1 **Climate biogeography of *Arabidopsis thaliana:* linking distribution models and individual**
2 **variation**
3
4 Christina Yim[1], Emily S. Bellis[1,2], Victoria L. DeLeo[1], Diana Gamba[1], Robert Muscarella[3],
5 Jesse R. Lasky[1,*]
6
7 [1]Department of Biology
8 Pennsylvania State University
9 University Park
10 USA
11
12 [2]Department of Computer Science
13 Arkansas State University
14 Jonesboro
15 USA
16
17 [3]Plant Ecology and Evolution
18 Evolutionary Biology Centre
19 Uppsala University
20 Uppsala
21 Sweden
22
23 *Corresponding Author
24 lasky@psu.edu
25
26 **Biosketch**
27 The research team is interested in understanding how individual-level processes influence
28 evolutionary and ecological dynamics at continental scales.
29
30 **Author contributions**
31 All authors contributed to study design. CY, ESB, VLD, DG, and JRL conducted analyses. CY
32 and JRL led the writing with contributions from all authors.

Abstract

AIM

The role of environmental conditions in limiting species distributions is often hypothesized in biogeography, but it is challenging to gather large-scale data to demonstrate environmental impacts on individual performance. The past and present biogeography of model organisms is key context to understanding how environment shapes species' genetic and phenotypic diversity.

LOCATION

Global

TAXON

Arabidopsis thaliana ("Arabidopsis")

METHODS

We fit occurrence records to climate data, and then projected the distribution of Arabidopsis under last glacial maximum, current, and future climates. We confronted model predictions with individual performance measured on 2,194 herbarium specimens, and we asked whether predicted suitability was associated with life-history and genomic variation measured on 898 natural accessions.

RESULTS

The most important climate variables constraining the Arabidopsis distribution were winter cold in northern and high elevation regions and summer heat in southern regions. Herbarium specimens from regions with lower habitat suitability in both northern and southern regions were smaller, supporting the hypothesis that the distribution of Arabidopsis is constrained by climate-associated factors. Climate anomalies partly explained interannual variation in herbarium specimen size, but these did not closely correspond to local limiting factors identified in the distribution model. Late-flowering genotypes were absent from the lowest suitability regions, suggesting slower life histories are only viable closer to the center of the realized niche. We identified glacial refugia farther north than previously recognized, as well as refugia concordant with previous population genetic findings. Lower latitude populations, known to be genetically distinct, are most threatened by future climate change. The recently colonized range of Arabidopsis was well-predicted by our native-range model applied to certain regions but not others, suggesting it has colonized novel climates.

MAIN CONCLUSIONS

Integration of distribution models with performance data from vast natural history collections is a route forward for testing biogeographical hypotheses about species distributions and their relationship with evolutionary fitness across large scales.

2

Introduction

A major goal in biology and biogeography is to understand how environmental conditions limit the performance of individuals and species distributions. First, the environment-distribution relationship can help project species distributions under past climates to understand their ecological and evolutionary history (Forester et al., 2013). Second, the environment-distribution relationship can identify regions with currently suitable habitat but that are unoccupied for a variety of reasons (e.g. dispersal limitation) (Elith et al., 2010). Third, the environment-distribution relationship can predict how distributions will shift under future environments (Thomas et al., 2004), which is an urgent task due to rapid anthropogenic climate change.

One avenue to advance these goals in future biogeographic studies is through greater integration of biogeography with organismal biology to test hypotheses about the organismal and population mechanisms controlling distributions. For example, it is becoming possible to collect molecular genetic and phenotypic data on large numbers of organisms across a species range. These approaches are particularly accessible (and valuable) in studies on model organisms like *Arabidopsis thaliana*. With model systems researchers can link detailed information on genetics and organismal biology with population and community processes (Rudman et al., 2019; Takou et al., 2019) to contribute ecological context and advance integrative biology. At the same time, advances in the organization and digitization of museum specimens are rapidly expanding the available data on range-wide variation among individuals from nature. In this spirit we focus here on *Arabidopsis thaliana,* a small annual plant (hereafter referred to as "Arabidopsis") (Koornneef & Meinke, 2010). Arabidopsis has been key to understanding how molecular biology and physiology are linked to ecology, but past biogeographic studies did not use newer distribution modeling tools (Hoffmann, 2002) or overlooked large parts of its range (Banta et al., 2012; Zou et al., 2017).

Environment-distribution relationships fundamentally arise from processes acting on individuals (Clark, 2010). In general, transplant experiments show that individual performance tends to decline outside species' natural geographic range (Hargreaves et al., 2014) and efforts to integrate information at the individual level into distribution models are emerging (Buckley et al., 2011; Elith et al., 2010; Lasky et al., 2020; Merow et al., 2014; Samis & Eckert, 2007). Where populations inhabit harsh environments (e.g. at range margins), local adaptations can emerge, such as life history changes to tolerate or escape harsh periods (Bontrager et al., 2021). In Arabidopsis, there is evidence that local adaptation to environment involves genetic changes in life history (e.g. flowering time) (Lovell et al., 2013; Martínez-Berdeja et al., 2020). However, it is challenging to determine how environment shapes individual performance and life history variation across the ranges of broadly distributed species, as well as the consequences for distributions.

A vast resource of individual-level information can be found in natural history collections (Heberling, 2021; Lopez et al., 2019). For example, Bontrager & Angert (2015) showed with herbarium specimens of *Clarkia* that fecundity decreased with drier summers, and toward the western species range margin both summer precipitation and individual fecundity declined, suggesting a mechanism limiting its distribution. In Arabidopsis, DeLeo et al. (2020) found decadal shifts in traits of herbarium specimens. For many species, seed banks house great diversity from across their ranges that can also be used to study their biogeography (Estarague

120   et al., 2021; Scholl et al., 2000). Here we estimate performance and life history variation from
121   museum and seed bank accessions to interrogate predictions from distribution models.
122   Arabidopsis is native across Eurasia and Africa and with human assistance has
123   colonized the Americas and Australia. Other species in the genus *Arabidopsis* are more
124   restricted to cool temperate climates, with Arabidopsis having expanded to a broader range, *e.g.*
125   Mediterranean habitats (Hoffmann, 2005). Arabidopsis can behave as a spring annual with a
126   rapid life cycle, germinating in the spring and flowering in the late spring and summer. However,
127   many individuals are longer-lived winter annuals, germinating in the fall and flowering in early
128   spring (Wilczek et al., 2009). Studies of large-scale environmental response in Arabidopsis have
129   focused on the role of local adaptation in genetic diversity in the species (e.g. Hancock et al.,
130   2011; Lasky et al., 2018; Martínez-Berdeja et al., 2020; Toledo et al., 2020), but less is known
131   about the determinants of the species' distribution. The last overview of the climate
132   biogeography of Arabidopsis was Hoffmann (2002), who considered Arabidopsis native to
133   western and central Eurasia, but non-native in China and sub-Saharan Africa. However, studies
134   show these latter populations are likely native, being genetically diverse, and with many unique
135   genetic variants (Durvasula et al., 2017; Zou et al., 2017). Advances in distribution modeling
136   (Elith et al., 2010; Muscarella et al., 2014) and updates to climate and occurrence datasets
137   highlight the need to update to our understanding of Arabidopsis biogeography.
138   Studies suggest that multiple climatic factors limit performance and distribution of
139   Arabidopsis. Experiments have shown that winter cold is a major factor limiting performance
140   (Ågren & Schemske, 2012; Korves et al., 2007). Additionally, Arabidopsis lacks physiological
141   traits for dealing with severe water deficit so it is likely individual performance in nature is limited
142   by drought (Clauw et al., 2015). Based on occurrence data, Hoffmann (2002) concluded that
143   spring and fall cold limited Arabidopsis in northern Europe, winter temperatures in Russia, heat
144   and drought toward the southern range margins, and a lack of snow in central Asia. However,
145   these conclusions were based on visual inspection of climate isotherms and range margins, not
146   statistical inference. Distribution models fit to occurrence data allow an explicit quantitative
147   statement of environment-distribution relationships and the potential to infer the environmental
148   factors limiting distributions (Elith et al., 2010; G. Li et al., 2015).
149   Distributions are dynamic through time due to environmental change, dispersal, and
150   demographic stochasticity. Studies have used genetic data to infer Arabidopsis refugia during
151   the last glacial maximum (LGM), where subpopulations (sometimes referred to as "relicts")
152   persisted locally before subsequent admixture with an expanding, now widely distributed "non-
153   relict" lineage (Lee et al., 2017). Whether these refugia corresponded to suitable climates is less
154   clear. Future climate change projected impacts on Arabidopsis have focused on *relative* climate
155   impacts on different genotypes (Exposito-Alonso et al., 2018; Fournier-Level et al., 2016), rather
156   than distribution dynamics. Additionally, many species, have colonized new regions due to
157   human introduction, sometimes exhibiting traits that appear distinct from native range
158   populations, potentially due to plastic or genetic responses to new environments (Turner et al.,
159   2015). Arabidopsis has colonized many regions, but it is unclear to what degree these places
160   represent novel environments.
161   Here we sought to demonstrate an approach to integrate distribution modeling with
162   organismal biology and intraspecific variation, and to revise our understanding of Arabidopsis's
163   biogeography. To do so, we combined occurrences (including many outside Europe poorly

164   represented in previous work) with climate data to build distribution models, and then test how
165   model predictions correspond to performance estimated from herbarium specimens and genetic
166   variation in natural accessions. We ask the following questions:

167   1. What climate factors constrain the distribution of Arabidopsis? We hypothesize that
168   winter cold and summer drought are the most important constraints, depending on
169   region.
170   2. Are individual performance and genetic variation associated with model-estimated
171   habitat suitability? We hypothesize that individuals reach larger sizes in regions with
172   greater suitability and that populations adapt along gradients in suitability through
173   changes in traits such as flowering time, a key component of life history.
174   3. Are occurrences outside the native range predicted by a native range model, suggesting
175   stable realized niches following colonization? Or is there evidence Arabidopsis has
176   colonized novel environments?
177   4. Where did Arabidopsis persist during the Last Glacial Maximum (LGM)? And where will
178   Arabidopsis move in future climates?

179

180   **Methods**

181   *Occurrence data*
182   We developed a set of high-quality occurrence data (*i.e.* species ID verified and location
183   checked, N=4,024) from published research (Durvasula et al., 2017; Hsu et al., 2019;
184   Mandáková et al., 2017; Zeng et al., 2017; Zou et al., 2017), publicly available herbarium and
185   germplasm accessions with known collection locations (Alonso-Blanco et al., 2016; DeLeo et
186   al., 2020), and some of our recent field collections in East Africa (Gamba et al., 2022). These
187   span a period of 1794 - 2018. The herbarium specimens and new collections include little-
188   studied populations in Saudi Arabia, Somalia, Djibouti, Eritrea, Rwanda, Ethiopia, Uganda,
189   Sudan, and Nepal. Duplicate occurrence points were eliminated (samples are often split and
190   sent to different herbaria). For model fitting, we excluded occurrences from putative non-native
191   regions (the Americas, New Zealand, Japan).
192   We also used occurrence data (with coordinates and without flagged problems
193   N=115,226) from the Global Biodiversity Information Facility (GBIF) to test model predictions in
194   regions outside of the native range of Arabidopsis (downloaded 30 Dec 2020, Gbif.Org, 2020).
195   We deem these occurrences as lower quality given that many have not had the species identity
196   and location checked (DeLeo et al., 2020).

