36-year study reveals stability of a wild wheat population across microhabitats

Tal Dahan-Meir[1], Thomas James Ellis[2], Fabrizio Mafessoni[1], Hanan Sela[3,4], Ori Rudich[1], Jacob Manisterski[4], Naomi Avivi-Ragolsky[1], Amir Raz[1,5], Moshe Feldman[1], Yehoshua Anikster[4], Magnus Nordborg[2*], Avraham A. Levy[1*]

[1]Department of Plant and Environmental Sciences, Weizmann Institute of Science; Rehovot, Israel.

[2]Gregor Mendel Institute, Austrian Academy of Sciences, Vienna BioCenter; Vienna, Austria.

[3]Institute of Evolution, University of Haifa; Haifa, Israel.

[4]The Institute for Cereal Crops Improvement, Tel-Aviv University; Tel Aviv, Israel.

[5]Migal, Galilee Technology Center; Kiryat Shmona, Israel.

*Corresponding authors: Avraham A. Levy, Magnus Nordborg.

Emails: avi.levy@weizmann.ac.il; magnus.nordborg@gmi.oeaw.ac.at

Abstract

Long-term genetic studies of wild populations are very scarce, but are essential for connecting ecological and population genetics models, and for understanding the dynamics of biodiversity. We present a study of a wild wheat population sampled over a 36-year period at high spatial resolution. We genotyped 832 individuals from regular sampling along transects during the course of the experiment. Genotypes were clustered into ecological microhabitats over scales of tens of metres, and this clustering was remarkably stable over the 36 generations of the study. Simulations show that it is difficult to explain this spatial and temporal stability using only limited dispersal, suggesting a role for fine-scale local adaptation to ecological parameters. Using a common-garden experiment, we showed that the genotypes found in distinct microhabitats differ phenotypically, further supporting the hypothesis of local adaptation. Our results provide a rare insight into the population genetics of a natural population over a long monitoring period.

Introduction

The spatial and temporal scale of population genetic dynamics are key to understanding questions of local adaptation and the response to selection. For example, how does the scale of genetic variation correspond to that of environmental variation, and how quickly can populations respond to natural selection? To address these questions, we need to describe the spatial structure of genetic variation, and track how this changes over many generations.

Numerous studies have examined the spatio-temporal dynamics of wild plant populations over multiple years (Horovitz & Feldman 1991; Volis *et al.* 2004; Franks *et al.* 2007; Ozbek *et al.* 2007; Hendrick *et al.* 2016; Monnahan & Kelly 2017; Frachon *et al.* 2017; Troth *et al.* 2018; Fu *et al.* 2019; Todesco *et al.* 2020; Kolis *et al.* 2022). A few studies have examined the temporal scale of allele frequency change over decades, especially in long-lived vertebrates (Aguillon *et al.* 2017; Chen *et al.* 2019; Ashraf *et al.* 2021; Stoffel *et al.* 2021; Kelly 2022). However, there are essentially no genetic studies examining both spatial and temporal dynamics of wild plants' populations over decades.

We examined the genetic diversity of a population of wild emmer wheat (*Triticum turgidum ssp. dicoccoides*) in Ammiad, Israel, which has been monitored for 36 years (Figure 1A). Wild emmer wheat is an annual, self-pollinating, allotetraploid cereal (genome BBAA). It is the progenitor of durum wheat (*T. turgidum ssp. durum*) and the maternal donor of the A and B sub-genomes of bread wheat (*T. aestivum*, genome BBAADD) (Levy & Feldman 2022). Its biodiversity is an invaluable resource for food

56 security, due to the fact that major effect loci from wild emmer wheat can be
57 transferred to the background of domesticated wheat. The Ammiad population
58 shows high genetic diversity compared to other wild wheat populations (Levy &
59 Feldman 1988; Felsenburg *et al.* 1991), together with variable topography (Figure 1,
60 Movie S1) and a diversity of ecological niches within the scale of tens of metres
61 (Anikster & Noy-Meir 1991; Noy-Meir *et al.* 1991b). The site is within a protected
62 nature reserve, allowing repeated sampling of the population along transects that
63 encompass considerable environmental variation. In this study we genotyped
64 samples from nine harvests of these transects between 1984 and 2020, which
65 allowed us to characterise the spatial and temporal dynamics of genetic change at
66 Ammiad. Previous works suggesting that plants in Ammiad are adapted to
67 microhabitats (Anikster *et al.* 1991; Noy-Meir *et al.* 1991b) were later disputed by Volis
68 (Volis *et al.* 2015, 2016). Our dataset provides an almost unique insight into genetic
69 change in a natural population of a predominantly selfing plant over time and space.
70 This is especially relevant, because wild progenitors of modern crops such as wild
71 emmer wheat serve as an immense reservoir of biotic and abiotic stress-resistance
72 genes (Khoury *et al.* 2021) that could be utilised to improve commercial varieties.
73 Sustained studies of progenitor species in their natural ecological niches can provide
74 a better understanding of their survival in different ecological niches, and inform
75 strategies for conservation of such resources *in-situ*, as well as in gene banks (Wintle
76 *et al.* 2019).
77 Here, we describe patterns of genetic variation at Ammiad in time and space. We
78 then investigate whether an apparent association with microhabitat might be
79 biologically meaningful by comparing the observed pattern of variation to an explicit

model accounting for the limited seed and pollen dispersal of a predominantly self-fertilising plant. Finally, we discuss the implications of these results for local adaptation and conservation.

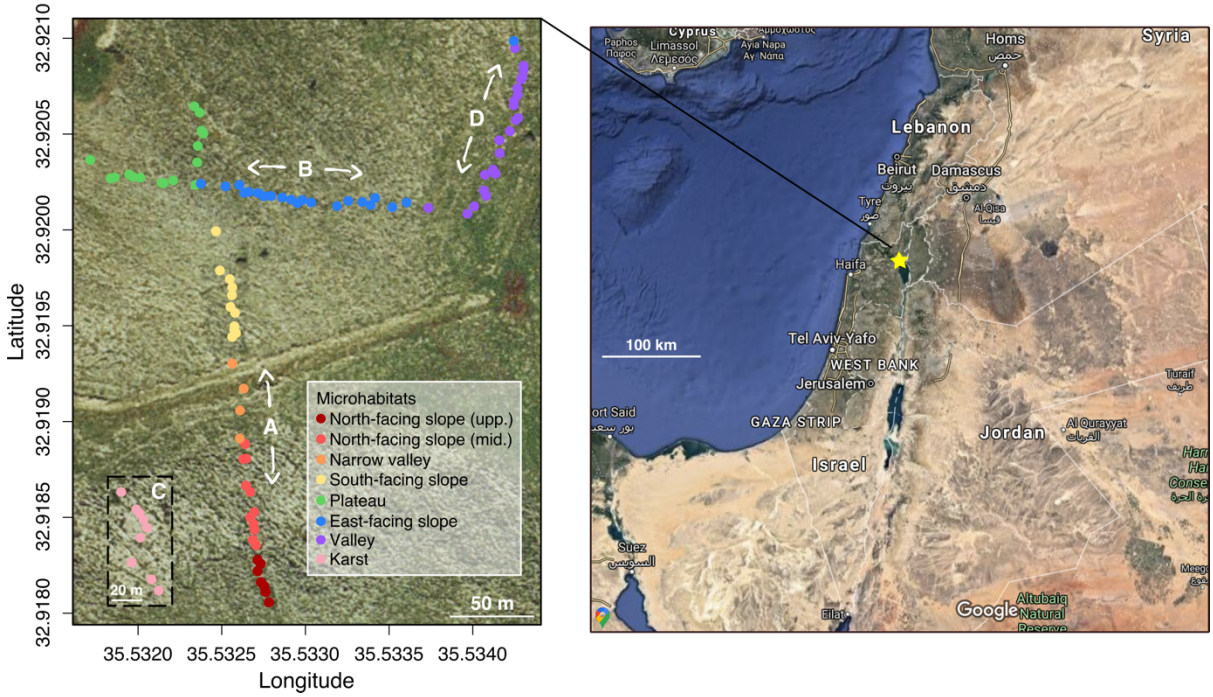


Figure 1. Ammiad wild-wheat population study. Location of Ammiad near the lake of Galilee (star, right and geographic coordinates, left). Sampling points of transects A, B, C, D collected in the years 1984, 1988, 1992, 1996, 2002, 2014, 2016, 2018 and 2020. Sampling points are shown on-site for transects A, B, D, but transect C, shown in the dashed rectangle, is located ~500 metres north-west of A, B, D. Colours correspond to the different microhabitats labelled in the legend. The map on the left is taken from gov.map.co.il, and the map on the right is taken from Google Maps.

Results

The population is composed of groups of near-identical genotypes

We catalogued the genetic diversity at Ammiad by repeated sampling over 36 years. We marked 100 sampling locations with pegs and recorded GPS coordinates along four transects traversing seven ecologically distinct microhabitats (Anikster & Noy-Meir 1991); Figure 1, Dataset S1, Movie S1). We sampled seeds from the closest plant to each sampling location at each of nine time points, and sowed them in the greenhouse for genotyping. In total, we genotyped 832 individuals using a genotyping-by-sequencing (GBS) protocol (Poland *et al.* 2012), resulting in 5196 single copy, single-nucleotide-polymorphism (SNP) markers that were unambiguously positioned on the wild emmer wheat reference genome (Avni *et al.* 2017).

