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 an agent of directional selection that forces populations to change and adapt, or face extinction. Current assessments of the risks associated with climate change[1,2], however, do not typically take into account that natural selection can dramatically impact the genetic makeup of populations[3]. We made use of extensive genome information in *Arabidopsis thaliana* and measured how rainfall-manipulation affected the fitness of 517 natural lines grown in Spain and Germany. This allowed us to directly infer selection at the genetic level[4]. Natural selection was particularly strong in the hot-dry Spanish location, killing 63% of lines and significantly changing the frequency of ~5% of all genome-wide variants. A significant proportion of this selection over variants could be predicted from climate (mis)match between experimental sites and the geographic areas of where variants are found (R^2=29-52%). Field-validated predictions across the species range indicated that Mediterranean and Western Siberia populations — at the edges of the species' environmental limits — currently experience the strongest climate-driven selection, and Central Europeans the weakest. With rapidly increasing droughts and rising temperatures in Europe[5], we forecast a wave of directional selection moving North, putting many native *A. thaliana* populations at evolutionary risk.

To predict the future impact of climate change on biodiversity, the typical starting point has been climatic tolerances inferred from the current species distributions. These tolerances are usually treated as static, and risks are assessed based on whether species' environmental niches will shrink[1,2] or shift faster than the species can migrate[1,6]. However, these approaches do not account for within-species genetic variation, and for natural selection causing species to genetically change and adapt over time[3]. To predict the "evolutionary impact" of climate change on a species, i.e. how much genetic change is required for adaptation to climate change, we thus need to quantify and model environment-driven natural selection at the genetic level. Thanks to species-wide genome scans[7–9], as well as genome associations with climate of origin[10–14], we increasingly understand the genomic basis of past selection and climate adaptation, which has been used to estimate future adaptation debt or "genomic vulnerability" [10,11].

Natural selection, however, is only indirectly inferred in the types of analyses discussed above. The best way to directly quantify selection in a specific environment is provided by field experiments in which multiple genotypes of a species are grown together in a common environment [15,16]. With such experiments, relative fitness can be directly associated with genetic variation across populations[4,17–19]. Ideally, one would carry out such field experiments at many different sites throughout the species range, but this is rarely practical[20]. Nevertheless, an emergent finding is that individuals are normally locally adapted and that local genotypes are often positively selected over

foreigners in their "home" environment, while negatively selected in their "away" environments[21,22]. An intuitive conclusion is that it should be possible to derive a metric of natural selection from the extent of change in local climate at an individual's home. Here we combine high-throughput associations of genome and current climate variation with experimentally quantified *in situ* natural selection in the plant *Arabidopsis thaliana*. We exploit these associations to forecast natural selection driven by future climate change, and how it impacts the genomic variation of a species across its geographic range — what we interpret as a new metric of evolutionary risk of populations.

To study climate change-driven 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 in Madrid (Spain, 40.40805ºN -3.83535ºE), and at the distribution center in Tübingen (Germany, 48.545809ºN 9.042449ºE) (for details see ref. 20). 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. 20). In fall of 2015 we sowed over 300,000 seeds of 517 natural lines 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 exactly 30 seeds that were allowed to germinate and grow without intervention. At the end of the experiment in June 2016, we had collected data from 23,154 pots, 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 was generally higher in the most stressful environment, which was defined by reduced survival ($0.00<H^2<0.551$; Table S3), i.e., in Spain under low precipitation and at high density. In this environment, only 193 of the 517 accessions survived, whereas in Germany at least a few plants of each accession reproduced.

In each experimental environment, we quantified genome-wide selection at the genetic level based on the difference in relative fitness of lines with the minor and the major allele at each genomic position (1,353,386 biallelic SNPs across 515 lines with high-quality genome information) (Fig. 1). Our approach identifies both causal variants, as well as many more variants that are in significant linkage disequilibrium (LD) with causal variants[24,25] — due to so-called background selection or genetic hitchhiking. We use the term *allelic selection differentials* (s, called total selection by Thurman and Barrett [4]), to denote the realized selection affecting each SNP resulting from the combination of selection acting directly on the focal variant, and the indirect effects due to selection on causal SNPs that are in LD with the focal variant. Calculating allelic selection differentials using Linear Models (LM-GEMMA, ref. 26, see Supplemental Methods V.3), we found a total of

421,962 SNPs with allelic selection differentials below a 0.05 significance threshold (Benjamini & Hochberg FDR correction) in at least one of the eight environments (see Table S2). 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[27] prevails in this climate-manipulation experiment. These high numbers are not surprising, given that we expect to capture many SNPs that are only indirectly selected. Thinking about our experiment as studying a population of plants with multiple genotypes, the change of allele frequencies in response to one generation of selection would be up to 10% in Spain and low precipitation, while it would not exceed 2% in the benign high-precipitation environment in Germany (see Supplemental Methods V, Fig. S9). While variants inferred to be under positive or negative selection after Bonferroni-correction were overall more likely to be located in intergenic regions than in genes (Fisher's Exact test Odds ratio [Odds]=1.11, $P=7x10^{-30}$), such variants were enriched for nonsynonymous mutations (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 [28,29], and a potentially significant demographic decimation — what Haldane called "the demographic cost of natural selection"[30].

Changes in allele frequency are not only determined by the adaptive value of a variant but also the alleles it is linked to. We therefore improved the detection of direct targets of selection by correcting for LD-driven effects[25,31] using Bayesian Sparse Linear Mixed Model associations with relative fitness (BSLMM-GEMMA, ref. 31), see Supplemental Methods V). This analysis indicated that the fraction of the genome likely to be a target of selection was only $8x10^{-5}$—$5x10^{-6}$. This fraction was much smaller than what we had identified with significant allelic selection differentials ($2x10^{-5}$—0.001), confirming that selection must be mostly indirect[4,32,33].

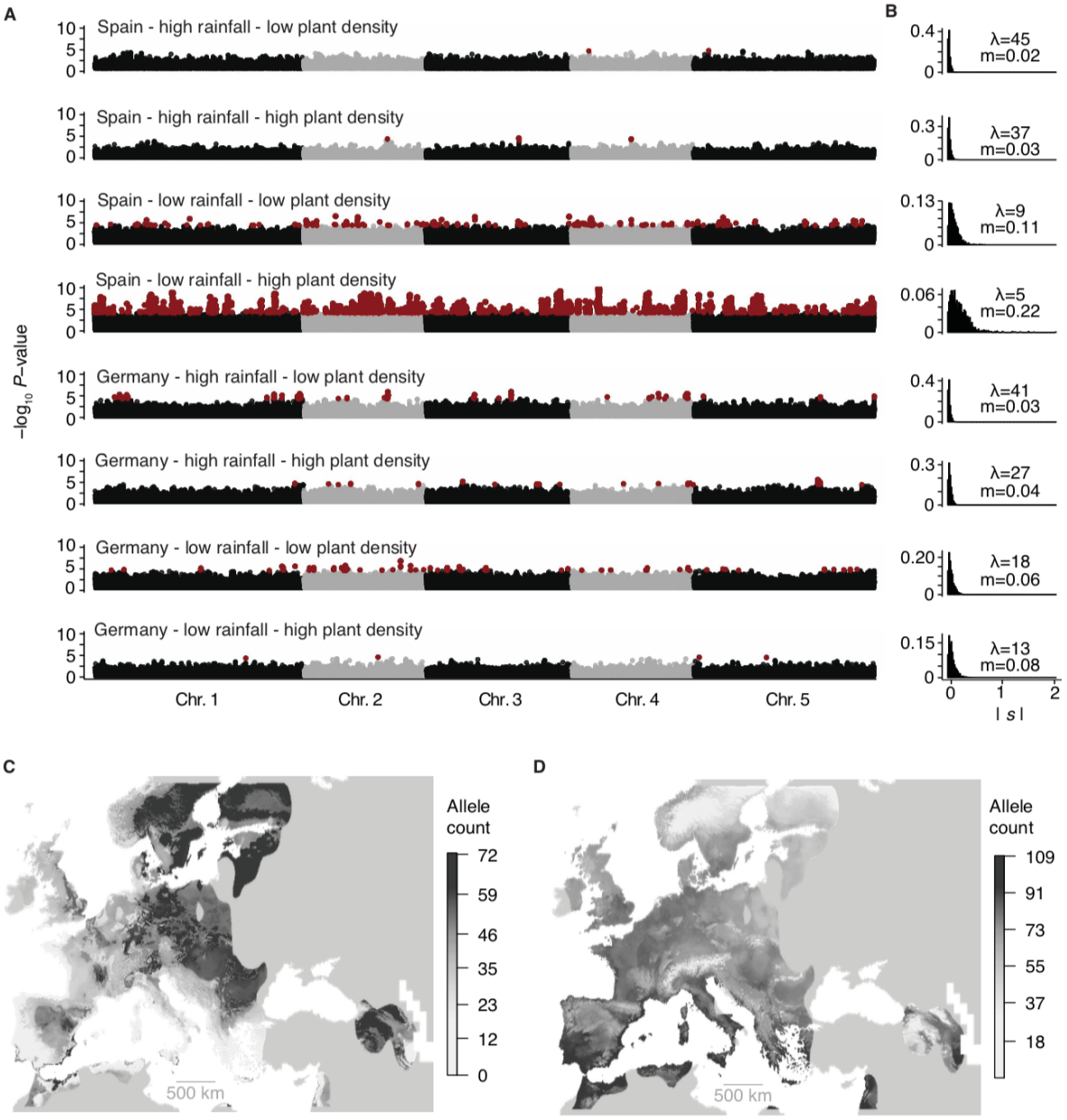
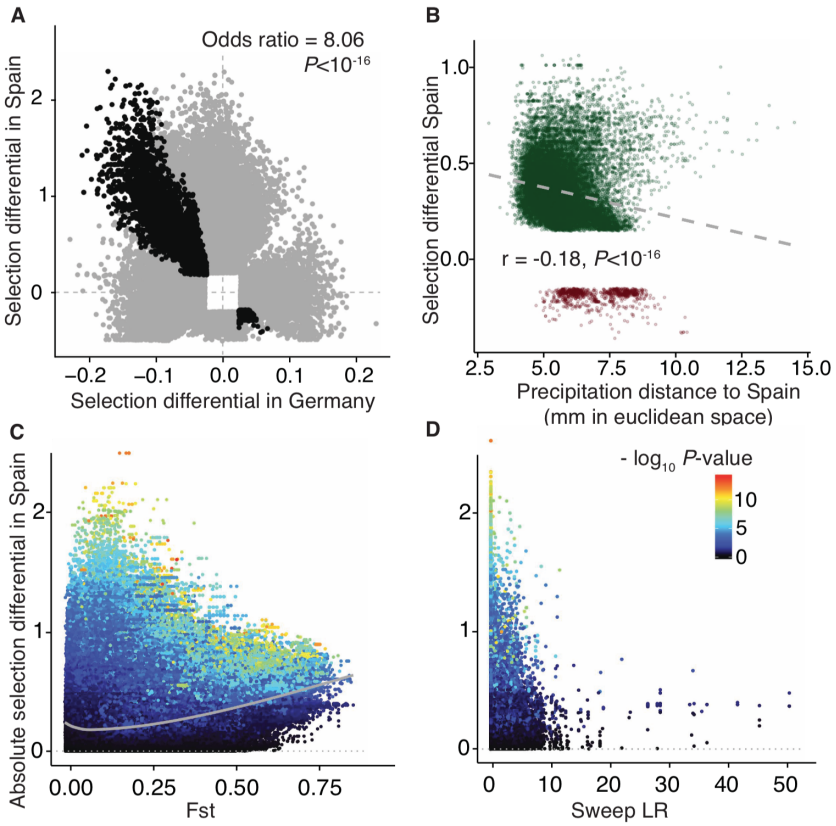


Fig. 1 A genome map of allelic selection differentials. (A) Manhattan plots of SNPs significantly associated with relative lifetime fitness in eight different environments. SNPs significant after FDR (black and grey) or Bonferroni correction (red) are shown. For genome-wide scans of survival and fecundity 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, D) Environmental niche models for the most significant SNPs in each 0.5 Mb window of the genome. (C) 424 windows had significant SNPs in high-precipitation experiments. (D) 279 windows had significant SNPs in low-precipitation experiments.

We studied whether the direction and intensity of selection was dependent on the experimental environment. Alleles that were positively selected under low precipitation tended to be

negatively selected under high precipitation, and vice versa, so called antagonistic pleiotropy[18] (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, $P<10^{-16}$). Together, this indicates opposite selection across precipitation but not temperature gradients. This is an important observation, because it tells us that nature cannot select for generalist genotypes that are successful in wide range of precipitation environments.