197

198   *Environmental data*
199   Climate data were extracted from CHELSA (Climatologies at High resolution for the Earth's
200   Land Surface Areas) v1.2 at 30 arc second spatial resolution (Karger et al., 2017). Current
201   conditions are the average of 1979-2013 estimates. We selected the following climate variables
202   based on hypothesized importance (Gienapp et al., 2017; Hancock et al., 2011; Lasky et al.,
203   2014, 2018) and relatively low inter-correlation (Pearson correlation coefficients among
204   variables at occurrences < 0.75): isothermality (Bio3), minimum temperature of coldest month
205   (Bio6), temperature annual range (Bio7), mean temperature of wettest quarter (Bio8), mean
206   temperature of the warmest quarter (Bio10), precipitation seasonality (Bio15), precipitation of

207   wettest quarter (Bio16), and precipitation of driest quarter (Bio17). We also included altitude
208   from Hijmans et al. (2005).
209   For projecting past distributions, we obtained climate estimates from the last glacial
210   maximum (LGM) at 21k yrs before present from CHELSA PMIP3 (Karger et al., 2017). We used
211   the global altitude and bathymetry map with 15 arc second resolution from Tozer et al. (2019)
212   with a sea level 134 m lower than present (Lambeck et al., 2014) to project potential suitable
213   habitat at the LGM on land in areas currently submerged. For projecting future distributions, we
214   used climate projections five divergent global climate models for 2050 from CHELSA v1.2 using
215   the RCP 4.5 emissions scenario (Karger et al., 2017). We also show RCP 6.0 in the supplement
216   for context (Figure S13), though it is highly similar to 4.5 in the target time period.
217   To characterize temporal variation in climate (climate anomalies), we used the Climate
218   Research Unit (CRU) TS 4.01 dataset, providing a global time series of monthly temperature
219   and precipitation for the period 1900-2010 at a 0.5º resolution (Harris et al., 2014). From the
220   CRU data we calculated the same bioclimatic parameters that we used from CHELSA, but in
221   the CRU data these bioclimatic variables were specific to each herbarium specimen in the time
222   period it was collected (Supplemental Methods). We then calculated local anomalies for each of
223   these variables by taking the observed value, subtracting mean across the entire time series,
224   and dividing by the standard deviation (DeLeo et al., 2020).
225
226   *Performance estimates from herbarium specimens*
227   We estimated fecundity on a subset of herbarium specimens using two traits. First, we
228   measured the length of the longest inflorescence, reasoning that longer inflorescences would
229   have more fruits and seeds. Second, we measured maximum rosette leaf length, reasoning that
230   larger rosettes would support later reproductive investment if these collected plants were
231   allowed to continue growth *in situ* (see Supplement for a validation). We used ImageJ on 2,194
232   specimen images to estimate the tallest point of each inflorescence (N=2,188) and the
233   maximum rosette leaf length (N=1,264; see Supplement).
234
235   *Range-wide genetic variation in life history*
236   Many late-flowering Arabidopsis genotypes require cold cues (vernalization) to flower and also
237   show slower growth and more stress tolerance, delineating a life history axis (Lovell et al., 2013;
238   Vasseur et al., 2018). To assess life history variation, we used published experimental data on
239   flowering time for 898 whole-genome resequenced accessions from the native range with
240   reliable geographic coordinates grown at 10 and 16ºC (The 1001 Genomes Consortium 2016).
241   To estimate vernalization sensitivity, we calculated the difference between flowering time at 10
242   and 16ºC.
243
244   *Species Distribution Modeling*
245   We thinned the original 4,024 high quality occurrence points to one sample per 1 km grid cell to
246   reduce sampling bias (N=662) using the 'sp' package (Bivand et al., 2008). To characterize
247   potentially inhabited sites, we generated pseudoabsence background points using the 'dismo'
248   package (Hijmans & Elith, 2013) by sampling 10,000 random points within a 500 km buffer
249   around occurrence points.

250   We used Maxent version 3.4.0 to generate a species distribution model (Phillips et al.,
251   2006). MaxEnt was implemented with the R package 'ENMeval' v2, and parameters were
252   optimized using the 'checkerboard2' method for cross validation (Muscarella et al., 2014).
253   Among the tested settings (ENMeval defaults), we chose the model with lowest AICc value and
254   used this to project habitat suitability under recent conditions across the globe. For all models
255   we used the logistic output of MaxEnt that scales suitability from zero to unity.
256   We used permutation importances to determine which climatic factors drove predictions
257   in the distribution model. We also used the 'limiting' function in the R package 'rmaxent' to
258   determine the most limiting climatic factors in each location, defined as the environmental
259   covariate that has the largest decrease in suitability at a given location relative to the suitability
260   achieved if the covariate had its value equal to the global mean (Baumgartner et al., 2017; Elith
261   et al., 2010). State another way, the local limiting factor is the environmental condition most
262   limiting suitability, compared to an alternative scenario where that condition takes its global
263   mean.
264   In non-native regions, we evaluated whether Arabidopsis is limited from further
265   expansion at range edges by climate, *i.e.* whether there were no more unoccupied suitable
266   environments near existing populations. To do so, we calculated suitability in a zone 50-100 km
267   from existing GBIF occurrences and compared occupancy in these buffers in the native range to
268   invaded regions.
269   We also projected the MaxEnt model using past (LGM) and future climate conditions.
270   For future conditions, we calculated the mean and standard deviation of habitat suitability
271   projected for the five climate models. To assess whether model predictions were extrapolating
272   into poorly characterized or novel climates, we compared the present-day model training
273   climates to each predicted climate conditions, calculating multivariate environmental similarity
274   surfaces (MESS) following Elith et al. (2010). Higher values on a MESS map indicate conditions
275   in a location (or point in time) are similar to the reference environmental conditions used to fit
276   the model. Negative values indicate that at least some variables are outside the range of
277   environments used to fit the model, signifying extrapolation into novel environments (Elith et al.,
278   2010).
279
280   *Performance and habitat suitability*
281   We asked whether suitability corresponded to plant size. We first tested these relationships with
282   Pearson's and Spearman's correlations. We also fit Generalized Additive Models (GAMs) where
283   herbarium specimen sizes were the response variable. The model included covariates for
284   suitability at the collection location (square-root transformed to reduce the lower-tail influence)
285   and (as a nuisance variable) the year of collection (scaled to mean zero and unit variance) to
286   account for potential changes in size over time. We used GAMs to allow smooth spatial
287   variation in parameters, allowing us to capture any regional variation. The model can be
288   represented as:
289
290    $$\text{Phenotype}_{i,year} = \beta_0 + \beta_{i,year}\text{Year} + \beta_{i,HS}\text{Habitat Suitability}_i + \varepsilon_{i,year}$$
291
292   where i represents the specimen location. Errors were modeled as normally distributed.
293   Spatially-varying parameters were constrained to smooth spatial variation using the 'mgcv'

294   package for R, which we used with REML to fit all GAMs (Wood, 2011). We considered

295   covariates to be significant at a given location if their 95% confidence interval (CI) excluded 0,

296   and the Moran's I (Cooper, 2021) for model residuals did not indicate spatial autocorrelation.

297   We also asked whether local limiting factors corresponded to climate variables with

298   significant effects on plant performance. We hypothesized that for a given region, temporal

299   fluctuations in local limiting factors would be associated with temporal variation in plant size in

300   herbarium specimens. To address this hypothesis, we tested if specimen sizes were correlated

301   with anomalies in climate conditions corresponding to the climate variables used in the MaxEnt

302   model. We used the yearly climate anomalies we calculated from CRU in GAMs with spatially

303   varying coefficients, where for a phenotype measured at location i:

304

305              $\text{Phenotype}_{i,year} = \beta_{i,0} + \beta_{i,year}\text{Year} + \beta_{i,climate}\text{BioClim Variable Anomaly} + \varepsilon_{i,year}$

306

307   We then qualitatively assessed whether local limiting factors inferred by MaxEnt had anomalies

308   that were correlated with plant size (local $\beta_{i,climate}$ coefficients in the GAMs).

309

310   *Genetic variation and habitat suitability*

311   We used two strategies to test how genetic variation in Arabidopsis varies with suitability. First,

312   we focused on genetic variation in a key life history trait, flowering time. We estimated linear

313   models relating suitability and flowering time. To account for population structure and neutral

314   processes that can affect spatial variation in flowering time, we also tested whether suitability

315   was associated with flowering time after using random effects for genome-wide similarity among

316   accessions. A significant suitability-flowering time association in this model would suggest

317   selection linked to suitability acts on flowering time. This test is akin to Q_{ST}-F_{ST} contrasts

318   (Whitlock & Guillaume, 2009), except that an explicit environmental gradient is tested

319   (suitability). We implemented this mixed model the function 'lmekin' from the R package 'coxme'

320   (Therneau & Therneau, 2015), along with 'kinship2' (Sinnwell et al., 2014). The kinship matrix

321   was obtained from 2,027,463 published whole genome resequencing SNPs (Alonso-Blanco et

322   al., 2016). Additionally, to test for geographic variation in suitability-flowering time relationships,

323   we fitted GAMs of flowering time with spatially varying coefficients:

324

325              $\text{FloweringTime}_{i,year} = \beta_0 + \beta_{i,HS}\text{Habitat Suitability} + \varepsilon_i$.

326

327   Our second strategy was to scan the Arabidopsis genome for genes where different

328   alleles were found in high versus low suitability locations. Such a change in allele frequency

329   across gradients in suitability would suggest that this genetic variation was involved in local

330   adaptation to low versus high suitability environments. We used univariate linear mixed-effects

331   models in GEMMA (v 0.98.3) (Zhou & Stephens, 2012) to perform genome wide association

332   studies (GWAS) in a set of 2,053,939 SNPs filtered for MAF=0.05 from 1003 native-range

333   ecotypes part of the 1001 genomes panel.

334

335

336   **Results**

337

338   *Q1. Climate constraints on the distribution of Arabidopsis*

339   Our optimization of MaxEnt models (with AICc) resulted in a selection of a model with linear and

340   quadratic effects, with the regularization multiplier of 0.5. The training AUC value was 0.78 and

341   the average test AUC with checkerboard2 cross validation was 0.78. The omission rate of the

342   10th percentile of suitability for training points was 0.10, suggesting our models were not overfit

343   as they were able to predict low probability occurrences as well as expected (Fielding & Bell,

344   1997; Peterson et al., 2011). The areas of high suitability overall correspond well to the

345   documented Arabidopsis distribution, with one notable exception being tropical lowland sites

346   (Congo basin) in sub-Saharan Africa (Figure 1A). This region was near zero multivariate

347   environmental similarity to training data, indicating the model may have been poorly constrained

348   there, while most of the regions of high predicted suitability where Arabidopsis is documented

349   have positive similarity (MESS, Figure S2).