As expected for a highly selfing species (The 1001 Genomes Consortium 2016), the population is almost entirely composed of genetically distinct groups of nearly identical, highly homozygous individuals, clustered in space. Pairwise genetic identity follows a multimodal distribution that is characterised by pairs of plants that are either effectively identical, or share 70-90% of SNP alleles (Figure 2A). We unambiguously assigned all individuals to distinct genotype groups (DGGs) using a cut-off of >98% allele-sharing. This resulted in 118 DGGs over all nine time points, including 61 singletons (DGGs with a single member). The 10 largest DGGs encompass 65% of the samples, while the number of DGGs per year varied between 32 in 2014 to 49 in 1988 (Dataset S1). Rarefaction curves indicate that the sampling scheme was sufficient to detect almost all of the DGGs present close to the transects (Figure S1).

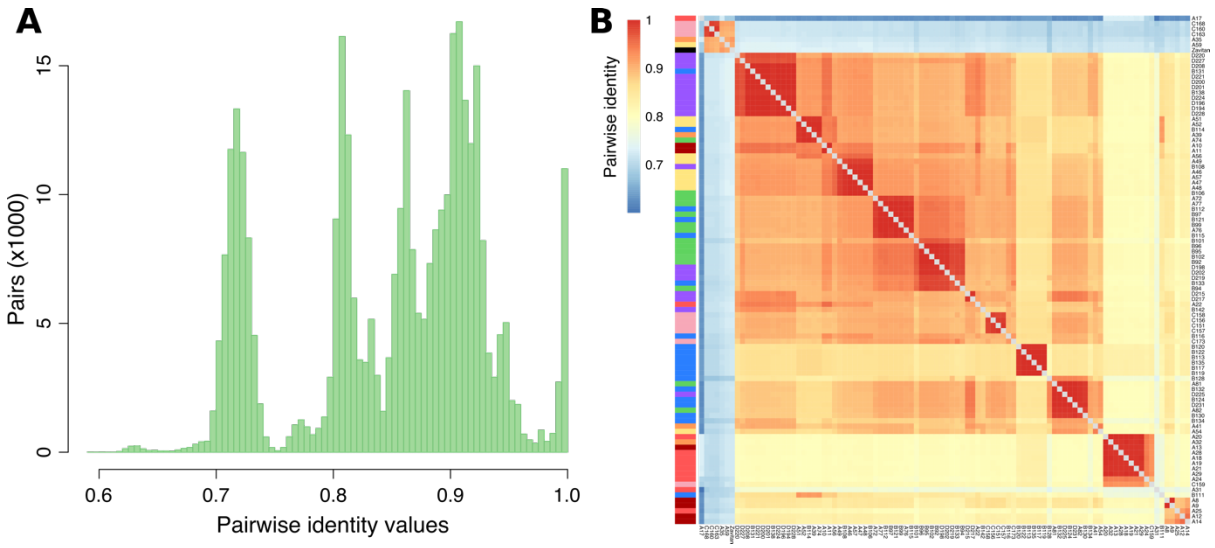


Figure 2. Ammiad population structure (A) Distribution of pairwise identity values between the plants in the collection. (B) Population structure of Ammiad in the first year of collection (1984). Heatmap colours represent identity value for each pairwise comparison, annotation colours (on the left) correspond to the list of microhabitats in Figure 1.

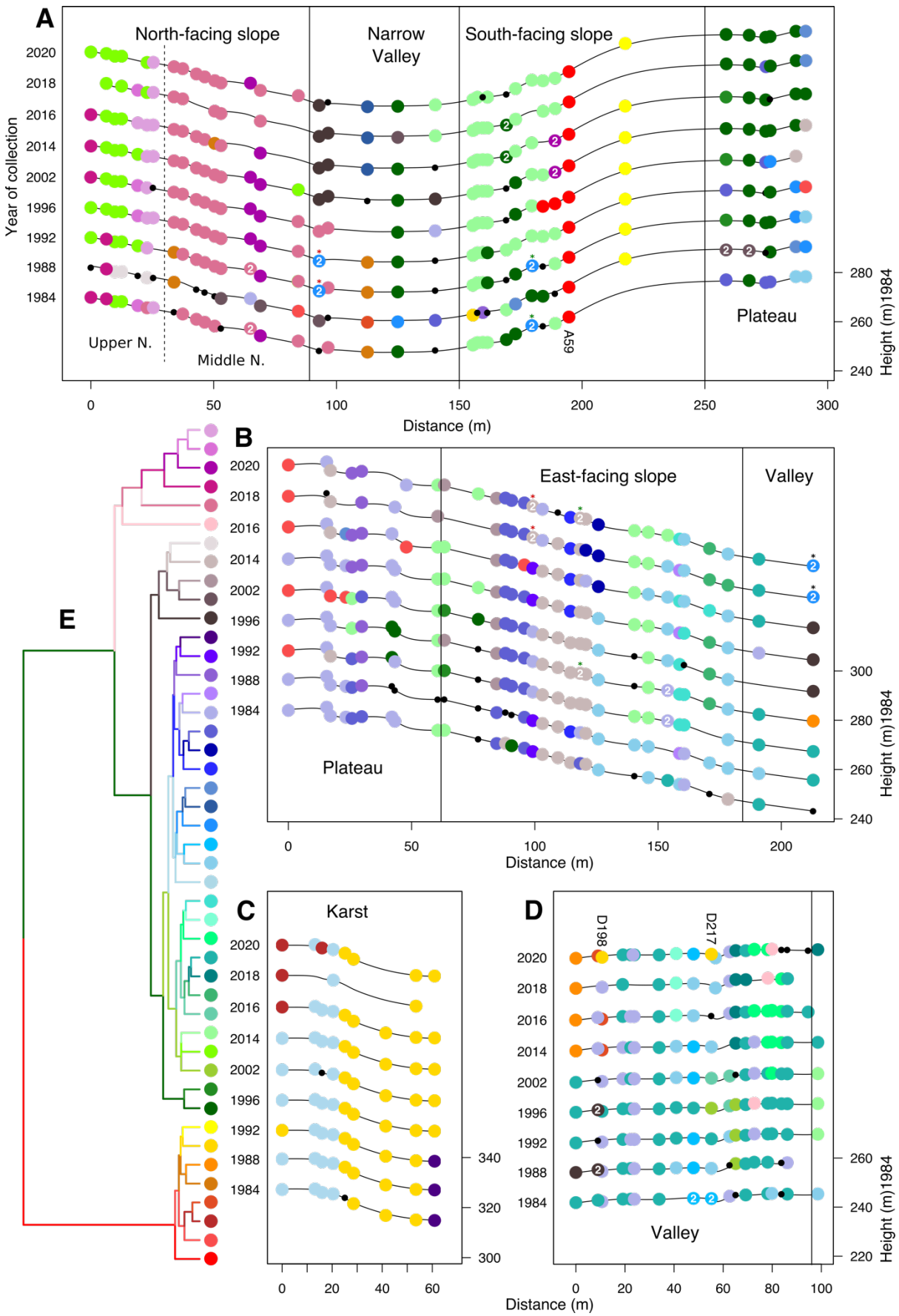
We also identified 15 (1.8%) highly heterozygous plants, which is consistent with outcrossing in recent ancestors. Using SNPMatch (Pisupati *et al.* 2017), we were able to identify their parental DGGs (Figure S2B), suggesting that outcrossing happens between plants that are nearby. We estimated an average outcrossing rate of 3.7% across years in Ammiad, with detectable variation between the different habitats (1.8-8.8%, Table S1). These values are compatible with estimates in bread wheat (*Triticum aestivum*) (Enjalbert & David 2000).

Genotypes are clustered by microhabitats and are stable through time

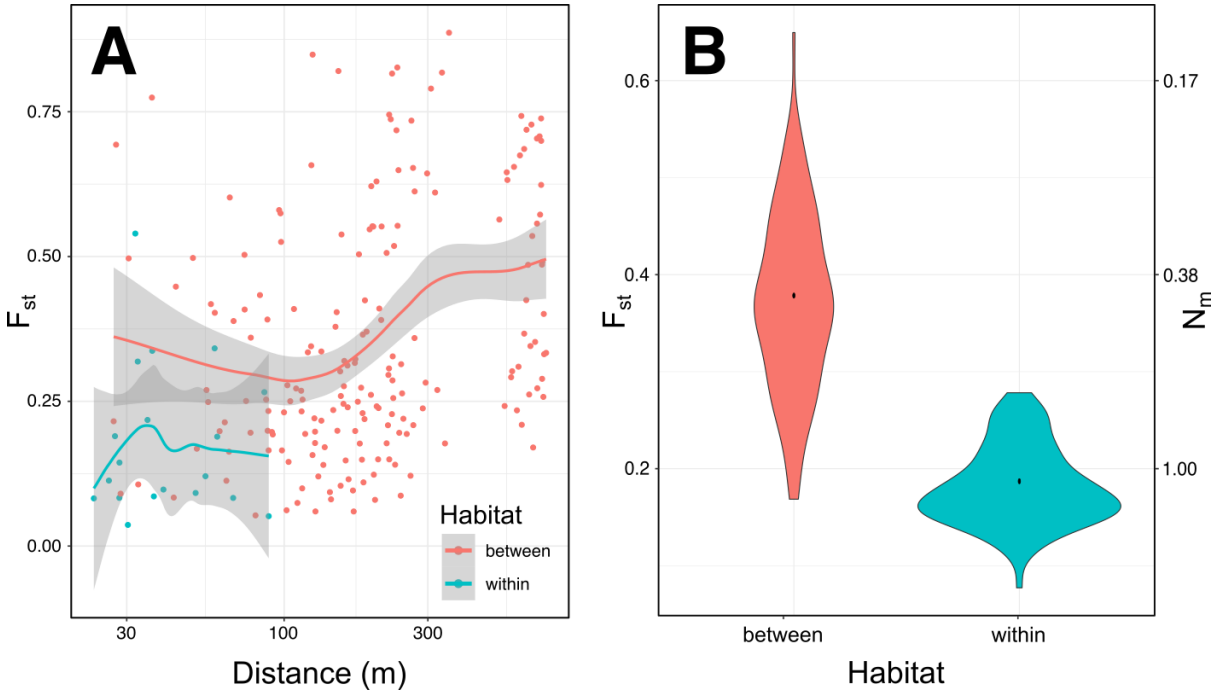
The spatio-temporal distribution of DGGs indicates two clear patterns (Figure 3). First, DDGs tend to cluster in space, which is expected for a self-compatible organism with limited dispersal. However, the clustering (Figure 3) and a principal component analysis of molecular variation (Figures S3A and S3B) show a striking concordance

