A map of climate change-driven natural selection in *Arabidopsis thaliana*

Moises Exposito-Alonso[1], Hernán A. Burbano[2], Oliver Bossdorf[3], Rasmus Nielsen[4], Detlef Weigel[1*]

[1]Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany.
[2]Research Group of Ancient Genomics and Evolution, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany.
[3]Institute of Evolution and Ecology, University of Tübingen, 72076 Tübingen, Germany.
[4]Departments of Integrative Biology and Statistics, University of California Berkeley, Berkeley, CA 94720, USA. Natural History Museum of Denmark, Øster Voldgade 5-7, 1350 København K, Denmark

*correspondence to: weigel@weigelworld.org

Keywords: *Arabidopsis thaliana*, climate change, environmental niche models, field experiments, genetic natural selection, selection scan.

Running title: A map of climate change-driven natural selection

Through the lens of evolution, climate change is a directional selection acting on populations, forcing them to change and adapt, or face extinction. We studied climate-driven selection at the genetic level by measuring fitness of 517 *Arabidopsis thaliana* lines grown in rainfall-manipulation experiments in Spain and Germany. Natural selection in the hot-dry Mediterranean location was the strongest, killing 63% of the lines and significantly changing the frequency of about 5% of all genome-wide variants. Rainfall treatment dictated the direction and magnitude of selection: variants that increased in frequency in a low-rainfall environment, decreased in frequency under high rainfall, and vice versa. A significant proportion of this selection was predictable from the annotation of variants and from the climate in the areas where different variants are found (R^2=29-52%). Extrapolating our predictions to other areas of the species range suggested that selection driven by local climate would be strongest in the Mediterranean and Western Siberia — the edge of the species' environmental limits — and weakest in Central Europe. With rapidly increasing droughts and rising temperatures in Europe, we forecast a wave of directional selection moving North, consequently decimating and causing a genetic turnover of some native *A. thaliana* populations.

Introduction

The impacts of climate change on species are typically either assessed based on how climate exceeds the environmental tolerance range of a species[1], or on how fast the species can migrate to escape unfavorable conditions[1,2]. Because of the fast rate of current climate change, species survival will often depend on whether populations can genetically adapt and evolve in situ[3]. In order to predict the "evolutionary impact" of climate change on populations, we thus need to quantify and model the direction and magnitude of selection at the genetic level. Thanks to species-wide genome scans[4–6], as well as genome associations with climate of origin[7–10] and ecologically-relevant traits[7,11,12], we increasingly understand the genomic basis of selection and adaptation, which in turn can be useful to infer plausible scenarios for past adaptation. Natural selection, however, is only indirectly inferred in these types of analyses.

The gold standard for direct quantification of selection in a specific environment is in the form of common garden experiments in which multiple populations of a species are grown together[13,14]. With such direct observations, variation in traits – or more rarely in genetics – across a population[15–17] can be directly correlated with relative fitness. A disadvantage of such field experiments is that they are laborious and can often be carried out in a few locations within a species' range. Nevertheless, an emergent finding is that individuals are normally locally adapted and they are positively selected over foreigners when grown in their "home environment", while negatively selected in their "away environments"[18,19]. For these reasons, it appears natural to

combine high-throughput genome scans of climate of origin, with experimentally quantified natural selection *in situ*[20,21]. Here, we present the results from such an integrated approach, and transfer the knowledge we derive to forecast how present and future climate changes impact the genomic variation of the plant *Arabidopsis thaliana* across the species' range.

Natural selection under different precipitation levels

To study short-term natural selection in the annual plant *A. thaliana*, we performed two common garden experiments for one generation in two climatically distinct field stations, at the warm edge of the species distribution, Madrid (Spain, 40.40805ºN -3.83535ºE), and at the distribution core, Tübingen (Germany, 48.545809ºN 9.042449ºE) (for details see ref. 22). Under the rainout shelters at each site, we simulated high precipitation typical of a wet year in Germany, and low precipitation typical of a dry year in Spain (see Fig. 2 of ref. 22). In fall of 2015 we sowed over 300,000 seeds of 517 genotypes capturing species-wide genomic diversity[23]. For each line, we prepared seven pots in which only a single plant was retained after germination, and five pots with 30 seeds that were allowed to germinate and grow without intervention. At the end of the experiment in June 2016, we had collected 23,154 fitness observations, consisting of survival to the reproductive stage, the number of seeds per surviving plant (fecundity), and lifetime fitness (the product of survival and individual fecundity). Heritability of fitness traits were higher in more stressful environments ($0.00<H^2<0.551$; Table S3), i.e., under low vs. high precipitation, in high vs. low density, and in Spain vs. Germany. In the most stressful environment, low precipitation and high density in Spain, only 193 of the 517 accessions reproduced, whereas in Germany at least a few plants of each accession survived. Fitness also correlated with the precipitation at the geographic origin of each accession — As an example, survival in Spain was correlated with the driest annual quarter at the location of origin (Spearman's r=0.13, $P=2.54\times10^{-4}$; see Table S8,9), confirming previous observations of local adaptation to precipitation gradients in *A. thaliana*[7].

A genome map of selection

To quantify selection at the genetic level, we estimated the difference in relative fitness between genotypes with the minor and the major allele for each of the 1,353,386 informative SNPs across the 515 genomes (Fig. 1). This analysis, carried out with Linear Models (LM) implemented in GEMMA [24], is agnostic regarding causality of each tested SNP. As many variants will be in linkage disequilibrium (LD) with causal variants, they will experience indirect selection proportional to their association with the causal loci[25,26] — so-called background selection or genetic hitchhiking. This is particularly likely given the high genome-wide LD in *A. thaliana*[27]. We use the term *allelic selection differentials* (*s*, see Thurman and Barrett [17] for a similar discussion), to denote the realized selection affecting a SNP

resulting from the combination of selection acting directly on the SNP, and the indirect effects due to selection on causal SNPs that are in LD with the focal variant, We found a total of 421,962 SNPs with allelic selection differentials below an 0.05 significance threshold in at least one of the eight environments (for marginal values and overlaps see Table S2) (Benjamini & Hochberg FDR correction). Using more stringent Bonferroni correction ($<7x10^{-7}$), we still detected 6,538 SNPs distributed throughout the genome, suggesting that the polygenic model of natural selection[28] prevails in this climate-manipulation experiment. When we calculated the changes in allele frequency between the 515 parents and the seeds they produced based on the starting allele frequency and the allelic selection differentials, we found these were of considerable magnitude in the experiment in Spain and low precipitation, up to 10%, while they did not exceed 2% in the benign high-precipitation environment in Germany (see Supplementary Methods V, Fig. S9). Calculating allelic selection differentials for survival and fecundity separately, we found no correlation between survival-only and fecundity-only estimates themselves ($r<0.07$, Fig. S4-6), consistent with different stages of a plant being differentially affected by environmentally imposed selection[29]. Variants detected after Bonferroni-correction were more likely to be located in intergenic regions (Fisher's Exact test Odds ratio [Odds]=1.11, $P=7x10^{-30}$) and less likely in exons (Odds=0.91, $P=1x10^{-26}$). An exception were nonsynonymous mutations, which were also more likely to be significant (Odds=1.05, $P=2x10^{-4}$). The large number of variants affected by selection implies a strong turnover of variation across the entire genome as a response to the environment [30–32], and a potentially significant demographic decimation — what Haldane called "the demographic cost of natural selection" [33–35].

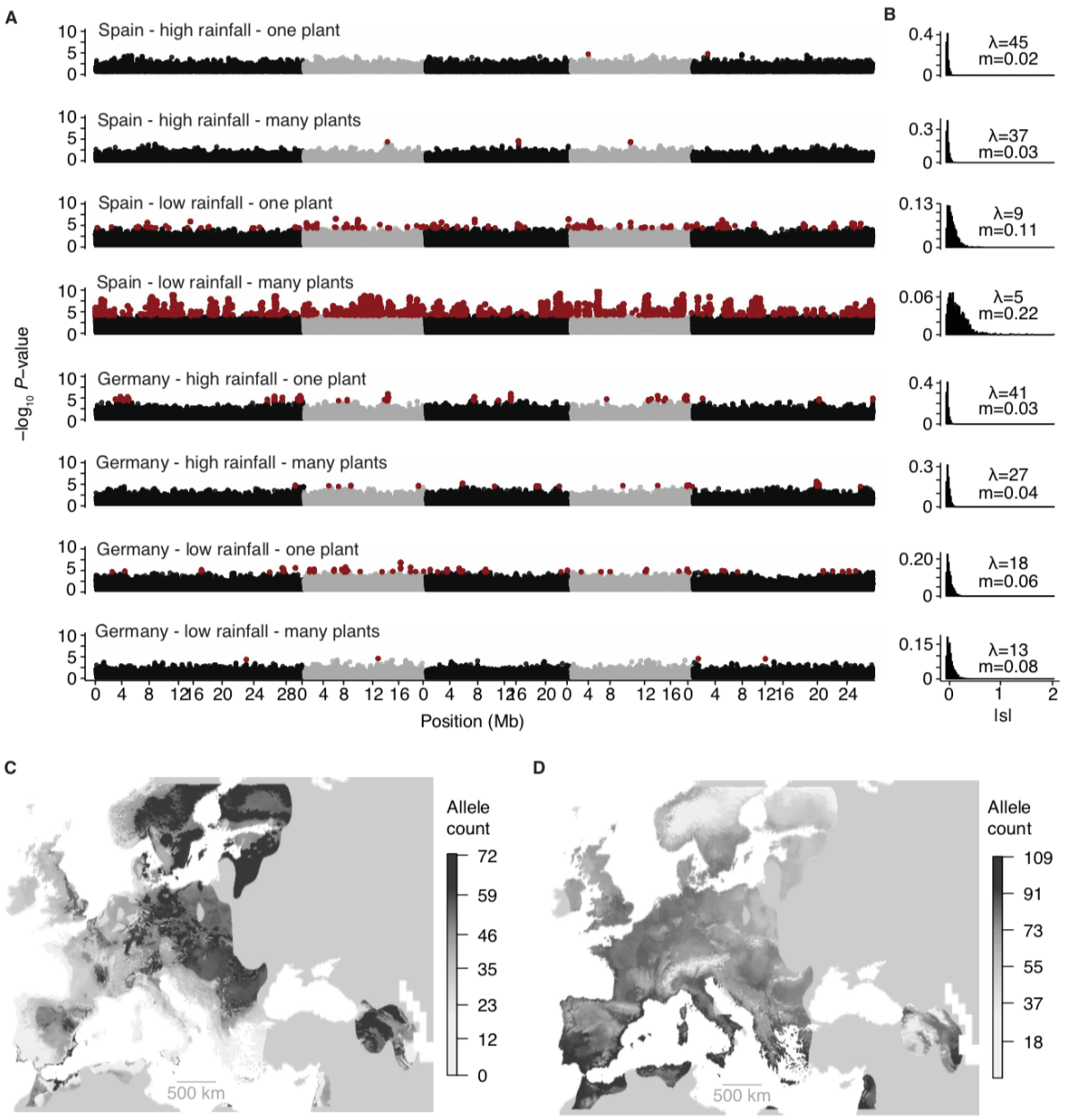


Fig. 1 A genome map of allelic selection differentials. (A) Manhattan plots of SNPs significantly associated with relative lifetime fitness in eight different environments. Only SNPs significant after FDR (black and grey) or Bonferroni correction (red) are shown. For genome-wide scans of survival and fecundity fitness see Figures S4 and S5. (B) Distribution of absolute allelic selection differentials $|s|$ per experiment. λ denotes maximum likelihood-inferred parameter of an exponential distribution, and *m* denotes the mean allelic selection differential. (C) Environmental niche models for 424 SNPs in high-precipitation experiments, and (D) for 279 SNPs in low-precipitation experiments. We used the most significant SNP in each 0.5 Mb window when there was any.

Because a change in allele frequency does not necessarily reflect whether a variant is adaptive or not, particularly when there is high LD, we aimed to better locate potential direct targets

of selection using a Bayesian Sparse Linear Mixed Model (BSLMM) with GEMMA, another GWA model[26,36] (see Supplementary Methods V). By jointly fitting all SNPs (n=1,353,386), this method statistically corrects for LD arising from population structure and/or low recombination. It also provides an estimate of the proportion of the genome, π, that is very likely to have a major effect on the studied trait and thus likely being the target of selection. This proportion was similar across experiments: $8x10^{-5} < \pi < 5x10^{-6}$, and it was much smaller than what we had identified with allelic selection differentials ($2x10^{-5}$—0.001), which indicates that selection must be mostly indirect[17,37,38]. This is likely the case in species where many mutations co-occur due to a shared population history and local adaptation[7,26,39], thus we advocate the use of allelic selection differentials to capture the total magnitude of selection over a variant, which ultimately will better reflect the frequency change in response to climate-driven selection.

Past and present natural selection along heterogeneous environments

Past events of climate-driven natural selection might have or have not left a statistical signature in local polymorphisms along the genome. To explore whether short-term, ecological selection differentials align with selection experienced over longer evolutionary time[40], we scanned the genome for two signatures of selection, the Composite Likelihood Ratio of a selective sweep using SweepFinder2 (ref. 41), and F_{ST} as a metric of divergence between eleven *A. thaliana* lineages[7,23]. Comparing frequency-matched background SNPs with Bonferroni-corrected significant SNPs, we found that the latter had higher average F_{ST} values (0.39 compared to 0.14, Wilcoxon test, $P<10^{-16}$), but were not any more likely to have experienced a selective sweep (P=0.2) (Fig. 2, Fig. S7-8). This observation indicates that adaptation to climate does not typically result in species-wide selective sweeps, different from the conclusions obtained from sweep analyses and climate associations that did not have access to SNP effects in the field [10]. All in all, the signature that allelic selection differentials coincide with allele frequency shifts across population lineages are better reconciled with a polygenic model of natural selection[31].

We hypothesized that the F_{ST} signatures of population divergence that we found could have arisen from opposing past selection across populations in geographic locations with contrasting climates. Alleles that were positively selected under low precipitation tended to be negatively selected under high precipitation, suggesting antagonistic pleiotropy[16] (Fig. 2, Fisher's exact test Odds Ratios >1.31, $P<4x10^{-24}$; Table S5) — an observation particularly clear when comparing the two most "natural" conditions, low precipitation in Spain and high precipitation in Germany (Odds Ratio=6.72). In contrast, when we compared the same precipitation condition between the two locations, selection was either in the same direction (0.23<Pearson's r<0.51), or there was selection in one environment and neutrality in the other, displaying conditional neutrality (All Odds ratio<1,

$P<10^{-16}$). We conclude that selection acted in opposite directions across precipitation gradients, but not across temperature gradients.