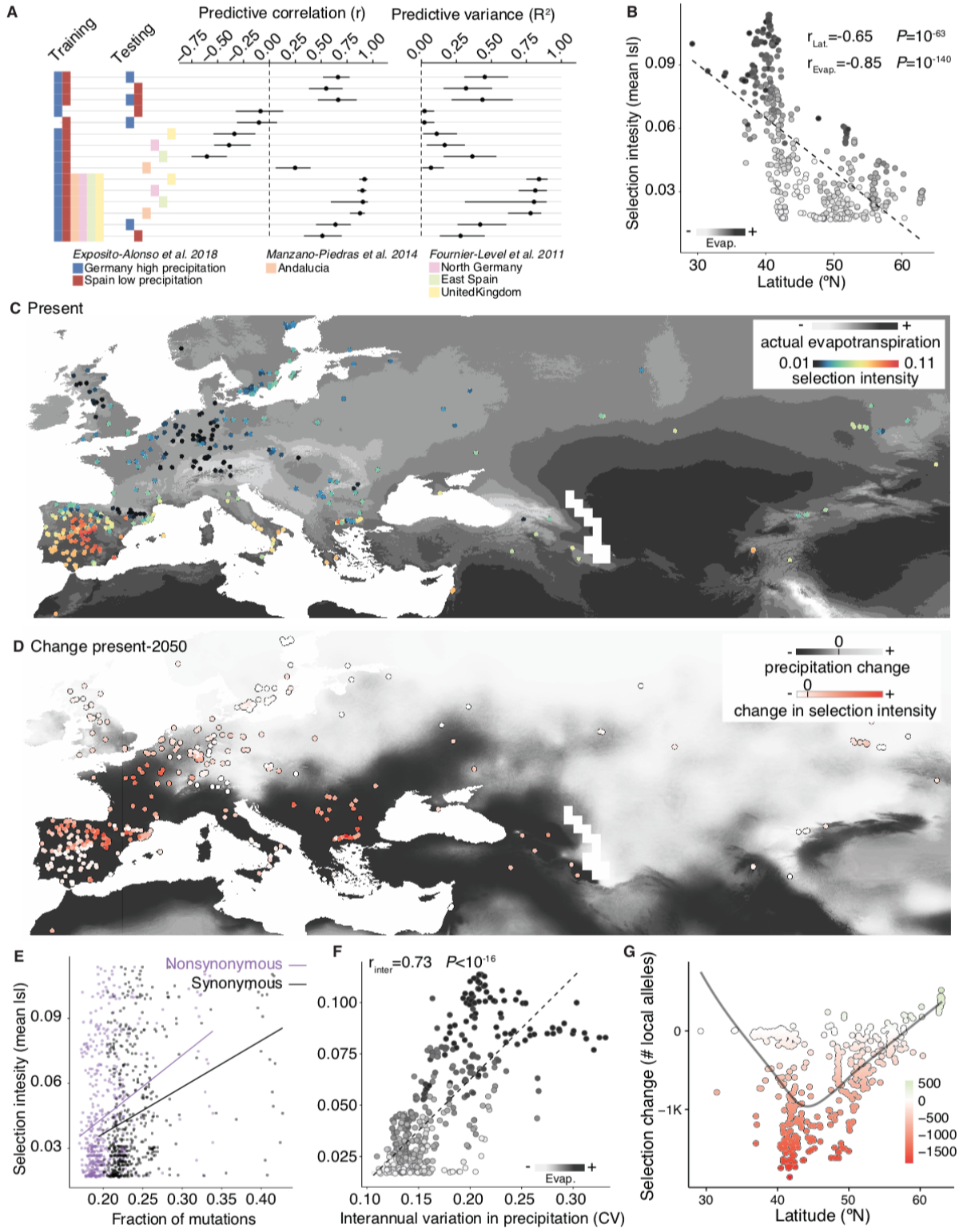
To study whether short-term selection in our experiments aligns with genomic footprints of past selection (see Supplemental Methods III and VII), we searched for selective sweeps (ref. 34), for outlier allele frequency differentiation (F_{ST}) between eleven previously defined *A. thaliana* genetic groups[10,23], and for climate-genome associations (GWA with 1960-1990 climate averages, worldclim.org v.1, ref. 35). Comparing frequency-matched background SNPs with Bonferroni-corrected significant SNPs for allelic selection differentials, 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). Absolute values of allelic selection differentials were significantly higher for strongly climate-correlated SNPs (e.g. annual precipitation [bio1] and temperature [bio12]: Spearman's r=0.12, p<16). The 1% top hits for climate associations also had higher F_{ST} values than frequency-matched background SNPs (e.g. bio1 and bio12: $P<10^{-5}$), but no differences in sweep likelihood (P=0.9). Implementing genome-wide environmental niche models[10] (see Supplemental Methods VII), we found that alleles selected in Germany and high precipitation were more likely to come from higher latitudes (Fig. 1C), while the opposite was true for alleles selected in Spain and low precipitation (Fig. 1D). In agreement, similarity in precipitation regime was a good predictor for alleles selected in Spain (bio 12-19; Spearman's r=-0.18, $P<10^{16}$; (Fig. 2D) (for other experiments, see Fig. S10). All in all, the signature that allelic selection differentials coincide with allele frequency shifts across population lineages are most easily reconciled with a polygenic model of natural selection[28].



142 **Fig. 2 Selection trade-offs and the signal of environmental local adaptation.** (A) Empirical selection
143 coefficients across two contrasting environments; Spain with low precipitation and high population density, and
144 Germany with high precipitation and low population density. Grey dots indicate SNPs with significant
145 conditional neutrality, black dots indicate SNPs with antagonistic pleiotropy. SNPs that behave in the same
146 direction or that were non-significant in both environments are not plotted. (B) Correlation between field
147 selection differentials and climate distances between test site and the centroids of SNPs´ geographic ranges (n=
148 54,146 FDR significant SNPs). (C) Relationship between field absolute allelic selection differentials and F_{ST}
149 values across 11 lineages, and (D) the likelihood ratio of selective sweeps. In all panels except (B), all 1,353,386
150 SNPs are plotted.

151 We finally aimed to build an environmental model that can predict allelic selection
152 differentials based on the climate and diversity patterns. We used as predictors the per-allele
153 associations with climate of origin from climatic GWA as well as the signatures of past selection at
154 each SNP, F_{ST}, π, and sweep likelihood, and their genome annotations — all predictors were
155 derived from public databases (worldclim.org, 1001genomes.org, arabidopsis.org). We used a
156 regression with decision trees using Random Forests to build Genome-wide Environment Selection
157 ("GWES") models. Conceptually, GWES models are similar to concepts related to Environmental
158 Niche Models (ENMs), but instead of training them with presence/absence data of a genetic
159 variant[10,11], we trained them with our measured allelic selection differentials. This provided a means
160 to predict whether alleles should increase/decrease in frequency in a certain climate, instead of

merely an indication of whether alleles are likely to be present, which is the indirect ENMs' version. By training models jointly with experimental data from Spain and Germany and using cross-validation, we confirmed that inferred and measured selection differentials were correctly predicted, with a high correlation accuracy (0.56 < Pearson's r < 0.7) and explaining a large proportion of variance (R^2= 29-52%) (Fig. 3A, variable importance Table S5) (further details in Supplemental Methods VIII). To further validate the predictive accuracy of our models, we made use of published fitness data for different sets of natural lines that had been grown at different locations in Spain, Germany and England[3637]. Using these data and GWES predictions, we confirmed moderate predictability (7%<R^2<36%)[38]. Predictability of GWES models increased when including data from all six experiments (17%<R^2<84%) (Fig. 3A, Fig. S12 Table S10) (for further details on predictability analyses and null expectations, see Fig. S12-13, Supplemental Methods IX.4-5).



172 **Fig. 3 A geographic map of climate-driven selection and its predictability.** (A) Genome-wide Environment
173 Selection (GWES) models trained with a combination of environments, to infer allelic selection differentials
174 throughout the species range. Mean predicted allelic selection differentials ("selection intensity"; n= 10,752
175 SNPs, one random SNP per 10 kb windows) in known locations of *A. thaliana* populations in relationship to (B)
176 latitude, and (C) evapotranspiration in summer (C) (ref. 35). (D) Predicted changes in selection intensity by 2050

(2050 MP rcp 8.5, ref. 35). (E) Relationship between selection intensity and synonymous and nonsynonymous polymorphisms. (F) Relationship between selection intensity and interannual variation in precipitation from 1958-2017 (ref. 39). (G) Number of local alleles whose selection predicted to positively or negative change (>5% in relative fitness) in 2050 across the latitudinal range.

Using the trained GWES models, we then predicted genetic natural selection at hundreds of locations, simulating field experiments in which the same set of diverse natural lines is challenged by different local climates (Fig. 3). The intensity of selection, i.e. genome-wide average allelic selection differentials, was strongest towards the environmental limits of the species, i.e. in hot (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 lower-than-average ratio of nonsynonymous to synonymous polymorphisms (Fig. 3E, r=-0.276, $3x10^{-10}$). This depletion of nonsynonymous substitutions, which are expected to be on average more deleterious than synonymous substitutions, provides independent evidence for natural selection having acted more efficiently in these local populations. 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, $P=3x10^{-4}$). This could be explained by the differential age of populations[10,40] and/or by the direction of selection having fluctuated over time, with alternative polymorphisms having been selected in each period[41]. Inspection of precipitation data from 1958 to 2017 (ref. 39) 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 how important temporal resolution in climate databases is for predictions, but also that climate stochasticity is a source of evolutionary constraints over species. Increased drift due to small population sizes is often thought to be the major force in shaping genetic diversity at range edges[42]. The patterns in *A. thaliana* are more consistent with proposals that a species' environmental niche[43] and its populations' genetic diversity[44,45] are primarily shaped by increasing natural selection towards the range edges.

A sudden change in climate or increased climate variability[46,47] will obviously 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. 5,35), 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, Fig. S3; Fig. S11). To enable comparability across locations, our metric of selection intensity is standardized based on the same set of diverse accessions (Fig. 3C-D). Local populations typically consist of more closely related lines that harbor only a subset of genetic variants, which may put these populations either in

a better or worse position to respond to future climate than our global set of more diverse lines. We therefore looked for SNPs predicted to change most strongly in selection by 2050 (fitness advantage or disadvantage changed over 5%), and evaluated whether the allele positively changing in selection is locally present (Fig. 3G, Fig. S14). We found that most local alleles will become more negatively selected; we therefore predict that the degree of local adaptation will decrease for many native populations (Fig. 3G), leading to an adaptation deficit.

Conclusion

The expected changes in climate during the 21st century will threaten the survival of many species. Because the distribution of genetic diversity is so well characterized in *A. thaliana*, we have used it to address the challenge of predicting the effects of climate-driven natural selection genomic variation across a species' range. Integration of genome-climate associations with direct fitness observations allowed us to build models that directly predict selection at the genetic level rather than mere probability of presence/absence of variants. This information enabled us to infer range-wide evolutionary vulnerability in the face of rapid climate change. The first two steps in our project, assembling a range-wide collection and genome sequencing of a number of diverse lines, are in reach for many species of plants. A greater challenge is the generation of fitness data, but this can be partially solved by identifying particularly informative field sites — as we have done in our study – and by exploiting the immense progress in field phenotyping at different scales[48,49]. Combining such observations with our new genome-wide environment modeling approach will help us to fully incorporate evolution into predicting the impacts of climate change on biodiversity.

ADDITIONAL INFORMATION

Accession codes. Phenotypic datasets are available as supplemental material of ref. 20 at [update link to journal] with doi: [update]. 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. The GWA scans for fitness and climate variables will be deposited at aragwas.1001genomes.org.

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 and Angela Hancock for 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.

REFERENCES

1. Urban, M. C. Climate change. Accelerating extinction risk from climate change. *Science* **348,** 571–573 (2015).
2. Warren, R., Price, J., Graham, E., Forstenhaeusler, N. & VanDerWal, J. The projected effect on insects, vertebrates, and plants of limiting global warming to 1.5°C rather than 2°C. *Science* **360,** 791–795 (2018).
3. Hoffmann, A. A. & Sgrò, C. M. Climate change and evolutionary adaptation. *Nature* **470,** 479–485 (2011).
4. Thurman, T. J. & Barrett, R. D. H. The genetic consequences of selection in natural populations. *Mol. Ecol.* **25,** 1429–1448 (2016).
5. Intergovernmental Panel on Climate Change. *Climate Change 2013 - The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2014). doi:10.1017/CBO9781107415324
6. 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).
7. Nielsen, R. *et al.* Genomic scans for selective sweeps using SNP data. *Genome Res.* **15,** 1566–1575 (2005).
8. 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).
9. Bonhomme, M. *et al.* Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. *Genetics* (2010). doi:10.1534/genetics.110.117275
10. Exposito-Alonso, M. *et al.* Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. *Nat Ecol Evol* **2,** 352–358 (2018).
11. Bay, R. A. *et al.* Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359,** 83–86 (2018).
12. Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J. K. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185,** 1411–1423 (2010).
13. Hancock, A. M. *et al.* Adaptation to climate across the Arabidopsis thaliana genome. *Science* **334,** 83–86 (2011).
14. Lasky, J. R. *et al.* Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Mol. Ecol.* **21,** 5512–5529 (2012).
15. Kingsolver, J. G. *et al.* The strength of phenotypic selection in natural populations. *Am. Nat.* **157,** 245–261 (2001).