135 with the boundaries of ecological microhabitats, which were defined based on
136 topographic and floristic data without knowledge of the genetic structure of the
137 population (Noy-Meir *et al.* 1991b), Figure 3). For example, in transect A, there is a
138 sharp transition between DGGs between the upper and middle north-facing slope
139 across the 36 years of study (Figure 3A). Another example is the genotype at
140 collection point A59 in a unique location under the shade of a *Ziziphus spina-christi*
141 tree (Figure 3A). In contrast, in transect D, where there is negligible microhabitat
142 differentiation, there is little spatial clustering of DGGs (Figure 3D). We found little
143 evidence for a decay of relatedness that would be expected under a model of simple
144 isolation by distance (Figure S3C and S4). Population differentiation measured by F_{st}
145 was consistently higher between microhabitats than within, even when taking
146 distance into account (Figures 4A, B), and genetic distances are strongly correlated
147 with microhabitat. We also found greater differentiation among DGGs between
148 microhabitats than would be expected by chance with the same level of spatial
149 autocorrelation using a test based on Moran-spectral-randomization Mantel statistics
150 (MSR-Mantel) that explicitly account for spatial distance (Figure S5). In summary,
151 DGGs are associated with ecological microhabitats at the scale of tens of metres.
152 Second, the spatial clustering of DGGs was stable through time. For example, 33%
153 of plants sampled in 1984 belonged to the same DGG as plants sampled at the same
154 position in 2020. An analysis of molecular variance (AMOVA) showed that 28% of the
155 genetic variation between samples is explained by differences between
156 microhabitats, and that the habitat effect is highly significant (Table S2A, B and C).
157 Conversely, time did not explain a significant proportion of variance (variance
158 explained<1%), suggesting a strong stability of the genetic structure over time. These

159 results suggest that space, and particularly microhabitats, might constrain the
160 genetic structure of Ammiad to be stable over the scale of tens of generations.


A
North-facing slope
Narrow Valley
South-facing slope
Year of collection
Upper N.
Middle N.
A59
Plateau
Height (m)1984
Distance (m)
B
East-facing slope
Valley
Plateau
C
Karst
D
D198
D217
Valley
E

163 **Figure 3. Spatio-temporal distribution of genotypes.** (A-D) Distribution of samples
164 coloured by distinct genotype group (DGG) in transects A, B, C and D. Y-axes indicate year
165 of collection (left) and altitude for 1984 (right). Singleton samples are marked in small black
166 dots, while doubleton groups are marked with "2" in the middle with the colour of their closest
167 DGG. Doubleton groups that are closest to the same DGG were marked with coloured
168 asterisks. (E) Dendrogram of DGGs with ≥ 3 samples, hierarchically clustered by UPGMA.



169
170 **Figure 4. Spatial clustering at Ammiad. (A)** Genetic differentiation between and within
171 habitats. Plants within each habitat were grouped in 30m demes. Pairwise F_{st} was calculated
172 among all demes, within and between habitats and plotted as a function of distance,
173 measured as the average distance of plants in the two demes. Credibility intervals of a loess
174 fit with degree 1 are shown in grey. Nm values (i.e., the theoretical number of migrants
175 between parcels in an infinite island model) calculated are shown on the right axis. (B) Plants
176 within each habitat were grouped in 30m demes. Pairwise F_{st} was calculated among all
177 demes with average distance <50m between each other, within and between habitats.
178 Average observed values are shown as a black dot. Uncertainty was estimated with 1000
179 bootstraps of the parcels considered for the comparisons. Average within-habitat Fst was
180 higher than between-habitat F_{st} in 3% of the bootstraps. Nm values (i.e., the theoretical
181 number of migrants between parcels in an infinite island model) calculated are shown on the
182 right axis.

183 **Neutral processes are unlikely to explain spatio-temporal stability**

184 Our findings indicate that the genetic structure at Ammiad is more clustered into
185 environmental microhabitats than would be expected under a null hypothesis of no
186 clustering. However, it is clear that such a strict null hypothesis is not a good
187 reference because it ignores the underlying biology of pollen and seed dispersal. To
188 generate a more biologically realistic alternative hypothesis we tested whether the
189 patterns at Ammiad can be explained using explicit demographic models that
190 incorporate self-fertilisation and limited dispersal. The question is whether these
191 neutral demographic processes - on their own - could generate the clustering into
192 microhabitats observed at Ammiad.

193

194 We performed individual-based simulations of populations of plants that initially
195 reflect the genetic structure at Ammiad, and subsequently evolve by seed dispersal
196 and outcrossing only (see *Materials and Methods*). We simulated populations in a
197 two-dimensional continuous space using the genetic structure at Ammiad in 1984
198 along one of those dimensions. We also recorded the microhabitat structure of those
199 genotypes but did not model selection for those microhabitats. At each of 40
200 subsequent generations seeds dispersed, with some proportion of seeds being
201 drawn from a two-year seed bank. In addition, some proportion of those seeds were
202 assigned to be outcrossed and given a new genotype. At each generation we
203 sampled a transect through the population and sample plants closest to 30 evenly
204 spaced sampling points five metres apart. In this way we aimed to simulate a
205 population evolving in the absence of selection and to sample it in a way that mimics
206 the sampling of the study site at Ammiad.

207 We ran simulations using different values of outcrossing rate, plant density, seed
208 dormancy and mean seed dispersal distance that approximate a range of biologically
209 realistic values. Empirical values of outcrossing rates in wild wheat have been
210 estimated to be as low as 0.5% (Golenberg 1988), although our estimates point to
211 outcrossing rates of between 2% and 8% (mean = 3.7%; table S1). This species has
212 a substantial seed bank with up to 30% of germinating seedlings coming from the
213 previous year, but very few from previous years (Horovitz 1998). Average plant
214 density in the Ammiad population was measured in 1984 as 4 plants/m^2 (Noy-Meir *et*
215 *al.* 1991a). We repeated this measurement in 2020 for transect A and found a similar
216 density of 3.8 plants/m^2 (Dataset S1). Seed dispersal occurs through a combination
217 of gravity, ants, birds, grazing mammals and rodents, enabling both short and long-
218 distance dispersal. We lack direct estimates of seed dispersal distances, but average
219 gene flow, a composite of seed and pollen dispersal, was estimated to be 1.25 m per
220 year (Golenberg 1987); since most dispersal is via seeds, the average seed dispersal
221 distance is likely to be around 1 m per year. We aimed to simulate values above and
222 below these estimates to allow for a range of biologically plausible values.
223
224 We compared the change in genetic differentiation among microhabitats in observed
225 and simulated populations by tracking the change in F_{st} through time (Figure 5). At
226 Ammiad F_{st} fluctuated between 0.08 in 1988 and 0.26 in 1996, but remained stable
227 at around 0.15 on average throughout the period of study. When outcrossing rates in
228 simulations were at their lowest value (0.5% per generation) F_{st} was similarly stable
229 across time, especially when dispersal is also very limited (Figure 5, top row).
230 However, for all other parameter combinations we observed a steady decrease in F_{st}

in simulated populations over the 40 generations. The rate of decay was markedly faster as outcrossing rates increased, and depended surprisingly on dispersal distance, seed dormancy or plant density. This indicates that in the absence of selection, limited outcrossing of a few percent combined should be sufficient to erase any apparent sorting into microhabitats within the period of study examined here. Unless the outcrossing rate is much lower than we estimated, we cannot explain the observed microhabitat differentiation via neutral demographic processes.