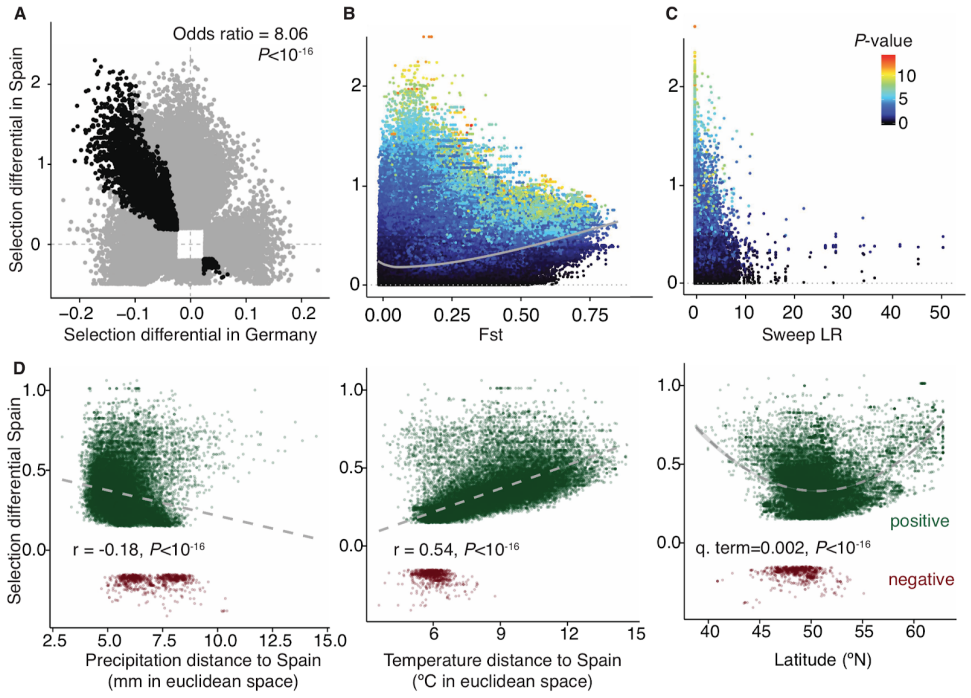


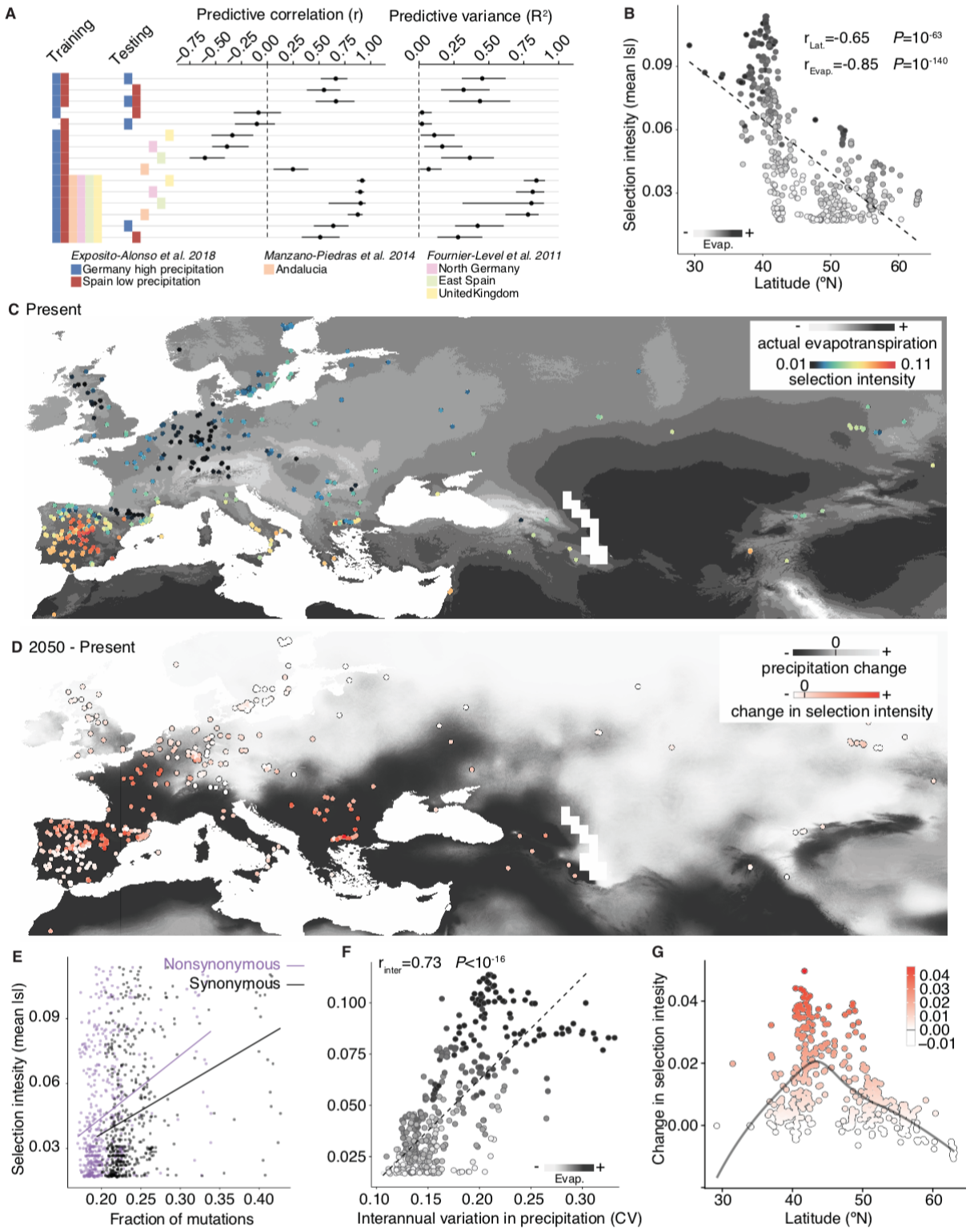
Fig. 2 Selection trade-offs and a signal of local adaptation. (A) Empirical selection coefficients across two contrasting environments; Spain with low precipitation and high population density, and Germany with high precipitation and low population density. Grey dots indicate SNPs with significant conditional neutrality, black dots indicate SNPs with antagonistic pleiotropy. SNPs that behave in the same direction or that were non-significant in both environments are not plotted. (B) Relationship between field allelic selection differentials and F_{ST} values across 11 lineages, and (C) the likelihood ratio of selective sweeps. (D) Correlation between field selection differentials and climate distances between test site and the centroids of SNPs´ geographic ranges.

If selection had acted in the past in opposite directions across environmental gradients, it should have generated geographically-restricted distributions of genetic variants. We implemented genome-wide environmental niche models[7] fitted with available climate data (1960-1990 averages, worldclim.org v.1, ref. 42) (see Supplementary Methods VII), which revealed that alleles selected in Spain and low precipitation are more likely to come from Mediterranean locations (Fig. 1D) . The opposite was true for the alleles selected in Germany and high precipitation(Fig. 1C) — indicating that environmental proximity of a SNP's location of origin to the experimental stations is broadly related to the experienced selection in the field. In local adaptation theory, at the level of individuals, this is often called "home advantage"[18]. We directly tested this hypothesis at the genetic level,

correlating SNPs' selection differentials measured in Spain and low precipitation with the SNPs' mean environment of origin. We confirmed that alleles found in locations with more similar precipitation to the Spanish experiment were more likely to be positively selected in Madrid (bio 12-19; Spearman's r=-0.18, $P<10^{16}$) (Fig. 2). For temperature, we observed an opposite pattern (bio1-11; Spearman's r=0.54, $P<10^{16}$) (for other experiments, see Fig. S10). The apparent reason is that not only Mediterranean alleles, but also ones from cold, but dry regions such as Scandinavia or Russia, are equally positively selected in Spain and low precipitation. That is, alleles both at the southern and northern edges of the distribution found themselves "at home" in Spain (quadratic relationship = 0.002, $P<10^{-16}$; Fig. 2B) [7].

Forecasting climate-driven selection across the species' geographic distribution

Because the home environment of alleles strongly correlated with selection in our experiments, we wondered to what degree allelic selection differentials might be predictable[43]. To predict selection differentials from climatic and genomic information, we first conducted a series of climate GWAs (see Table S6 for climate variables). We not only used as predictors the estimates of how strongly the presence of an allele was associated with climate of origin [10], but also included the SNP annotations, F_{ST}, π, and sweep likelihood summary statistics. We used a regression decision tree approach with Random Forests to build Genome Environment Selection ("GES") models. Conceptually, GES models are similar to Environmental Niche Models (ENMs), but instead of training them with presence/absence data of a genetic variant[7], we trained them with our measured allelic selection differentials. This provided a means to predict whether alleles should increase/decrease in frequency in a certain climate, instead of whether alleles are "likely to be present", the ENMs' version. By training models with both experiments jointly and using cross-validation, we confirmed that inferred and measured selection differentials were correctly predicted, with a modest to high correlation accuracy (0.56 < Pearson's r < 0.7) and explaining a large proportion of variance (R^2= 0.29-52%) (Fig. 3A, variable importance Table S5) (further details in Supplementary Methods VIII).



192 **Fig. 3 A geographic map of climate-driven selection and its predictability.** (A) Genome Environment
193 Selection (GES) models trained with a combination of environments, to infer allelic selection
194 differentials throughout the species range. Mean predicted allelic selection differentials ("selection
195 intensity"; n= 10,752 SNPs, one random SNP per 10 kb windows) in known locations of *A. thaliana*
196 populations was related to (B) latitude, and (C) evapotranspiration in summer (C) (ref. 42). (D)

Predicted changes in selection intensity by 2050 (2050 MP rcp 8.5, ref. 42). (E) Relationship between selection intensity and synonymous and nonsynonymous polymorphisms. (F) Relationship between selection intensity and interannual variation in precipitation from 1958-2017 (ref. 44). (G) Relationship between selection intensity change (see panel D) and latitude.

To further validate the predictive accuracy of our models, we made use of published fitness data from common garden experiments conducted with diverse *A. thaliana* lines in Eastern Spain, North Germany, and the United Kingdom[12], and in Southern Spain[45]. We calculated allelic selection differentials using fitness and genomic data for each experiment as before (#SNPs=1,353,386, #Genomes=515, see Supplementary Methods V & IX). Using the worldclim database[42], we extracted climate variables for the four locations and inferred allelic selection differentials. Using our two experiments and the four published ones, we could build GES models with high prediction power in any of the six experiments, with variance explained of $17\%<R^2<65\%$ and accuracies of $0.53<r<0.76$ (Fig. 3A, Fig. S12 Table S10) (all-to-all analysis: r= 0.761 , Bootstrap CI= 0.55—0.90, R^2= 0.517; for further details on predictability analyses and null expectations, see Supplementary Methods IX.4-5).

Using the trained models, we predicted allelic selection at hundreds of locations across the species range. The strongest predicted selection were towards the environmental limits of the species (Euclidean distances from the mean environment in 98-dimension climate space, Spearman's r=0.42, $P<10^{-16}$), i.e. in warm (annual temperature, r=0.62, $P<10^{-16}$), dry (annual precipitation, r=-0.457, $P=10^{-27}$), and high evapotranspiration locations (actual evapotranspiration in August, r=0.86, $P<10^{-16}$) (Fig. 3B-C). High selection intensity coincided with locations where natural lines have a low ratio of nonsynonymous to synonymous polymorphisms (Fig. 3E, r=-0.276, 3×10^{-10}). The depletion of nonsynonymous substitutions, which are expected to be on average more deleterious than synonymous substitutions, suggests that natural selection has acted efficiently in the past. High selection intensity also correlated with high local genetic diversity π (r=-0.01, P=0.85) and elevated Tajima's D (r=0.161, $P=3\times10^{-4}$). This could be explained by the direction of selection having alternated over time, with multiple polymorphisms having been selected in each period[46]. Inspection of precipitation data from 1958 to 2017 (ref. 44) revealed that locations where we had inferred strong selection and high diversity often suffered high year-to-year climatic variation (r=-0.73, $P<10^{-16}$; Fig. 3F). This finding not only highlights the importance of temporal resolution of climate databases for predictions, but also climate stochasticity as a source of evolutionary constraints over species. Although the complex population history of adaptation in *A. thaliana* can complicate interpretations[7,23,39], our geographic predictions align with the theory that a species' environmental niche[47] and its populations' genetic diversity[48,49] are shaped by increasing natural

selection towards the range edges, rather than increasing low population size and high drift at the range margins[50].

The obvious implication of our findings is that a sudden change in climate or increased climate variability[51,52] could increase the magnitude of natural selection. Using climate projections of 2050 as a proxy for potentially abrupt changes in local climate (Intergovernmental Panel on Climate Change, www.ipcc.ch, ref. 42), we predict that selection intensity will likely increase in much of Southern-Central Europe, with an expected decrease in annual precipitation and increase in annual temperatures (Fig. 3D,G, Fig. S3; Fig. S11).

Conclusions

The expected changes in climate during the 21st century will threaten the survival of many species. We have addressed the challenge of predicting the effects of climate-driven natural selection on *A. thaliana* genomic variation across the species' range. Experimentally quantifying natural selection over standing genetic variation currently present in the species, we showed that selection was highly polygenic, with many co-occurring variants in the genome being co-selected, thus leading to a dramatic genetic turnover in our experimental populations. Integration of genome-climate associations with direct fitness observations in strategically located field stations allowed us to build predictive models that could also be validated at other geographic locations with results from published experiments. Finally, we have illustrated how this information can be used to predict range-wide evolutionary vulnerability in the face of rapid climate change. This brings us one step further to fully incorporating evolution into predicting the impacts of climate change on biodiversity.

REFERENCES

1. Urban, M. C. Climate change. Accelerating extinction risk from climate change. *Science* **348,** 571–573 (2015).
2. Jezkova, T. & Wiens, J. J. Rates of change in climatic niches in plant and animal populations are much slower than projected climate change. *Proc. R. Soc. B* **283,** 20162104 (2016).
3. Hoffmann, A. A. & Sgrò, C. M. Climate change and evolutionary adaptation. *Nature* **470,** 479–485 (2011).
4. Nielsen, R. *et al.* Genomic scans for selective sweeps using SNP data. *Genome Res.* **15,** 1566–1575 (2005).
5. Horton, M. W. *et al.* Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. *Nat. Genet.* **44,** 212–216 (2012).
6. Bonhomme, M. *et al.* Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. *Genetics* (2010). doi:10.1534/genetics.110.117275
7. Exposito-Alonso, M. *et al.* Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. *Nat Ecol Evol* **2,** 352–358 (2018).