16. Savolainen, O., Lascoux, M. & Merilä, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14,** 807–820 (2013).
17. Gompert, Z. *et al.* Experimental evidence for ecological selection on genome variation in the wild. *Ecol. Lett.* **17,** 369–379 (2014).
18. 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).
19. 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
20. Exposito-Alonso, M. *et al.* A rainfall-manipulation experiment with 517 Arabidopsis thaliana accessions. *bioRxiv* (2017). doi:10.1101/186767
21. Hereford, J. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173,** 579–588 (2009).
22. Leimu, R. & Fischer, M. A meta-analysis of local adaptation in plants. *PLoS One* **3,** e4010 (2008).
23. 1001 Genomes Consortium. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166,** 481–491 (2016).
24. 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
25. 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).
26. Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **11,** 407–409 (2014).
27. Berg, J. J. & Coop, G. A Population Genetic Signal of Polygenic Adaptation. *PLoS Genet.* **10,** e1004412–e1004412 (2014).
28. 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).
29. 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).
30. Haldane, J. B. S. The Cost of Natural Selection. *Genetics* **55,** 511–524 (1957).
31. Zhou, X., Carbonetto, P. & Stephens, M. Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet.* **9,** e1003264 (2013).
32. Lewontin, R. C. *The genetic basis of evolutionary change*. **560,** (Columbia University Press New York, 1974).
33. Charlesworth, B. The effects of deleterious mutations on evolution at linked sites. *Genetics* **190,** 5–22 (2012).
34. DeGiorgio, M., Huber, C. D., Hubisz, M. J., Hellmann, I. & Nielsen, R. SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics* **32,** 1895–1897 (2016).
35. 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).
36. Fournier-Level, A. *et al.* A map of local adaptation in Arabidopsis thaliana. *Science* **334,** 86–89 (2011).
37. 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).

38. Nosil, P. *et al.* Natural selection and the predictability of evolution in Timema stick insects. *Science* **359,** 765–770 (2018).
39. 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).
40. Lee, C.-R. *et al.* On the post-glacial spread of human commensal Arabidopsis thaliana. *Nat. Commun.* **8,** 14458 (2017).
41. Levene, H. Genetic Equilibrium When More Than One Ecological Niche is Available. *Am. Nat.* **87,** 331–333 (1953).
42. Henry, R. C., Bartoń, K. A. & Travis, J. M. J. Mutation accumulation and the formation of range limits. *Biol. Lett.* **11,** 20140871 (2015).
43. 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
44. Stebbins, G. L. Aridity as a Stimulus to Plant Evolution. *Am. Nat.* **86,** 33–44 (1952).
45. Leffler, E. M. *et al.* Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biol.* **10,** e1001388 (2012).
46. 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).
47. Samaniego, L. *et al.* Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Clim. Chang.* **8,** 421–426 (2018).
48. Shakoor, N., Lee, S. & Mockler, T. C. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. *Curr. Opin. Plant Biol.* **38,** 184–192 (2017).
49. Asner, G. P., Nepstad, D., Cardinot, G. & Ray, D. Drought stress and carbon uptake in an Amazon forest measured with spaceborne imaging spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6039–6044 (2004).

Supplemental Information Guide for

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

Table of Content

SUPPLEMENTAL METHODS

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

The field experiment is described in detail in Exposito-Alonso et al. "A rainfall-manipulation experiment with 517 *Arabidopsis thaliana* accessions" (ref. 20). Data and processing code are available at https://github.com/MoisesExpositoAlonso/dryAR update doi: xxx. Information of the 517 accessions can be found in Table S1. The experiment resulted in observations from 23,154 pots. Three measurements of fitness were produced: survival from seed to reproductive adult (proportion 0-1) and the average fecundity per reproductive adult (inflorescence skeleton lengths ranged from 18,400 to 1,622,000 pixels, which approximately corresponds to 1 to 6,127 seeds per plant). Fecundity was only measured for plants with at least one fruit. We finally calculated an integrated lifetime fitness value by multiplying the survival proportion to adulthood with the total offspring produced. Data from only 515 accessions were used for subsequent analyses, because 2 accessions had insufficient genome information.

II. 1001 Genomes Project data

We used VCFtools v.0.1.12b (ref. 50) to subset and filter the 1001 Genomes VCFv4.1 (available at: http://1001genomes.org/data/GMI-MPI/releases/v3.1/). We used vcftools with the flags: --maf 0.01 --max-alleles 2 --min-alleles 2 --max-missing 0.95. The resulting high-quality dataset was a genome matrix of 515 individuals by 1,353,386 variants for which we did not impute the small number of missing data points.

We annotated the 1001 Genomes VCF using the package SnpEff 4.3p (ref. 51). We then manually curated a set of eight categories of variants: intergenic, intron, UTR3, UTR5, exon, synonymous, nonsynonymous, exon noncoding.

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

We used the genetic groups previously defined for the same accessions [10] and computed F_{ST} using PLINK version 1.9 (ref. 52). We also used PLINK to calculate π and Tajima's D using PLINK in windows of 100 SNPs across the genome.

We used SweepFinder2 (ref. 34) to scan the genome for deviations of the Site Frequency Spectrum (SFS) that might be caused by selective sweeps. We used all 11,769,920 biallelic SNPs from the 1001 Genomes Project (without the filters of 1% MAF and maximum missing data of 5%, which were applied to generate the variants used in the GWA [see section V]).

III.1 Geographic proxies of diversity metrics

In order to estimate π and a proxy of Tajima's D at a regional scale, we used the 4 closest neighbouring accessions in our set (same patterns were observed with three neighbours within a geographic area of 5° 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 D is normally divided by the standard error, we only wanted to rank our natural lines so we used the difference between π and θ as a proxy of D.

IV. Heritability of fitness

To estimate how much variance in fitness is related to the genotypes of the lines, we used generalized linear mixed models using the R package MCMCglmm (ref. [53]). We used fitness estimates per replicate and, appart from including the natural line ID, we controlled for block (growing tray) and position within the block (longitudinal, latitudinal, and the interaction). As this is a Bayesian approach, we used flat priors, we used 10,000 MCMC steps, a burnin of 10%, and confirmed that this was sufficient for convergence of the chain. 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. 31) 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 the selection differential s from populations genetics and evolutionary biology [4], which estimates the difference in fitness between the genotypes carrying the alternative (1) or reference allele (0) in a haploid model: $w1 - w0 = s$. The correspondence with allele frequency change Δp is dependent on the variance in allele frequency in addition to s: $\Delta p = p(1-p)s$. Low frequency variants are "less often seen" by selection.

We also run in GEMMA a Bayesian Sparse Linear Mixed model (BSLMM), to more accurately pinpoint casual positions. This model accommodates both poly- and oligogenic architectures and by jointly fitting all SNPs (n=1,353,386), it statistically corrects for LD arising from population structure and/or low recombination. 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), for which 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, $r_{x_1x_2}$, which in the case of genotypes is called linkage disequilibrium, as in the form of:

$$\beta^*_{x_1} = \frac{\beta_{x_1} - r_{x_1x_2} \times \beta_{x_2}}{\sqrt{(1-\beta^2_{x_1})(1-\beta^2_{x_2})}}$$

Thus we find an analogy between β and β^* and s and s^* from population genetics (eq. 4 from [54] simplified as conceptual model),where a selection differential is dependent on the true selection and an indirect term dependent on all other n SNPs in the genome.

V.1 Trade-offs of selection

V.1.1 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). We also asked how often a selected variant in one environment was neutral in another environment (conditional neutrality). Only for the purpose of two-environment comparisons, to calculate the odds ratio, we considered SNPs whose selection differential *P*-value was lower than 0.01 in one environment as conditionally neutral, while antagonistic pleiotropy SNPs were ones whose *P*-values in both environments were < 0.02 (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.1.1 Across life history stages

Calculating allelic selection differentials for survival and fecundity separately, we found no correlation between survival-only and fecundity-only estimates (r<0.07, Fig. S4-6), consistent with different stages of a plant being differentially affected by environmentally imposed selection[55].

V.2 Intensity of selection

The distribution of absolute allelic selection differentials, $|s|$, has a shape resembling that of an exponential. We calculated the expected rate using Maximum Likelihood optimization in R, which can also can be approximated as the inverse of the mean:

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

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

V.3 Important notes on population structure correction in a wild species

The goal of genome-wide association studies in humans is the identification of individual SNPs that are causal for traits such as disease susceptibility. It is therefore imperative to penalize SNPs that are correlated with population structure, as the lack of controlled experiments can generate spurious associations between genetic variants more common in some human ancestry groups with regional measurement errors or different cultural and nutritional environments[56]. When a population includes multiple ancestries with different disease susceptibilities, population structure correction can obscure real signals, which has in turn led to the development of admixture mapping[57]. Similarly, in animal or plant breeding, SNPs conferring an advantageous trait but that are highly associated with a particular group of breeds or varieties (the equivalent of geographic populations) are also avoided, as selection of these SNPs can drag along undesirable traits from such parents [58].

The situation is very different in a wild species such as *A. thaliana*. Deliberately ignoring selection over SNPs linked to natural population history would be inappropriate, as it is known that populations may have adapted to different climates as they migrated and became isolated, with causal SNPs therefore showing strong geographic patterns of distribution (see refs. [10,40]). This is of particular importance when mapping relative fitness, where the goal is to quantify the total selection over a genetic variant, as this includes both direct selection as well as experienced indirect selection from other causal variants that are in linkage disequilibrium.

VI. Climate Genome-Wide Association

Similarly to our GWA with relative fitness, we used each climate variable m (see Section VII.1) as response variable y_m in a LM model using GEMMA (ref. 31, see Section V):

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

This β coefficient for SNP i, which reflects the correlation of the alternative allele's presence and a climate variable, was used later on in our predictive models (Section VIII). As this is a raw correlation between allele presence and climate variables, it will capture both past signatures of climate adaptation and historic population migration and differentiation.

VII. Climate and modeling

VII.1. Climate layers

We used the classic bioclim variables (n=19), plus monthly data of minimum and maximum temperature, and precipitation (n=12 x 3) (worldclim.org). From these we estimated monthly evapotranspiration rates using the R package EcoHydRology v. 0.4.12 (ref. 59) and actual monthly evapotranspiration using a bucket model [60] (n=12 x 2). Based on ref. 14 we calculated whether *A. thaliana* can grow in a given month based on temperature and precipitation (n=12), and derived from this the length of the potential growing season (n=1). Over the potential growing season, 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 used an environmental niche model to generate a map surface of the most likely plant lifespan (n=2). This provides an estimate of the actual growing season, which we subtracted from 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 fit using decision trees with presence/absence of SNPs as response variable and the climate variables described in the previous section and latitude and longitude a predictors; as described [10]. To fit the models we used an Stochastic Gradient Boosting approach with the R package caret (ref. 61). 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 was determined after running our GEMs for some exemplary SNPs and confirming that this set of parameters was typically optimal for reducing 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 between 0 and 1 that the alternative allele was in a map cell. We judge this as a more appropriate output than a discrete 0/1 outcomes, as sometimes alleles were widespread 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, we extracted climate variables (Table S8) in a 50 Km buffer where they were originally collected from and calculated the coefficient of variation (CV) across grid cells.

To study temporal variability, we used climate data from 1958-2017 [39] to calculate annual precipitation values for each population, from which we in turn derived the inter-annual CV..

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. 62,63) 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} (Table S8, see section VII). 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 (see section III). In addition, we included as predictors a 1,353,386 x 8 matrix θ of non mutually exclusive variables taking values of 0 or 1 indicating genomic annotations: intergenic, intron, UTR3, UTR5, exon, synonymous, nonsynonymous, exon noncoding (see section II). A total of 112 variables were thus used as predictors: $s = f(\beta_{clim}, \mu, \theta)$. In the cases where we trained models with two environments, we also included the 2 x 98 x_{clim} climate variables 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 (with smaller and larger SNP sets, we had determined that training with more than 10,000 observations did not improve predictions). To test accuracy and bias we used a different set of 10,000 SNPs, divided into 100 bootstrap samples, and we report the intervals of the 95% bootstrap distribution. The results presented in Fig. 3 were produced with 10,000 randomly drawn SNPs across the genome. To confirm that there was no confounding from non-independent samples in the training and testing SNPs, we repeated all analyses, training with 10,000 random SNPs from chromosome 1 and testing with 10,000 random SNPs from the four other chromosomes. There were no substantial changes in predictability.

Several combinations of training and testing were performed 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.

VIII.2 A simple visualization of environmental distance

To visualize more directly the relationship between allelic selection differentials at a location and the overall environment where the alleles are found, we calculated the distances between the field station's climate and the allele's home environment (as defined below). We started using the locations of the 515 data points and the genome matrix (515 individuals, 1,353,386 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), resulting in a 515 x 19 matrix E. Then we computed the mean of environments at locations where there are alternative SNPs: $(E^T X)/X^T X$; this 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 of them as: $\sqrt{\sum_{i=1}^{19} d_j^2}$. The same approach can be used to calculate the distance to the optimal environment of the species. One could assume that the density of geographic locations where *A. thaliana* has been sampled would be the "realized optimum".

VIII.3 Note on limitations and interpretations

As in any predictive exercise, our projections have limitations (discussed below). We nevertheless firmly believe that they are indispensable to move forward in the field of forecasting climate impacts. Models such as ours are tremendously useful for subsequent experimental validation (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). This iterative prediction ↔ validation process will be key to advancing the complex field of predicting the effects of climate change on biodiversity.