350

351

352 **Figure 1. (A)** Habitat suitability (green) under current climate conditions with thinned
353   occurrences (n = 662) used in fitting shown as black circles and (B) limiting factors from the
354   MaxEnt model fit to current climates. Regions far (>500 km) from known occurrence have a gray
355   mask in (A). Regions of low suitability (less than 0.25) in (B) are marked with gray 'x' symbols.
356   Equal Earth projection was used. Abbreviations in (B) as follows: precipitation of the driest
357   quarter "Prec driest Q", precipitation of the wettest quarter "Prec wettest Q", precipitation
358   seasonality "Prec seas", mean temperature of the warmest quarter "Mean temp wrm Q", mean
359   temperature of the wettest quarter "Mean temp wettest Q", temperature annual range "Temp
360   ann range", minimum temperature of the coldest month "min temp coldest month.
361   **A.**

380   **B.**

10
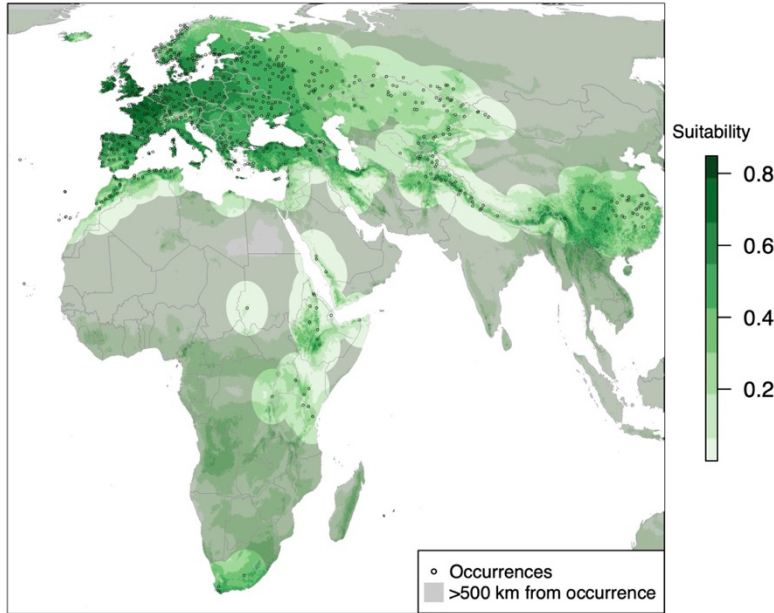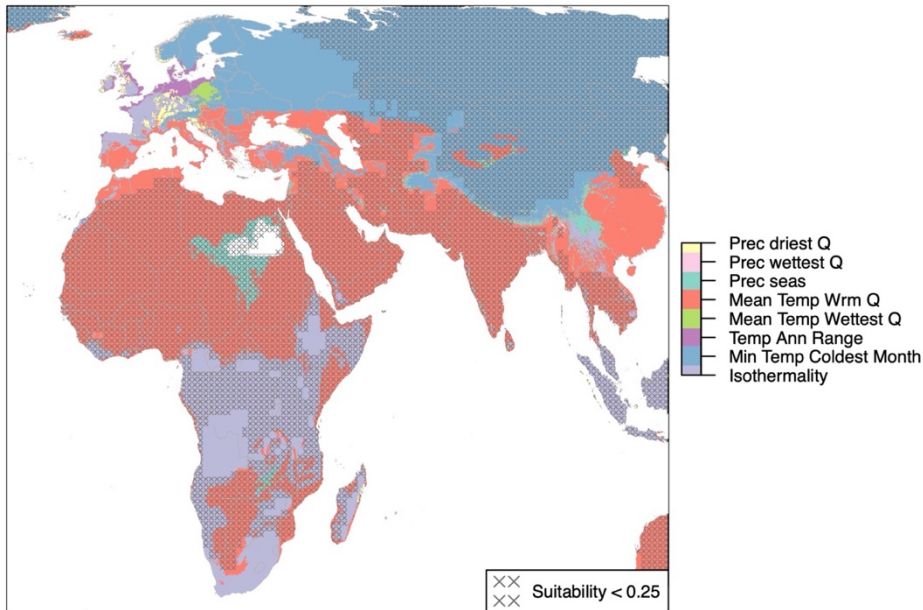
386 The permutation importance of the model covariates across the entire native range revealed
387 that the minimum temperature of the coldest month (PI = 40%) and the mean temperature of the
388 warmest quarter (PI = 32%) were the two most important variables, suggesting winter cold
389 stress and summer heat stress are most important in constraining Arabidopsis's distribution
390 (Table S1). The next most important variable was isothermality (PI =14%), as Arabidopsis tends
391 to be found in regions with low isothermality (e.g. most temperate zones).
392   To identify spatial variation in climate constraints, we also identified local limiting factors
393 (Figure 1B). Across northern Eurasia, minimum temperature of the coldest month limited habitat
394 suitability. Across the Mediterranean and tropical/subtropical regions, the temperature of the
395 warmest quarter was limiting. Across Eastern Europe and Central Asia, winter cold was limiting
396 adjacent to other regions where summer heat was limiting, highlighting the multiple stressors in
397 this region.
398   We hypothesized that for a given region, temporal fluctuations in local limiting factors
399 would be associated with temporal variation in plant size in herbarium specimens. We found
400 that years with warmer minimum winter temperatures were associated with significantly taller
401 inflorescences from Turkey to central Asia (Figure S3). Concordantly, in much of this region the
402 Maxent model indicated that minimum temperature of the coldest of the month was the most
403 limiting factor (namely in the Caucuses and from Kazakhstan to northern India, Figure 1B). We
404 also found that years with greater seasonality of precipitation were associated with shorter
405 inflorescences in central Asia, but taller inflorescences in central Europe (Figure S3). Partly
406 consistent, in the southern part of this central Asian region the MaxEnt model indicated
407 precipitation seasonality was limiting (Figure 1B). Years with higher isothermality were positively
408 associated with inflorescence height in Eastern Europe but this was not a limiting factor in this
409 region (Figure S3). Other climate anomalies were not significantly associated with temporal
410 variation in plant size.
411
412 *Q2. Declining fitness and stress escape life history in regions of low predicted habitat suitability*
413 As habitat suitability increased, so did the inflorescence height of individual plants in herbarium
414 specimens (Pearson's r = 0.11, p < 10^{-7}; Spearman's rho = 0.09, p < 10^{-5}; n = 2053). The
415 relationship with suitability was even stronger for maximum rosette leaf length (r = 0.23, p < 10^{-16}
416 ; rho = 0.24, p < 10^{-16}; n = 1179, Figure 2), which was also more correlated with total fruit +
417 flower number (Supplement), suggesting fecundity was greater in regions of high predicted
418 suitability. We tested size-suitability relationships using GAMs with spatially varying suitability
419 coefficients (but not including climate anomalies – distinct from the previous section Q1). We
420 found a consistently positive relationship between suitability and size that was significant for
421 rosette leaf length in western Europe and for inflorescence height across most of Eurasia (Figs
422 S5 and S6). Unexpectedly, in these models where year of collection was considered primarily
423 as a nuisance variable, plant size significantly declined over time in northwest Europe (Figs S5
424 and S6; e.g. in Scandinavia, maximum rosette leaf length vs year, Spearman's rho = -0.17).
425
426 **Figure 2**. Predicted current habitat suitability compared with individual plant size (A & B) and
427 genetic variation in a measure of life history (C). Size measures include inflorescence height (A,
428 relationship with suitability: Pearson's r = 0.11, Spearman's rho = 0.09) and maximum rosette
429 leaf length (B, r = 0.23, rho = 0.24) from herbarium specimens. Flowering time of natural

430   accessions (C, r = 0.21) was taken from published data on a growth chamber experiment at
431   10ºC (Alonso-Blanco et al., 2016). Linear model fits are shown. N = 2053 for inflorescence
432   height, N = 1179 for maximum rosette leaf length, and N=953 for flowering time.
433

Herbarium specimens   **Natural accessions**

434
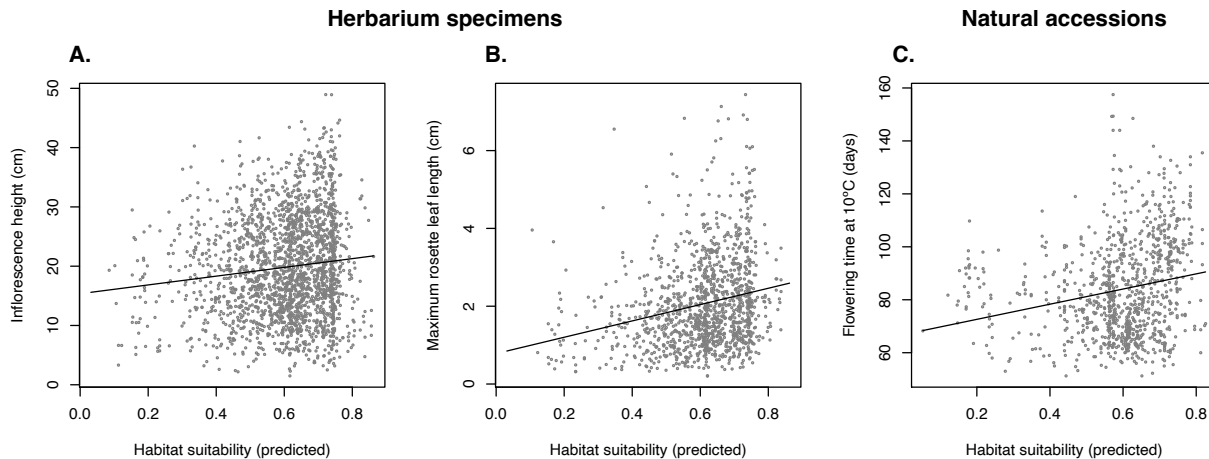
435   We compared habitat suitability with published data on genetic variation in flowering
436   time. We found that flowering time at 10ºC and flowering time plasticity were significantly
437   positively associated with suitability (r=0.21, p<10^{-11} and r=0.10, p=0.0021, respectively), and
438   also for days to flower at 16 ºC (r=0.08, p=0.0166, Figures 2 & S7). In GAMs with spatially-
439   varying suitability coefficients, the suitability-flowering time pattern were largely consistent
440   across Eurasia (Figure S8). The suitability-flowering time association was significant even when
441   accounting for genomic similarity among accessions, suggesting selection associated with
442   suitability could maintain geographic clines in flowering time. Specifically, linear mixed-effects
443   models found a positive association for flowering time at 10ºC (p<10^{-6}, n=953), and also at
444   16ºC, the latter of which may have been obscured by population structure that was unaccounted
445   for in the simple linear regression model (p<10^{-5}, n=920). The suitability association was not
446   significant for plasticity in flowering time when accounting for genomic similarity (p=0.38, n=920)
447   (Tables S1-S3). Consistent with this potential obscuring of population genetic structure, we
448   found that 10 ADMIXTURE genetic clusters in Arabidopsis (Alonso-Blanco et al., 2016) were
449   significantly different in their habitat suitability (F(9,993)=201.6, p<10^{-16}, Table S5, Figure S9).
450   We scanned the genome for genes that showed allele frequency correlations with
451   suitability. The most strongly associated SNP was in the putative promoter region (878 bp from
452   the start) of ERF53 (AT2G20880), a transcription factor that regulates response to drought, salt,
453   and heat (Figure S11) (Cheng et al., 2012; B. Li et al., 2019). This SNP showed a strong allele
454   frequency cline from Europe to Asia, where the alternate allele was nearly fixed in accessions
455   east of the Ural Mountains, which is a region of low estimated suitability (Fig S10). Furthermore,
456   we found that this alternate SNP allele was associated with higher expression of ERF53 (Wilcox
457   test, p = 0.0071) in published transcriptome data (Kawakatsu et al., 2016), suggesting a *cis*-
458   regulatory variant is locally adapted to low suitability parts of Asia.
459
460
461   *Q3. Using the native range model to predict outside the native range*