8. Bay, R. A. *et al.* Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359,** 83–86 (2018).
9. Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J. K. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185,** 1411–1423 (2010).
10. Hancock, A. M. *et al.* Adaptation to climate across the Arabidopsis thaliana genome. *Science* **334,** 83–86 (2011).
11. Jones, F. C. *et al.* The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484,** 55–61 (2012).
12. Fournier-Level, A. *et al.* A map of local adaptation in Arabidopsis thaliana. *Science* **334,** 86–89 (2011).
13. Kingsolver, J. G. *et al.* The strength of phenotypic selection in natural populations. *Am. Nat.* **157,** 245–261 (2001).
14. Savolainen, O., Lascoux, M. & Merilä, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14,** 807–820 (2013).
15. Gompert, Z. *et al.* Experimental evidence for ecological selection on genome variation in the wild. *Ecol. Lett.* **17,** 369–379 (2014).
16. Anderson, J. T., Lee, C.-R. & Mitchell-Olds, T. Strong selection genome-wide enhances fitness trade-offs across environments and episodes of selection. *Evolution* **68,** 16–31 (2014).
17. Thurman, T. J. & Barrett, R. D. H. The genetic consequences of selection in natural populations. *Mol. Ecol.* **25,** 1429–1448 (2016).
18. Hereford, J. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173,** 579–588 (2009).
19. Leimu, R. & Fischer, M. A meta-analysis of local adaptation in plants. *PLoS One* **3,** e4010 (2008).
20. Lasky, J. R., Forester, B. R. & Reimherr, M. Coherent synthesis of genomic associations with phenotypes and home environments. *Mol. Ecol. Resour.* (2017). doi:10.1111/1755-0998.12714
21. Price, N. *et al.* Combining population genomics and fitness QTLs to identify the genetics of local adaptation in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* 201719998 (2018). doi:10.1073/pnas.1719998115
22. Exposito-Alonso, M. *et al.* A rainfall-manipulation experiment with 517 Arabidopsis thaliana accessions. *bioRxiv* (2017). doi:10.1101/186767
23. 1001 Genomes Consortium. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166,** 481–491 (2016).
24. Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **11,** 407–409 (2014).
25. Kojima, K. & Lewontin, R. C. Evolutionary Significance of Linkage and Epistasis. in *Mathematical Topics in Population Genetics* 367–388 (Springer, Berlin, Heidelberg, 1970). doi:10.1007/978-3-642-46244-3_12
26. Gompert, Z., Egan, S. P., Barrett, R. D. H., Feder, J. L. & Nosil, P. Multilocus approaches for the measurement of selection on correlated genetic loci. *Mol. Ecol.* **26,** 365–382 (2017).
27. Long, Q. *et al.* Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. *Nat. Genet.* **45,** 884–890 (2013).
28. Berg, J. J. & Coop, G. A Population Genetic Signal of Polygenic Adaptation. *PLoS Genet.* **10,** e1004412–e1004412 (2014).
29. Mitchell-Olds, T. Pleiotropy Causes Long-Term Genetic Constraints on Life-History Evolution in Brassica rapa. *Evolution* **50,** 1849–1858 (1996).
30. Jain, K. & Stephan, W. Rapid Adaptation of a Polygenic Trait After a Sudden Environmental Shift.

Genetics **206,** 389–406 (2017).
31. Pritchard, J. K., Pickrell, J. K. & Coop, G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20,** R208–15 (2010).
32. Wittmann, M. J., Bergland, A. O., Feldman, M. W., Schmidt, P. S. & Petrov, D. A. Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proc. Natl. Acad. Sci. U. S. A.* **114,** E9932–E9941 (2017).
33. Haldane, J. B. S. The Cost of Natural Selection. *Genetics* **55,** 511–524 (1957).
34. Orr, H. A. & Unckless, R. L. The population genetics of evolutionary rescue. *PLoS Genet.* **10,** e1004551 (2014).
35. Crow, J. F. Genetic Loads and the Cost of Natural Selection. in *Mathematical Topics in Population Genetics* 128–177 (Springer, Berlin, Heidelberg, 1970). doi:10.1007/978-3-642-46244-3_5
36. Zhou, X., Carbonetto, P. & Stephens, M. Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet.* **9,** e1003264 (2013).
37. Lewontin, R. C. *The genetic basis of evolutionary change*. **560,** (Columbia University Press New York, 1974).
38. Charlesworth, B. The effects of deleterious mutations on evolution at linked sites. *Genetics* **190,** 5–22 (2012).
39. Lee, C.-R. *et al.* On the post-glacial spread of human commensal Arabidopsis thaliana. *Nat. Commun.* **8,** 14458 (2017).
40. Eyre-Walker, A. & Keightley, P. D. The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8,** 610–618 (2007).
41. DeGiorgio, M., Huber, C. D., Hubisz, M. J., Hellmann, I. & Nielsen, R. SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics* **32,** 1895–1897 (2016).
42. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25,** 1965–1978 (2005).
43. Nosil, P. *et al.* Natural selection and the predictability of evolution in Timema stick insects. *Science* **359,** 765–770 (2018).
44. 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).
45. Manzano-Piedras, E., Marcer, A., Alonso-Blanco, C. & Picó, F. X. Deciphering the adjustment between environment and life history in annuals: lessons from a geographically-explicit approach in Arabidopsis thaliana. *PLoS One* **9,** e87836–e87836 (2014).
46. Levene, H. Genetic Equilibrium When More Than One Ecological Niche is Available. *Am. Nat.* **87,** 331–333 (1953).
47. Lee-Yaw, J. A. *et al.* A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecol. Lett.* (2016). doi:10.1111/ele.12604
48. Stebbins, G. L. Aridity as a Stimulus to Plant Evolution. *Am. Nat.* **86,** 33–44 (1952).
49. Leffler, E. M. *et al.* Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biol.* **10,** e1001388 (2012).
50. Henry, R. C., Bartoń, K. A. & Travis, J. M. J. Mutation accumulation and the formation of range limits. *Biol. Lett.* **11,** 20140871 (2015).
51. Giorgi, F., Bi, X. & Pal, J. Mean, interannual variability and trends in a regional climate change experiment over Europe. II: climate change scenarios (2071–2100). *Clim. Dyn.* **23,** 839–858 (2004).
52. Samaniego, L. *et al.* Anthropogenic warming exacerbates European soil moisture droughts. *Nat.*

Clim. Chang. **8,** 421–426 (2018).

ADDITIONAL INFORMATION

Accession codes. Phenotypic datasets are available as supplementary material of ref. 22 at [update journal web] with doi: [update with journal link]. Genomes are available at http://1001genomes.org/data/GMI-MPI/releases/v3.1/. The seed collection can be obtained from the Arabidopsis Biological Resource Center (ABRC) under accession CS78942.

Author contribution MEA, HAB and DW conceived the project outline. MEA designed, implemented and coordinated the project. MEA carried out the experiment in Tübingen and in Madrid with technical support. MEA carried out statistical analyses. HAB, OB, RN, and DW supervised the project and discussed analyses interpretation. MEA prepared the first draft and the final manuscript was written by MEA, HAB, OB, RN, and DW.

Acknowledgements We gratefully thank all field helpers, Patricia Lang for thorough comments on the manuscript, and the Weigel and Burbano labs for discussions.

Funding statement This work was funded by an EMBO ST fellowship (MEA), ERC Advanced Grant IMMUNEMESIS and the Max Planck Society (DW).

Disclosure statement The authors declare no competing financial interests. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary Information Guide for

Exposito-Alonso et al.: A map of climate change-driven natural selection in *Arabidopsis thaliana*

Table of Content

SUPPLEMENTARY METHODS 16
- **I. Fitness data from: "A rainfall-manipulation experiment with 517 Arabidopsis thaliana ecotypes"** 16
- **II. 1001 Genomes Project data** 16
- **III. Classic Fst and selective sweep signatures from polymorphism data** 16
  - **III.1 Geographic proxies of diversity metrics** 17
- **IV. Heritability of fitness** 17
- **V. Genome-Wide Association, selection differentials, and direct selection estimates** 18
  - **V.1 Trade-offs across field experiments** 19
  - **V.2 Intensity of selection** 19
  - **V.3 Population structure correction in a wild species** 20
- **VI. Climate Genome-Wide Association** 20
- **VII. Climate and modeling** 20
  - **VII.1. Climate layers** 20
  - **VII.2. Environmental Niche Models** 21
  - **VII.3. Climate variability** 21
- **VIII. Predictions of selection differentials with summary statistics** 21
  - **VIII.1 The model** 21
  - **VIII.2 A simple visualization of environmental distance** 22
  - **VIII.3 Limitations and interpretations** 23
- **IX. Re-analysis of published data from common garden experiments** 24
  - **IX.1 Manzano-Piedras et al. 2014** 24
  - **IX.2 Fournier-Level et al. 2011** 24
  - **IX.3 1001 Genomes x RegMap panel phenotype imputation** 24
  - **IX.4 Sanity check analysis for imputation and geographic predictions** 25
  - **IX.5 An explanation for "inverse predictability"** 25

SUPPL. FIGURES & MEDIA 27
- **Figure S1. Accession locations** 27
- **Figure S2. Extent of native environments of A. thaliana** 28
- **Figure S3. Map of environmental change** 29
- **Figure S4. Genome maps of survival** 30
- **Figure S5. Genome maps of fecundity** 31
- **Figure S6. Trade-offs in survival and fecundity** 32
- **(Fig. S6 continued)** 33

412 **Figure S7. Fst and empirical selection** **34**
413 **(Fig. S7 continued)** **35**
414 **Figure S8. Sweeps and empirical selection** **36**
415 **(Fig. S8 continued)** **37**
416 **Figure S9. MAF and relative fitness** **38**
417 **(Fig. S9 continued)** **39**
418 **Figure S10. Environmental distance and selection differentials** **40**
419 **Figure S11. Future change in selection for different climate change scenarios** **41**
420 **Figure S12. Null expectation of predictability** **42**
421 **SUPPL. TABLES & DATASETS** **43**
422 **Table S1. List of accessions** **43**
423 **Table S2. Summary of fitness data** **43**
424 **Table S3. Heritability of traits** **43**
425 **Table S4. Number of SNPs with significant selection differentials** **43**
426 **Table S5. Expected allele frequency changes in response to selection** **43**
427 **Table S6. Variable importance of predictive models** **43**
428 **Table S7. Predictability of environmental models** **44**
429 **Table S8. Description of climate variables** **44**
430 **Table S9. Spearman correlations among climate variables and phenotypes** **44**
431 **Table S10. GBLUP heritability and imputation accuracy of data from published field**
432 **experiments** **44**
433 **Table S11. Correlation between inferred natural selection intensity and other variables** **44**
434 **SUPPLEMENTARY REFERENCES** **45**

435 **SUPPLEMENTARY METHODS**

436 I. Fitness data from: "A rainfall-manipulation experiment with 517 *Arabidopsis* 437 *thaliana* ecotypes"

438 The field experiment is described in detail in Exposito-Alonso et al. "A rainfall-manipulation
439 experiment with 517 *Arabidopsis thaliana* ecotypes" (ref. 22). The data and processing code is
440 available at https://github.com/MoisesExpositoAlonso/dryAR update doi: xxx. Information of the 517
441 accessions can be found in Table S1. The experiment comprised 23,154 observations. Three
442 measurements of fitness were produced: survival from seed to reproductive adult (proportion 0-1)
443 and the average fecundity per reproductive adult (inflorescence skeleton length ranged
444 18,400—1,622,000 pixels, which approximately corresponds to 1—6,127 seeds per plant). Fecundity
445 was only measured for those that had at least one fruit, thus it does not incorporate information on
446 survival, similarly as Anderson et al. [16]. We finally calculated an integrated lifetime fitness value by
447 multiplying the survival proportion to adulthood with the total offspring produced.

448 II. 1001 Genomes Project data

449 We used VCFtools v.0.1.12b (ref. 53) to subset and filter the 1001 Genomes VCFv4.1 (available at:
450 http://1001genomes.org/data/GMI-MPI/releases/v3.1/). We used vcftools with the flags: --maf 0.01
451 --max-alleles 2 --min-alleles 2 --max-missing 0.95. The resulting dataset was a genome matrix of 515
452 individuals by 1,353,386 variant. Two accessions from the experiments were not sequenced for the
453 1001 Genomes but were originally included for a better geographic representation, these could not
454 be included in subsequent analyses.

455 III. Classic F_{st} and selective sweep signatures from polymorphism data

456 We used the genetic groups previously defined for the same accessions [7] and computed Fst using
457 PLINK version 1.9 (ref. 54). We also used PLINK to calculate π and Tajima's D using PLINK in a
458 windows of 100 SNPs across the genome.

459 We used SweepFinder 2 (ref. 41) to scan the genome for deviations of the Site Frequency
460 Spectrum (SFS) that might be caused by selective sweeps. To not bias the SFS, we used all 11,769,920
461 SNPs biallelic SNPs (without a filter of 1% MAF).

We annotated the 1001 Genomes VCF using the package SnpEff 4.3p (ref. 55). We then manually curated a set of eight categories of variants: intergenic, intron, UTR3, UTR5, exon, synonymous, nonsynonymous, exon noncoding.

III.1 Geographic proxies of diversity metrics

In order to estimate local π and Tajima's D at a regional scale, we used the 4 closest neighbours (same patterns were observed with three neighbours within a geographic area of 5º Latitude-Longitude radius), and computed the total number of polymorphisms P in the subset and the sum of all pairwise Hamming differences, H. Then we calculated θ, π and D as:

$$\pi = \frac{H}{6 \times G} \frac{N_{full}}{N_{all}}$$

$$\theta = \frac{P}{1.8666 \times G} \frac{N_{full}}{N_{all}}$$

$$\hat{D} = \pi - \theta$$

Where G is the genome size, N_{full} are all SNPs with full information that were used to count polymorphisms and distances, and N_{all} are all SNPs of the genome matrix. In the denominators, 6 is the number of pairwise comparisons of four genomes, and 1.8666 is the harmonic number of 4. Although the sd of D should be divided by the standard error, because we are only interested in relative differences between populations, we only get the difference between π and θ and make a rank.

IV. Heritability of fitness

To estimate how much variance in fitness is related to the genotype, we used generalized linear mixed models using the R package MCMCglmm (ref. 56). We used fitness estimates per replicate and, appart from including the genotype identity, we controlled for block (growing tray) and position within the block (longitudinal, latitudinal, and the interaction). As this is a Bayesian approach, we used 10,000 MCMC steps, a burning of 10%, and checked for convergence of the chain (good in all cases, as it is a simple model). For survival proportion we used a Binomial link, for number of seeds we used a Poisson link, and for the combined lifetime relative fitness we used a Gaussian link. The mode and 95% Highest Posterior Density of the posterior distribution of each random effect were extracted (Table S3).

V. Genome-Wide Association, selection differentials, and direct selection estimates

We used GEMMA (ref. 36) to run regular linear models (LM) of the form:

$$y = \mu + \beta_i x_i + \epsilon;$$

which was repeated for every SNP in the genome. This provided us with allele effects on relative fitness per SNP. This is no more and no less than the selection differential s from populations genetics and evolutionary biology [17], that estimate the difference in fitness between the genotypes carrying the alternative (1) or reference allele(0) allele in a haploid model: $w1 - w0 = s$. The correspondence with allele frequency change Δp was already defined by RT, and is also dependent on the variance in allele frequency, as low frequency variants are "less seen" by selection:

$$\Delta p = p(1-p)\frac{w1 - w0}{\bar{w}}$$

$$\Delta p = p(1-p)s$$

We also run in GEMMA a Bayesian Sparse Linear Mixed model (BSLMM), to more accurately pinpoint casual positions, as this model accommodates both poly- and oligogenic architectures. It models two effect hyperparameters, a basal effect, α, that captures the fact that many SNPs contribute to the phenotype, and an extra effect, β, that captures the stronger effect of only a subset of SNPs. An internal parameter measuring the probability of having another extra effect, γ, can be used to prioritize SNPs. In BSLMM the total effect of an allele is $= \alpha + \gamma\beta$.

$$y = 1_n\mu + X\beta + X\alpha + \epsilon;$$

$$\beta_i \sim \pi N(0, \sigma_a^2\tau^{-1}) + (1-\pi)\delta_0\ ;$$

$$\alpha_i \sim N(0, \sigma_b^2/(p\tau));$$

$$\epsilon \sim MVN_n(0, \tau^{-1}I_n)\ .$$

The BSLMM model is also useful also used to calculate the proportion of variance explained ('chip heritability'). To do this we used the last 1,000 samples of the MCMC chain and calculated the 95% Highest Posterior Density Interval (95% HPD). To do that we report the median and the 2.5% and 97.5% percentiles.

