation in seasonal life-history strategies in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **117**, 2526–2534 (2020).
101. Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. *cluster: Cluster Analysis Basics and Extensions*. R package version 2.1.6 <https://CRAN.R-project.org/package=cluster> (2019).
102. Hijmans, R. J. *Geosphere: Spherical Trigonometry* (2019).
103. Hijmans, R. J. *Raster: Geographic Data Analysis and Modeling* (2021).
104. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
105. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
106. Dong, J., MacAlister, C. A. & Bergmann, D. C. BASL controls asymmetric cell division in *Arabidopsis*. *Cell* **137**, 1320–1330 (2009).
107. Hunt, L. & Gray, J. E. The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. *Curr. Biol.* **19**, 864–869 (2009).
108. Yang, M. & Sack, F. The *too many mouths* and *four lips* mutations affect stomatal production in *Arabidopsis*. *Plant Cell* **7**, 2227–2239 (1995).
109. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
110. Lang, P. L. M. et al. Supplementary information for: Century-long timelines of herbarium genomes predict plant stomatal response to climate change. *figshare* <https://doi.org/10.6084/m9.figshare.25996414> (2024).
111. Copley, T. B. et al. New guidelines for delta13C measurements. *Anal. Chem.* **78**, 2439–2441 (2006).

Acknowledgements

We are grateful to members of the Bergmann and Exposito-Alonso labs for support, suggestions, discussion, input for experiments, analysis and reviewing the manuscript, L. Czech for bioinformatics and J. P. Spence for statistics support. This work would not be possible without historical herbarium collections and their curators. We are grateful to C. Krause, A. Rosenbauer and M. Thiv from the Stuttgart State Museum of Natural History and to O. Bossdorf from the Herbarium Tubingense for their introduction to and help in the herbaria and the kind permission to sample specimens; V. Schuenemann and E. Reiter for access to clean-room facilities and technical support; and D. Weigel and the Max Planck Society for generously funding the Research Group for Ancient Genomics and Evolution. P.L.M.L. is supported by a Human Frontiers Science Fellowship (LT000330/2019-L) and the University of California Berkeley. J.M.E. is supported by the Cell and Molecular Biology Training Grant from the National Institutes of

Health (T32GM007276). H.F.F. is supported by a graduate fellowship award from Knight-Hennessy Scholars at Stanford University. L.L. is partially supported by California State University, San Bernardino, and by National Science Foundation award BIO-BRC 2217793. G.A. is supported by funds from the National Institutes of Health (T32 5T32GM007790), the National Science Foundation (DGE-1656518) and a Stanford Graduate Fellowship. H.A.B. and S.M.L. were partially supported by Max Planck Society funding to the Department of Molecular Biology of the Max Planck Institute for Biology led by D. Weigel. J.R.L. was supported by NIH award R35 GM138300. H.A.B. is supported by a Royal Society Wolfson Fellowship (RSWF\R1\191011) and a Philip Leverhulme Prize from The Leverhulme Trust. M.E.-A. is funded by the Carnegie Institution for Science, a Department of Energy, Office of Biological and Environmental Research Grant (DE-SC0021286), the National Institutes of Health's Early Investigator Award (1DP5OD029506-01) and the University of California Berkeley and Howard Hughes Medical Institute. M.E.-A. is a Freeman Hrabowski Scholar of the Howard Hughes Medical Institute. Computation for this project was performed on the Calc, Memex and Moi Node clusters from the Carnegie Institution for Science. D.C.B. is an investigator of the Howard Hughes Medical Institute.

Author contributions

P.L.M.L. conceived and designed the project with input from M.E.-A., H.A.B. and D.C.B. P.L.M.L. designed and performed analyses with contributions from M.E.-A., C.L.W. and J.M.E. L.L., J.R.L., S.M.L. and H.A.B. contributed data. S.M.L. helped with historical sample collection and processing. H.F.F. and J.M.E. contributed microscopy. G.A. and H.A.B. contributed to discussion of results. M.E.-A. provided computational resources. D.C.B. supervised research. P.L.M.L. and J.M.E. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41559-024-02481-x>.

Correspondence and requests for materials should be addressed to Patricia L. M. Lang.

Peer review information *Nature Ecology & Evolution* thanks Eva-Maria Geigl, Arthur Korte and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

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

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

© The Author(s) 2024

¹Department of Biology, Stanford University, Stanford, CA, USA. ²Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. ³Department of Biological Sciences, California State University San Bernardino, San Bernardino, CA, USA. ⁴Department of Biology, Pennsylvania State University, University Park, PA, USA. ⁵Department of Genetics, Stanford University, Stanford, CA, USA. ⁶Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA, USA. ⁷Centre for Life's Origins and Evolution, Department of Genetics, Evolution and Environment, University College London, London, UK. ⁸Research Group for Ancient Genomics and Evolution, Department of Molecular Biology, Max Planck Institute for Biology, Tübingen, Germany. ⁹Department of Plant Biology, Carnegie Institution for Science, Stanford, CA, USA. ¹⁰Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA. ¹¹Present address: Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA. ¹²Present address: Department of Integrative Biology, University of California, Berkeley, CA, USA. ¹³Present address: Howard Hughes Medical Institute, University of California, Berkeley, CA, USA. ✉ e-mail: patriciamlang@gmail.com

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Modern genomic data was downloaded from <http://1001genomes.org/data/GMI-MPI/releases/v3.1/>.

The Arabidopsis thaliana reference genome was downloaded from https://www.arabidopsis.org/download_files/Genes/TAIR10_genome_release/TAIR10_chromosome_files/TAIR10_chr_all.fas.gz, and gene annotations from https://www.arabidopsis.org/download_files/Genes/TAIR10_genome_release/TAIR10_gff3/TAIR10_GFF3_genes_transposons.gff.