462 Australia and New Zealand occurrences were largely in areas predicted to be highly suitable
463 (Figure 3). By contrast, North America includes occurrences in highly suitable areas (Pacific
464 Northwest) but many in low suitability areas (interior east, where summer heat was predicted
465 limiting, Figure S12). The eastern US also has near zero similarity on the MESS map,
466 suggesting a novel set of conditions compared to the native range (Figure S2). Similarly, the
467 regions where Arabidopsis occurred in Korea and Japan were of lower suitability than in the
468 native range, also with summer heat as the predicted limiting factor (Figure S12), suggesting
469 Arabidopsis in eastern North America and East Asia inhabits climates with distinctly warm
470 summers.
471
472 **Figure 3.** Distribution of predicted habitat suitability (based on our native-range MaxEnt model,
473 underlying map surfaces in green) for GBIF occurrences in various regions (black circles). We
474 do not include a histogram for South America because there are too few occurrences. Equal
475 Earth projection was used.
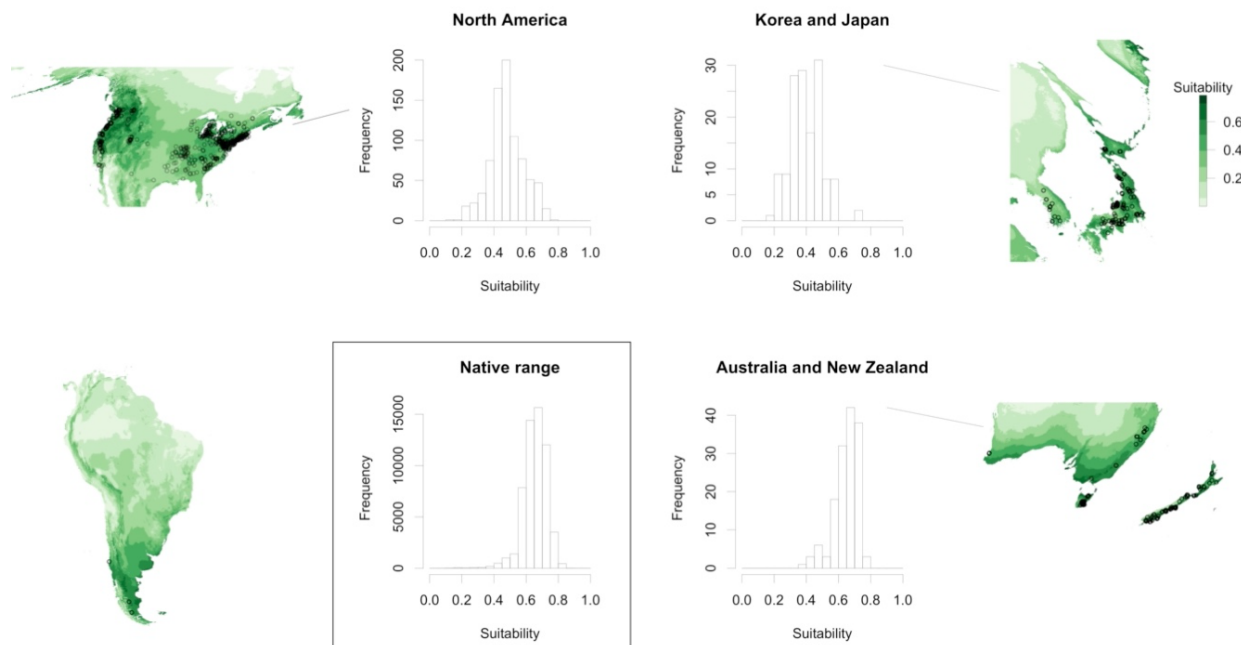
476
477   We investigated whether non-native Arabidopsis occupy all available habitat in their
478 regions or whether suitable habitat remains unoccupied. We found that locales 50-100 km from
479 occurrences included many areas of high suitability in North America, suggesting suitable sites
480 remain unoccupied (11.8% of grid cells in these 50-100 km regions had suitability > 0.6). In
481 comparison there were fewer such sites in the native range (5.1% of these 50-100 km regions
482 had suitability > 0.6). There were extensive areas of high suitability along the Pacific coast of
483 North America to around 60ºN, but no occurrences north of 50º, which we confirmed with an
484 expert botanist (pers comm. Matthew Carlson). The timing of the invasion is likely not a factor
485 given that occurrences from near Vancouver date at least to 1939. This region has near zero
486 similarity on the MESS map, potentially indicating that the native range model is not well
487 constrained there (Figure S2). Similarly, the southern coast of Australia is highly suitable but
488 Arabidopsis is apparently absent (31.9% of these 50-100 km regions with suitability > 0.8).

13

489   Records are restricted to the southeast and southwest (confirmed by botanists, Shelley James,
490   Tim Entwistle, Neville Walsh pers. comm.), even though records date at least to 1959 in SE
491   Australia. In South America there are a few Arabidopsis records from the Southern Cone, and it
492   appears to be rare in the region (pers. comm. Diego Salariato), while apparently suitable
493   environments occur throughout Patagonia and high elevation Andean sites that are apparently
494   unoccupied (pers comm. Gwendolyn Peyre, Santiago Madriñán). Large areas of southern
495   Australia and South America show positive similarity on the MESS map suggesting the model is
496   well constrained in those regions. By contrast in Korea and Japan there are very few sites
497   expected to be highly suitable that are not already occupied by Arabidopsis (4.9% of these 50-
498   100 km regions with suitability > 0.6).
499
500   *Q4. The distribution of Arabidopsis during the Last Glacial Maximum and in the future*
501   We projected suitability onto the climate at the LGM and found several putative refugia, where
502   past environmental conditions could have supported Arabidopsis persistence. In particular, the
503   Mediterranean/Caucuses/south Caspian Sea, much of sub-Sharan Africa > 1000 m (current)
504   asl, and China and SE Asia appear as refugia (Figure 4). North Africa, the Atlantic European
505   coast, and the islands of Sicily, Corsica, and Sardinia also appear as highly suitable potential
506   refugia.
507
508

14

509 **Figure 4.** (A) Predicted distribution during the Last Glacial Maximum and (B) areas in black with
510 suitability > 0.3 during both the LGM and current conditions as well as within 500 km of known
511 current occurrences. Because of their greater level at the LGM, in (A) we masked the LGM
512 Caspian and Aral Seas (including regions of high putative suitability) from the map (Prentice et
513 al., 1993). Lake Victoria was left unmasked as it was likely very low during the LGM (Johnson et
514 al., 1996). Equal Earth projection was used.
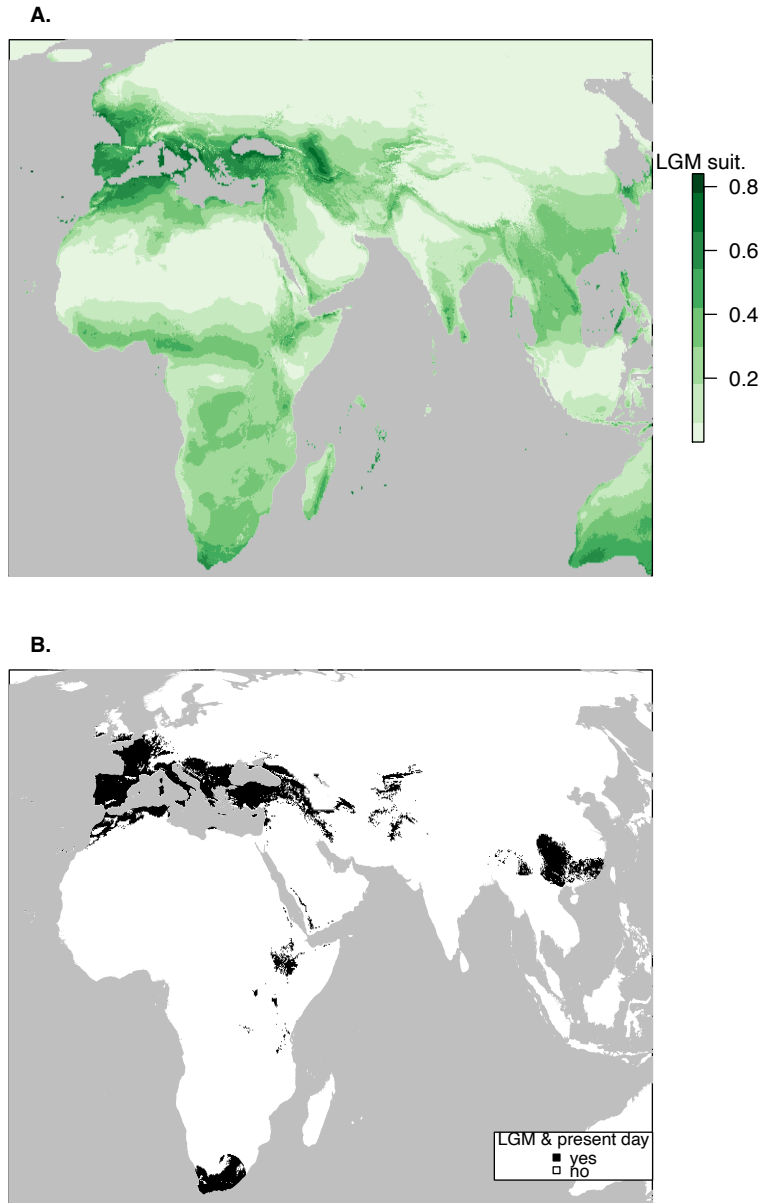
515
516
517 We projected climate suitability for Arabidopsis in the year 2050. Some current high suitability
518 regions will remain so, such as in Northern Europe. Nevertheless, we found poleward and up-
519 elevation shifts in regions of high suitability and retreats at lower latitudes and elevations. In the