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 exact allelic selection differentials we observed are contingent on the accessions studied. For example, if in another GWA panel, a specific site is not variable, one could not calculate an allelic selection differential for that site. It is therefore important to carefully select a set of

accessions that represents geographic and genetic diversity of the species[20]. This is what we have done, and our selection estimates should thus be 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* where outcrossing is rare (species selfing rate average=97% [64]).

C. Short-term selection differentials (over ecological times) do not necessarily reflect long-term selection coefficients (i.e. over evolutionary times).

D. Demographic dynamics are ultimately determined both by natural selection and stochastic demographic forces. Therefore, the knowledge of selection coefficients is necessary but not sufficient to determine the fate of a population.

E. Bet-hedging strategies such as seedbank demographic dynamics buffer allele frequency changes over time.

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

For a conceptual diagram of predictability validation with external common garden datasets, see Fig. S12

IX.1 Manzano-Piedras et al. 2014

Manzano-Piedras and colleagues [37] planted exactly 60 seeds per line in pots. They monitored how many plants established at the rosette stage and later on became reproductive adults (survival proportion). From these, 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 [36] 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 transplanted 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 partial survival and reproduction.

We excluded the experiment in Finland in downstream analyses because only 58 natural lines were planted there in the original publication[36] 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 with 250 (http://arabidopsis.gmi.oeaw.ac.at:5000/DisplayResults, ref. 8) included 1,307 natural lines with 214,051 biallelic SNPs from array hybridization. The two populations shared 413 lines. Of these, 185 were shared with the 515 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 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 all of the 2,029 natural lines belonging to 1001 Genomes and RegMap 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 yielded 154,090 biallelic SNPs. Given the linkage disequilibrium and genome size of *A. thaliana*, this easily suffices for generating a relationship matrix A, which we computed using the R package rrBLUP (ref. 65). The data of survival, reproduction, and lifetime fitness was an 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 a 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 given in Table S9.

IX.4 Sanity checks for imputation and geographic predictions

We carried out sanity checks to ensure that the imputed fitness from other experiments was not just an artifactual 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 artifactual phenotype would depend on population structure and thus would likely be predictable from climate structure alone.

We shuffled the genotype identities from Fournier-Level et al. and Manzano-Piedras et al. with their fitness values. Then we repeated the GBLUP analysis with 50 rounds of shuffling and computed heritabilities and prediction accuracies. We confirmed that heritability with shuffled data

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 true heritable variation, imputation of fitness would be random and not dependent on the relationship matrix.

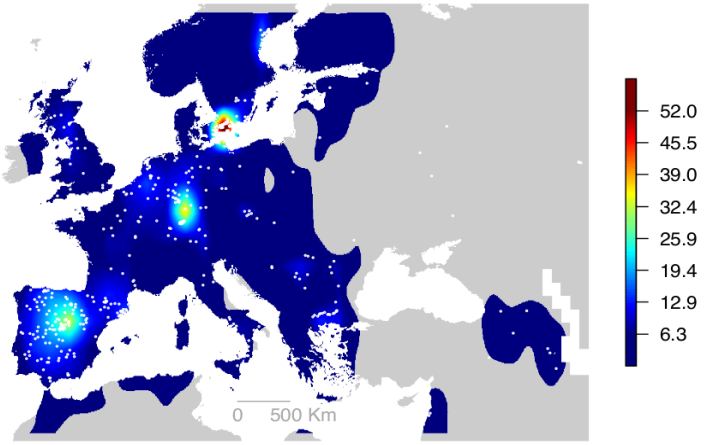
We also were concerned that geographic predictions could be driven by some underlying bias in our analyses, i.e. bias inherent to geographic sampling, population history of genotypes chosen, etc. In other words, we were concerned that the null expectation of predictability would be non-zero. As before, we randomized fitness values with genotypes for all six datasets. Then, we repeated the GWA to estimate allelic selection differentials (as Fig. 1), and trained different combinations of GWES models to re-predict allelic selection differentials at teach location based on climate (as Fig. 3). We confirmed that, differently from the analyses of real data presented in 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. 36. 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 for several reasons. First, the worldclim.org climate averages (1960-1990) at 2.5 arc-minutes resolution might strongly deviate from the truly experienced environmental conditions in the years the experiments were conducted. Such climate variability can exert opposite selection in different years[66]. Second, differences in experimental design could lead to different lifetime fitness estimates. In ref. 36, early survival of seedlings was not measured at all, as only seedlings that had survived for two weeks in the greenhouse were transplanted into the field. In the Southern Spain experiment[37], seeds were sown directly in the field, as in our own experiments, and accordingly, we had "positive predictability" (r=0.24, Bootstrap CI=0.09—0.41). In further support of this experimental design confounder, when we trained GWES models with only reproduction-based allele selection differentials in our experiment in Southern Germany, i.e., excluding early survival from lifetime fitness, we correctly predicted the sign of selection differentials in Fournier's Northern Germany experiment (r= 0.392, Bootstrap CI= 0.20—0.57) (for null expectations see Supplemental Methods IX.4). The differences in predictions between two- and six-environment-trained models did not yield differences in downstream analyses and conclusions (correlation between predictions, r=0.56, $P<10^{-16}$), but predictability increased with the number of experiments included in the training set (r= 0.761 , Bootstrap CI= 0.55—0.90, R^2= 0.517, Fig. 3A).

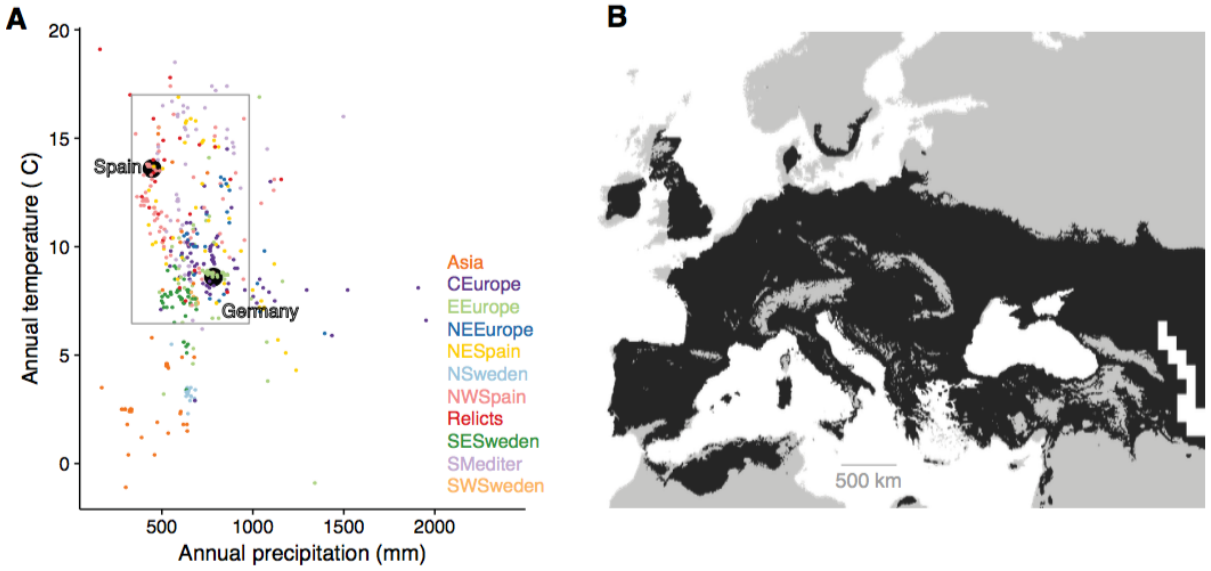
SUPPLEMENTAL FIGURES

Figure S1. Accession locations



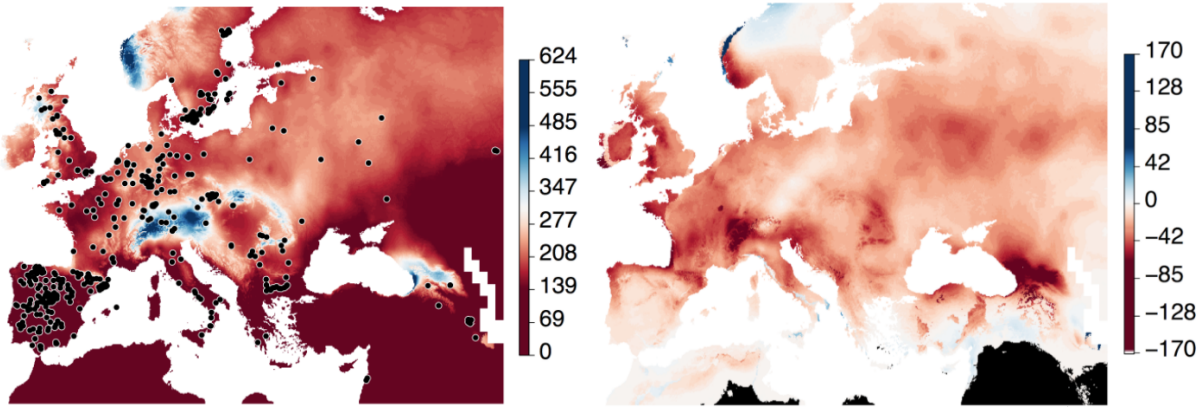
Points indicate the locations where the 517 *A. thaliana* accessions were collected. The color gradient is the density of samples from our study in squares of approximately 200 km x 200 km. The limits of the colored area were determined using a combined density grid from gbif.org and 1001 Genomes records. The density was generated in a grid of 125 min resolution and by applying a bilinear and then Gaussian smoothing. The threshold was chosen to be the 50% of the upper distribution, which roughly corresponds to 10 records per 200 km x 200 km square. Regions outside the colored were excluded from future climate change predictions, as we prefer to make predictions only in regions where the presence of *A. thaliana* is rather likely and continuous (Fig. 1).

747 **Figure S2. Environment ranges**



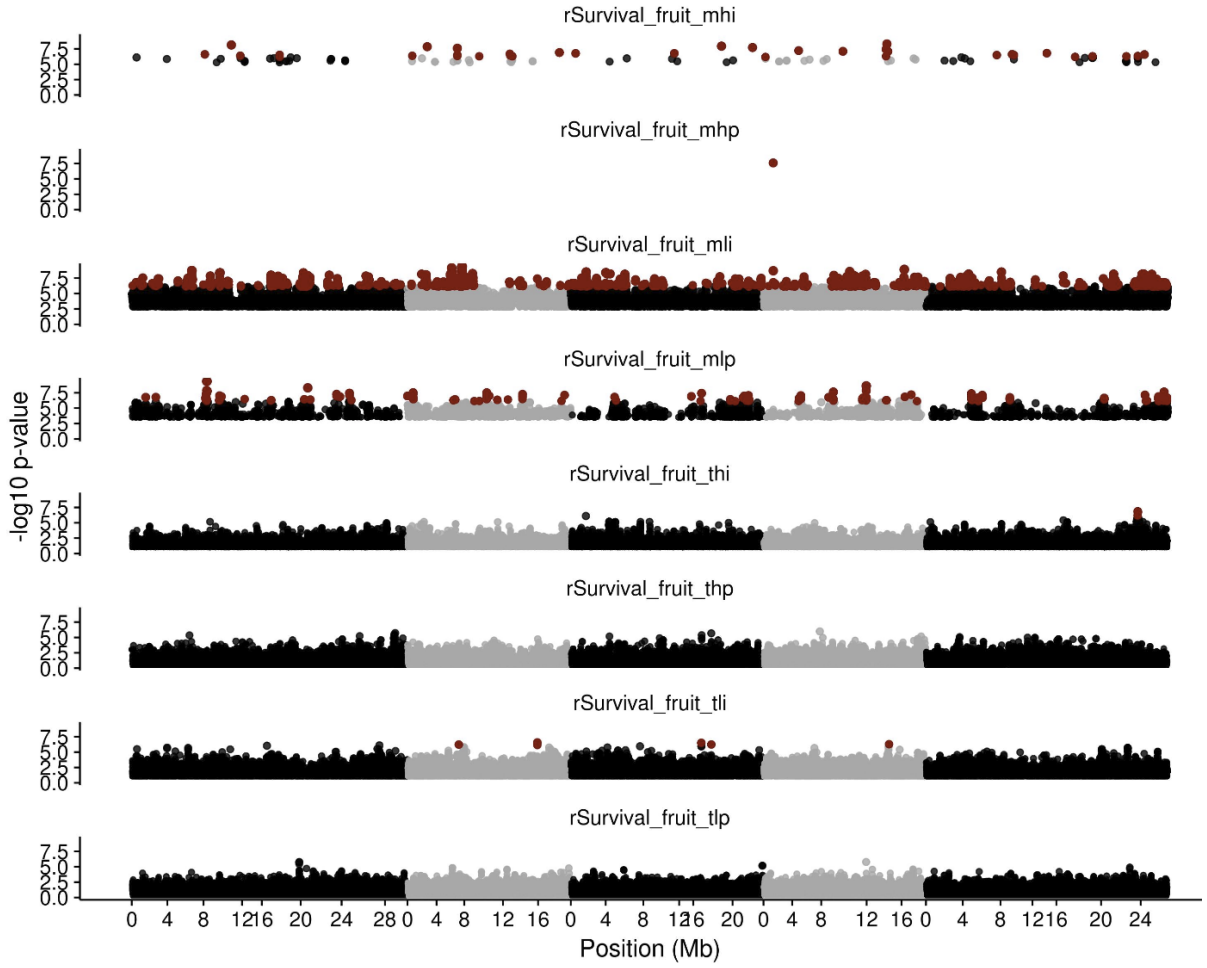
748 (**A**) Classic biplot of precipitation *vs.* temperature of origin of accessions (black dots) and field
749 experiment of Spain (sepia) and Germany (green). Grey box indicates locations where precipitation
750 was at least 70% of Spain and no more than 130% of Germany, and where temperature was no less
751 than 70% of Germany and no more than 130% of Spain. (**B**) Areas that would be within the grey box
752 in (A). Compare to Fig. S1.