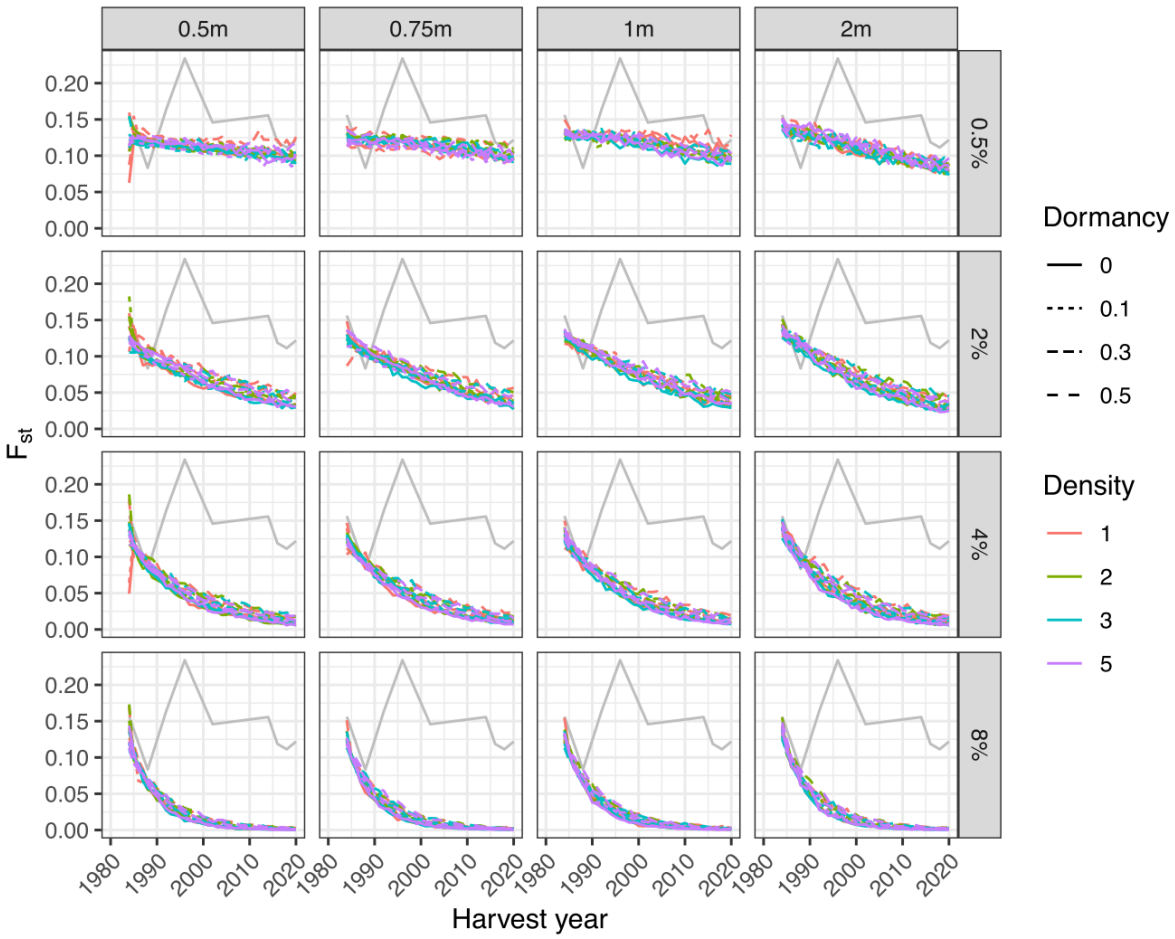


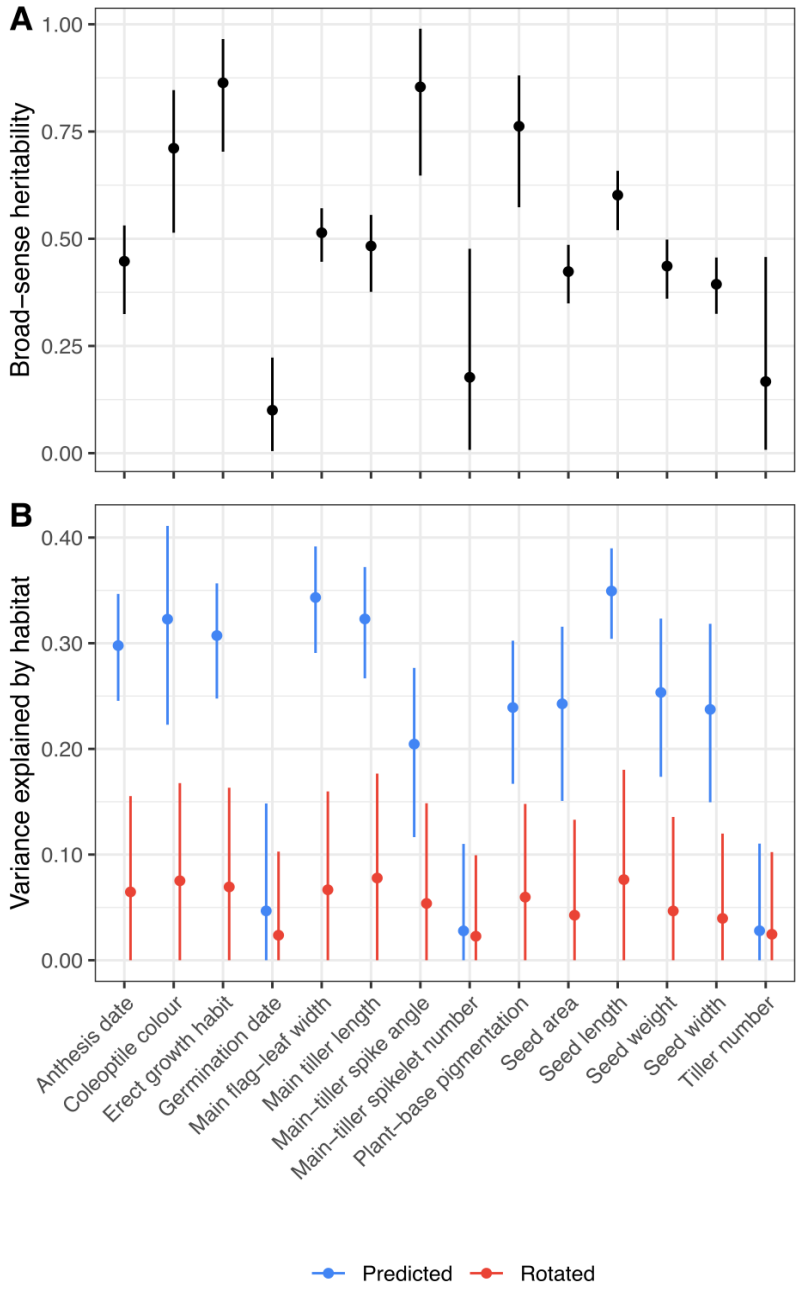
Figure 5. Genetic differentiation between microhabitats at Ammiad and in simulated populations. Grey lines show F_{st} between microhabitats across all transects at Ammiad. Coloured lines show average F_{st} over 100 replicate simulations that begin with the genetic structure at Ammiad in 1984 but evolve neutrally after that. Curves indicate different combinations of mean dispersal distance (grey boxes, panels left to right), outcrossing rate (grey boxes, panels top to bottom), proportion of seeds coming from the seed bank (dashed lines) and plant density per square metre (colours).

Phenotypes measured in a common garden predict phenotypic differentiation between habitats

If there is a functional link between population structure and microhabitat differentiation, there is a prediction that habitats should be associated with differences in plant phenotypes. Owing to the challenging terrain and the status of

252 the site as a protected conservation area it is not possible to estimate phenotypes
253 with replication in a common-garden experiment under field conditions. We instead
254 estimated phenotypes in a controlled net house experiment and used these data to
255 predict phenotypic differentiation in the field.
256
257 To characterise the phenotypic variation between DGGs we grew six replicates of 95
258 DGGs in fully randomised blocks in a net house at the Weizmann Institute in
259 conditions at Ammiad as closely as possible. We measured 14 phenotypes related to
260 growth and reproduction (Figure S6) and estimated genetic values for each DGG
261 using Bayesian generalised-linear models, accounting for differences between
262 blocks. With the exceptions of germination date, number of spikes and number of
263 tillers, broad-sense heritabilities (on the scale of the link functions; (de Villemereuil *et*
264 *al.* 2016) were moderate to high (Figure 6A), indicating that substantial genetic
265 variation exists for these traits.
266
267 We then used phenotype estimates from the net house to predict phenotypic
268 differentiation in the field. We assigned the 832 field samples a phenotype value for
269 each trait from each of 1000 posterior draws of the corresponding linear models to
270 generate 1000 sets of predicted phenotypes across time and space. A linear
271 discriminant analysis that accounts for all traits jointly revealed clear clustering by
272 microhabitat, with seed traits explaining the most separation (Figure S7; Table S3).
273 For each trait separately, we then calculated the variance in each trait explained by
274 habitat and year. Differences in habitat explained a very small proportion of the
275 variation in the three traits that showed low heritability (tiller number, spikelet number

276 and germination time; Figure 6B). For other traits, differences between habitats
277 explained 19-33% of the phenotypic variation for all traits (based on posterior means;
278 Figure 6B), whereas variation between years explained less than 10% of the overall
279 phenotypic variance in the population. There thus seems to be differentiation in
280 heritable traits between habitats which remains stable through time, consistent with
281 the spatial and temporal patterns observed among DGGs.
282 We next asked whether spatial structure in phenotypes corresponds to the structure
283 of microhabitats. We compared the predicted phenotypes to a 'rotated' dataset
284 where the vector of phenotypes is offset by a random number of positions relative to
285 the vector of habitat labels. This preserves the structure of phenotypes and habitat
286 labels, but breaks any covariation between phenotype and habitat. In rotated
287 datasets habitat typically explained less than 20% of the variation in phenotype. This
288 is similar to that explained by the three predicted phenotypes with low heritability, but
289 less than that explained by the eleven remaining phenotypes for more than 99% of
290 predicted datasets (Figure 6B). Nevertheless, our results indicate that the genetic
291 clustering by microhabitat is matched by phenotypic differentiation between
292 microhabitats.


A
Broad−sense heritability
1.00
0.75
0.50
0.25
0.00
B
Variance explained by habitat
0.40
0.30
0.20
0.10
0.00
Anthesis date
Coleoptile colour
Erect growth habit
Germination date
Main flag−leaf width
Main tiller length
Main−tiller spike angle
Main−tiller spikelet number
Plant−base pigmentation
Seed area
Seed length
Seed weight
Seed width
Tiller number
Predicted
Rotated

293

294 **Figure 6. Heritability and variance explained by habitat for in predicted phenotypes at**
295 **Ammiad.** (A) Posterior mean and 95% credible intervals for broad-sense heritability of 14
296 phenotypes measured on six replicates each of 95 homozygous DGGs in a common-garden
297 net house experiment. Because phenotypes display a range of error distributions, heritability
298 is calculated on the scale of the corresponding link functions for comparison. (B) Plots show
299 the posterior mean and 95% credible intervals for 801 field samples predicted (blue) from
300 phenotypes measured in the net house experiment. Variance components are shown for the
301 raw predictions based on real microhabitat positions, and for 'rotated' data (red) where
302 microhabitat labels were offset by a random number of positions to break any correlation
303 between phenotypes and microhabitats.