520   native range, all African sites show declining suitability, as did most Mediterranean sites,
521   highlighting vulnerability of these populations. By contrast, Arctic Europe, and the mountains of
522   central Asia/Tibetan plateau show increased suitability, highlighting potential range expansions
523   (Figure 5).
524
525   **Figure 5.** Future predicted suitability for Arabidopsis (A), change in suitability from future
526   compared with present (blue indicates improving suitability and red decreasing, B), regions of
527   potential colonization (blue) continued occupancy (gray), and extinction (red) based on a
528   threshold suitability of 0.25 for occupancy (C), and the standard deviation in suitability among
529   the 5 tested climate models giving uncertainty (D). Equal Earth projection is used. RCP 4.5
530   emissions scenario is shown, see Fig S13 for highly similar patterns under RCP 6.0.
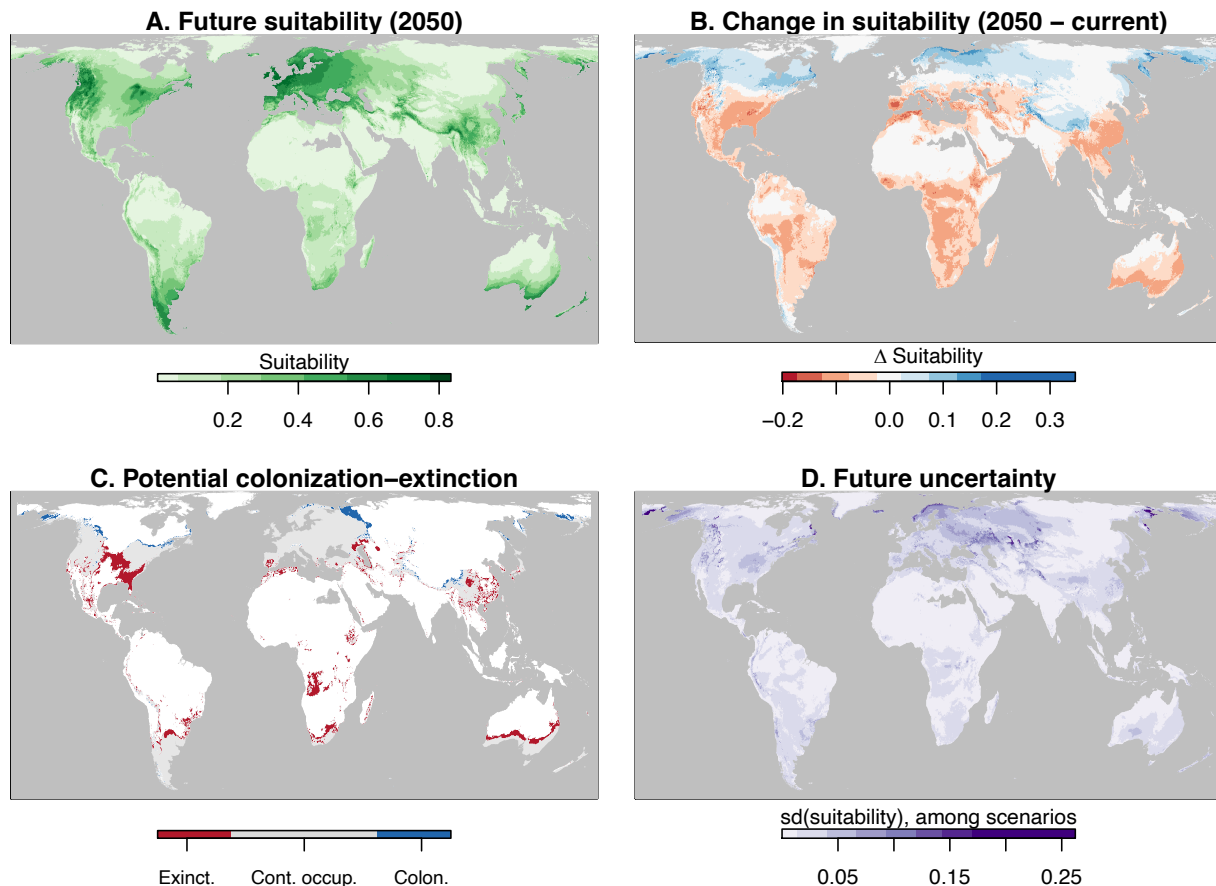
531
532
533
534
535   **Discussion**
536
537   We used the model plant Arabidopsis in a case study of integrative climate biogeography of a
538   species' past, present, and future distributions. This species is key to a large body of plant
539   biology research, and an in-depth study connecting its biogeography to genetic and phenotypic
540   variation provides important context for the biology of Arabidopsis. The size of individual plants

541   in herbarium specimens, as well as genetic variation in flowering time, suggested that regions
542   with lower predicted suitability harbored populations with reduced fitness and altered life history.
543   The consistency between model predictions and individual variation provides partial model
544   validation, bolstering our confidence in conclusions from model projections. Using these model
545   projections, we identified new glacial refugia, and looking to the future, we found that genetically
546   distinct lower latitude populations are most threatened by climate change.
547
548   *The current distribution of Arabidopsis and its limiting factors*
549   The fitted native-range model largely corresponded to occurrences and indicates Arabidopsis
550   broadly distributed across Europe, moister regions of central and eastern Asia, and mountains
551   across Africa. Our results advance beyond the most recent climate biogeographical study of
552   Arabidopsis by Hoffmann (2002), who used monthly climate data (but notably no synthetic
553   bioclimatic variables) from Leemans & Cramer (1991) mostly to qualitatively describe the
554   regions occupied by Arabidopsis. Our study included populations in sub-Saharan Africa and
555   Asia that were overlooked or considered non-native by Hoffmann (2002). We also included
556   occurrences from regions apparently missed by Hoffmann (2002). Zou et al. (2017) more
557   recently built models of Arabidopsis's distribution using Eurasian populations and default
558   MaxEnt settings, but did not include occurrences from Arabia, sub-Saharan Africa, and much of
559   the Himalayas, Russia, and central Asia, and did not subsample occurrence to reduce bias.
560   Likely as a result of these issues, the predictions of Zou et al. (2017) show a pronounced peak
561   in suitability in Germany, which was densely sampled in the genomics studies used for
562   occurrences by Zou et al. (2017), but predicted low suitability in much of the core European
563   range and near zero suitability in sub-Saharan Africa, Arabia, and most of the Russian part of
564   the species range.
565   The two dominant limiting factors in our model were winter cold (at lower latitudes and
566   elevations) and summer heat (at higher latitudes and elevations). Winter cold is recognized to
567   limit Arabidopsis performance, in particular winter cold appears to be a dominant force in local
568   adaptation of Arabidopsis (Ågren & Schemske, 2012; Gienapp et al., 2017; Monroe et al.,
569   2016). In southern Europe, where summer heat was inferred to be limiting, Arabidopsis flowers
570   in early spring (DeLeo et al., 2020) and thus summer heat is not usually directly experienced. In
571   spring in these regions, warm temperatures might not reach consistently stressful levels (e.g. to
572   induce fruit abortion) (Warner & Erwin, 2005) but it may be that moisture deficit driven by
573   evaporative demand is directly limiting in late spring. Where suitability was highest for
574   Arabidopsis, including the British Isles and a belt along the coast from France to Poland,
575   temperature annual range was identified as limiting. However, interpreting limiting factors in an
576   area of extremely high suitability (near one) is not meaningful given that suitability can scarcely
577   be increased.
578   Despite inference of winter cold and summer heat as primarily limiting, these were only
579   partly reflected by temporal fluctuations in individual plant performance from herbarium
580   specimens. Winter cold and precipitation variability anomalies were significantly associated with
581   specimen size in much of Asia, where these were also the MaxEnt modeled limiting factors from
582   Iran and Kazakhstan to Afghanistan and the Himalayas, suggesting these climate factors truly
583   limit Arabidopsis populations. However, there were discrepancies between individual
584   performance and limiting factors, likely for several reasons. First, the size of specimens is an

585   imperfect fitness proxy. Usually only reproductive individuals are collected in herbaria, and
586   individuals that would die before reproduction are excluded. Second, individuals are not
587   randomly sampled from populations (Daru et al., 2018). Third, MaxEnt models face limitations
588   due to misspecification, problems with occurrence data, or a mismatch between covariates and
589   the true ecological factors limiting populations. Nonetheless, the limiting factors we identified
590   here fit well with our knowledge of Arabidopsis ecophysiology and natural history.
591
592   *Habitat suitability, individual performance, and life history*
593   We found a decrease in plant size with decreasing habitat suitability across the range of
594   Arabidopsis, suggesting that our model suitability captured a substantial part of environmental
595   effects on individual fitness. Given the great variability in individual plant performance within
596   populations obvious to casual observers, it is unsurprising that suitability explains a minority of
597   total variation in individual size, leaving most variation unexplained (Figure 2). Furthermore,
598   fecundity and fecundity response to treatments often have low heritability in many species
599   (Price & Schluter, 1991), even in controlled experiments (Lasky et al., 2015).
600   There have still been few studies of individual performance with the geographic scope
601   that allows inference across a species range (Angert & Schemske, 2005; Csergő et al., 2017;
602   Greiser et al., 2020; Samis & Eckert, 2007). In a synthesis of studies of 40 species, Lee-Yaw et
603   al. (2016) found that, on average, individual performance and distribution-model inferred
604   suitability decline beyond range margins. From 42 studies Lee-Yaw et al. (2021) found that 38%
605   identified some predictive ability of distribution models for individual performance. However,
606   many previous studies relied on intensive observations of a small number of populations, while
607   our estimates of performance from thousands of herbarium specimens allowed us to cover most
608   of the species range. The increased availability of digitized museum specimens with trait data
609   indicates an opportunity to estimate performance across distributions for many species
610   (Bontrager & Angert, 2015).
611   We found that low suitability regions had earlier flowering genotypes. This relationship
612   was noisy, with early flowering genotypes frequent in all levels of suitability, but later flowering
613   genotypes lacking from the least suitable regions. The suitability-flowering time associations
614   were significant when accounting for genome-wide similarity between accessions, suggesting
615   they reflect selection associated with suitability. We interpret the direction of the relationship as
616   indicating that when suitability is low, Arabidopsis employs stress escape strategies, *i.e.* a rapid
617   life cycle during favorable conditions (Ludlow, 1989). Nevertheless, our findings suggest that
618   low suitability central Asian populations have some stress tolerating mechanisms, as they
619   harbor distinct allele at a transcription factor (ERF53, AT2G20880) that regulates response to
620   abiotic stressors (Cheng et al., 2012; B. Li et al., 2019). The restriction of late flowering
621   genotypes to more suitable regions is counterintuitive given physiological work showing these
622   are more stress tolerant (Lovell et al., 2013). However, it may be true that more favorable
623   conditions make possible a slow growing, slow flowering, freezing-tolerant winter annual
624   strategy.
625
626   *The distribution of Arabidopsis outside its native range*
627   Arabidopsis has spread across the globe, largely to climates well-predicted by native range
628   models, but to some regions predicted to be less suitable. Populations in western North