753 **Figure S3. Map of predicted environmental change**



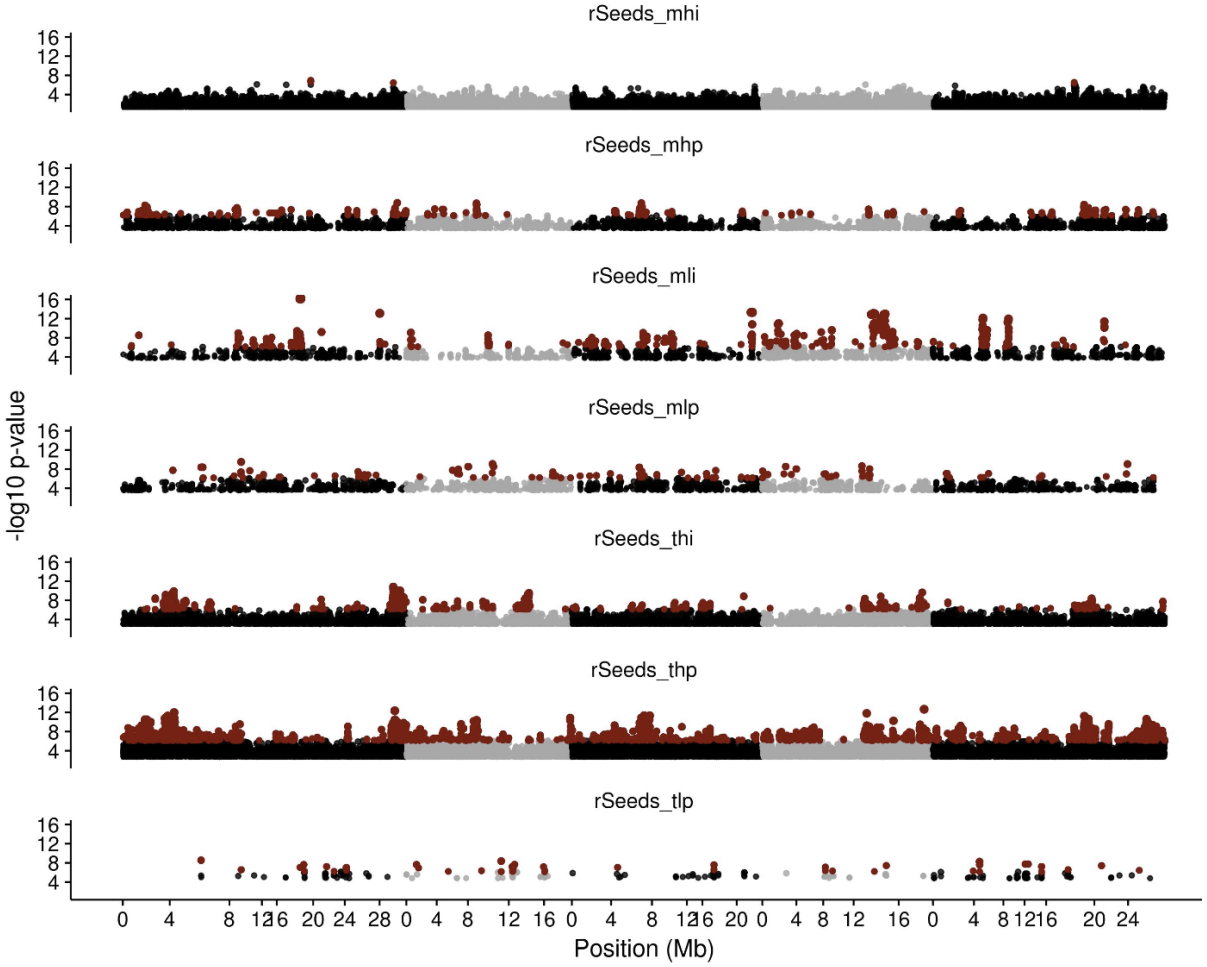
754 Precipitation during the warmest quarter (bio18, left), and its change predicted for 2070 (rcp 8.5)
755 (right). Black areas indicate regions where precipitation will be lower than any area where *A. thaliana*
756 has been currently sampled (black dots, left).

Figure S4. Genome maps of survival



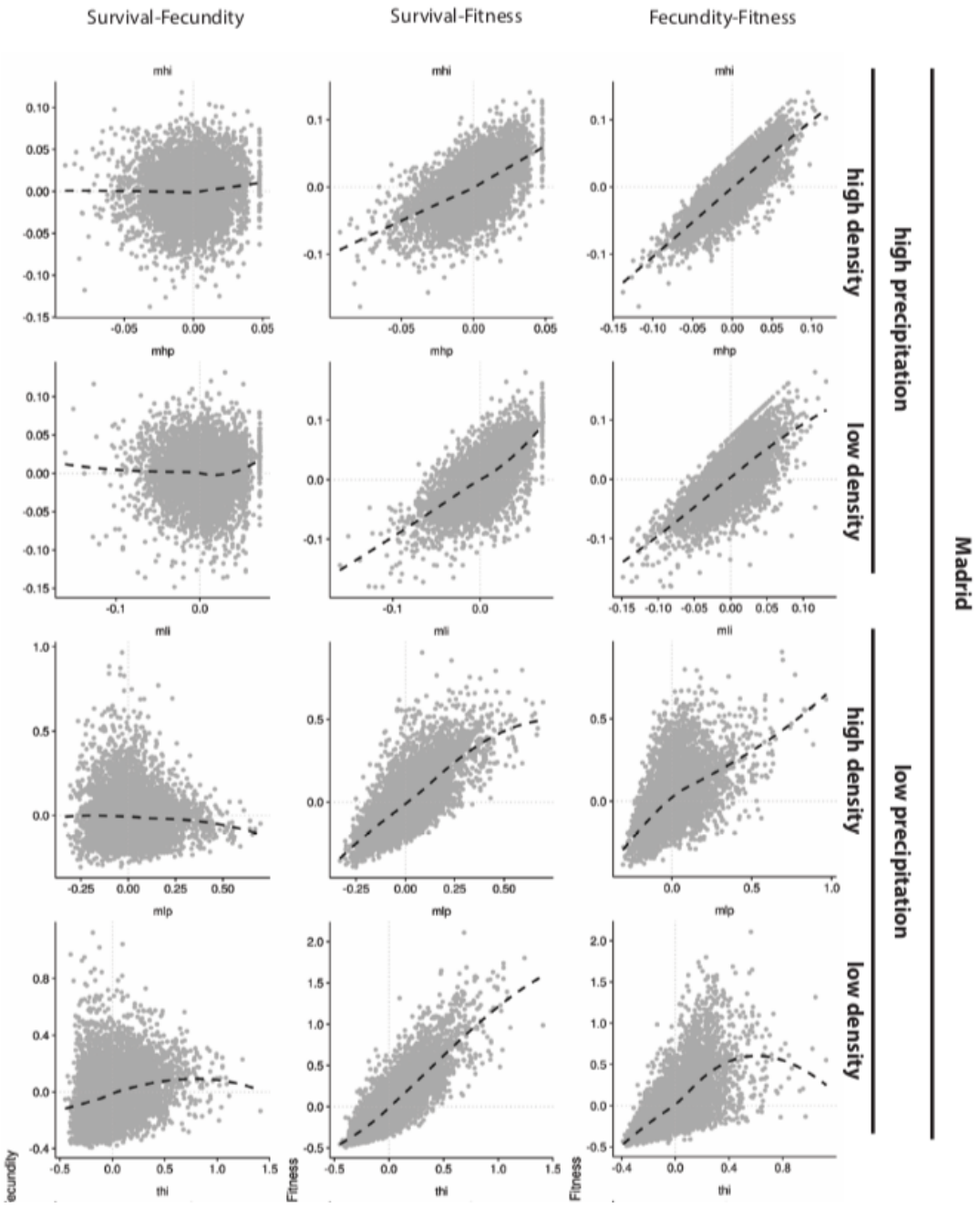
Same as Fig. 1, but only using the survival component of fitness.

Figure S5. Genome maps of fecundity

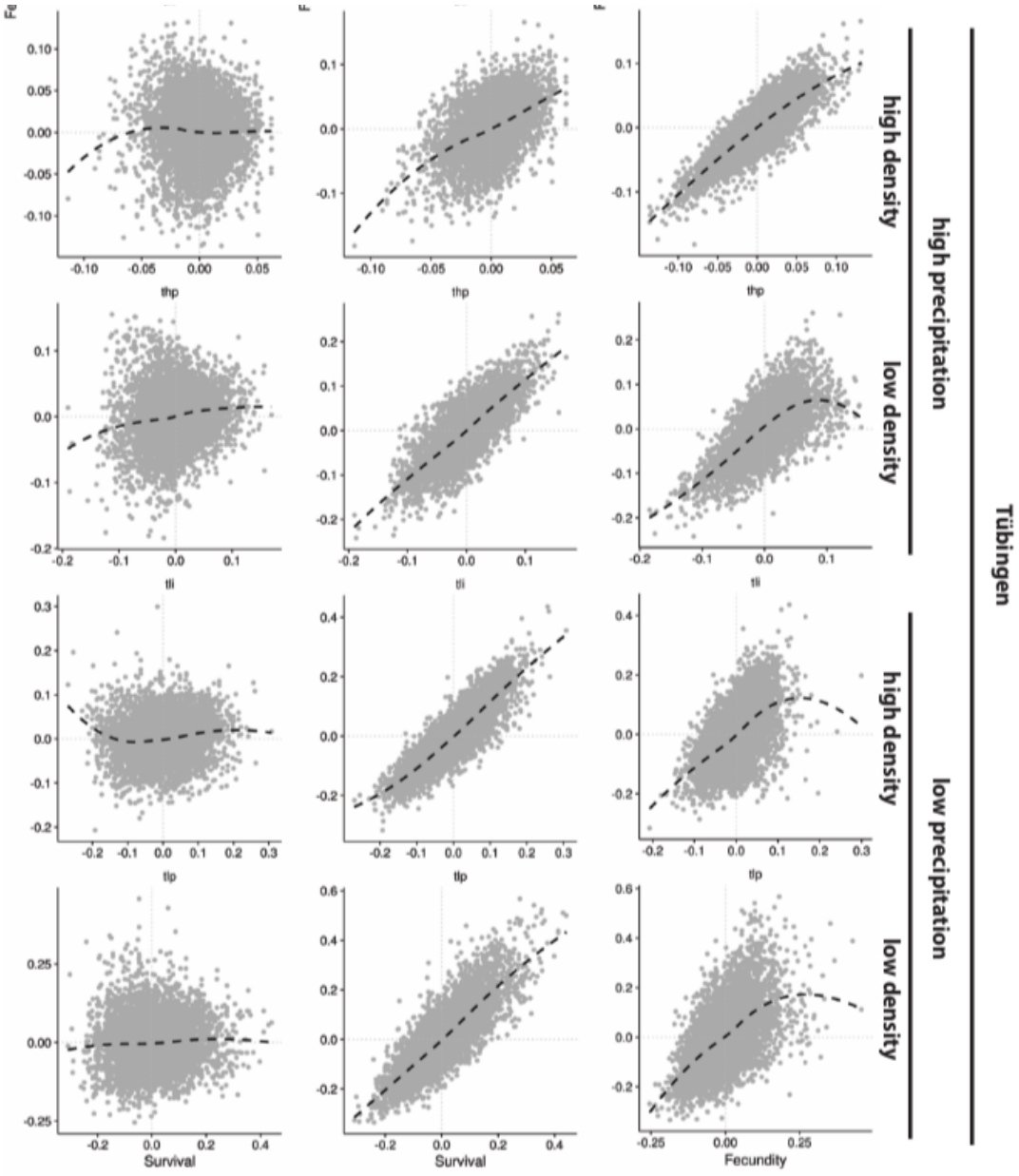


Same as Fig. 1, but only using the fecundity component of fitness.