To illustrate the differences between a univariate (LM) and multivariate regressions (similar to BSLMM), we show an example of two SNP predictors, x_1 and x_2. For mathematical convenience we assume that the response variable fitness, y, as well as the predictors, are mean centered and variance scaled. From the univariate approach, where the effect of a SNP is estimated marginally or independently, β would be:

$$\beta_{x_1} = \frac{cov(x_1, y)}{var(x_1)}$$

This would be estimated separately for SNP one and two. In a multivariate regression framework, the regression coefficient, called conditional or partial coefficient, β^*, is corrected by the correlation between the two predictors, $r_{x_1 x_2}$, which in the case of genotypes is called linkage disequilibrium, as in the form of:

$$\beta^*_{x_1} = \frac{\beta_{x_1} - r_{x_1 x_2} \times \beta_{x_2}}{\sqrt{(1-\beta^2_{x_1})(1-\beta^2_{x_2})}}$$

Thus we find an analogy between β and β^* and s and s^* from population genetics (eq. 4 from [57] simplified as conceptual model),where a selection differential is dependent on the true selection and a indirect term dependent on all other n SNPs in the genome.

V.1 Trade-offs across field experiments

We looked for genetic variants with a positive selection differential in one experimental environment that had a negative differential in another (antagonistic pleiotropy) or whether it was more prevalent that a selected variant in one environment was neutral in another environment (conditional neutrality). Only for the purpose of two environment comparisons, we took those SNPs whose selection differential p-value was lower than 0.01 in one environment as conditionally neutral, while SNPs whose p-value in both environments was < 0.02 were considered to act in antagonistic pleiotropy (because antagonistic pleiotropy requires two tests, one in each environment, the significance threshold should be 2 x p-value in each test) (see Fig. 2).

V.2 Intensity of selection

The distribution of absolute values of selection differentials, $|s|$, have a shape similar distribution to an exponential. We calculated the expected rate using Maximum Likelihood optimization in R but also can be approximated as the reciprocal of the mean:

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}.$$

For this matter, we use $\hat{\lambda}$ or the mean of $|s|$ as metrics of the overall intensity of selection (Fig. 1B, Fig. 3D).

V.3 Population structure correction in a wild species

In human genetics, penalizing SNPs that are correlated with population structure is particularly important, as the lack of controlled experiments can generate spurious associations between genetic variants more common is some human ancestries with regional measurement errors or different growing cultural environments and nutrition[58]. Even in human genetics, when populations are comprised for multiple ancestries that also have different disease sensitivities, population structure correction has been deemed as faulty, what lead to the development of admixture mapping[59].

In animal or plant breeding, SNPs conferring some advantageous trait but highly associated to different races or varieties (the equivalent of geographic populations) are also avoided, as during marker assisted breeding these SNPs would also drags along undesirable traits related to such races [60]. In wild species such as our case, deliberately ignoring selection happening over SNPs associated to natural population history would be faulty, as it is known that adaptation to different climates in Arabidopsis happened as populations migrated and isolated (see refs. [7,39]).

VI. Climate Genome-Wide Association

Similarly to our GWA with relative fitness, we use each climate variable m (see Section VII.1) as response variable y_m in a similar a LM model as for relative fitness using GEMMA (ref. 36, see Section V):

$$y_m = \mu + \beta_i x_i + \epsilon;$$

This β coefficient for SNP i reflects the correlation of the alternative allele's presence and a climate variable, was used later on in our predictive models (Section VIII).

VII. Climate and modeling

VII.1. Climate layers

We used the classic 19 bioclim variables, monthly data of minimum and maximum temperature, and precipitation (n=12 x 3) (worldclim.org). From those we estimated evapotranspiration using the R package EcoHydRology v. 0.4.12 (ref. 61) and actual evapotranspiration using a bucket model [62] (n=12 x 2). Based on ref. 63 we calculated whether *Arabidopsis thaliana* can grow a given month based on temperature and precipitation (n=12), and calculated the total number of months, i.e. potential growing season (n=1). Within the growing months, we calculated minimum and maximum temperature, and total precipitation (n=3). Finally, using the mean and variance flowering time

(=lifespan) across all our field experiments per accession, and based on their climate of origin using the above variables, we carried out an environmental niche model to generate a map surface of the most likely plants lifespan (n=2). This provides an estimate of the actual growing season, wich we subtracted to the potential growing season to generate one more composite variable (n=1). Each variable is further described in Table S6. A total of 98 raster layers are available as .gri/.grd files (native R format) from: github.com/MoisesExpositoAlonso/araenv, with doi: update.

VII.2. Environmental Niche Models

Genome Environmental niche Models (GEMs) were performed using decision trees to fit models with presence/absence of SNPs as response variable and the climate variables described in the previous section and latitude and longitude a predictors; as previously described [7]. To fit the models we used an Stochastic Gradient Boosting approach with the R package caret (ref. 64). The parameters used to fit the model were: 50 decision trees, an interaction depth of 2, a shrinkage of 0.1, and a minimum of observations at end nodes of 10. This set of parameters were manually determined after carrying our GEMs for some exemplary SNPs and confirming that this set of parameters were typically optimal to reduce residual mean squared error in a Repeated Cross-Validation approach.

We used these models to predict from raster maps of the climate layers a probability ranging from 0 to 1 to have the alternative allele in a map cell. We believe this is a more appropriate output than a discrete 0/1 outcomes, as sometimes alleles were widespreads or at intermediate frequencies in many regions and thus their environment niche was not strictly defined.

VII.3. Climate variability

To study spatial climate variability, for each Arabidopsis natural line, the climate in a 50 Km window from where it was collected. [42], and calculated the coefficient of variation across grid cells.

To study temporal variability, we used 60 years of climate data, 1958-2017 [44] to calculate how much precipitation varied inter-annually across arabidopsis populations. Using annual values for precipitation, we calculated the inter-annual coefficient of variation.

VIII. Predictions of selection differentials with summary statistics

VIII.1 The model

We used a decision tree approach with Random Forest using the R package randomForest (ref. 65,66) to predict the vector (n=1,353,386) of GWA results with relative fitness in one environment, which we call allelic selection differentials *s*, from a 1,353,386 x 98 matrix of GWA associations with climate variables, β_{clim}. We also included as predictors a 1,353,386 x 5 matrix μ of genetic diversity and

frequency metrics: minimum allele frequency, π diversity, Tajima's D, selective sweep likelihood ratio, and selective sweep alpha value. In addition,a we included also as predictors a 1,353,386 x 8 matrix θ of variables taking values of 0 or 1 indicating genomic annotations: intergenic, intron, UTR3, UTR5, exon, synonymous, nonsynonymous, exon noncoding. We ended with a total of 112 variables were used as predictors: $s = f(\beta_{clim}, \mu, \theta)$. In the cases where we trained models with two environments, we also included the climate variables 2 x 98 x_{clim} at our field stations: $s = f(x_{clim}, \beta_{clim}, \mu, \theta)$.

Because training a Random Forest with the full dataset would be computationally expensive, we only trained with 10,000 observations (by using smaller and bigger sets of SNPs, we manually determined that training with more than 10,000 observations did not improve predictions). To test the accuracy and bias we used a different set of 10,000 SNPs, divided in 100 bootstrap samples and report the intervals of the 95% bootstrap distribution.

Several combinations of training and testing were done to validate the predictions of "unobserved" environments. We did four model fittings, each time using training observations from three environments only, and then used observations from a fourth environment for testing. The final model, used for production maps, was trained with observations of four experiments and tested with the same four environments but different observations.

To train the models we used data from either one environment or from the two experimental environments. Because SNPs typically found in warm/hot environments were generally positively selected in the warm environment but negatively in the cold environment, when data from the two field experiments were used for training together, we also needed to include as predictors the values of the local climate at the experimental stations. We used the four experiments where precipitation treatment and location were naturally meaningful, i.e. Germany high precipitation and Spain low precipitation (excluding the combinations of Spain high precipitation and Germany low precipitation). We also tested predictions in either one environment or both (Fig. 3).

VIII.2 A simple visualization of environmental distance

To visualize more directly the relationship between allelic selection differentials at a location and environment where the alleles are found, we calculated the distances between the field station's climate and the allele's home environment. We started using the locations of the 515 data points and the genome matrix (515 individuals, 1,5 Million SNPs). This matrix is called X and the x_{ji} represent the genotype, 0 for reference and 1 for alternative, for the j individual and i SNP. For all 515 locations we extract 19 climate variables available from raster databases (worldclim.org), geting a 515 x 19 matrix E. Then we computed the mean of environments at locations where there are

alternative SNPs: $(E^T X)/X^T X$, what would reflect the alleles' "home environment". Instead of calculating the distance d_j in a single dimension between the field station and the natural population j, we computed the euclidean distance of the 19 bioclim means (variance and mean scaled) or a subset m of them as: $\sqrt{\sum_{i=1}^{19} d_j^2}$. The same approach can be used to calculate the distance to the optimal species' environment. One could assume that the density of geographic locations where Arabidopsis has been sampled would be the "realized optimum".

VIII.3 Limitations and interpretations

As in predictive exercise, our projections have many limitations (discussed below). We believe such limitations should not discourage us to present our projections, and that to move forward in the field of forecasting climate impacts on species, we should aim to build multiple models under multiple scenarios, and later validate those experimentally (as we are currently doing through an experimental evolution network: GrENE-net.org) or with *in situ* observations collected as we move into the future (e.g. iNaturalist.org, iSpot.org). Through this iterative prediction ↔ validation process, we will make serious advances in this complex, yet promising field.

The limitations, enumerated and discussed:

A. Selection is a "relative force". The selection of a genetic variant, depends on what other alternative genetic variants are in the population and at what frequency they are. Thus, the allelic selection differentials are a property of our 515 accessions GWA panel. For example. if in another GWA panel, one SNP is not variable, selection cannot act upon it. Because we carefully selected a global set of accessions representing geographically and genetically the diversity of the species[22], we think our selection estimates are informative of relative trends of the species.

B. Even with an equally diverse GWA panel, if outcrossing and recombination was more common the co-occurrence of two mutations would be different. This would then change allelic selection differentials, as a large amount of selection is driven by linkage disequilibrium in *A. thaliana* because outcrossing is rare (selfing rate 98% [67]).

C. Short-term selection differentials do not necessarily reflect long-term selection coefficients (i.e. over evolutionary times). However, we think short-term selection is more important to understand the demographic fate of a population in the near term.

D. Selection differentials cannot be readily translated into long-term demographic projections. It could be the case that even though there are positively selected genetic variant in the population, the population dies out because of stochastic demographic forces.
E. Bed-hedging strategies such as seedbank demographic dynamics, buffer allele frequency changes over time.

IX. Re-analysis of published data from common garden experiments

IX.1 Manzano-Piedras et al. 2014

Manzano-Piedras and colleagues [45] planted 60 counted seeds per genotype in pots. They monitored how many rosettes established and became reproductive adults (survival proportion). Then they counted the number of fruits per pot and divided them by the number of reproductive adults (reproduction, seed set). We computed lifetime fitness as the product of survival and reproduction.

IX.2 Fournier-Level et al. 2011

Fournier-Level and colleagues [12] germinated seeds in greenhouses, and two weeks after germination (established seedling stage), they transplanted seedlings to outdoor field stations where one plant was transplanted in one pot. They counted how many seedlings survived to reproduction (partial survival proportion), and the number of fruits per plant (reproduction, seed set). We again computed lifetime fitness as the product of survival and reproduction.

We excluded the experiment of Finland in downstream analyses because only 58 natural lines were planted there in the original publication[12] and because later we verified the imputation accuracy was very low (Pearson's $r<0.008$).

IX.3 1001 Genomes x RegMap panel phenotype imputation

The 1001 Genomes panel (http://1001genomes.org/, ref. 23) includes 1,135 natural lines with 11,769,222 biallelic SNPs from Illumina sequencing. The RegMap panel (http://arabidopsis.gmi.oeaw.ac.at:5000/DisplayResults, ref. 5) included 1,307 natural lines, of which 413 lines overlapped. Of those, 185 also overlapped with our 517 lines used in the field experiments.

Of the 157 accessions of Fournier-Level et al., all were part of the RegMap panel, 89 were part of the 1001 Genomes, and 50 overlapped with our field lines. Of the 279 accessions of Manzano-Piedras et al., 150 were part of the 1001 Genomes, and 131 overlapped with our field lines.

Because fitness is heritable, we tried to impute missing data based on the overall genomic relationships among the total 2029 natural lines belonging to either or both panels. After downloading and transforming the RegMap dataset to PLINK format, we overlapped genome-wide

SNPs and filtered them for a genotyping rate of 95%, which ending up with 154,090 biallelic SNPs. This is more than enough to calculate a relationship matrix A which we computed using the R package rrBLUP (ref. 68). The data of survival, reproduction, and lifetime fitness was of average per genotype, so we fit a classic GBLUP: $y = Zg + \epsilon$; where y is the fitness trait of interest, Z is a design matrix of genotypes and g is random effect factor with covariance matrix equal to the relationship matrix $g \sim MVN(0, A\sigma_g^2)$. Heritability of traits and imputation accuracy from the Manzano et al. and Fournier-Level experiments is in Table S9. We noticed that imputation accuracy was only very low for fitness in the Finland experiment of Fournier-Level et al (2011).

IX.4 Sanity check analysis for imputation and geographic predictions

We carried out a sanity check analysis in order to be sure that the imputed fitness from other experiments was not just an artifact phenotype with the same structure as the relationship matrix. This would mislead us to think there is predictability, as we would expect that allele selection differentials calculated in such artifact phenotype would depend on population structure and thus would likely predictable from climate structure.

For sanity check, we shuffled the genotype identities from Fournier-Level and Manzano-Piedras with their fitness values. Then we repeated the GBLUP analysis doing 50 rounds of shuffling and computed heritabilities and prediction accuracies. We confirmed that heritability was negligible ($1x10^{-9} < h^2 < 1.6^{-3}$) and so was the accuracy of imputation ($-0.047310 < r < 0.070380$). This indicated that in the absence of a biologically true heritable variation, imputation of fitness would be random and not dependent on the relationship matrix.

We also were concerned that geographic predictions could be driven by some underlying bias in our analyses, i.e. bias inherent to geographic sampling, population history of genotypes chosen, etc. In other words, that the null expectation of predictability in predictability would be non-zero. As before, we randomized fitness values with genotypes for all six datasets. Then, repeated GWA to estimate allelic selection differentials (as Fig. 1), and trained different combinations of GES models to re-predict allelic selection differentials at teach location based on climate (as Fig. 3). We confirmed that, differently from the real analyses of Fig. 3, there was no significant predictability (Fig. S15).