629   America, Australia, and New Zealand occur in climates well-predicted by the native range,
630   suggesting stable realized niches following colonization. But in eastern North America, Korea,
631   and Japan, these climates tend to be of low predicted suitability. Whether these are truly of
632   lower suitability is unknown without performance data. It may be that Arabidopsis has colonized
633   novel environments in these regions, or that the low suitability is only a model artefact. Studies
634   often find limited transferability among native and introduced range models (Early & Sax, 2014;
635   Liu et al., 2020).
636   In southern Australia, the Southern Cone, and Alaska there are widespread seemingly
637   suitable habitats where Arabidopsis is absent. The lack of more populations may partly owe to
638   the strict quarantine on imports to Australia to avoid species invasions (Pheloung, 1999).
639   Additionally, a lack of human disturbance in some regions (e.g. Alaska) may limit the potential
640   for colonization given Arabidopsis often is found in disturbed sites.
641
642   *The distribution of Arabidopsis at the last glacial maximum*
643   Using population genetic patterns, Lee et al. (2017) hypothesized five glacial refugia for Eurasia
644   to be in Iberia, Sicily, Balkans, the Levant, and Turkmenistan. These are consistent with our
645   range reconstruction, although we do not find clear barriers of unsuitable climates that would
646   have isolated Italian, Balkan, and Levant LGM populations. Furthermore, we note that
647   Sardinia/Corsica appears as a refugium, and we note that narrow, highly suitable areas along
648   the southern and eastern Caspian Sea could be the location of the hypothesized Turkmenistan
649   refugium. Lee et al. (2017) further hypothesized that the bulk of current European genotypes
650   derive from an expansion originating on the northwest coast of the Black Sea, where we found a
651   strip of highly suitable conditions for Arabidopsis during the LGM. Additional genetic studies
652   hypothesized a long history of existence in sub-Saharan Africa (Durvasula et al., 2017) and
653   conditions during the LGM suggest the species could have been much more widespread than
654   currently. Concordantly, Chala et al., (2017) used models of the distribution of Afroalpine habitat
655   generally during the LGM to find expanded areas in East Africa compared to present day,
656   though some locations such as Jebel Marra, Sudan were still surrounded by landscapes of
657   unsuitable habitat even at the LGM.
658   We found that the North Atlantic European coast was highly suitable (including areas
659   currently inhabited) during the LGM, but current British Isle and French populations do not show
660   genetic signatures of refugia (Lee et al., 2017). In both cases, local, distinct genotypes that
661   survived the LGM may have gone extinct following expansion by the now dominant European
662   genetic cluster ("non-relicts") (Fulgione & Hancock, 2018; Lee et al., 2017). The previous
663   Arabidopsis distribution modeling of the LGM (Zou et al., 2017) did not identify France and NE
664   Iberia as highly suitable during the LGM and did not mask the Caspian Sea (thus overinflating
665   that refugium). Zou et al. (2017) also did not identify the African refugia (except for a small
666   region in SE Africa) and potential refugia in mountains of SE Asia.
667
668   *The future distribution of Arabidopsis*
669   Over the next decades climate conditions are predicted to worsen for Arabidopsis across large
670   areas of its range. Central Spain and mountains in north and east Africa and Arabia may see
671   the most worsening. These climate change effects may already be emerging: our intensive field
672   search in the Dai Forest of Djibouti in 2018 failed to yield any Arabidopsis, despite the presence

19

673 of abundant (but likely more drought tolerant) annual mustards (Brassicaceae) *Erucastrum*
674 *arabicum* and *Sisymbrium erysimoides* in otherwise appropriate habitat for Arabidopsis*.*
675 Arabidopsis was collected in this juniper woodland in 1954 but recent decades have seen
676 increased drought and tree mortality (Witsen, 2012), and the potential extinction of this isolated
677 marginal Arabidopsis population. Given the unique genetic diversity of the model plant
678 Arabidopsis housed in its lower latitude populations (Durvasula et al., 2017; Hsu et al., 2019;
679 Lee et al., 2017; Zou et al., 2017) the conservation of these populations could benefit plant
680 biology research.
681   By contrast, conditions are expected to improve for the northernmost populations in
682 Europe, suggesting a potential current colonization front of Arabidopsis, in addition to higher
683 elevation locations in Tibet, the Caucuses, Ural, Alps, and Hengduan mountains adjacent to
684 currently inhabited regions. Similar currently unoccupied higher elevations are available in many
685 African mountains though these are usually small mountaintop areas and may be poorly
686 characterized by CHELSA climate data (Karger et al., 2017). Future distribution models would
687 benefit from improved environmental data for these high elevation tropical sites.
688
689 *Conclusion*
690 Species distribution models provide potentially powerful windows into past, present, and future
691 macroecology, but they are rarely confronted with individual performance data. Here we showed
692 that lower suitability habitats inferred from distribution models had smaller plants with distinct life
693 history, suggesting a stress escape strategy. While the relationships were noisy it may still be
694 remarkable that they emerge against the many microsite contributors to individual level variation
695 in a habitat generalist annual plant. Arabidopsis populations are distributed across diverse
696 climates, but genetically distinct populations in lower latitudes that are potentially valuable for
697 research are also highly threatened by anthropogenic climate change in the next few decades.
698 We believe that combining distribution models with individual data on traits and genotypes
699 across a species range can be a useful approach to validate distribution models and to dissect
700 the organismal mechanisms underlying distributions.
701

711
712

References

Ågren, J., & Schemske, D. W. (2012). Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range. *New Phytologist*, *194*(4), 1112–1122. https://doi.org/10.1111/j.1469-8137.2012.04112.x

Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., Borgwardt, K. M., Cao, J., Chae, E., Dezwaan, T. M., Ding, W., Ecker, J. R., Exposito-Alonso, M., Farlow, A., Fitz, J., Gan, X., Grimm, D. G., Hancock, A. M., Henz, S. R., Holm, S., … Zhou, X. (2016). 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell*, *166*, 481–491. https://doi.org/10.1016/j.cell.2016.05.063

Angert, A. L., & Schemske, D. W. (2005). The Evolution of Species' Distributions: Reciprocal Transplants Across the Elevation Ranges of Mimulus Cardinalis and M. Lewisii. *Evolution*, *59*(8), 1671–1684. https://doi.org/10.1111/j.0014-3820.2005.tb01817.x

Banta, J. A., Ehrenreich, I. M., Gerard, S., Chou, L., Wilczek, A., Schmitt, J., Kover, P. X., & Purugganan, M. D. (2012). Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in Arabidopsis thaliana. *Ecology Letters*, *15*(8), 769–777. https://doi.org/10.1111/j.1461-0248.2012.01796.x

Baumgartner, J., Wilson, P., & Esperón-Rodríguez, M. (2017). rmaxent: Tools for working with Maxent in R. *R Package Version 0.8*, *3*.

Bivand, R. S., Pebesma, E. J., Gómez-Rubio, V., & Pebesma, E. J. (2008). *Applied spatial data analysis with R* (Vol. 747248717). Springer.

Bontrager, M., & Angert, A. L. (2015). Effects of range-wide variation in climate and isolation on floral traits and reproductive output of Clarkia pulchella. *American Journal of Botany*, *103*, 10–21. https://doi.org/10.3732/ajb.1500091

738 Bontrager, M., Usui, T., Lee-Yaw, J. A., Anstett, D. N., Branch, H. A., Hargreaves, A. L., Muir, C.

739   D., & Angert, A. L. (2021). Adaptation across geographic ranges is consistent with strong

740   selection in marginal climates and legacies of range expansion. *Evolution*, *75*(6), 1316–

741   1333. https://doi.org/10.1111/evo.14231

742 Buckley, L. B., Waaser, S. A., MacLean, H. J., & Fox, R. (2011). Does including physiology

743   improve species distribution model predictions of responses to recent climate change?

744   *Ecology*, *92*(12), 2214–2221. https://doi.org/10.1890/11-0066.1

745 Chala, D., Zimmermann, N. E., Brochmann, C., & Bakkestuen, V. (2017). Migration corridors for

746   alpine plants among the 'sky islands' of eastern Africa: Do they, or did they exist? *Alpine*

747   *Botany*, *127*(2), 133–144. https://doi.org/10.1007/s00035-017-0184-z

748 Cheng, M.-C., Hsieh, E.-J., Chen, J.-H., Chen, H.-Y., & Lin, T.-P. (2012). Arabidopsis RGLG2,

749   Functioning as a RING E3 Ligase, Interacts with AtERF53 and Negatively Regulates the

750   Plant Drought Stress Response. *Plant Physiology*, *158*(1), 363–375.

751   https://doi.org/10.1104/pp.111.189738

752 Clark, J. S. (2010). Individuals and the Variation Needed for High Species Diversity in Forest

753   Trees. *Science*, *327*(5969), 1129–1132.

754 Clauw, P., Coppens, F., Beuf, K. D., Dhondt, S., Daele, T. V., & Maleux, K. (2015). Leaf

755   responses to mild drought stress in natural variants of Arabidopsis. *Plant Physiol.*, *167*.

756   https://doi.org/10.1104/pp.114.254284

757 Cooper, M. (2021). *moranfast: Calculate Moran's I quickly with low memory footprint for Big*

758   *Data.* [C++]. https://github.com/mcooper/moranfast

759 Csergő, A. M., Salguero-Gómez, R., Broennimann, O., Coutts, S. R., Guisan, A., Angert, A. L.,

760   Welk, E., Stott, I., Enquist, B. J., McGill, B., Svenning, J.-C., Violle, C., & Buckley, Y. M.

761   (2017). Less favourable climates constrain demographic strategies in plants. *Ecology*

762   *Letters*, *20*(8), 969–980. https://doi.org/10.1111/ele.12794

763   Daru, B. H., Park, D. S., Primack, R. B., Willis, C. G., Barrington, D. S., Whitfeld, T. J. S.,

764   Seidler, T. G., Sweeney, P. W., Foster, D. R., Ellison, A. M., & Davis, C. C. (2018).

765   Widespread sampling biases in herbaria revealed from large-scale digitization. *New*

766   *Phytologist*, *217*(2), 939–955. https://doi.org/10.1111/nph.14855

767   DeLeo, V. L., Menge, D. N. L., Hanks, E. M., Juenger, T. E., & Lasky, J. R. (2020). Effects of

768   two centuries of global environmental variation on phenology and physiology of

769   Arabidopsis thaliana. *Global Change Biology*, *26*(2), 523–538.

770   https://doi.org/10.1111/gcb.14880

771   Durvasula, A., Fulgione, A., Gutaker, R. M., Alacakaptan, S. I., Flood, P. J., Neto, C.,

772   Tsuchimatsu, T., Burbano, H. A., Picó, F. X., Alonso-Blanco, C., & Hancock, A. M.

773   (2017). African genomes illuminate the early history and transition to selfing in

774   Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 201616736.