304

305 ## Discussion

306 Studies of natural populations rarely have fine-scale geographic resolution or
307 temporal resolution beyond the duration of a PhD or postdoc (Siepielski *et al.* 2009).
308 Our study, spanning over 36 years of dense sampling, has both, and thus provides a
309 unique dataset that could improve our understanding of gene flow and local
310 adaptation on an ecological scale.
311 Our analysis revealed that the Ammiad population varies more in space than in time.
312 We found unexpectedly strong long-term genetic stability, demonstrating that gene
313 flow can effectively be highly limited (unless it is via pollen dispersal) in a wild cereal
314 despite its wide distribution throughout the fertile crescent. Remarkably, the stable
315 genetic clusters observed in Ammiad matched ecological microhabitats previously
316 identified at the site, suggesting that both natural selection and limited dispersal have
317 contributed to shaping population structure. Previous studies of wild populations
318 have demonstrated local adaptation (Schemske 1984; Volis *et al.* 2015, 2016; Abebe

319 *et al.* 2015; Sedlacek *et al.* 2016; Hendrick *et al.* 2016; Frachon *et al.* 2017; Gardiner
320 *et al.* 2018) over large distances, but not at this spatial scale nor using long-term
321 datasets. We also showed that the differences in genotypic compositions between
322 habitats correspond to heritable differences in phenotypes, and that some of these
323 phenotypes consistently differ among habitats. This further supports the hypothesis
324 of local adaptation at the micro-environment level. Although we tried to replicate
325 environmental conditions as closely as possibly, it is important to keep in mind that
326 phenotypes were measured in a common garden rather than field conditions, and as
327 such we cannot capture the precise environmental conditions at Ammiad, nor
328 account for ways in which the phenotypes of individuals may vary when grown in
329 different across microhabitats.
330 Our simulations showed that while the highly stable and structured population at
331 Ammiad could be explained by very low outcrossing and seed migration, it is
332 incompatible with the realistic parameters we estimated. This, together with the
333 correlation between phenotypes-genotypes and microhabitats over a long
334 experimental period, suggest that the population structure reflects adaptation to
335 microhabitats, at least in part. The possibility that local adaptation occurs at such a
336 fine spatial scale as in our study is intriguing and invites further experimental
337 confirmation.
338
339 Although climate change was not yet an established fact in 1984, it became part of
340 the motivation of Ammiad experiment. In retrospect, the increase in local ambient
341 temperatures by ~1.5 degrees Celsius (Figure S8A) and in atmospheric CO_2
342 concentration by 70 ppm (Figure S8B) (Keeling *et al.* 2001) during the course of the

experiment, could have led to selection that affected the population structure. However, only minor temporal changes were seen. This genetic temporal stability could be interpreted as the population not having had enough time to adapt due to the rapid pace of climatic changes, in which case we would expect a reduction in fitness of the population. However, plant density was similar in 1984 and in 2020, suggesting that, so far, this has not occurred. Wild emmer wheat might have evolved to be resilient to temperature fluctuations since its speciation 0.8 MYA ago (Levy & Feldman 2022) and that the recent climate changes might not have yet reached the threshold that endangers the species.

Our results are relevant to conservation strategies as well as to the assessment of biodiversity resilience under climate change. First, our study shows that *in-situ* conservation can be quite effective in preserving diversity, without the need to frequently sample seeds and store them in seed banks (*ex-situ*). Second, our results highlight that effective *in-situ* and *ex-situ* conservation strategies of wild plants biodiversity, particularly for wild grasses and crop progenitors such as wild emmer wheat, should consider habitat diversity as a potential reservoir of adaptive diversity. In fact, our long-term study would not have been possible if the site had not been declared a natural reserve for wild-wheat conservation by the Israel nature and parks authority. Indeed, this population is fenced and undergoes moderate grazing. Unrestricted grazing, or urbanisation, as seen in other parts of the Levant can reduce and sometimes eliminate wild grasses. The Ammiad reserve is relatively small (~10 hectares) but contains much diversity. Establishing natural reserves of crop

366 progenitors, partially protected from anthropogenic activities, does not necessarily
367 require large areas and is a feasible goal that is critical for biodiversity conservation.
368
369 In conclusion, we provide a unique example of strong temporal stability, a
370 counterpoint to recent studies demonstrating dramatic changes in local populations
371 even between seasons (Bergland *et al.* 2014; Wittmann *et al.* 2017; Machado *et al.*
372 2021; Rudman *et al.* 2022). Fully explaining this stability will require further work, but,
373 whatever the causes, such robustness in a rapidly changing world demonstrates the
374 importance of conservation of wild relatives of crops in nature, as gene pools for
375 breeding (Zsögön *et al.* 2021).
376

377 ## Materials and Methods

378 ### Sample collection and DNA extraction

379 Wild wheat spikes were collected in four different linear transects every 3 to 5 metres
380 from sampling points marked with pegs and GPS coordinates (Dataset S1). Plants
381 were assigned to a sampling point when collected in successive years in a radius of
382 ~ 1m from the peg. A single seed was selected for each plant and DNA was extracted
383 from young seedlings. Tissue was collected and ground using mortar and pestle with
384 the addition of Merck sea-sand in liquid nitrogen. DNA extraction was done using
385 Macherey-Nagel NucleoSpin® Plant II extraction kit, according to the manufacturer's
386 protocol, eluting in 50 ul EB total. Concentrations and DNA quality were estimated
387 using a NanoDrop® ND-1000 Spectrophotometer.

388 **Genotyping by sequencing (GBS) and variant calling**

389 Plant genotyping was done according to the Genotyping-By-Sequencing method
390 (Elshire *et al.* 2011; Poland *et al.* 2012), using *MspI* and *PstI-HF* restriction enzyme,
391 and 96 barcodes of plate 384A (Poland *et al.* 2012). Samples were sequenced using
392 Illumina NextSeq 550 mid-output 150 base-pairs single-end kits. A reference sample,
393 Zavitan accession (Avni *et al.* 2017), was added to each run, undergoing the same
394 plate GBS library preparation and sequencing. Samples were divided by barcodes
395 using a python 2.7 code. Mapping to the WEW_v2.0 genome assembly (Zhu *et al.*
396 2019) was done using *bwa-mem (Li 2013)*. Conversion to binary alignment map (BAM)
397 format, filtering for uniquely mapped reads with minimum QUAL score of 30, sorting
398 and indexing was done using SAMtools (Li *et al.* 2009). Variant calling was done in
399 parallel running on all 845 (832 Ammiad samples and 13 Zavitan controls) samples
400 using GATK HaplotypeCaller 3.8 (Poplin *et al.* 2018). Filtering for quality (>30), read
401 depth (>6), genotype quality (>15), maximum missing sites (10%) and no-indels was
402 done using VCFtools (Danecek *et al.* 2011). Filtering for up to 10% heterozygous
403 SNPs per plant was done using TASSEL5 (Bradbury *et al.* 2007). Filtered VCF was
404 used for all further analyses (see Data and code availability).

405 **Pairwise genetic identity**

406 We measured genetic identity using R (see Data and code availability) by calculating
407 the pairwise identity values between all samples, divided by the number of
408 comparisons. The pipeline also divides the samples into distinct genotype groups
409 (DGGs) by setting a threshold of 98.1% identity of the variant sites, which is the lowest
410 threshold where no sample is assigned to two or more DGGs, and constructs the
411 dendrogram describing the relatedness between the DGGs. We performed

412 hierarchical clustering R by UPGMA (method "average" in *hclust* function). All
413 samples of the reference accession, Zavitan, were grouped together into one DGG,
414 indicating that we can reliably group identical genotypes. The heatmap of samples
415 from 1984 was done using *pheatmap* (https://cran.r-
416 project.org/web/packages/pheatmap/index.html) with default hierarchical clustering.

417 **Temperature and CO_2 concentration data**

418 We obtained data on daily minimum and maximum temperatures from 1984 to 2020
419 from the Israeli Meteorological Service at the Israel government portal
420 (https://ims.data.gov.il/he/ims/2). We obtained data on weekly atmospheric CO_2
421 concentrations from The Scripps Institution of Oceanography UC San Diego (Keeling
422 et al., 2001).

423 **Rarefaction, principal-component and identity=by-distance analyses**

424 We performed a rarefaction analysis using R (see *Data and code availability*) by taking
425 successively small subsamples of the DGGs present in each year, and fitting a
426 polynomial curve to the number of unique DGGs sampled across subsamples. We
427 repeated this 100 times for each year.