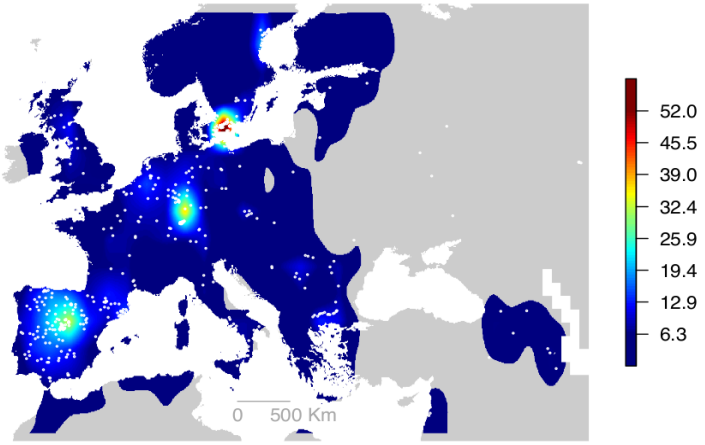
IX.5 An explanation for "inverse predictability"

We noticed that using only our two experiments for model training, there was "inverse predictability" for the three experiments from ref. 12. While the sign of inferred selection differentials was the opposite of the observed values ($-0.33<r<-0.51$, $P<0.001$), the magnitude of selection was correctly inferred ($15\%<R^2<25\%$, Fig. 3A). Such a phenomenon could arise from a

723 number of reasons. First, the worldclim.org climate averages (1960-1990) at 2.5 arc-minutes
724 resolution might vary from the truly experienced environmental conditions in the years the
725 experiments were conducted. Such climate variability can exert opposite selection in different
726 years[69]. Second, differences in experimental design could lead to different lifetime fitness estimates.
727 In ref. [12], early survival of seedlings was not measured at all, as only seedlings that had survived for
728 two weeks in greenhouses were transplanted into the field. In the Southern Spain experiment[45],
729 seeds were sown directly in the field, as in our own experiments, and accordingly, we had "positive
730 predictability" (r=0.24, Bootstrap CI=0.09—0.41). In further support of this experimental design
731 confounder, when we trained GES models with only reproduction-based allele selection differentials
732 in our experiment in Germany, i.e., excluding early survival from lifetime fitness, we correctly
733 predicted the sign of selection differentials in Fournier's Northern Germany experiment (r= 0.392,
734 Bootstrap CI= 0.20—0.57) (for null expectations see Supplementary Methods IX.4). The differences in
735 predictions between two- and six-environment-trained models did not yield differences in
736 downstream analyses and conclusions (correlation between predictions, r=0.56, $P<10^{-16}$), but
737 predictability increased with the number of experiments included in the training set (r= 0.761 ,
738 Bootstrap CI= 0.55—0.90, R^2= 0.517, Fig. 3A).

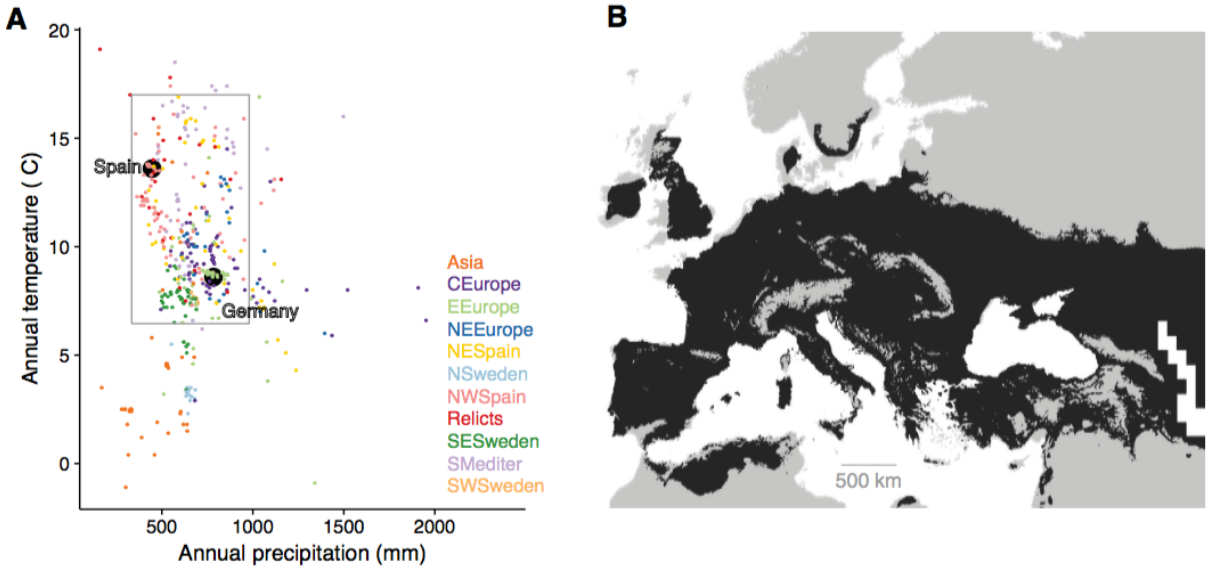
SUPPL. FIGURES & MEDIA

Figure S1. Accession locations



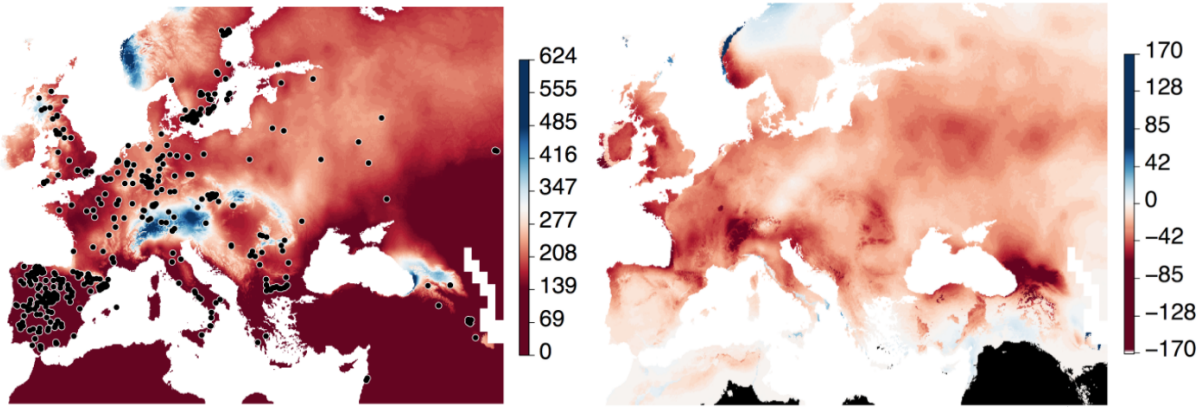
Points indicate where the 517 *A. thaliana* accessions were collected. The color gradient is the density of samples from our study roughly in squares of 200 km x 200 km. The limits of the colored area were determined using a combined density grid from gbif.org and 1001 Genomes records. The density was generated in a grid of 125 minute resolution and by applying a bilinear and then Gaussian smoothing. The threshold was chosen to be the 50% of the upper distribution, which roughly correspond to 10 records per 200 km x 200 km square. The color area was used to mask future climate change predictions, as we prefer to make predictions only in regions where the presence of *A. thaliana* is rather likely and continuous.

749 **Figure S2. Extent of native environments of *A. thaliana***



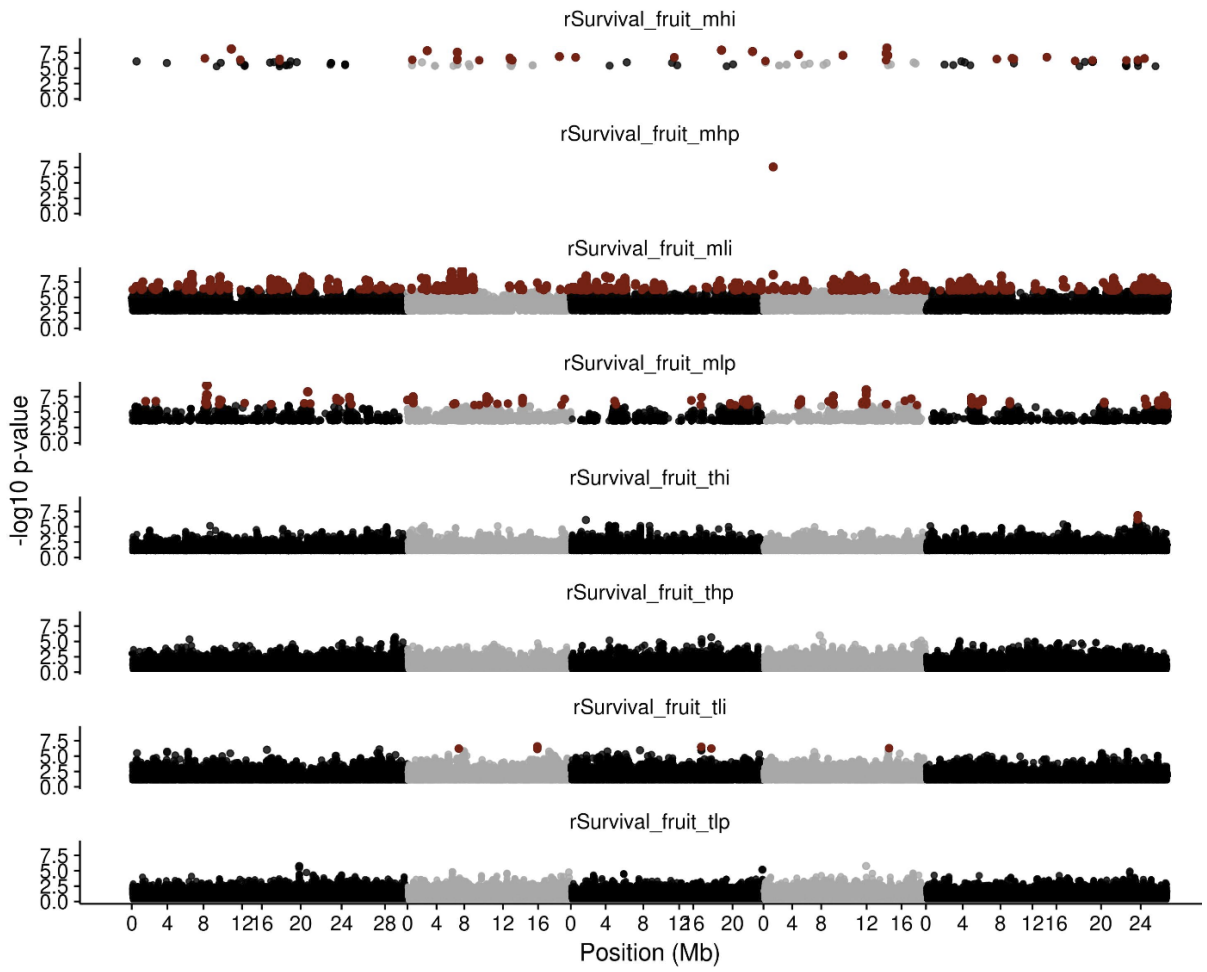
750 (**A**) Classic biplot of precipitation *vs.* temperature of origin of accessions (black dots) and field
751 experiment of Spain (sepia) and Germany (green). Grey box indicates a range of +/- 30% in
752 precipitation and temperature within the experiments ranges. (**B**) Areas that would be within the
753 grey box in (A).

Figure S3. Map of environmental change



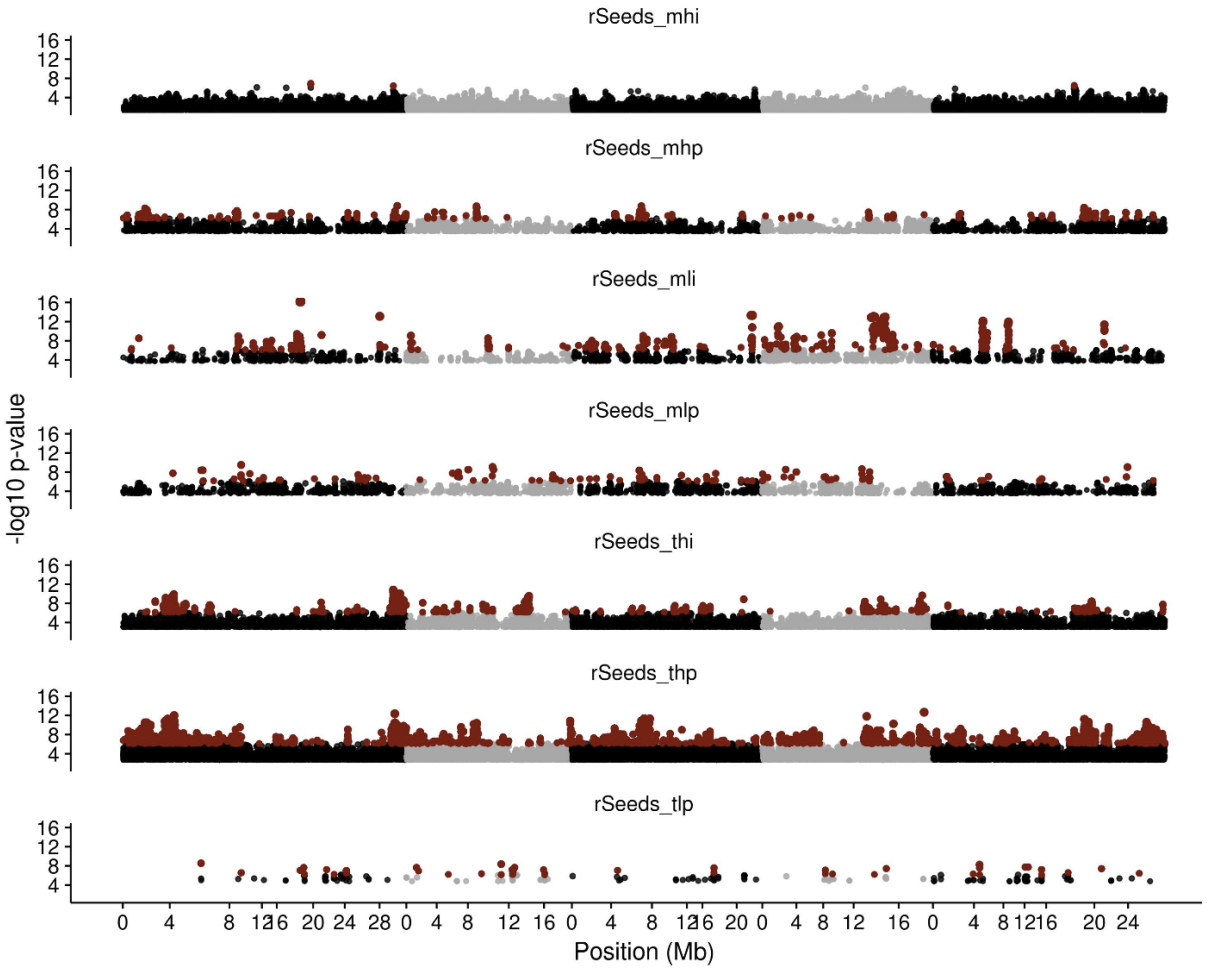
Precipitation during the warmest quarter (bio18, left), and change in the same predicted for 2070 (rcp 8.5) (right). Black areas indicate regions where precipitation will go below what is found for any currently sampled *A. thaliana* population (black dots, left).