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

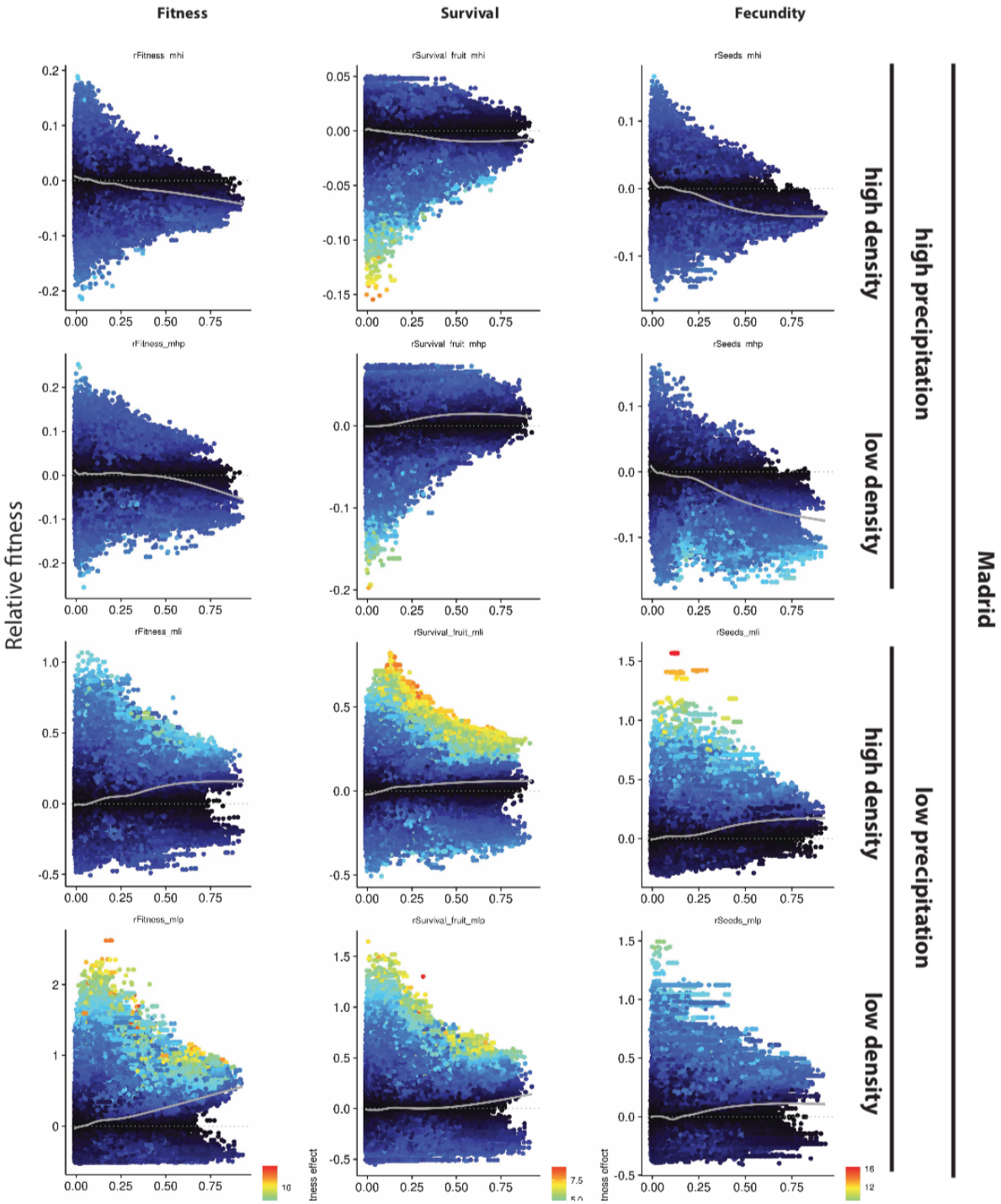


(Fig. S6 continued)

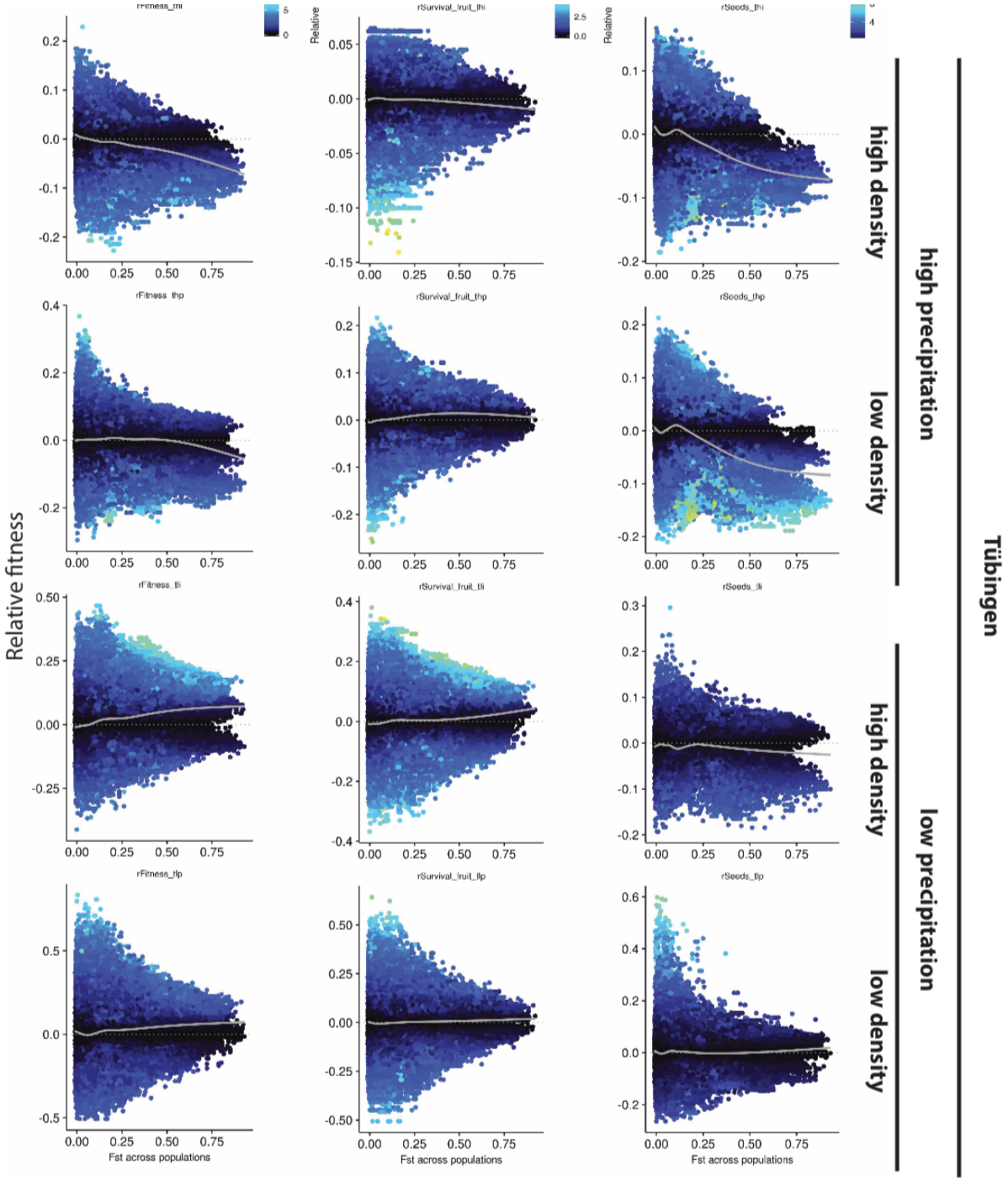


Comparisons of allele selection differentials computed only with the survival component, only with the fecundity component, and with both.

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

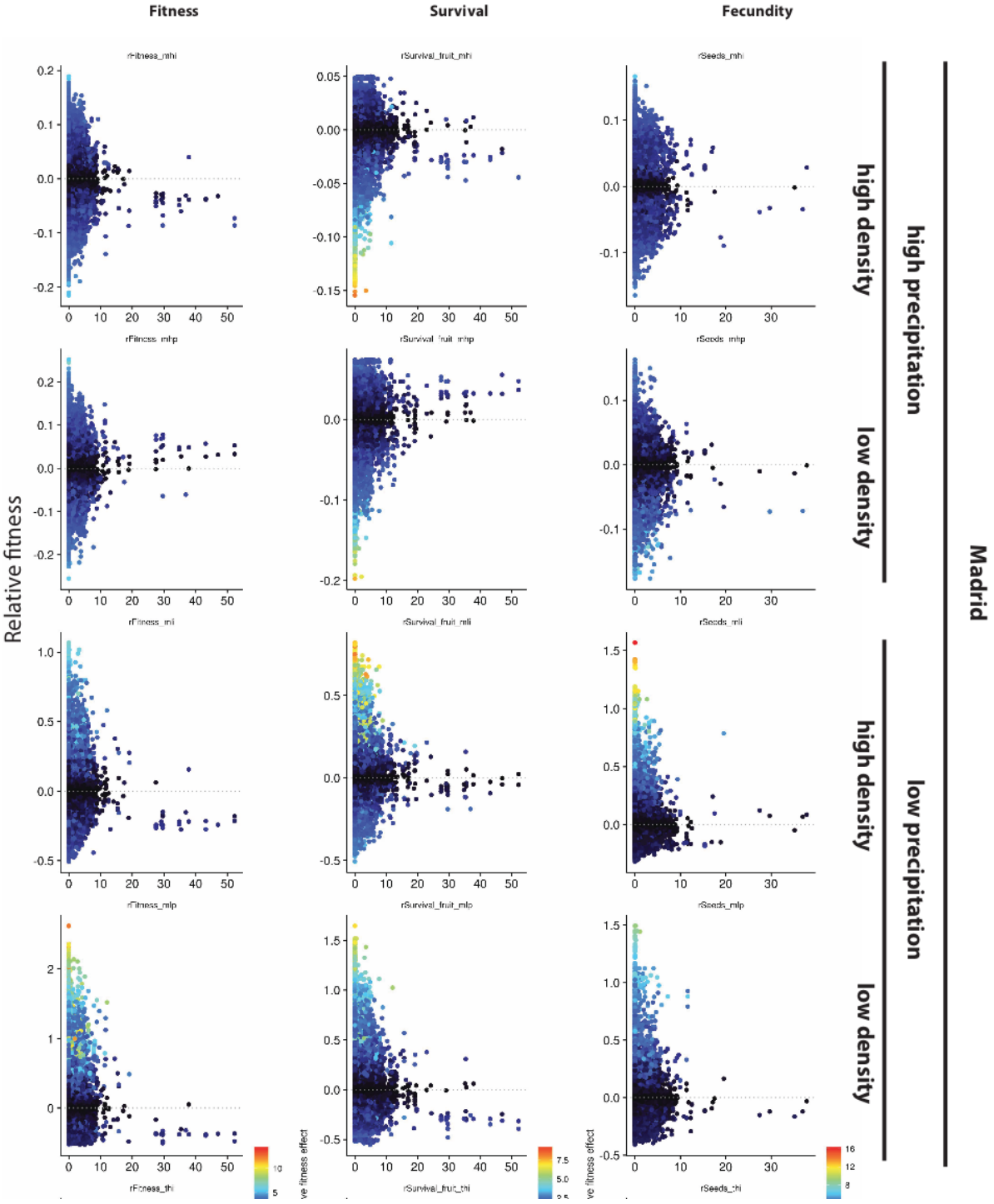


(Fig. S7 continued)

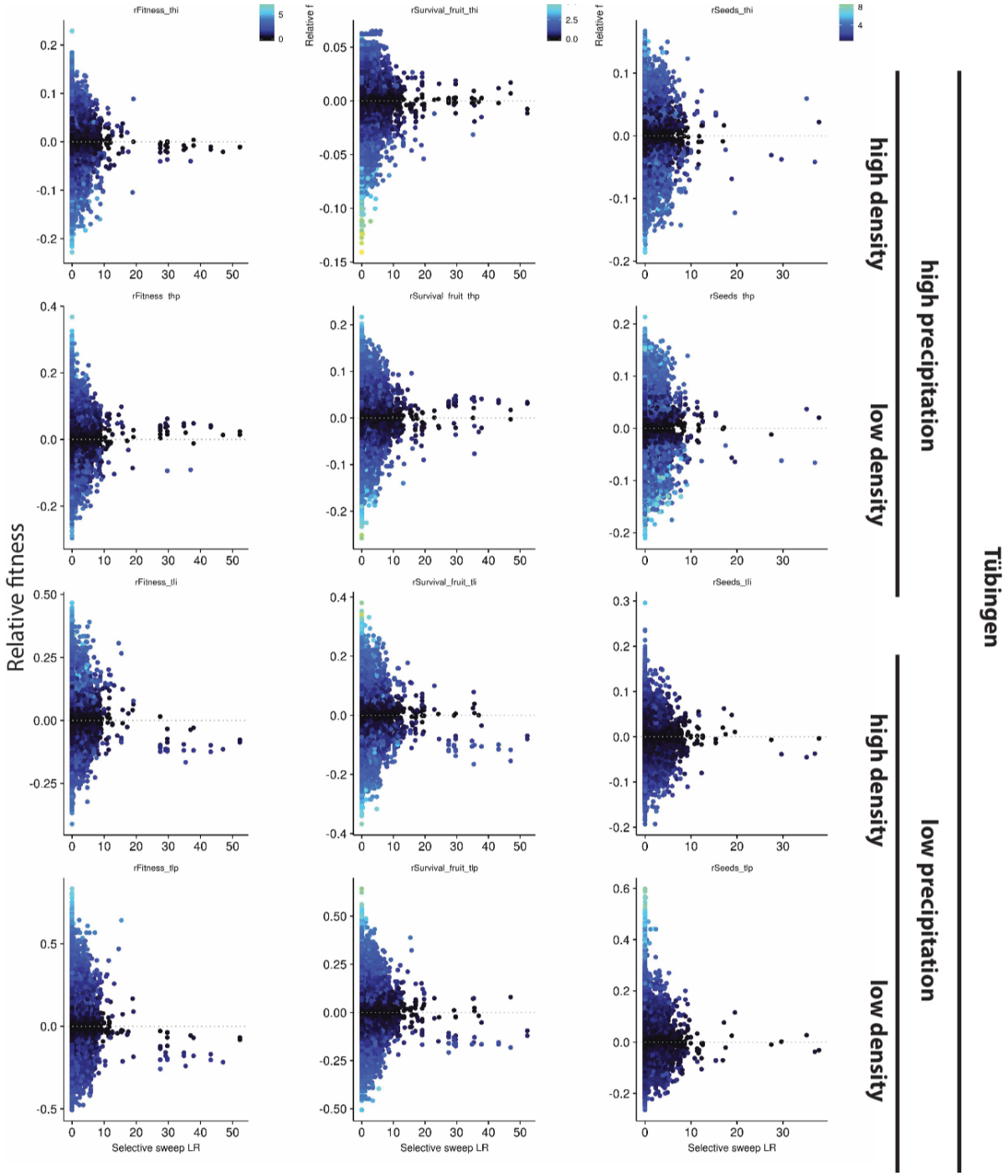


As Fig. 2C, for all environments.