428

429 We performed principal component analysis on the VCF file of the entire Ammiad
430 collection using *SNPRelate* R Package (Zheng *et al.* 2012). We calculated identity by
431 distance for each transect separately. We calculated pairwise Euclidean distances
432 using the *distHaversine* function from the *Geosphere* R package (Hijmans 2022). We
433 performed Mantel tests with 1000 permutations to calculate the correlation between

434 genetic identity and Euclidean distance matrices, using *mantel.test* in the *Ape* R
435 package (Paradis *et al.* 2004).
436
437 **Outcrossing estimation and parental lines detection**
438 To estimate outcrossing we calculated F_{IS} using the R package *hierfstat* (Goudet
439 2005). We calculated confidence intervals as the 2.5% and 97.5% quantiles of 100
440 bootstraps using the function boot.ppfis. We estimated outcrossing rates (*f*) using the
441 relationship $F_{IS} = (1-f)/(1+f)$ *(Wright 1969)*. To detect the parental DGGs of
442 heterozygous plants we used SNPMatch (Pisupati *et al.* 2017) by comparing each
443 plant to a database of all DGGs.
444
445 **Microhabitat clustering analysis**
446 We tested whether population differentiation, measured by F_{st}, is higher when
447 measured between plants occurring in different microhabitats than in the same
448 microhabitat. To this goal we subdivided plants into demes with a maximum distance
449 between plants of 30m. We then computed the pairwise F_{st} between demes using the
450 R package *pegas (Paradis 2010)*. To correct for the effects of isolation by distance
451 and spatial structure we stratified comparisons by distance, measured as the
452 distance between the average locations of plants in the two demes.
453
454 To account for spatial autocorrelation we first performed a partial Mantel test using
455 the function *multi.mantel* from the R package *phytools (Revell 2012)*. Specifically, we
456 considered as a response variable the pairwise genetic distances used to define
457 DGGs. We tested for the effect of habitat, represented as a distance matrix with value

458 “0” for plants found in the same microhabitat and “1” for plants found in two different
459 microhabitats. In addition, we included in the test a matrices of spatial and temporal
460 distance We estimated spatial distances based on latitude and longitude using the
461 function *earth.dist* from the R package *fossil* (https://palaeo-
462 electronica.org/2011_1/238/index.html). We detected a significant effect of space
463 and microhabitat (p-values <0.005 for both matrices) and a non-significant effect of
464 time (p-value = 0.863).
465
466 Mantel and partial Mantel tests have been severely criticised for their inaccuracy in
467 presence of high spatial autocorrelation (Legendre *et al.* 2015; Crabot *et al.* 2019). To
468 account for these known biases, we used the approach proposed by (Crabot *et al.*
469 2019), which uses a Moran spectral randomization procedure to generate random
470 samples which maintain the same autocorrelation properties of the original dataset.
471 To this goal, we used the R packages *sdpep* to process the spatial matrices and the
472 *msr* function of the package *adespatial (Guénard & Legendre 2022)*. The distances
473 above were transformed into Euclidean distances with function *cailliez* from the R
474 package *adegenet (Jombart & Ahmed 2011)*. Using this approach, we tested the
475 effect of habitat on both DGGs and pairwise genetic distances, using 1000
476 randomizations. In both cases the effect of microhabitat was significant (p-
477 value<0.05).
478 We computed an analysis of molecular variance using the R package *pegas (Paradis*
479 *2010)*, using plants sampled across all years. We included terms for microhabitat and
480 year of collection (nested within microhabitat) in the model.

481 **Individual-based simulations**

482 To investigate whether neutral demographic processes could maintain the degree of
483 clustering observed at Ammiad in 1984 over the period of study we performed
484 individual-based simulations of populations of plants evolving via seed dispersal and
485 outcrossing only. An R package, *simmiad*, is provided to perform these simulations
486 (DOI:10.5281/zenodo.8116105).

487

488 We simulate populations of *N* plants evolving in a uniform two-dimensional habitat
489 beginning with a structure that mimics that found at Ammiad with densities of 1, 2, 3
490 or 5 plants/m^2. One challenge is that the real population is sampled along a one-
491 dimensional transect, but simulations need to be seeded in two dimensions. To do
492 this we first concatenate the vector of genotypes from all transects in 1984, arrange
493 them along sampling points five metres apart along one axis of the habitat, then
494 copied this structure in the perpendicular direction to create bands of identical
495 genotypes in a grid. So as not to bias the measurement of the structure to any single
496 region of the populations we 'rotated' the concatenated transect by a random number
497 of positions in each simulation replicate. Sampling points are also assigned the same
498 microhabitat label as they had at Ammiad, but we do not model selection for
499 microhabitat. To ensure the population is at the correct density we then draw a
500 sample with replacement of size *N* of plants on the grid to generate a population at
501 generation 1.

502

503 At 40 subsequent generations, *N* plants are sampled with replication to produce *N*
504 seeds. Wild emmer wheat has a two-year seed bank, so from generation 3 onwards,

505 seeds are drawn from the previous generation with probability 1-*p*, and from two
506 generations previous with probability *p,* for values of *p* of 0.5, 0.3, 0.1 or 0. Seeds are
507 randomly assigned as outcrossed with probabilities 0.25%, 0.5%, 1%, or 2%. Since
508 we have no information about seed dispersal distance and are only concerned with
509 tracking unique genotypes, outcrossed seeds are assigned a new unique genotype,
510 just as outcrossed seeds in the real sample would appear as a unique DGG. Offspring
511 disperse from the mother in a random direction at a distance drawn from an
512 exponential distribution with a mean of 0.5, 0.75, 1 or 2m. To eliminate edge effects,
513 the habitat is modelled as a torus, such that if dispersal takes a seed over the edge
514 of the habitat it is reflected back to the opposite side. At each generation we then
515 draw a transect of 30 sampling points 5 metres apart through the middle of the
516 population, and sampled the plant closest to each sampling point, replicating the
517 sampling procedure in the field.
518
519 We measure microhabitat differentiation on each sampled transect as
520 $$F_{st} = (f_0 - \underline{f})/(1 - \underline{f})$$
521 where f_0 is the is the probability that two individuals in the same subpopulation are
522 the same DGG, and $\underline{f}$ is the probability that two individuals from the total population
523 are the same DGG. For each combination of plant densities, outcrossing rates, seed
524 bank parameter *p*, and seed dispersal distances we performed 100 replicate
525 simulations.

526 **Common-garden experiment**

527 We established a common garden experiment in a net-house in the Weizmann
528 Institute of Science. We grew plants from 95 distinct homozygous DGGs from DGGs

529 represented by at least two samples, as well as the reference genotype 'Zavitan', in
530 six fully randomised blocks in a net house. In order to keep conditions as similar as
531 possible to the conditions in Ammiad site, we sowed seeds in mid-November in Terra
532 Rossa soil collected at Ammiad. Plants were subject to natural winter rainfall, and we
533 provided additional irrigation on two occasions when plants were close to
534 desiccation.
535
536 We measured 4 phenotypes related to vegetative growth: (germination date;
537 presence/absence of coleoptile anthocyanin pigmentation; erect growth habit,
538 scored from 1 to 5; plant-base pigmentation during tillering, scored from 1 to 3), six
539 phenotypes related to anthesis (anthesis date; main tiller length (cm) from the soil to
540 the base of the spike; log main-tiller flag-leaf width (cm); main tiller number of
541 spikelets; main tiller spike angle, scored from 1 to 3; number of tillers) After seed
542 collection we measured four additional phenotypes based on 10 dry seeds per plant:
543 seed area (mm^2), seed length (mm), seed width (mm) and the weight of 10 seeds
544 (grams) using VIBE QM3 seed analyser. Raw phenotypic data are provided in Dataset
545 S2.
546
547 We estimated the posterior distribution of genetic values for each trait using a
548 generalised linear mixed model in the R package *brms (Bürkner 2017)*. We model the
549 phenotype z_{ijk} with link function *f(z)* (with inverse function $f`(z)$) of a plant *i* of genotype
550 *j* in block *k* as
551 $$z_{ijk} = f'(\eta_{ijk})$$
552 where

553 $$\eta_{ijk} = \alpha + \gamma_j + \beta_k + \epsilon_i$$

554 α is an intercept, γ_j is the mean phenotype of DGG *j*, β_{km} is the mean phenotype for
555 experimental block *k*, ϵ_i is a residual error term (for phenotypes with normally-
556 distributed errors), and:

557 $$\alpha \sim t_3(\mu, \sigma_\alpha)$$

558 $$\gamma \sim t_3(0, \sigma_{DGG})$$

559 $$\beta \sim t_3(0, \sigma_{block})$$

560 We used the default priors in *brms* for the mean and variance components. The 14
561 traits measured were a mixture of normally-distributed, Poisson, binary and ordinal
562 variables, so we used identity, log, logit, and cumulative-logit link functions
563 respectively, as appropriate. We standardised normally-distributed traits by mean-
564 centering and dividing by the standard deviation. We calculated broad-sense
565 heritability on the scale of the link function following (de Villemereuil *et al.* 2016) as

566 $$\frac{\sigma_{DGG}}{\sigma_{DGG} + \sigma_{block} + \sigma_\epsilon}$$

567 Note that for models with non-identity link functions σ_ϵ is only defined on the scale of
568 the data, and does not contribute to heritability estimates on the scale of the link
569 function (McCullagh & Nelder 1989; de Villemereuil *et al.* 2016).

570

571 We used the linear models to predict phenotypes in the field. We drew 1000 values
572 from the posterior distribution of each phenotype for each of the 832 samples in the
573 field data set (excluding 117 empty sampling plots) using the *brms* function
574 'posterior_predict'. This function simulates phenotype values based on the genetic
575 values for each DGG plus inter-individual error, accounting for uncertainty in both.
576 For the 26 samples that were not included in the net-house experiment, we used

577 *brms* to draw a phenotype value from the posterior distribution of genetic values of
578 phenotyped DGGs. Since these values vary across posterior draws this allows us to
579 assign phenotypes to the missing genotypes without biassing the distribution of
580 phenotypes, and renders tests of association with microhabitat conservative.
581
582 We performed a linear discriminant analysis of the predicted phenotypes in the field
583 as a function of microhabitat using the *MASS* library in R (Venables & Ripley 2002).
584 For this analysis we used posterior mean phenotypes for each DGG, using the 806
585 plants from DGGs that were included in the net-house experiment.
586
587 We then estimated the variance in each phenotype explained by habitat and year in
588 each predicted dataset using *lme4* (Bates *et al.* 2015). To generate a null distribution
589 for the association between phenotype and habitats we used a 'rotation' approach
590 similar to that used to test sorting by microhabitat. We concatenated transects within
591 each year into a single circular transect and rotated the vector of predicted
592 phenotypes by a random number of positions. This preserves spatial structure in
593 phenotypes and microhabitats but breaks any association between them. We
594 performed a single rotation for each of 1000 realisations of phenotype predictions.
595 For each realisation we recalculated the variance in rotated phenotype explained by
596 microhabitat and year, and compared the distribution of these variance estimates to
597 those for the corresponding predicted dataset.