Figure S4. Genome maps of survival



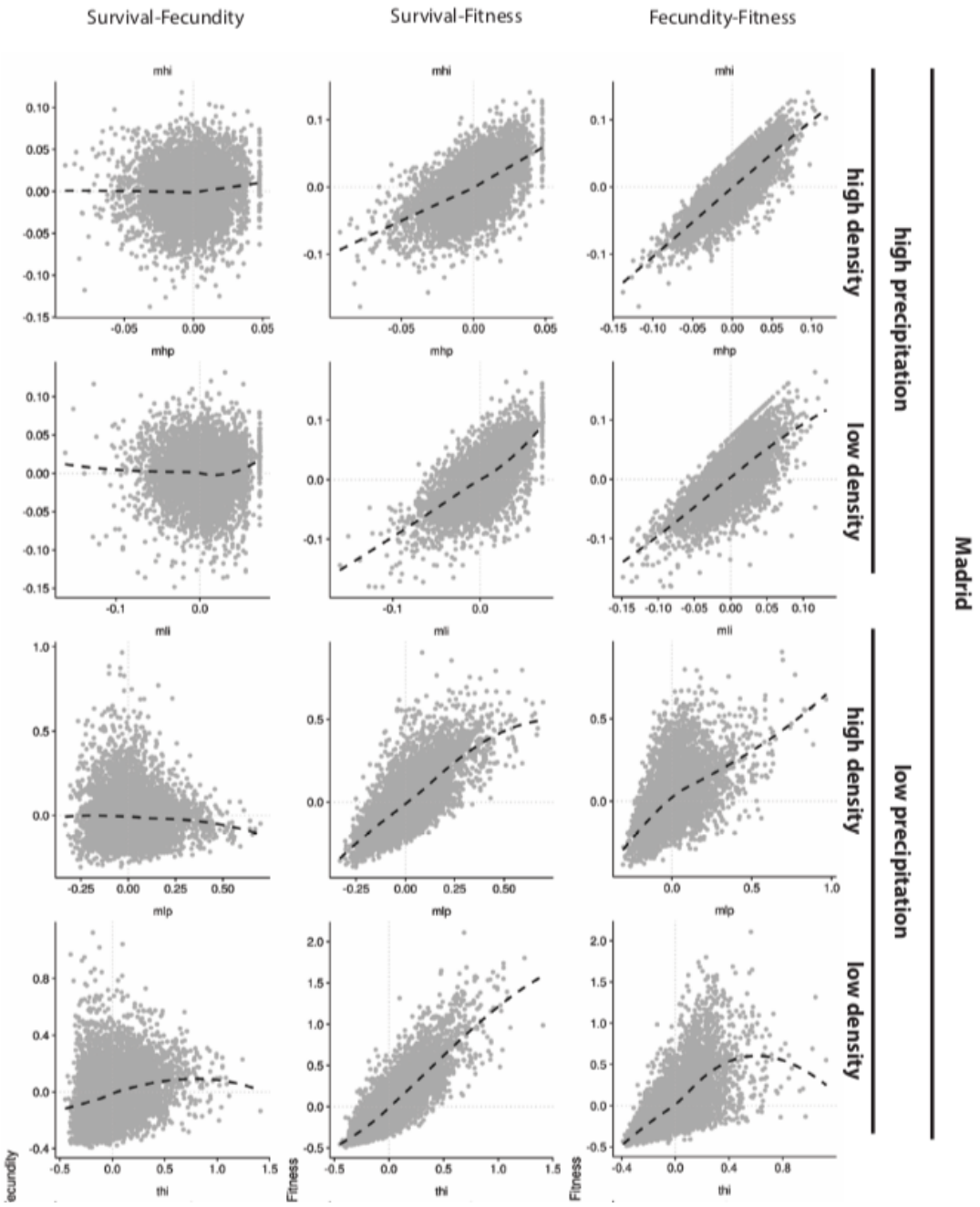
Same as Fig. 1 but only using the survival component of fitness.

760 **Figure S5. Genome maps of fecundity**

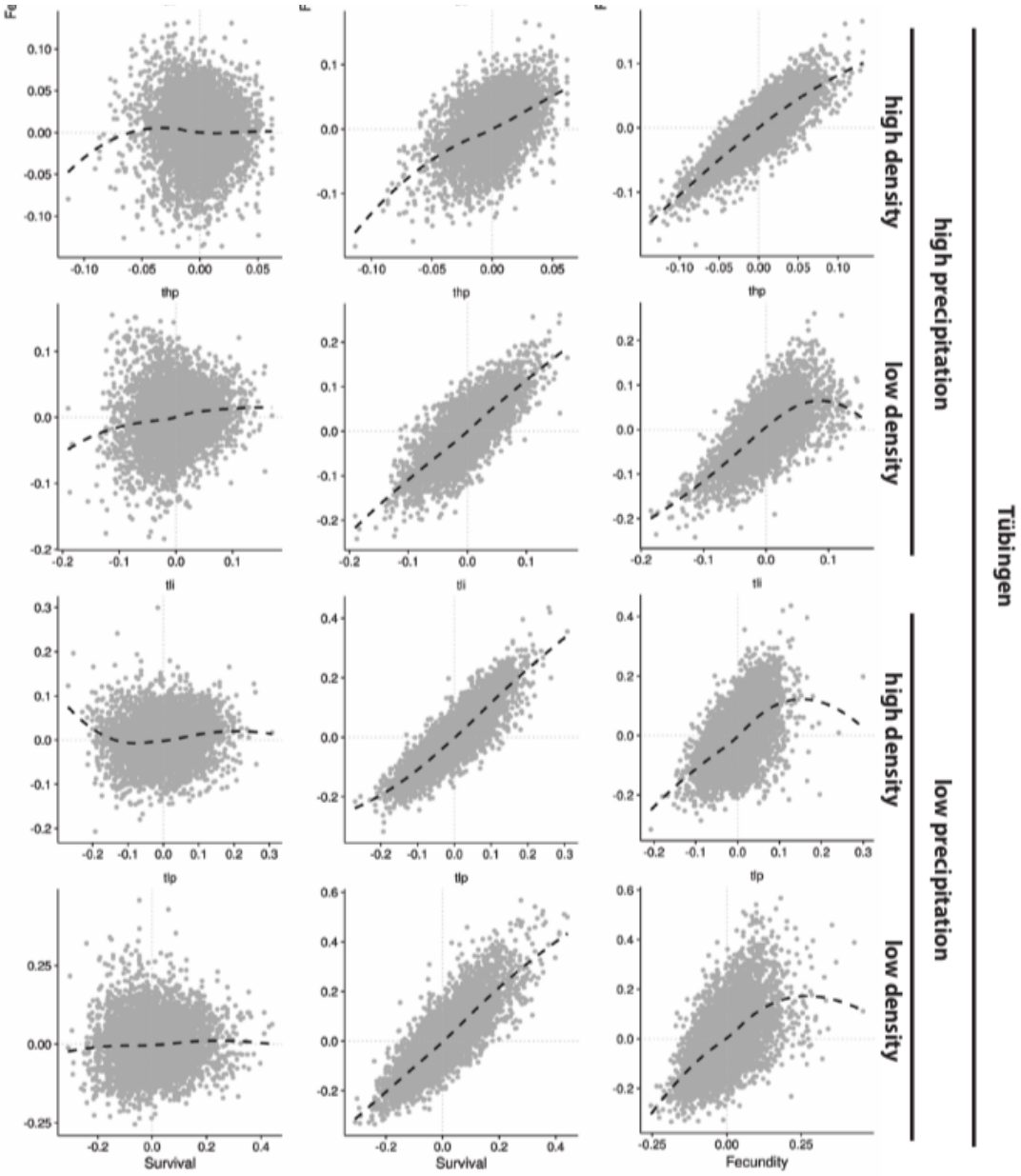


761 Same as Fig. 1 but only using the fecundity component of fitness.

762 **Figure S6. Trade-offs in survival and fecundity**

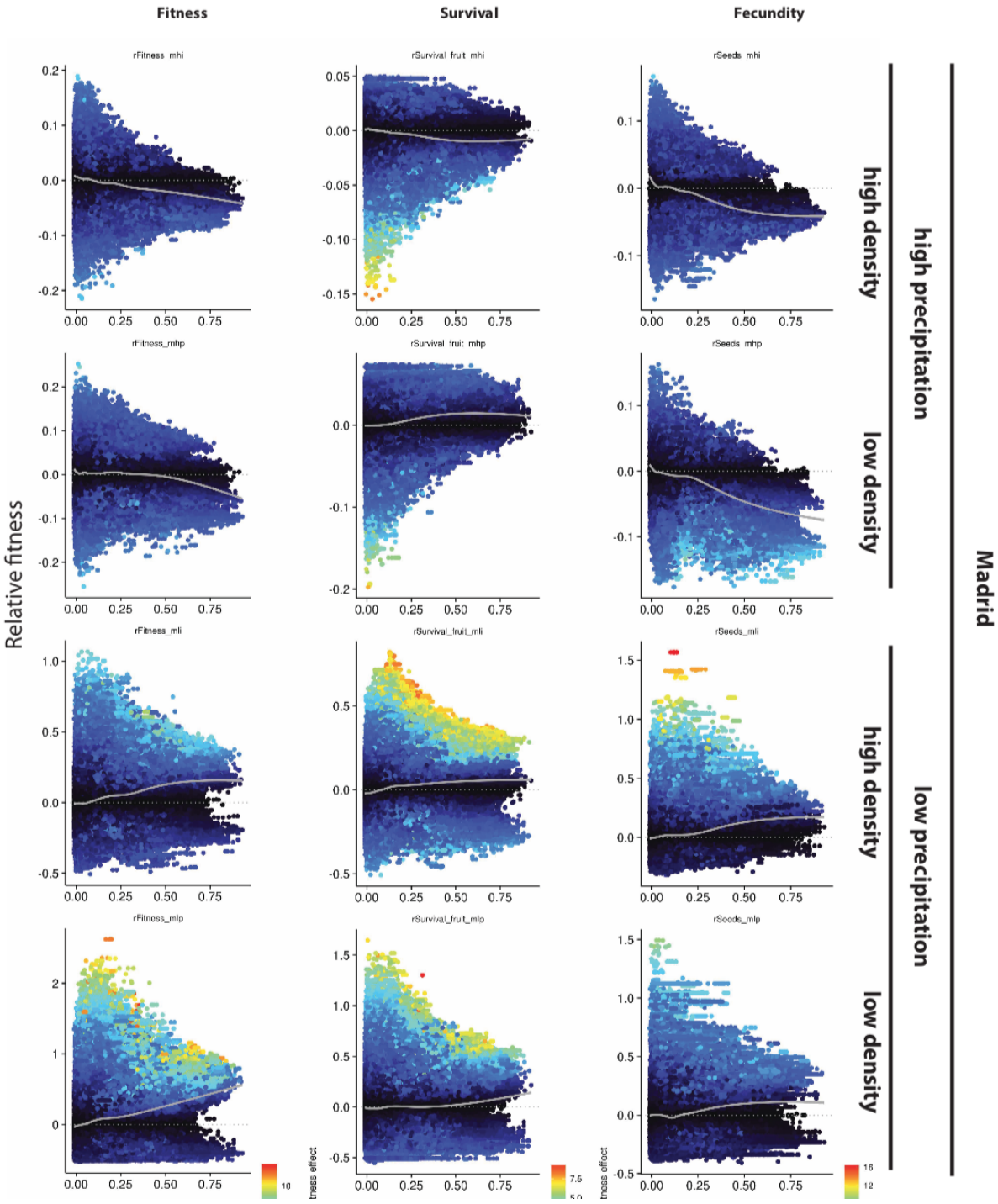


(Fig. S6 continued)

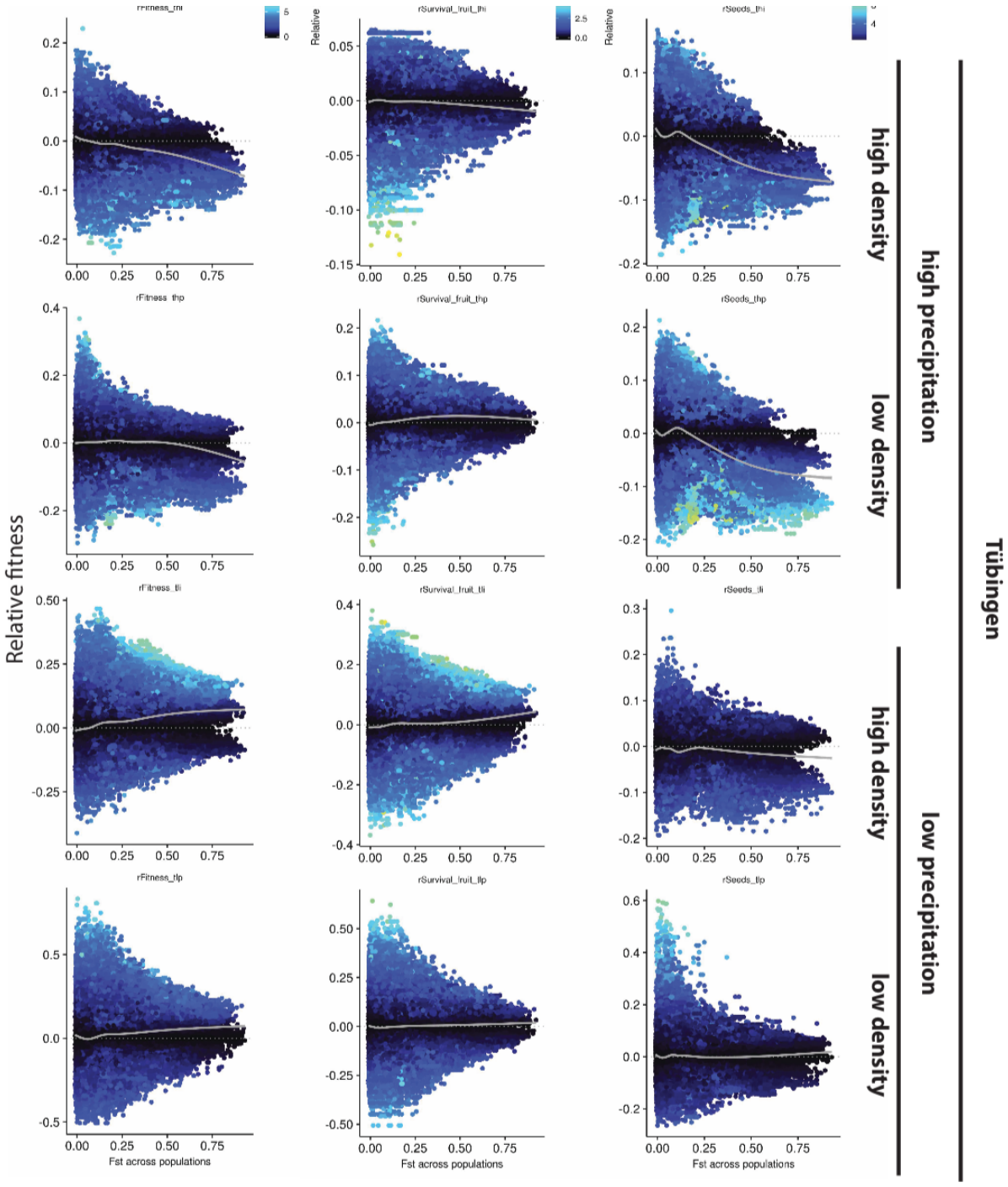


Comparisons of allele selection differentials computed only with the survival component, only with the fecundity component, and with both.