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

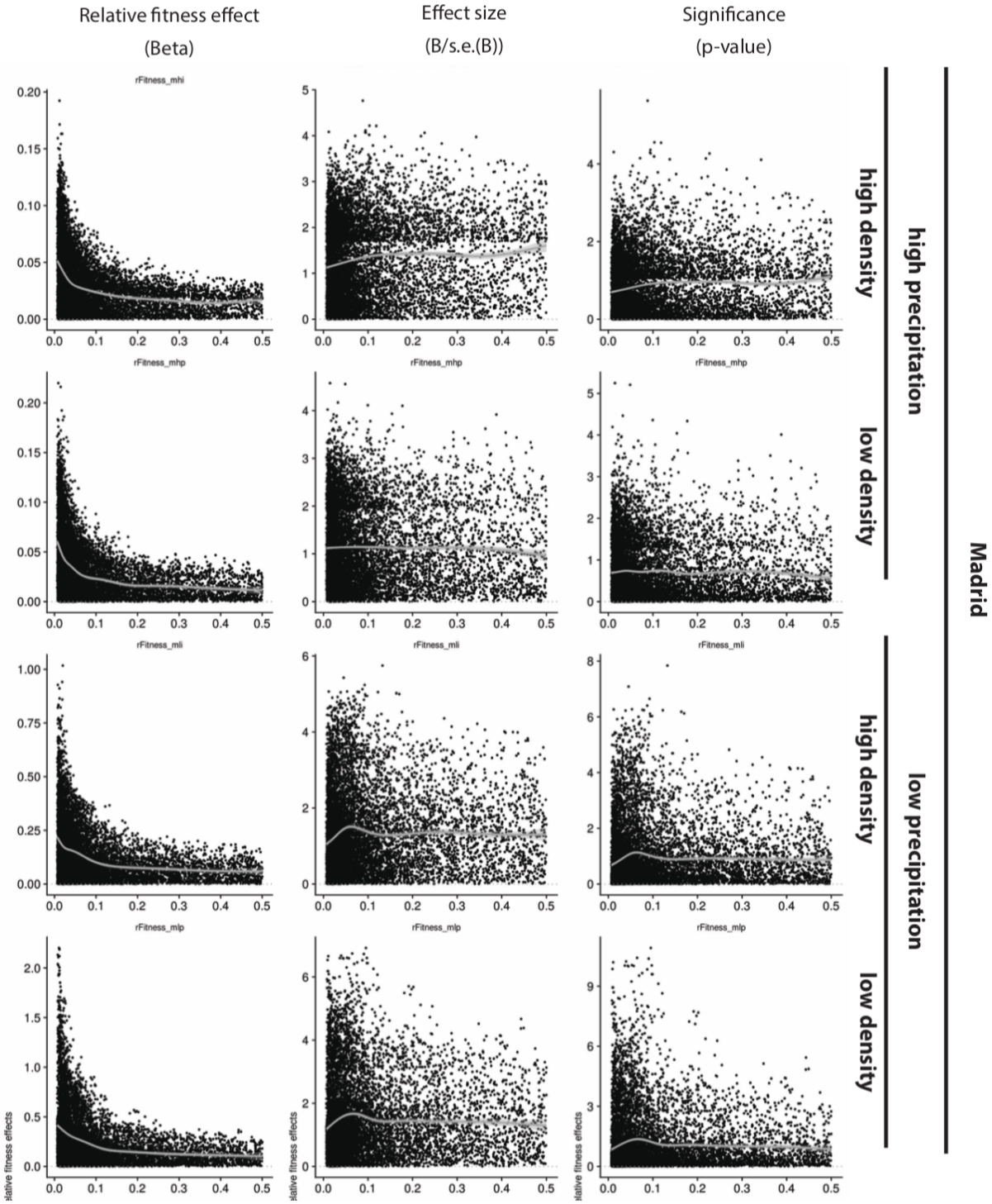


(Fig. S8 continued)

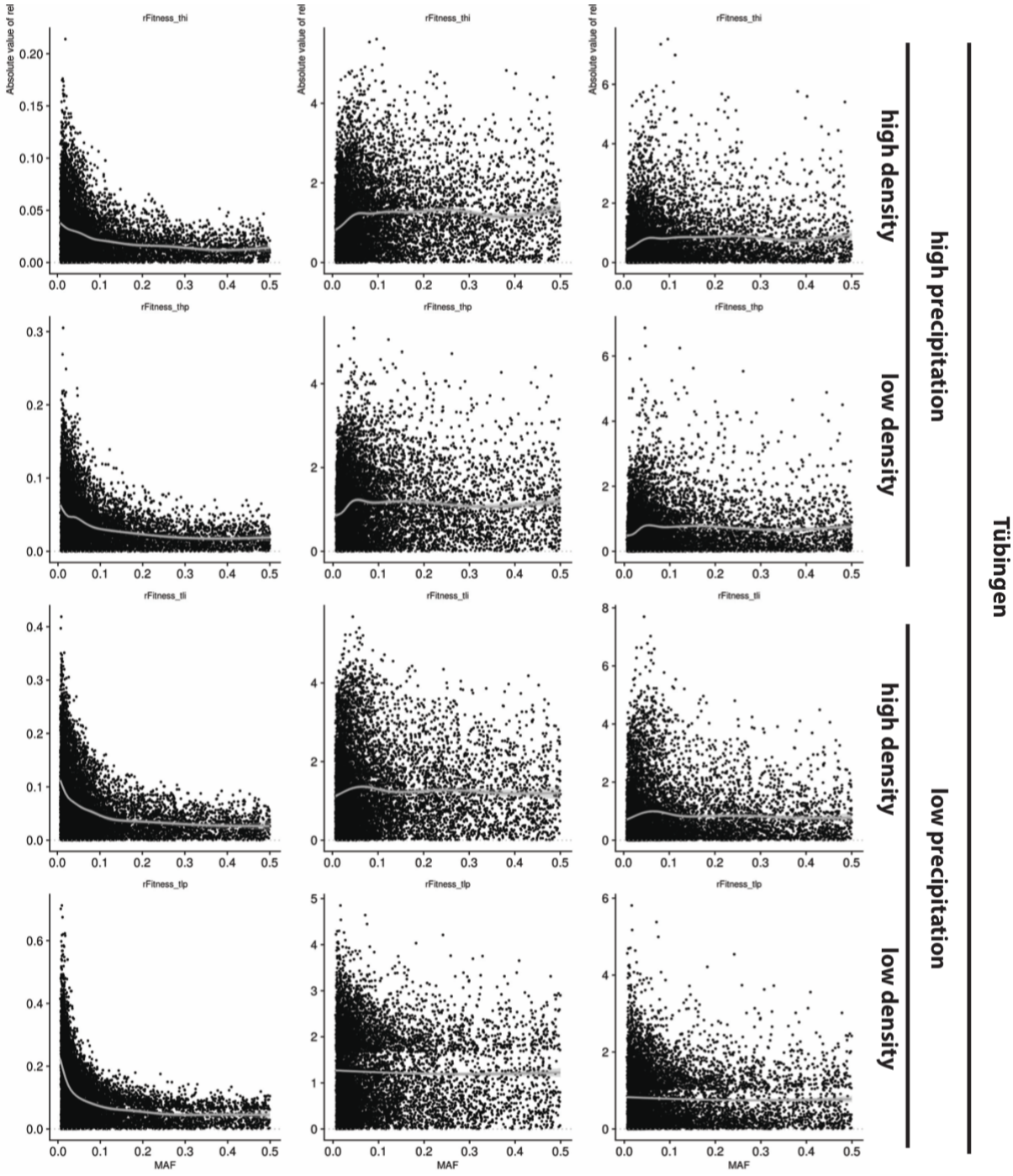


As Fig. 2D, for all environments.

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

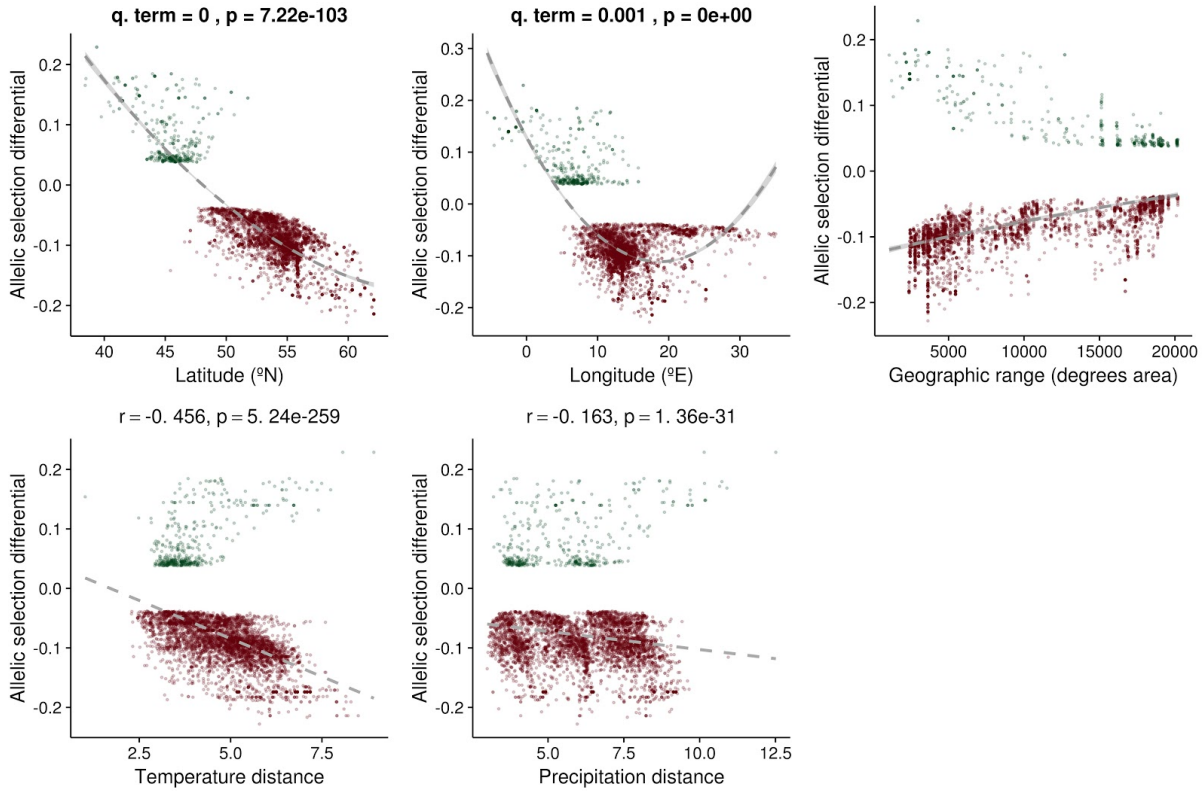


(Fig. S9 continued)