Acknowledgements

We thank T. Azulay, T. Grinevich, C. Melamed-Bessudo, O. Gross, R. Epstein, Z. Meir and members from the Institute for Cereal Crops Improvement at Tel-Aviv University for plant material collection. We thank A. Distelfeld and members of the Institute of Evolution, University of Haifa, for using their VIBE QR3 seed analyzer, and J. Müller for the drone movie. We also thank N. Barkai, Y. Eshed, D. Filiault, K. Swarts, S. Nowoshilow, N. De-Malach, R. Pisupati, R. Burns, Y. Voichek, P. Clauw, E. Kosman and the Levy and Nordborg labs for fruitful discussions.

606 ## References

607 Abebe TD, Naz AA, Léon J (2015) Landscape genomics reveal signatures of local
608 adaptation in barley (Hordeum vulgare L.). *Frontiers in plant science*, **6**, 813.
609 Aguillon SM, Fitzpatrick JW, Bowman R *et al.* (2017) Deconstructing isolation-by-distance:
610 The genomic consequences of limited dispersal. *PLoS genetics*, **13**, e1006911.
611 Anikster Y, Amram E, Smadar E, Auguste H (1991) Patterns of phenotypic variation in wild
612 tetraploid wheat at Ammiad. *Israel Journal of Botany*, **40**, 397–418.
613 Anikster Y, Noy-Meir I (1991) The wild-wheat field laboratory at Ammiad. *Israel Journal of*
614 *Botany*, **40**, 351–362.
615 Ashraf B, Hunter DC, Bérénos C *et al.* (2021) Genomic prediction in the wild: A case study
616 in Soay sheep. *Molecular ecology*.
617 Avni R, Nave M, Barad O *et al.* (2017) Wild emmer genome architecture and diversity
618 elucidate wheat evolution and domestication. *Science* , **357**, 93–97.
619 Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting Linear Mixed-Effects Models
620 Using lme4. *Journal of statistical software*, **67**, 1–48.
621 Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA (2014) Genomic evidence
622 of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. *PLoS*
623 *genetics*, **10**, e1004775.
624 Bradbury PJ, Zhang Z, Kroon DE *et al.* (2007) TASSEL: software for association mapping of
625 complex traits in diverse samples. *Bioinformatics* , **23**, 2633–2635.
626 Bürkner PC (2017) brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal*
627 *of statistical software*, **80**, 1–28.
628 Chen N, Juric I, Cosgrove EJ *et al.* (2019) Allele frequency dynamics in a pedigreed natural
629 population. *Proceedings of the National Academy of Sciences of the United States of*
630 *America*, **116**, 2158–2164.
631 Crabot J, Clappe S, Dray S, Datry T (2019) Testing the Mantel statistic with a spatially-

632 constrained permutation procedure. *Methods in ecology and evolution / British*
633 *Ecological Society*, **10**, 532–540.
634 Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools.
635 *Bioinformatics* , **27**, 2156–2158.
636 Elshire RJ, Glaubitz JC, Sun Q *et al.* (2011) A Robust, Simple Genotyping-by-Sequencing
637 (GBS) Approach for High Diversity Species (L Orban, Ed,). *PloS one*, **6**, e19379.
638 Enjalbert J, David JL (2000) Inferring recent outcrossing rates using multilocus individual
639 heterozygosity: application to evolving wheat populations. *Genetics*, **156**, 1973–1982.
640 Felsenburg T, Levy AA, Galili G, Feldman M (1991) Polymorphism of high-molecular-weight
641 glutenins in wild tetraploid wheat: spatial and temporal variation in a native site. *Israel*
642 *Journal of Botany*, **40**, 451–479.
643 Frachon L, Libourel C, Villoutreix R *et al.* (2017) Intermediate degrees of synergistic
644 pleiotropy drive adaptive evolution in ecological time. *Nature Ecology & Evolution 2017*
645 *1:10*, **1**, 1551–1561.
646 Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in
647 response to a climate fluctuation. *Proceedings of the National Academy of Sciences*,
648 **104**, 1278–1282.
649 Fu Y-B, Peterson GW, Horbach C *et al.* (2019) Elevated mutation and selection in wild
650 emmer wheat in response to 28 years of global warming. *Proceedings of the National*
651 *Academy of Sciences*, **116**, 20002–20008.
652 Gardiner LJ, Joynson R, Omony J *et al.* (2018) Hidden variation in polyploid wheat drives
653 local adaptation. *Genome research*, **28**, 1319–1332.
654 Golenberg EM (1987) Estimation of gene flow and genetic neighborhood size by indirect
655 methods in a selfing annual, Triticum dicoccoides. *Evolution; international journal of*
656 *organic evolution*, **41**, 1326–1334.
657 Golenberg EM (1988) Outcrossing rates and their relationship to phenology in Triticum

658 dicoccoides. *TAG. Theoretical and applied genetics. Theoretische und angewandte*
659 *Genetik*, **75**, 937–944.
660 Goudet J (2005) hierfstat, a package for r to compute and test hierarchical F-statistics.
661 *Molecular ecology notes*, **5**, 184–186.
662 Guénard G, Legendre P (2022) Hierarchical Clustering with Contiguity Constraint in R.
663 *Journal of statistical software*, **103**, 1–26.
664 Hendrick MF, Finseth FR, Mathiasson ME *et al.* (2016) The genetics of extreme
665 microgeographic adaptation: an integrated approach identifies a major gene underlying
666 leaf trichome divergence in Yellowstone Mimulus guttatus. *Molecular ecology*, **25**,
667 5647–5662.
668 Horovitz A (1998) The soil seed bank in wild emmer. In: *The Proceedings of International*
669 *Symposium of In Situ Conservation of Plant Genetic Diversity*, pp. 185–188.
670 Horovitz A, Feldman M (1991) Evaluation of the wild-wheat study at Ammiad. *Israel Journal*
671 *of Botany*, **40**, 501–508.
672 Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP
673 data. *Bioinformatics* , **27**, 3070–3071.
674 Keeling CD, Piper SC, Bacastow RB *et al.* (2001) Exchanges of Atmospheric CO2 and
675 13CO2 with the Terrestrial Biosphere and Oceans from 1978 to 2000. I. Global
676 Aspects. *I. Global Aspects, SIO Reference Series*, **Scripps In**, 88.
677 Kelly JK (2022) The genomic scale of fluctuating selection in a natural plant population.
678 *Evolution letters*, **6**, 506–521.
679 Khoury CK, Carver D, Greene SL *et al.* (2021) Crop wild relatives of the United States
680 require urgent conservation action. *Proceedings of the National Academy of Sciences*
681 *of the United States of America*, **117**, 33351–33357.
682 Kolis KM, Berg CS, Nelson TC, Fishman L (2022) Population genomic consequences of life-
683 history and mating system adaptation to a geothermal soil mosaic in yellow

monkeyflowers. *Evolution; international journal of organic evolution*, **76**, 765–781.

Legendre P, Fortin MJ, Borcard D (2015) Should the Mantel test be used in spatial analysis? *Methods in ecology and evolution / British Ecological Society*, **6**, 1239–1247.

Levy AA, Feldman M (1988) Ecogeographical distribution of HMW glutenin alleles in populations of the wild tetraploid wheat Triticum turgidum var. dicoccoides. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, **75**, 651–658.

Levy AA, Feldman M (2022) Evolution and origin of bread wheat. *The Plant cell*.

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* , **25**, 2078–2079.

Machado HE, Bergland AO, Taylor R *et al.* (2021) Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in Drosophila. *eLife*, **10**.

McCullagh P, Nelder J (1989) *Generalized Linear Models (2nd ed.)* (FL Boca Raton, Ed,). Chapman and Hall/CRC.

Monnahan PJ, Kelly JK (2017) The Genomic Architecture of Flowering Time Varies Across Space and Time in Mimulus guttatus. *Genetics*, **206**, 1621–1635.

Noy-Meir I, Agami M, Anikster Y (1991a) Changes in the population density of wild emmer wheat (Triticum turgidum var. dicoccoides) in a mediterranean grassland. *Israel Journal of Botany*, **40**, 385–395.

Noy-Meir I, Agami M, Cohen E, Anikster Y (1991b) Floristic and ecological differentiation of habitats within a wild wheat population at Ammiad. *Israel Journal of Botany*, **40**, 363–384.

Ozbek O, Millet E, Anikster Y, Arslan O, Feldman M (2007) Spatio-temporal genetic variation in populations of wild emmer wheat, Triticum turgidum ssp. dicoccoides, as

710 revealed by AFLP analysis. *TAG. Theoretical and applied genetics. Theoretische und*
711 *angewandte Genetik*, **115**, 19–26.
712 Paradis E (2010) pegas: an R package for population genetics with an integrated-modular
713 approach. *Bioinformatics* , **26**, 419–420.
714 Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R
715 language. *Bioinformatics* , **20**, 289–290.
716 Pisupati R, Reichardt I, Seren Ü *et al.* (2017) Verification of Arabidopsis stock collections
717 using SNPmatch, a tool for genotyping high-plexed samples. *Scientific Data*, **4**,
718 170184.
719 Poland JA, Brown PJ, Sorrells ME, Jannink J-L (2012) Development of High-Density
720 Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-
721 Sequencing Approach (T Yin, Ed,). *PloS one*, **7**, e32253.
722 Poplin R, Ruano-Rubio V, DePristo MA *et al.* (2018) Scaling accurate genetic variant
723 discovery to tens of thousands of samples. *bioRxiv*, 201178.
724 Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other
725 things). *Methods in ecology and evolution / British Ecological Society*, **3**, 217–223.
726 Rudman SM, Greenblum SI, Rajpurohit S *et al.* (2022) Direct observation of adaptive
727 tracking on ecological time scales in Drosophila. *Science*, **375**, eabj7484.
728 Schemske DW (1984) Population structure and local selection in Impatiens pallida
729 (balsaminaceae), a selfing annual. *Evolution; international journal of organic evolution*,
730 **38**, 817–832.
731 Sedlacek J, Cortés AJ, Wheeler J *et al.* (2016) Evolutionary potential in the Alpine: trait
732 heritabilities and performance variation of the dwarf willow Salix herbacea from
733 different elevations and microhabitats. *Ecology and evolution*, **6**, 3940–3952.
734 Siepielski AM, Dibattista JD, Carlson SM (2009) It's about time: The temporal dynamics of
735 phenotypic selection in the wild. *Ecology letters*, **12**, 1261–1276.