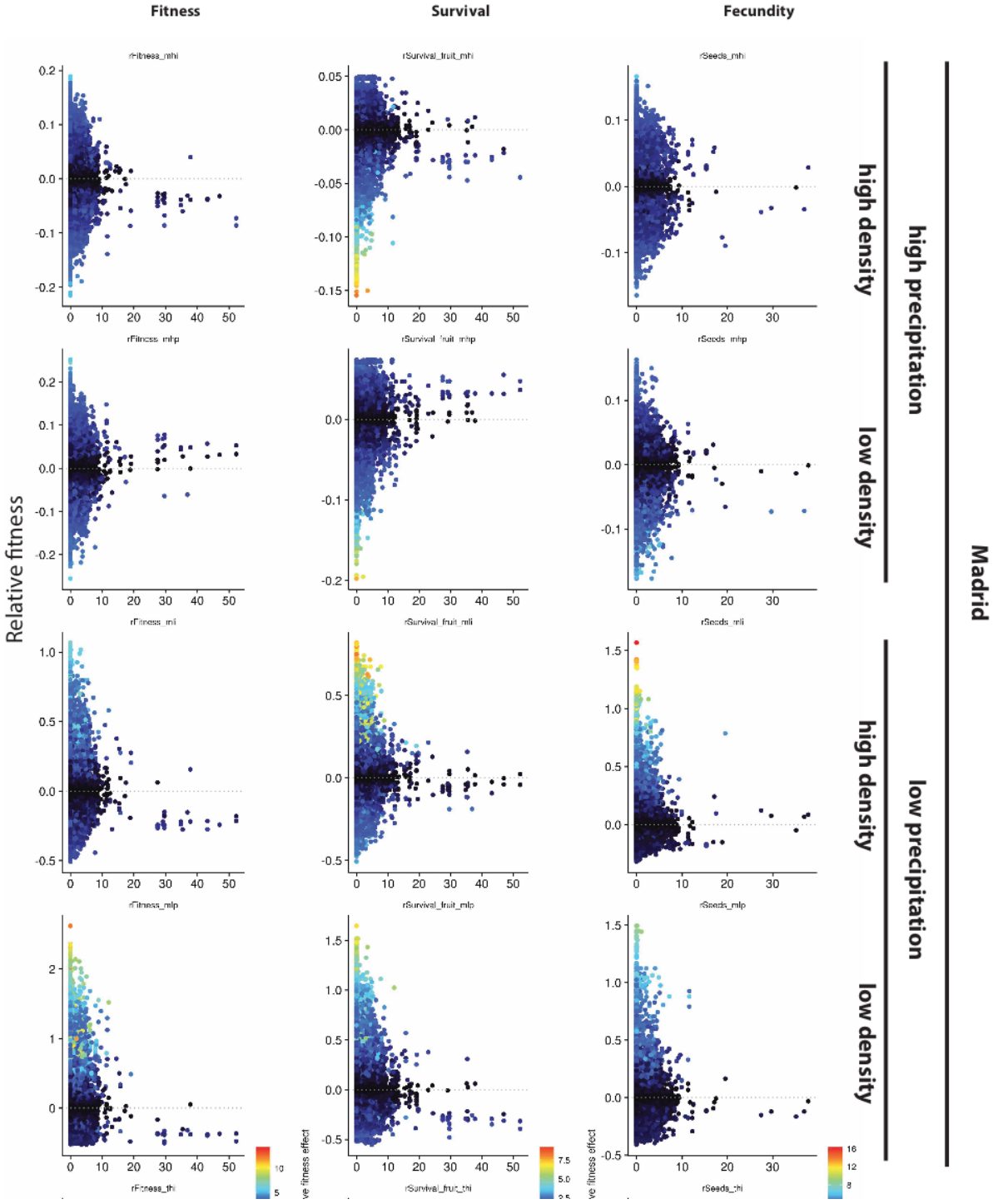
766 **Figure S7. F_{st} and empirical selection**



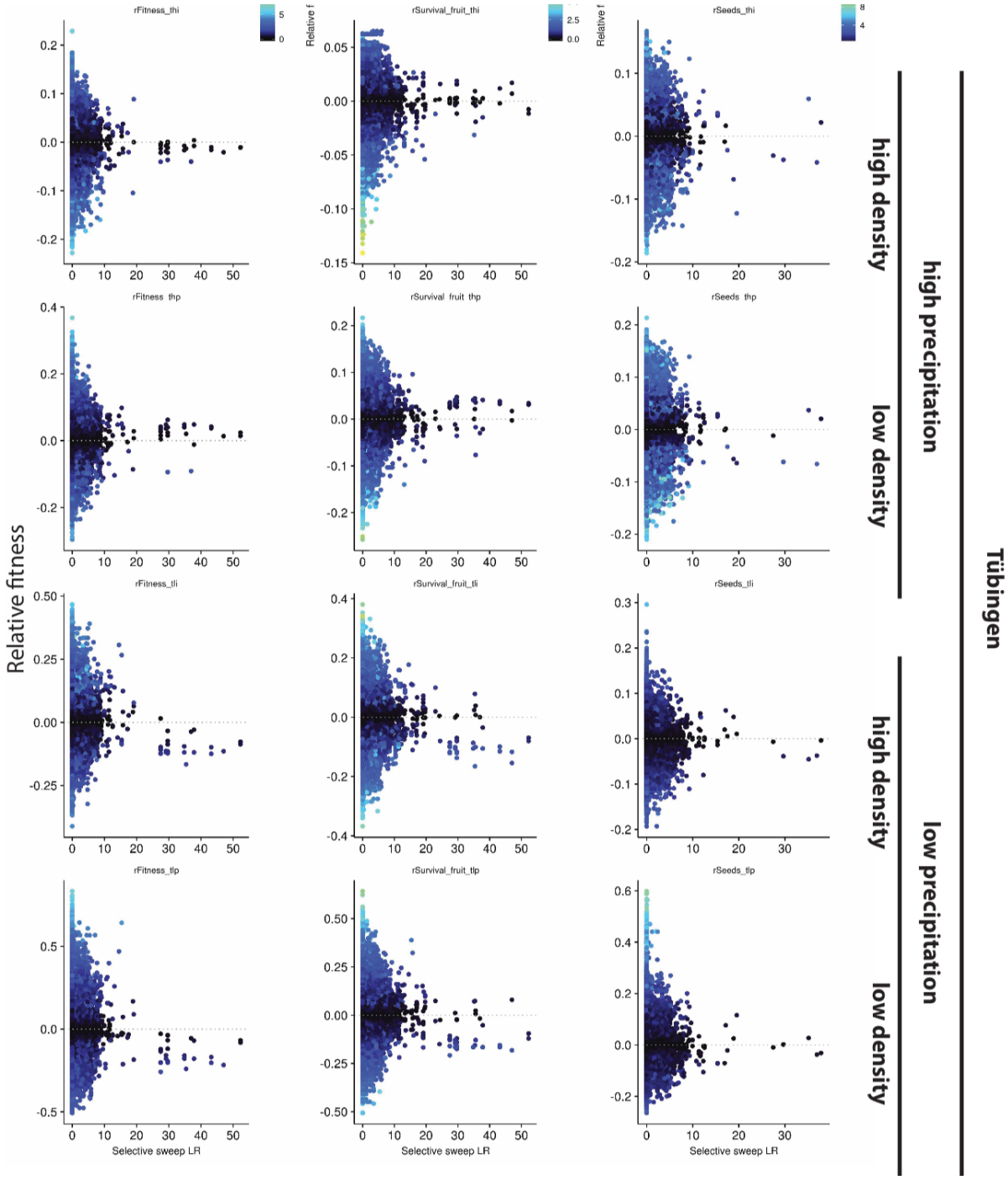
767 **(Fig. S7 continued)**



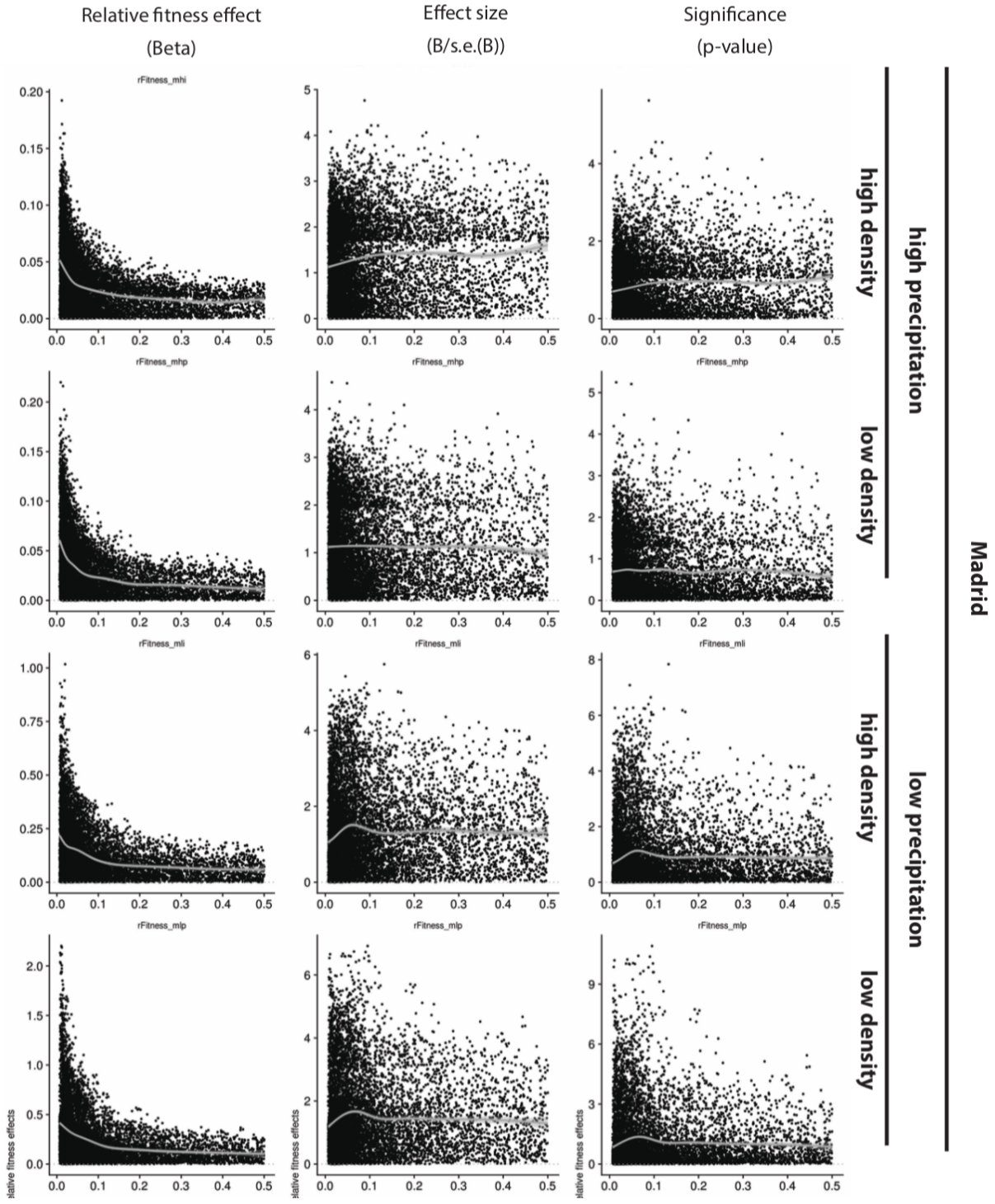
768 **Figure S8. Sweeps and empirical selection**



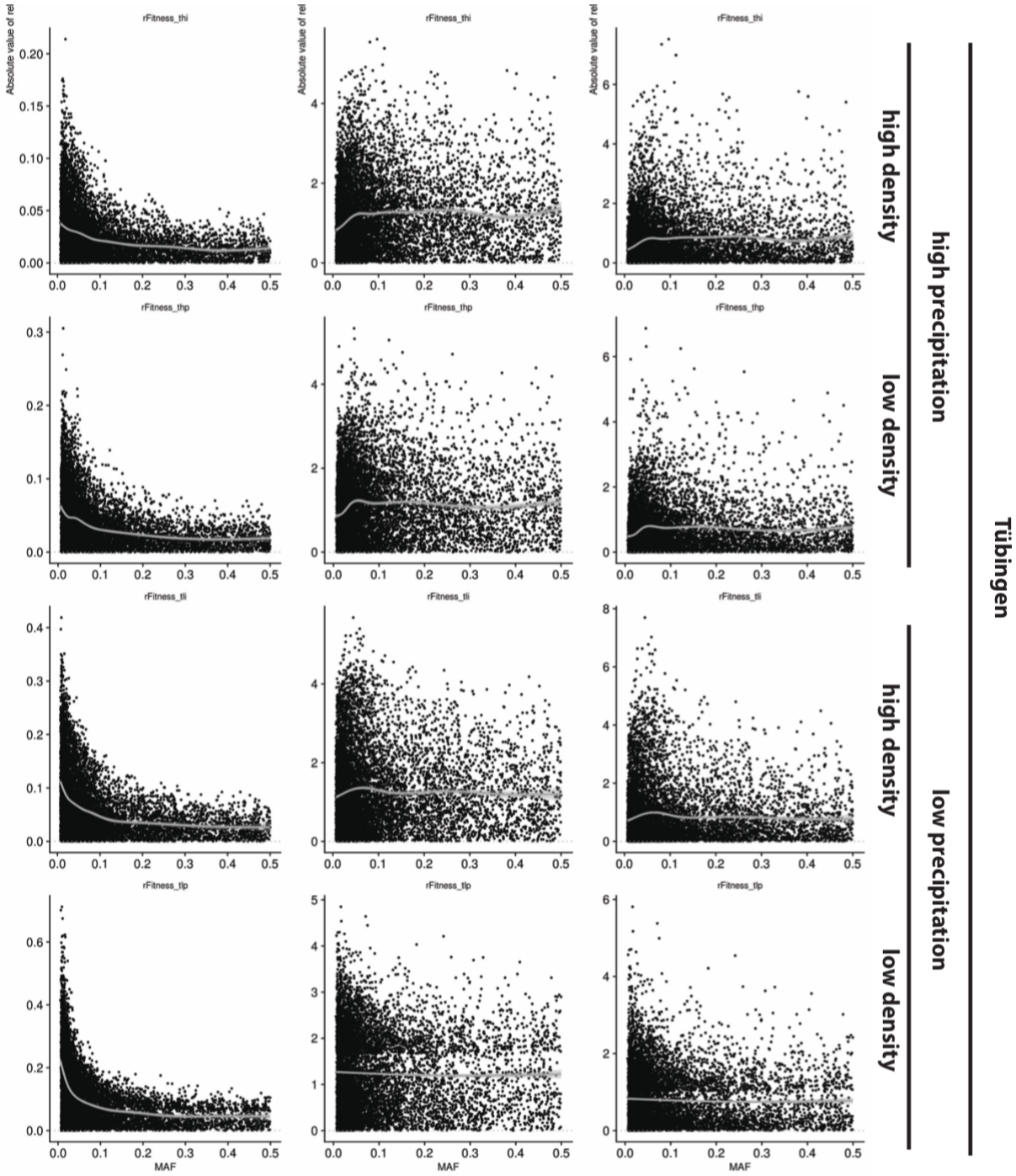
769 **(Fig. S8 continued)**



770 **Figure S9. MAF and relative fitness**



(Fig. S9 continued)