Novel German historical samples are available at GenBank, BioProject number PRJNA887392 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA887392>).

Novel global samples are available upon request (<https://zenodo.org/record/7187528>).

North American accessions were downloaded from the European Nucleotide Archive, study PRJEB24619 (<https://www.ebi.ac.uk/ena/data/view/PRJEB24619>).

African accessions were downloaded from the European Nucleotide Archive/Sequence Read Archive database, study PRJEB19780 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB19780>).

Climate data was downloaded from <https://climate.northwestknowledge.net/TERRACLIMATE-DATA/>.

Stomata phenotypes and water use efficiency data was downloaded from <https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fmec.14838&file=mec14838-sup-0003-Tables2.csv> (<https://onlinelibrary.wiley.com/doi/10.1111/mec.14838>).

Arabidopsis thaliana microscopy images were collected with a Leica SP5 confocal microscope and the Leica Application Suite Advanced Fluorescence 2.7.3.9723 and processed with Fiji v.2.9.0 and Adobe Illustrator 27.0.1.

Data analysis

All statistical analysis was conducted using R v1.2.1335 within RStudio v1.2.1335, with the packages cluster v2.1.0., geosphere v1.5-10, raster v3.4-10.

Phenotype data was processed using Python 3.7.9, with the packages pandas v1.1.2, numpy v1.19.2, and scikit-learn 1.2.0.

Genome sequence data was processed and analyzed using Adapterremoval v2.3.1, BCFtools v1.10.2, BWA v0.7.15-r1140, DeDup v0.12.8, gatk4-4.2.0.0-0, GEMMA v0.89.1, MapDamage v2.2.1, Picard v2.18.29-0, PLINK v1.90b6.16 64-bit, Samtools v1.9, SnpEff v5.0e, vcflib/20161123-git, VCFtools/0.1.16.

Microscopy images were post-processed with Fiji v2.1.0/1.53c and Adobe Illustrator v26.3.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data availability statement included.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.

Reporting on race, ethnicity, or other socially relevant groupings

Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status). Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.) Please provide details about how you controlled for confounding variables in your analyses.

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment

Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight

Identify the organization(s) that approved the study protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Depending on the analysis in question (see 'Data exclusions'), we used a full or partial panel of the 1001 Genomes Project. We combined all available whole genome sequences of historical <i>Arabidopsis thaliana</i> herbarium specimens and used all or parts of this combined data as described (see 'Data exclusions').
Data exclusions	For geographic analyses, only samples with known geographic coordinates were considered. For collection-time related analyses, only samples with known collection year were considered. To compare historical with modern samples, samples were matched by geographic proximity, excluding historical-modern sample pairs with a distance larger than 500km, excluding historical samples from the African continent, and excluding sample pairs between islands and mainland, or pairs matched across bodies of water or mountain ranges. Only sample pairs with the historical collection date predating the modern collection date were considered. For comparative analyses, only subsets of samples were used as indicated in the Materials and Methods. For the historical SNPcall, a number of quality filters were employed. Only samples with missing call frequencies smaller 50% were retained. Subsequently, we only considered called nucleotide variants with a minor allele present in at least three individual samples and with a site missingness <15%. Comparisons between historical and modern samples were conducted only on fully overlapping datasets.
Replication	Replication was not relevant in this study, as no new data was created experimentally. All analyses were conducted with analysis-specific, replicated control datasets. For each analyzed gene, 1000 control genes were selected (randomly, with replacement) to match gene lengths of original genes. For gene-group analyses, we calculated mean values for each of the 1000 random replicates to then compare to the gene-group mean of interest.
Randomization	Randomization was not relevant in this study, as no new data was created experimentally.
Blinding	Blinding was not relevant in this study, as now new data was created experimentally.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input type="checkbox"/>	<input checked="" type="checkbox"/> Plants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Public health
<input checked="" type="checkbox"/>	<input type="checkbox"/> National security
<input checked="" type="checkbox"/>	<input type="checkbox"/> Crops and/or livestock
<input checked="" type="checkbox"/>	<input type="checkbox"/> Ecosystems
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Demonstrate how to render a vaccine ineffective
<input checked="" type="checkbox"/>	<input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent
<input checked="" type="checkbox"/>	<input type="checkbox"/> Increase transmissibility of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Alter the host range of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable evasion of diagnostic/detection modalities
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable the weaponization of a biological agent or toxin
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

The following mutants and transgenic lines utilized for microscopy in this study were all reported previously: *basl-2* (Dong et al. 2009), *epf1-1* (Hara et al. 2007), *epf2-1* (Hunt & Gray 2009), *ice1-D* (*scrm-D*, Kanaoka et al. 2008), *ice1-2* (Kanaoka et al. 2008), *mute* (Pillitteri et al. 2007), *scrm2-1* (Kanaoka et al. 2008), *sdd1-1* (Berger & Altmann 2000), *spch-3* (MacAlister et al. 2007), *tmm-1* (Yang & Sack 1999), *uhf1-1* (Hara et al. 2007), *uhf1-2* (Berger & Altmann 2000), *epf1-1;epf2-1* (Hunt & Gray 2009), and *ice1-2;scrm2-1* (Kanaoka et al. 2008). Natural *A. thaliana* accessions were published previously (Col-0, e.g. (1001 Genomes Consortium 2016)).

Novel plant genotypes

Authentication

No novel plant genotypes were produced in this study, and all stable seed stock lines used for microscopy imaging were published previously. No authentication was required for any lines utilized.