775   https://doi.org/10.1073/pnas.1616736114

776   Early, R., & Sax, D. F. (2014). Climatic niche shifts between species' native and naturalized

777   ranges raise concern for ecological forecasts during invasions and climate change.

778   *Global Ecology and Biogeography*, *23*(12), 1356–1365.

779   https://doi.org/10.1111/geb.12208

780   Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods*

781   *in Ecology and Evolution*, *1*(4), 330–342. https://doi.org/10.1111/j.2041-

782   210X.2010.00036.x

783   Estarague, A., Vasseur, F., Sartori, K., Bastias, C., Cornet, D., Rouan, L., Beurier, G., Exposito-

784   Alonso, M., Herbette, S., Bresson, J., Vile, D., & Violle, C. (2021). *Into the range: A*

785   *latitudinal gradient or a center-margins differentiation of ecological strategies in*

786   *Arabidopsis thaliana?* (p. 2021.10.15.461205).

787   https://doi.org/10.1101/2021.10.15.461205

788   Exposito-Alonso, M., Vasseur, F., Ding, W., Wang, G., Burbano, H. A., & Weigel, D. (2018).

789   Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis

790   thaliana. *Nature Ecology & Evolution*, *2*(2), 352. https://doi.org/10.1038/s41559-017-

791   0423-0

792   Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors

793   in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38–49.

794   https://doi.org/10.1017/S0376892997000088

795   Forester, B. R., DeChaine, E. G., & Bunn, A. G. (2013). Integrating ensemble species

796   distribution modelling and statistical phylogeography to inform projections of climate

797   change impacts on species distributions. *Diversity and Distributions*, *19*(12), 1480–1495.

798   https://doi.org/10.1111/ddi.12098

799   Fournier-Level, A., Perry, E. O., Wang, J. A., Braun, P. T., Migneault, A., Cooper, M. D.,

800   Metcalf, C. J. E., & Schmitt, J. (2016). Predicting the evolutionary dynamics of seasonal

801   adaptation to novel climates in Arabidopsis thaliana. *Proceedings of the National*

802   *Academy of Sciences*, *113*(20), E2812–E2821.

803   https://doi.org/10.1073/pnas.1517456113

804   Fulgione, A., & Hancock, A. M. (2018). Archaic lineages broaden our view on the history of

805   Arabidopsis thaliana. *New Phytologist*, *219*(4), 1194–1198.

806   https://doi.org/10.1111/nph.15244

807   Gamba, D., Lorts, C., Haile, A., Sahay, S., Lopez, L., Xia, T., Kulesza, E., Elango, D., Kerby, J.,

808   Yifru, M., Bulafu, C. E., Wondimu, T., Glowacka, K., & Lasky, J. (2022). *The genomics*

809   *and physiology of abiotic stressors associated with global elevation gradients in*

810   *Arabidopsis thaliana* (p. 2022.03.22.485410). bioRxiv.

811   https://doi.org/10.1101/2022.03.22.485410

812   Gbif.Org. (2020). *Occurrence Download*. The Global Biodiversity Information Facility.

813   https://doi.org/10.15468/DL.H95GHR

814   Gienapp, P., Fior, S., Guillaume, F., Lasky, J. R., Sork, V. L., & Csilléry, K. (2017). Genomic

815   Quantitative Genetics to Study Evolution in the Wild. *Trends in Ecology & Evolution*,

816   *32*(12), 897–908. https://doi.org/10.1016/j.tree.2017.09.004

817   Greiser, C., Hylander, K., Meineri, E., Luoto, M., & Ehrlén, J. (2020). Climate limitation at the

818   cold edge: Contrasting perspectives from species distribution modelling and a transplant

819   experiment. *Ecography*, *43*(5), 637–647. https://doi.org/10.1111/ecog.04490

820   Hancock, A. M., Brachi, B., Faure, N., Horton, M. W., Jarymowycz, L. B., Sperone, F. G.,

821   Toomajian, C., Roux, F., & Bergelson, J. (2011). Adaptation to Climate Across the

822   Arabidopsis thaliana Genome. *Science*, *334*(6052), 83–86.

823   Hargreaves, A. L., Samis, K. E., Eckert, C. G., Schmitz, A. E. O. J., & Bronstein, E. J. L. (2014).

824   Are Species' Range Limits Simply Niche Limits Writ Large? A Review of Transplant

825   Experiments beyond the Range. *The American Naturalist*, *183*(2), 157–173.

826   https://doi.org/10.1086/674525

827   Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of

828   monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of*

829   *Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

830   Heberling, J. M. (2021). Herbaria as Big Data Sources of Plant Traits. *International Journal of*

831   *Plant Sciences*. https://doi.org/10.1086/717623

832   Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high

833   resolution interpolated climate surfaces for global land areas. *International Journal of*

834   *Climatology*, *25*(15), 1965–1978.

835   Hijmans, R. J., & Elith, J. (2013). Species distribution modeling with R. *R CRAN Project*.

836   Hoffmann, M. H. (2002). Biogeography of Arabidopsis thaliana (L.) Heynh. (Brassicaceae).

837   *Journal of Biogeography*, *29*(1), 125–134.

838   Hoffmann, M. H. (2005). EVOLUTION OF THE REALIZED CLIMATIC NICHE IN THE GENUS:

839   ARABIDOPSIS (BRASSICACEAE). *Evolution*, *59*(7), 1425-1436.

840 Hsu, C.-W., Lo, C.-Y., & Lee, C.-R. (2019). On the postglacial spread of human commensal

841   Arabidopsis thaliana: Journey to the East. *New Phytologist*, *222*(3), 1447–1457.

842   https://doi.org/10.1111/nph.15682

843 Johnson, T. C., Scholz, C. A., Talbot, M. R., Kelts, K., Ricketts, R. D., Ngobi, G., Beuning, K.,

844   Ssemmanda, I., & McGill, J. W. (1996). Late Pleistocene Desiccation of Lake Victoria

845   and Rapid Evolution of Cichlid Fishes. *Science*, *273*(5278), 1091–1093.

846   https://doi.org/10.1126/science.273.5278.1091

847 Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N.

848   E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's

849   land surface areas. *Scientific Data*, *4*(1), 1–20. https://doi.org/10.1038/sdata.2017.122

850 Kawakatsu, T., Huang, S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A., Castanon, R.,

851   Nery, J. R., Barragan, C., He, Y., Chen, H., Dubin, M., Lee, C.-R., Wang, C., Bemm, F.,

852   Becker, C., O'Neil, R., O'Malley, R. C., Quarless, D. X., … Ecker, J. R. (2016).

853   Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions. *Cell*,

854   *166*(2), 492–505. https://doi.org/10.1016/j.cell.2016.06.044

855 Koornneef, M., & Meinke, D. (2010). The development of Arabidopsis as a model plant. *The*

856   *Plant Journal*, *61*(6), 909–921. https://doi.org/10.1111/j.1365-313X.2009.04086.x

857 Korves, T. M., Schmid, K. J., Caicedo, A. L., Mays, C., Stinchcombe, J. R., Purugganan, M. D.,

858   & Schmitt, J. (2007). Fitness Effects Associated with the Major Flowering Time Gene

859   FRIGIDA in Arabidopsis thaliana in the Field. *The American Naturalist*, *169*(5), E141–

860   E157.

861 Lambeck, K., Rouby, H., Purcell, A., Sun, Y., & Sambridge, M. (2014). Sea level and global ice

862   volumes from the Last Glacial Maximum to the Holocene. *Proceedings of the National*

863   *Academy of Sciences*, *111*(43), 15296–15303. https://doi.org/10.1073/pnas.1411762111

864 Lasky, J. R., Des Marais, D. L., Lowry, D. B., Povolotskaya, I., McKay, J. K., Richards, J. H.,

865   Keitt, T. H., & Juenger, T. E. (2014). Natural variation in abiotic stress responsive gene

866   expression and local adaptation to climate in Arabidopsis thaliana. *Molecular Biology*

867   *and Evolution*, *31*(9), 2283–2296. https://doi.org/10.1093/molbev/msu170

868   Lasky, J. R., Forester, B. R., & Reimherr, M. (2018). Coherent synthesis of genomic

869   associations with phenotypes and home environments. *Molecular Ecology Resources*,

870   *18*(1), 91–106. https://doi.org/10.1101/051110

871   Lasky, J. R., Hooten, M. B., & Adler, P. B. (2020). What processes must we understand to

872   forecast regional-scale population dynamics? *Proceedings of the Royal Society B:*

873   *Biological Sciences*, *287*(1940), 20202219. https://doi.org/10.1098/rspb.2020.2219

874   Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., Juenger,

875   T. E., Hyma, K., Acharya, C., Mitchell, S. E., Buckler, E. S., Brenton, Z., Kresovich, S., &

876   Morris, G. P. (2015). Genome-environment associations in sorghum landraces predict

877   adaptive traits. *Science Advances*, *1*(6), e1400218.

878   https://doi.org/10.1126/sciadv.1400218

879   Lee, C.-R., Svardal, H., Farlow, A., Exposito-Alonso, M., Ding, W., Novikova, P., Alonso-Blanco,

880   C., Weigel, D., & Nordborg, M. (2017). On the post-glacial spread of human commensal

881   Arabidopsis thaliana. *Nature Communications*, *8*(1), 14458.

882   https://doi.org/10.1038/ncomms14458

883   Leemans, R., & Cramer, W. P. (1991). The CLIMATE data base of monthly normals for globally

884   gridded climatic variables (Version 2.1) [Data set]. In *Potsdam Institute for Climate*

885   *Impact Research*. PANGAEA. https://doi.org/10.1594/PANGAEA.759491

886   Lee-Yaw, J. A., Kharouba, H. M., Bontrager, M., Mahony, C., Csergő, A. M., Noreen, A. M. E.,

887   Li, Q., Schuster, R., & Angert, A. L. (2016). A synthesis of transplant experiments and

888   ecological niche models suggests that range limits are often niche limits. *Ecology*

889   *Letters*, *19*(6), 710–722. https://doi.org/10.1111/ele.12604

27

890 Lee-Yaw, J. A., McCune, J., Pironon, S., & Sheth, S. N. (2021). Species distribution models

891   rarely predict the biology of real populations. *Ecography*, *n/a*(n/a).