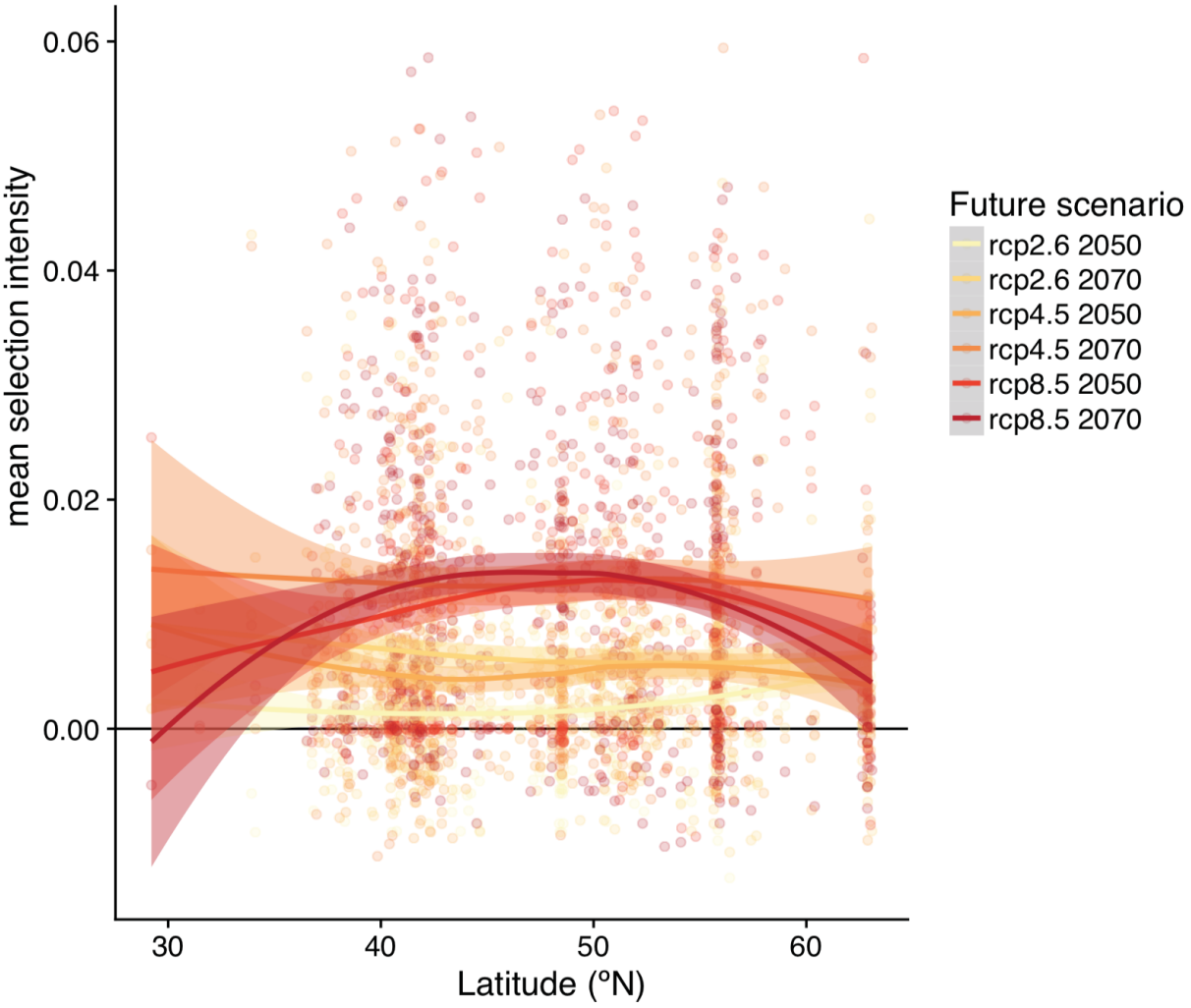
Relationships between relative fitness effect, relative fitness effect size, and *P*-values (calculated from GWA with relative fitness) and minor allele frequency of alleles in each environment.

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



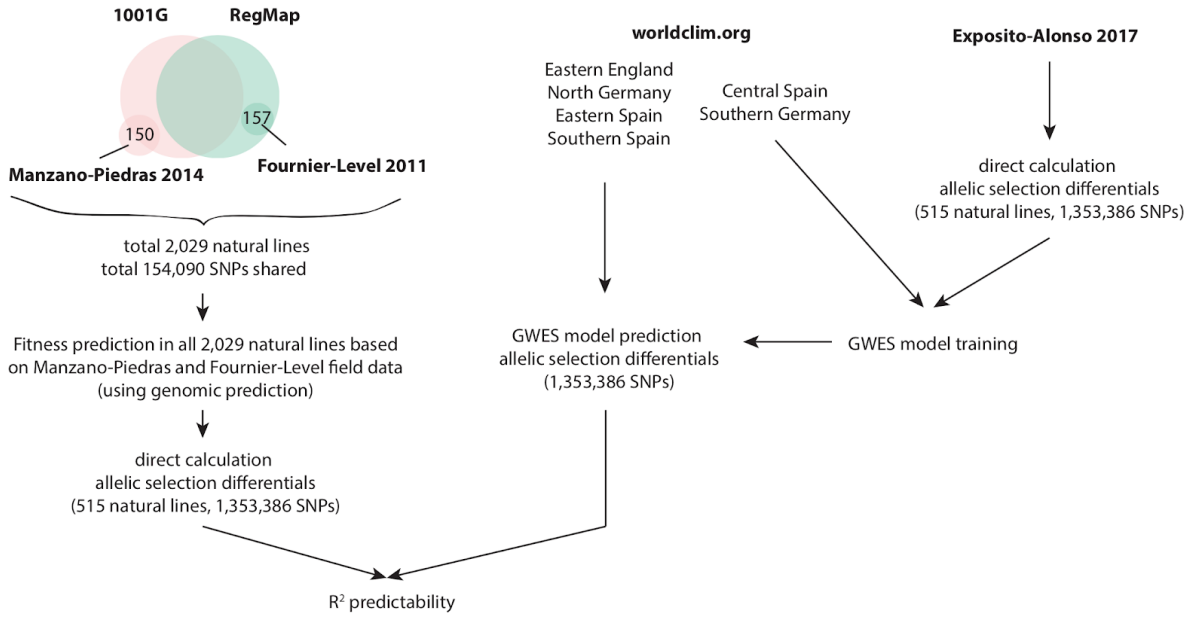
776 As Fig. 2B, but for allele selection differentials in Germany under high precipitation.

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



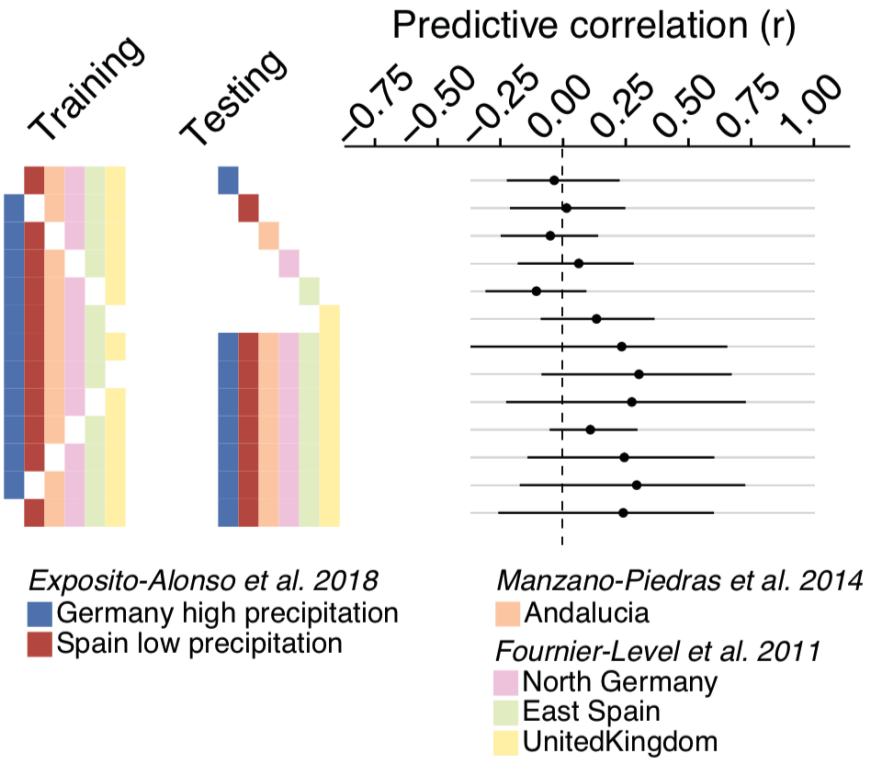
778 Same as Fig. 3G, but for different climate change scenarios. The higher the predicted CO_2 emissions,
779 the stronger the predicted increase in selection intensity.

780 **Figure S12. Field validation conceptual chart**



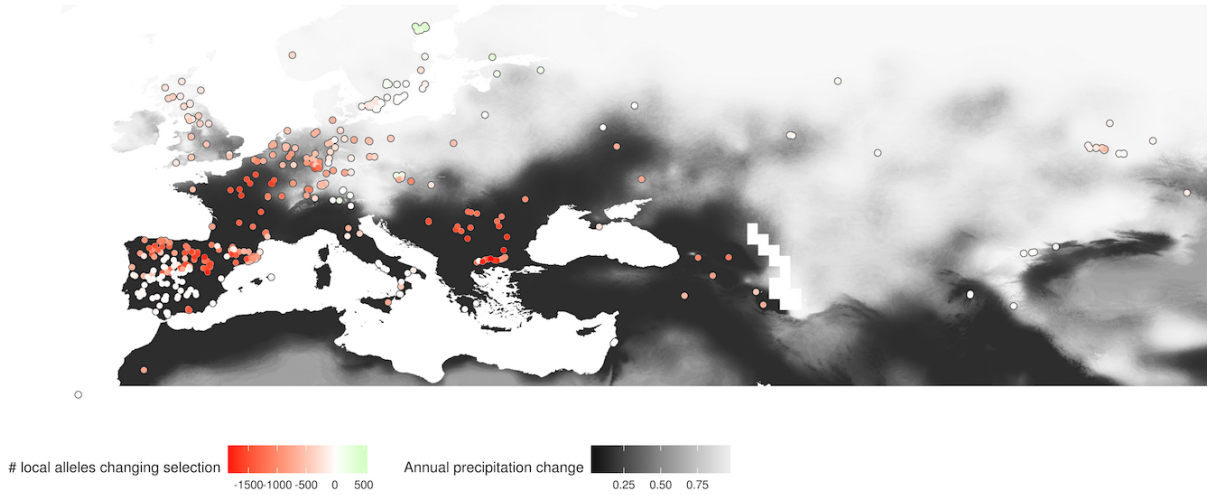
781 Conceptual workflow on field validation procedure with data from published experimens.

Figure S13. Null expectation of predictability



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

788 **Figure S14. Change in selection relative to local diversity**



789 Same as Fig. 3D, but counting the number of local alleles increasing or decreasing in selection (total
790 n=10,752 SNPs). Only changes with more than 5% advantage/disadvantage were considered (defined
791 *a posteriori* from Bonferroni-significant alleles, which generated at least 5% effect in fitness).

SUPPLEMENTAL TABLES & DATASETS

SupplementaL 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 latitude and longitude are reported.

Table S2. Summary of fitness data

Average survival, fecundity, and lifetime fitness. Total number of genotypes with at least one surviving replicate 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 line genotype), 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, and 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 GWES 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. 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 S10. Correlation between inferred natural selection intensity and other variables

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

SUPPLEMENTAL REFERENCES

50. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27,** 2156–2158 (2011).
51. 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).
52. 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).
53. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33,** 1–22 (2010).
54. Kirkpatrick, M., Johnson, T. & Barton, N. General models of multilocus evolution. *Genetics* **161,** 1727–1750 (2002).
55. Mitchell-Olds, T. Pleiotropy Causes Long-Term Genetic Constraints on Life-History Evolution in Brassica rapa. *Evolution* **50,** 1849–1858 (1996).
56. Novembre, J. & Barton, N. H. Tread Lightly Interpreting Polygenic Tests of Selection. *Genetics* **208,** 1351–1355 (2018).
57. Shriner, D. Overview of admixture mapping. *Curr. Protoc. Hum. Genet.* **1,** 1.23.1–1.23.6 (2013).
58. 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).
59. Dr, F., Mt, W., Ja, A., Ts, S. & Zm, E. EcoHydRology: A community modeling foundation for Eco-Hydrology. (2014).
60. Golicher, D. Implementing a bucket model using WorldClim layers. (2012). Available at: https://rpubs.com/dgolicher/2964.
61. Kuhn, M. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software, Articles* **28,** 1–26 (2008).
62. Breiman, L. Random Forests. *Mach. Learn.* **45,** 5–32 (2001).
63. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2,** 18–22 (2002).

857 64. Platt, A. *et al.* The scale of population structure in Arabidopsis thaliana. *PLoS Genet.* **6,** e1000843
858 (2010).
859 65. Endelman, J. B. Ridge Regression and Other Kernels for Genomic Selection with R Package
860 rrBLUP. *Plant Genome* **4,** 250–255 (2011).
861 66. Exposito-Alonso, M., Brennan, A., Alonso-Blanco, C. & Picó, F. X. Spatio-temporal variation in
862 fitness responses to contrasting environments in Arabidopsis thaliana. *Evolution (in press)*
863 (2018).