772 **Figure S10. Environmental distance and selection differentials**

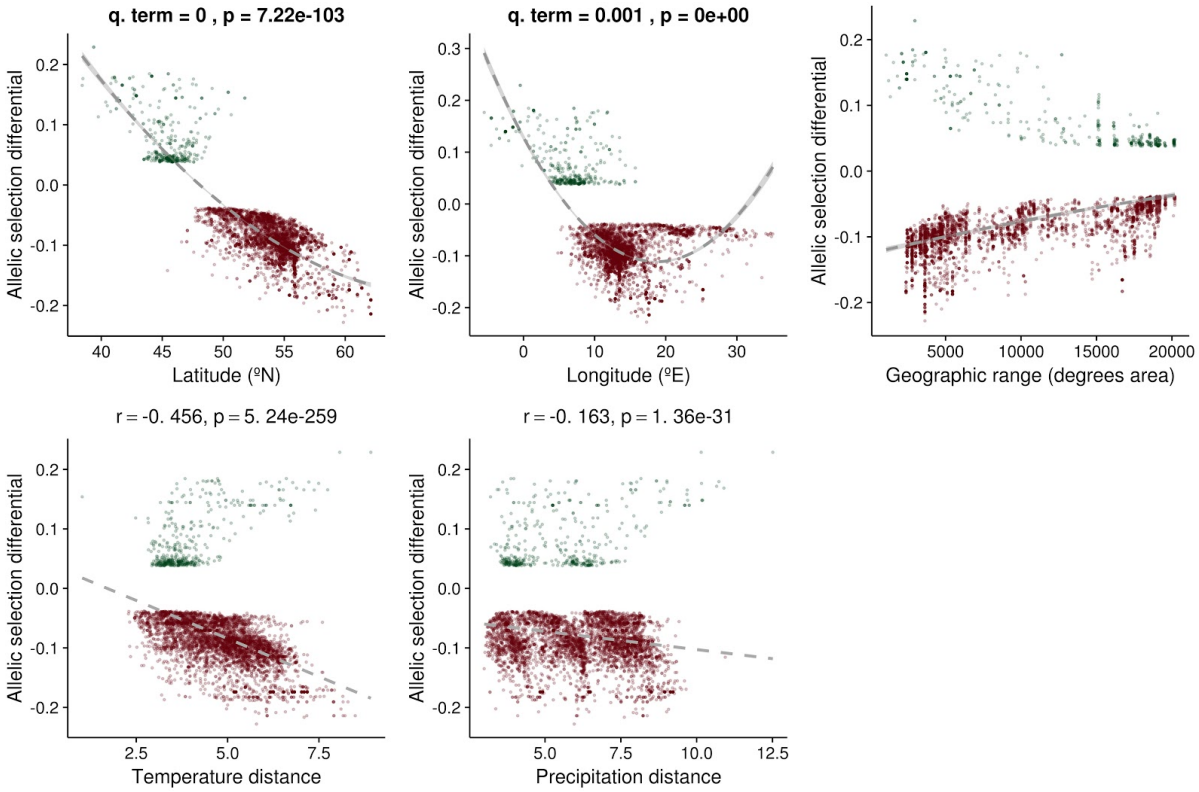
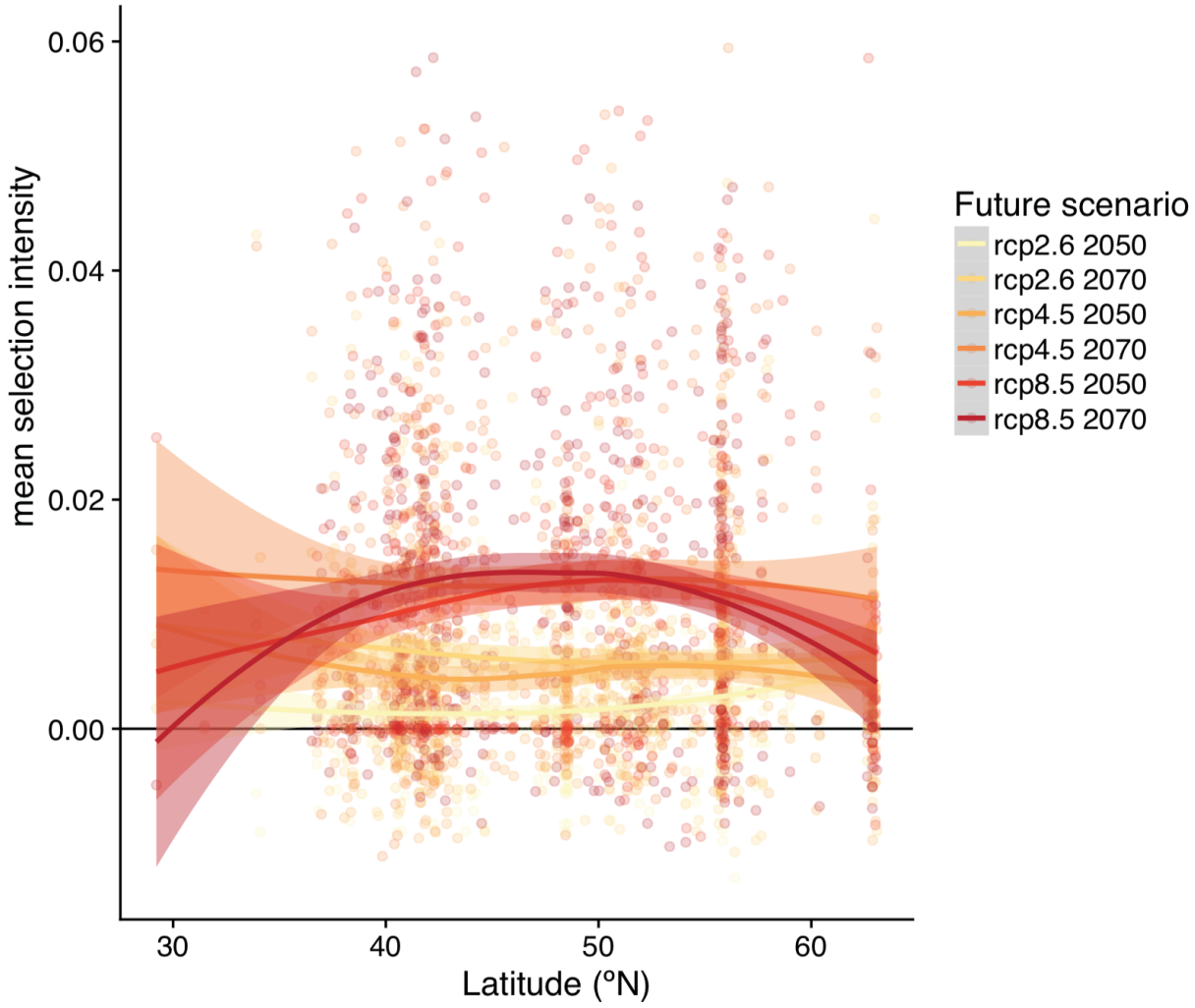
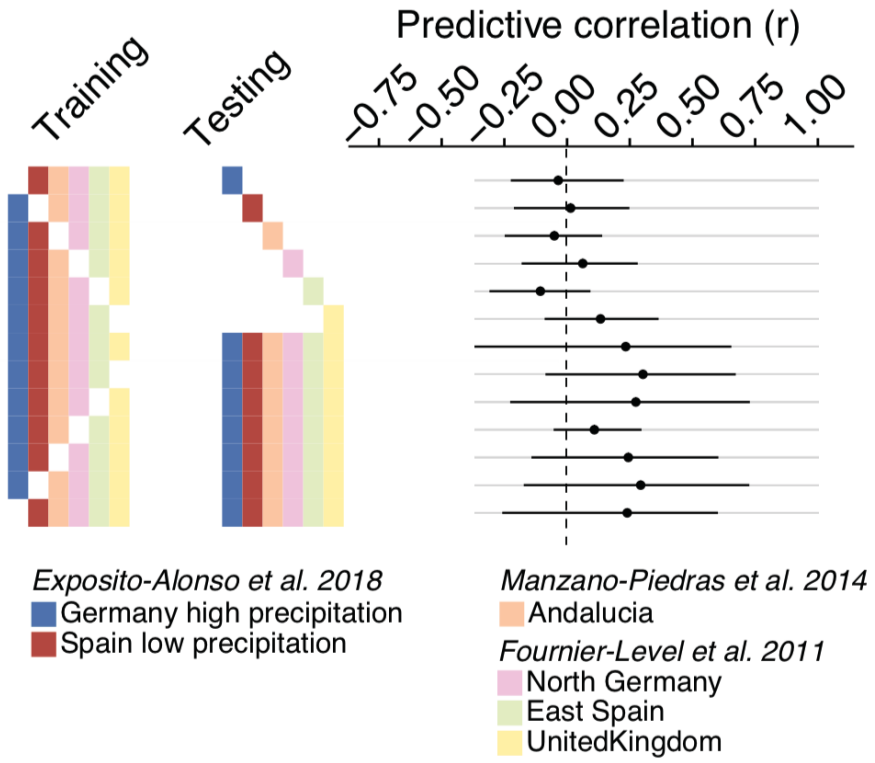


Figure S11. Future change in selection for different climate change scenarios



Same as Fig. 3G but for different climate change scenarios. The worse the scenario in CO_2 emissions, the stronger the predicted increase in selection intensity.

Figure S12. Null expectation of predictability



Same as Fig. 3, but with randomized fitness values associated to genotypes. We could not find any model combination that correctly inferred selection (95% bootstrap confidence overlaps with zero). This is a proof of concept that the predictability we show must have a biological basis, in which the combination of climate of origin of a genetic variant and the local climate allows to infer selection over such a variant.

SUPPL. TABLES & DATASETS

Supplementary tables are available in the online version of the paper with doi: xyz.

Table S1. List of accessions

The 1001 Genomes Project identifier, country of origin, and degrees latitude and longitude are reported.

Table S2. Summary of fitness data

Average survival, fecundity, and lifetime fitness. Total number of genotypes that at least one replicate survived per experiment. (Abbreviations: The three characters of the codes: MLI, MLP, MHI, MHP, TLI, TLP, THI, TLP; indicate M=Madrid (Spain), T=Tübingen (Germany), L=Low precipitation, H=High precipitation, I=Individual replicates (one plant per pot), P=Population replicates (up to 30 plants per pot).

Table S3. Heritability of traits

Broad sense heritability per trait (variance explained by genotype line), as calculated from a generalized linear mixed model, is reported as: σ_g/σ_{Total}. The proportion of variance explained by nuisance factors such as block (tray), position of the tray within a treatment block, and position of plant within a tray are reported in the same way. (Abbreviations: The three characters of the codes: MLI, MLP, MHI, MHP, TLI, TLP, THI, TLP; indicate M=Madrid (Spain), T=Tübingen (Germany), L=Low precipitation, H=High precipitation, I=Individual replicates (one plant per pot), P=Population replicates (up to 30 plants per pot).

Table S4. Number of SNPs with significant selection differentials

All significant variants from marginal GWA after FDR and Bonferroni correction and all variants with non-zero probability of inclusion from conditional GWA.

Sharing of significant variants across experiments.

Table S5. Expected allele frequency changes in response to selection

Summaries of allele frequency changes per experiment.

Table S6. Variable importance of predictive models

Sharing of significant variants across experiments.

Table S7. Predictability of environmental models

After training GES models with a set of experiments, we inferred allelic selection differentials on another set of experiments and compared those with the real allelic selection differentials. We calculated Pearson's product-moment correlation r and % of variance explained R^2 using a regression. 95% confidence intervals were calculated with 100 bootstrap replicates.

Table S8. Description of climate variables

Climate variables used for environmental models are described and their sources reported.

Table S9. Spearman correlations among climate variables and phenotypes

We calculated Spearman's r across phenotypic and climate variables.

Table S10. GBLUP heritability and imputation accuracy of data from published field experiments

We used GBLUP to impute fitness from Fournier-Level et al. (2011) and Manzano-Piedras et al. (2014) into our 517 global accessions. We report heritability, Pearson's r between GBLUP predicted fitness and real fitness, and the significance of the correlation test.

Table S11. Correlation between inferred natural selection intensity and other variables

Spearman's r between selection intensity and diversity metrics or climate metrics.

SUPPLEMENTARY REFERENCES

53. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27,** 2156–2158 (2011).
54. 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).
55. 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).
56. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33,** 1–22 (2010).
57. Kirkpatrick, M., Johnson, T. & Barton, N. General models of multilocus evolution. *Genetics* **161,** 1727–1750 (2002).
58. Novembre, J. & Barton, N. H. Tread Lightly Interpreting Polygenic Tests of Selection. *Genetics* **208,** 1351–1355 (2018).
59. Shriner, D. Overview of admixture mapping. *Curr. Protoc. Hum. Genet.* **1,** 1.23.1–1.23.6 (2013).
60. Yu, J. *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38,** 203–208 (2006).
61. Dr, F., Mt, W., Ja, A., Ts, S. & Zm, E. EcoHydRology: A community modeling foundation for Eco-Hydrology. (2014).
62. Golicher, D. Implementing a bucket model using WorldClim layers. (2012). Available at: https://rpubs.com/dgolicher/2964.
63. Lasky, J. R. *et al.* Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Mol. Ecol.* **21,** 5512–5529 (2012).
64. Kuhn, M. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software, Articles* **28,** 1–26 (2008).
65. Breiman, L. Random Forests. *Mach. Learn.* **45,** 5–32 (2001).
66. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2,** 18–22 (2002).
67. Platt, A. *et al.* The scale of population structure in Arabidopsis thaliana. *PLoS Genet.* **6,** e1000843 (2010).
68. Endelman, J. B. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome* **4,** 250–255 (2011).
69. Exposito-Alonso, M., Brennan, A., Alonso-Blanco, C. & Picó, F. X. Spatio-temporal variation in fitness responses to contrasting environments in Arabidopsis thaliana. *Evolution (in press)* (2018).