892   https://doi.org/10.1111/ecog.05877

893 Li, B., Gao, Z., Liu, X., Sun, D., & Tang, W. (2019). Transcriptional Profiling Reveals a Time-of-

894   Day-Specific Role of REVEILLE 4/8 in Regulating the First Wave of Heat Shock–

895   Induced Gene Expression in Arabidopsis. *The Plant Cell*, *31*(10), 2353–2369.

896   https://doi.org/10.1105/tpc.19.00519

897 Li, G., Du, S., & Guo, K. (2015). Evaluation of Limiting Climatic Factors and Simulation of a

898   Climatically Suitable Habitat for Chinese Sea Buckthorn. *PLoS ONE*, *10*(7), e0131659.

899   https://doi.org/10.1371/journal.pone.0131659

900 Liu, C., Wolter, C., Xian, W., & Jeschke, J. M. (2020). Species distribution models have limited

901   spatial transferability for invasive species. *Ecology Letters*, *23*(11), 1682–1692.

902   https://doi.org/10.1111/ele.13577

903 Lopez, L., Bellis, E. S., Wafula, E., Hearne, S. J., Honaas, L., Ralph, P. E., Timko, M. P.,

904   Unachukwu, N., dePamphilis, C. W., & Lasky, J. R. (2019). Transcriptomics of host-

905   specific interactions in natural populations of the parasitic plant purple witchweed (Striga

906   hermonthica). *Weed Science*, *67*(4), 397–411. https://doi.org/10.1017/wsc.2019.20

907 Lovell, J. T., Juenger, T. E., Michaels, S. D., Lasky, J. R., Platt, A., Richards, J. H., Yu, X.,

908   Easlon, H. M., Sen, S., & McKay, J. K. (2013). Pleiotropy of FRIGIDA enhances the

909   potential for multivariate adaptation. *Proceedings of the Royal Society B: Biological*

910   *Sciences*, *280*(1763), 20131043. https://doi.org/10.1098/rspb.2013.1043

911 Ludlow, M. M. (1989). Strategies of response to water stress. In K. H. Kreeb, H. Richter, & T. M.

912   Hinckley (Eds.), *Structural and Functional Responses to Environmental Stresses* (pp.

913   269–281). SPB Academic.

914 Mandáková, T., Thorbjörnsson, H., Pisupati, R., Reichardt, I., Lysak, M. A., & Anamthawat-

915   Jónsson, K. (2017). Icelandic accession of Arabidopsis thaliana confirmed with

28

916   cytogenetic markers and its orign inferred from whole-genome sequencing. *Icelandic*

917   *Agricultural Sciences*, *30*, 29–38. https://doi.org/10.16886/IAS.2017.03

918 Martínez-Berdeja, A., Stitzer, M. C., Taylor, M. A., Okada, M., Ezcurra, E., Runcie, D. E., &

919   Schmitt, J. (2020). Functional variants of DOG1 control seed chilling responses and

920   variation in seasonal life-history strategies in Arabidopsis thaliana. *Proceedings of the*

921   *National Academy of Sciences*, *117*(5), 2526–2534.

922   https://doi.org/10.1073/pnas.1912451117

923 Merow, C., Latimer, A. M., Wilson, A. M., McMahon, S. M., Rebelo, A. G., & Silander Jr, J. A.

924   (2014). On using integral projection models to generate demographically driven

925   predictions of species' distributions: Development and validation using sparse data.

926   *Ecography*, *37*(12), 1167–1183. https://doi.org/10.1111/ecog.00839

927 Monroe, J. G., McGovern, C., Lasky, J. R., Grogan, K., Beck, J., & McKay, J. K. (2016).

928   Adaptation to warmer climates by parallel functional evolution of CBF genes in

929   Arabidopsis thaliana. *Molecular Ecology*, *25*(15), 3632–3644.

930   https://doi.org/10.1111/mec.13711

931 Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., &

932   Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent

933   evaluations and estimating optimal model complexity for Maxent ecological niche

934   models. *Methods in Ecology and Evolution*, *5*(11), 1198–1205.

935   https://doi.org/10.1111/2041-210X.12261

936 Peterson, A., Soberón, J., Pearson, R., Anderson, R., Martínez-Meyer, E., Nakamura, M., &

937   Araújo, M. (2011). Ecological Niches and Geographic Distributions. In *Monographs in*

938   *Population Biology* (Vol. 49). https://doi.org/10.1515/9781400840670

939 Pheloung, P. (1999). Preventing the introduction of potential new weeds to Australia. *Practical*

940   *Weed Management: Protecting Agriculture and the Environment. 10th Biennial Noxious*

941   *Weeds Conference, Ballina, Australia, 20th-22nd July 1999.*, 18–24.

942 Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species

943   geographic distributions. *Ecological Modelling*, *190*(3), 231–259.

944   https://doi.org/10.1016/j.ecolmodel.2005.03.026

945 Prentice, I. C., Sykes, M. T., Lautenschlager, M., Harrison, S. P., Denissenko, O., & Bartlein, P.

946   J. (1993). Modelling Global Vegetation Patterns and Terrestrial Carbon Storage at the

947   Last Glacial Maximum. *Global Ecology and Biogeography Letters*, *3*(3), 67–76.

948   https://doi.org/10.2307/2997548

949 Price, T., & Schluter, D. (1991). On the Low Heritability of Life-History Traits. *Evolution*, *45*(4),

950   853–861. https://doi.org/10.2307/2409693

951 Rudman, S. M., Greenblum, S., Hughes, R. C., Rajpurohit, S., Kiratli, O., Lowder, D. B.,

952   Lemmon, S. G., Petrov, D. A., Chaston, J. M., & Schmidt, P. (2019). Microbiome

953   composition shapes rapid genomic adaptation of Drosophila melanogaster. *Proceedings*

954   *of the National Academy of Sciences*, *116*(40), 20025–20032.

955   https://doi.org/10.1073/pnas.1907787116

956 Samis, K. E., & Eckert, C. G. (2007). Testing the abundant center model using range-wide

957   demographic surveys of two coastal dune plants. *Ecology*, *88*(7), 1747–1758.

958   https://doi.org/10.1890/06-1153.1

959 Scholl, R. L., May, S. T., & Ware, D. H. (2000). Seed and Molecular Resources for Arabidopsis.

960   *Plant Physiology*, *124*(4), 1477–1480. https://doi.org/10.1104/pp.124.4.1477

961 Sinnwell, J. P., Therneau, T. M., & Schaid, D. J. (2014). The kinship2 R Package for Pedigree

962   Data. *Human Heredity*, *78*(2), 91–93. https://doi.org/10.1159/000363105

963 Takou, M., Wieters, B., Kopriva, S., Coupland, G., Linstädter, A., & De Meaux, J. (2019). Linking

964   genes with ecological strategies in Arabidopsis thaliana. *Journal of Experimental Botany*,

965   *70*(4), 1141–1151. https://doi.org/10.1093/jxb/ery447

966 Therneau, T. M., & Therneau, M. T. M. (2015). Package 'coxme.' *R Package Version*, *2*(5).

967   Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C.,

968   Erasmus, B. F. N., de Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B.,

969   van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Townsend

970   Peterson, A., Phillips, O. L., & Williams, S. E. (2004). Extinction risk from climate

971   change. *Nature*, *427*(6970), 145–148. https://doi.org/10.1038/nature02121

972   Toledo, B., Marcer, A., Méndez-Vigo, B., Alonso-Blanco, C., & Picó, F. X. (2020). An ecological

973   history of the relict genetic lineage of Arabidopsis thaliana. *Environmental and*

974   *Experimental Botany*, *170*, 103800. https://doi.org/10.1016/j.envexpbot.2019.103800

975   Tozer, B., Sandwell, D. T., Smith, W. H. F., Olson, C., Beale, J. R., & Wessel, P. (2019). Global

976   Bathymetry and Topography at 15 Arc Sec: SRTM15+. *Earth and Space Science*, *6*(10),

977   1847–1864. https://doi.org/10.1029/2019EA000658

978   Turner, K. G., Fréville, H., & Rieseberg, L. H. (2015). Adaptive plasticity and niche expansion in

979   an invasive thistle. *Ecology and Evolution*, *5*(15), 3183–3197.

980   https://doi.org/10.1002/ece3.1599

981   Vasseur, F., Sartori, K., Baron, E., Fort, F., Kazakou, E., Segrestin, J., Garnier, E., Vile, D., &

982   Violle, C. (2018). Climate as a driver of adaptive variations in ecological strategies in

983   Arabidopsis thaliana. *Annals of Botany*, *122*(6), 935–945.

984   https://doi.org/10.1093/aob/mcy165

985   Warner, R. M., & Erwin, J. E. (2005). Naturally occurring variation in high temperature induced

986   floral bud abortion across Arabidopsis thaliana accessions. *Plant, Cell & Environment*,

987   *28*(10), 1255–1266. https://doi.org/10.1111/j.1365-3040.2005.01361.x

988   Whitlock, M. C., & Guillaume, F. (2009). Testing for Spatially Divergent Selection: Comparing

989   QST to FST. *Genetics*, *183*(3), 1055–1063. https://doi.org/10.1534/genetics.108.099812

990   Wilczek, A. M., Roe, J. L., Knapp, M. C., Cooper, M. D., Lopez-Gallego, C., Martin, L. J., Muir,

991   C. D., Sim, S., Walker, A., Anderson, J., Egan, J. F., Moyers, B. T., Petipas, R.,

31

992   Giakountis, A., Charbit, E., Coupland, G., Welch, S. M., & Schmitt, J. (2009). Effects of

993   Genetic Perturbation on Seasonal Life History Plasticity. *Science*, *323*(5916), 930–934.

994   Witsen, D. (2012). *Djibouti Biodiversity and Tropical Forests: 118/119 Assessment* (p. 56).

995   United States Agency for International Development.

996   https://usaidgems.org/Documents/FAA&Regs/FAA118119/Djibouti2012.pdf

997   Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood

998   estimation of semiparametric generalized linear models. *Journal of the Royal Statistical*

999   *Society: Series B (Statistical Methodology)*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-

1000   9868.2010.00749.x

1001   Zeng, L., Gu, Z., Xu, M., Zhao, N., Zhu, W., Yonezawa, T., Liu, T., Qiong, L., Tersing, T., Xu, L.,

1002   Zhang, Y., Xu, R., Sun, N., Huang, Y., Lei, J., Zhang, L., Xie, F., Zhang, F., Gu, H., …

1003   Zhong, Y. (2017). Discovery of A high-altitude ecotype and ancient lineage of

1004   Arabidopsis thaliana from Tibet. *Science Bulletin*, *62*(24), 1628–1630.

1005   https://doi.org/10.1016/j.scib.2017.10.007

1006   Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association

1007   studies. *Nat Genet*, *44*(7), 821–824. https://doi.org/10.1038/ng.2310

1008   Zou, Y.-P., Hou, X.-H., Wu, Q., Chen, J.-F., Li, Z.-W., Han, T.-S., Niu, X.-M., Yang, L., Xu, Y.-C.,

1009   Zhang, J., Zhang, F.-M., Tan, D., Tian, Z., Gu, H., & Guo, Y.-L. (2017). Adaptation of

1010   Arabidopsis thaliana to the Yangtze River basin. *Genome Biology*, *18*(1), 239.

1011   https://doi.org/10.1186/s13059-017-1378-9

1012