736 Stoffel MA, Johnston SE, Pilkington JG, Pemberton JM (2021) Genetic architecture and
737 lifetime dynamics of inbreeding depression in a wild mammal. *Nature Communications*
738 *2021 12:1*, **12**, 1–10.
739 The 1001 Genomes Consortium (2016) 1,135 Genomes Reveal the Global Pattern of
740 Polymorphism in Arabidopsis thaliana. *Cell*, **166**, 481–491.
741 Todesco M, Owens GL, Bercovich N *et al.* (2020) Massive haplotypes underlie ecotypic
742 differentiation in sunflowers. *Nature*, **584**, 602–607.
743 Troth A, Puzey JR, Kim RS, Willis JH, Kelly JK (2018) Selective trade-offs maintain alleles
744 underpinning complex trait variation in plants. *Science*, **361**, 475–478.
745 Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*. Springer New York.
746 de Villemereuil P, Schielzeth H, Nakagawa S, Morrissey M (2016) General Methods for
747 Evolutionary Quantitative Genetic Inference from Generalized Mixed Models. *Genetics*,
748 **204**, 1281–1294.
749 Volis S, Anikster Y, Olsvig-Whittaker L, Mendlinger S (2004) The influence of space in
750 genetic-environmental relationships when environmental heterogeneity and seed
751 dispersal occur at similar scale. *The American naturalist*, **163**, 312–327.
752 Volis S, Ormanbekova D, Shulgina I (2016) Fine-scale spatial genetic structure in
753 predominantly selfing plants with limited seed dispersal: A rule or exception? *Plant*
754 *Diversity*, **38**, 59–64.
755 Volis S, Ormanbekova D, Yermekbayev K, Song M, Shulgina I (2015) Multi-Approaches
756 Analysis Reveals Local Adaptation in the Emmer Wheat (Triticum dicoccoides) at
757 Macro- but not Micro-Geographical Scale. *PloS one*, **10**, e0121153.
758 Wintle BA, Kujala H, Whitehead A *et al.* (2019) Global synthesis of conservation studies
759 reveals the importance of small habitat patches for biodiversity. *Proceedings of the*
760 *National Academy of Sciences of the United States of America*, **116**, 909–914.
761 Wittmann MJ, Bergland AO, Feldman MW, Schmidt PS, Petrov DA (2017) Seasonally

762 fluctuating selection can maintain polymorphism at many loci via segregation lift.
763 *Proceedings of the National Academy of Sciences of the United States of America*,
764 **114**, E9932–E9941.
765 Wright S (1969) *Evolution and the Genetics of Populations. Vol. 2: The Theory of Gene*
766 *Frequencies*. Chicago, University of Chicago Press.
767 Zheng X, Levine D, Shen J *et al.* (2012) A high-performance computing toolset for
768 relatedness and principal component analysis of SNP data. *Bioinformatics* , **28**, 3326.
769 Zhu T, Wang L, Rodriguez JC *et al.* (2019) Improved Genome Sequence of Wild Emmer
770 Wheat Zavitan with the Aid of Optical Maps. *G3* , **9**, 619–624.
771 Zsögön A, Peres LEP, Xiao Y, Yan J, Fernie AR (2021) Enhancing crop diversity for food
772 security in the face of climate uncertainty. *The Plant journal: for cell and molecular*
773 *biology*, **109**, 402–414.

774

Data and code availability

Sequencing data generated in this study are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the accession number PRJNA779576. All data processing and data analysis are publicly available on GitHub (https://github.com/TalDM/Ammiad) and Zenodo (DOI:10.5281/zenodo.8116655). An R package, *simmiad*, is provided to perform the individual-based simulations (DOI:10.5281/zenodo.8116105).

Author Contributions

T.D-M., M.F., Y.A. and A.A.L, conceived and designed the study. T.D-M. carried out the molecular work, SNP calling, genetic and phenotypic analyses with input from T.J.E., F.M., M.N. and A.A.L. T.J.E. performed the simulations and phenotypic analyses with input from T.D-M., M.N. and A.A.L. F.M. carried out the population differentiation analysis, with input from T.D-M. and A.A.L. T.D-M., O.R. and A.A.L, carried out the common-garden experiment. T.D-M., H.S., J.M., N.A-R., A.R., M.F., Y.A, and A.A.L carried out the collection. T.D-M., T.J.E., F.M, M.N. and A.A.L wrote the manuscript. M.N. and A.A.L supervised the project.

Competing Interest Statement

The authors declare no competing interests.

Funding

This work was supported by the Yeda-Sela Centre and Sustainability and Alternative Energy Research Initiative of the Weizmann Institute of Science. T.D-M. acknowledges a European Molecular Biology Organization scientific exchange grant 8368.

Supplementary Figures

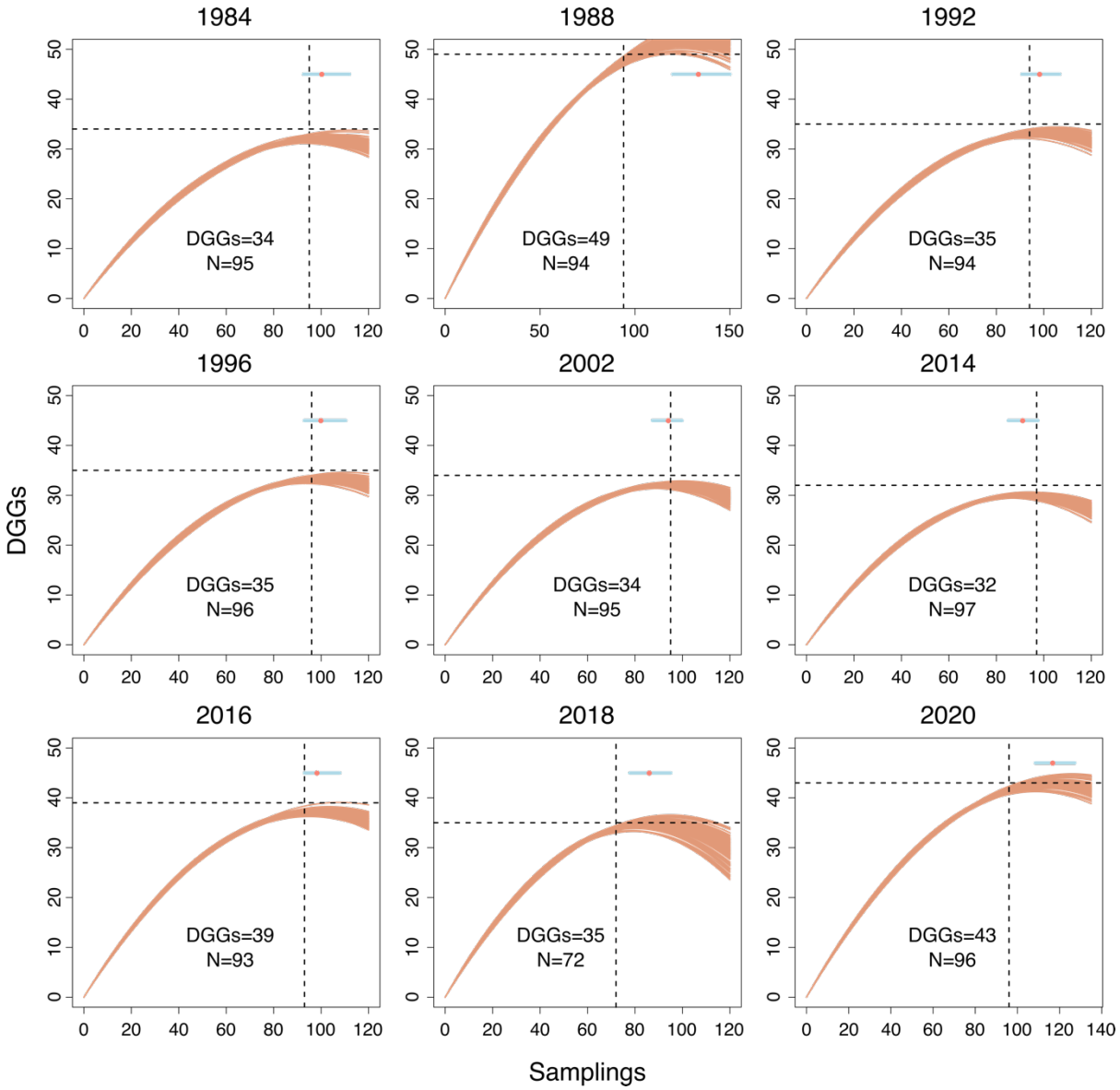


Figure S1. Rarefaction curves. Adjusted 2nd degree polynomials of 100x backwards sample of unique genotype number and number of samples collected per year. Dashed horizontal line represents the observed number of unique genotypes per year, dashed vertical lines represent number of samples per year. Light blue line and red dot represent range and mean of inflection points over replicate subsamples respectively.

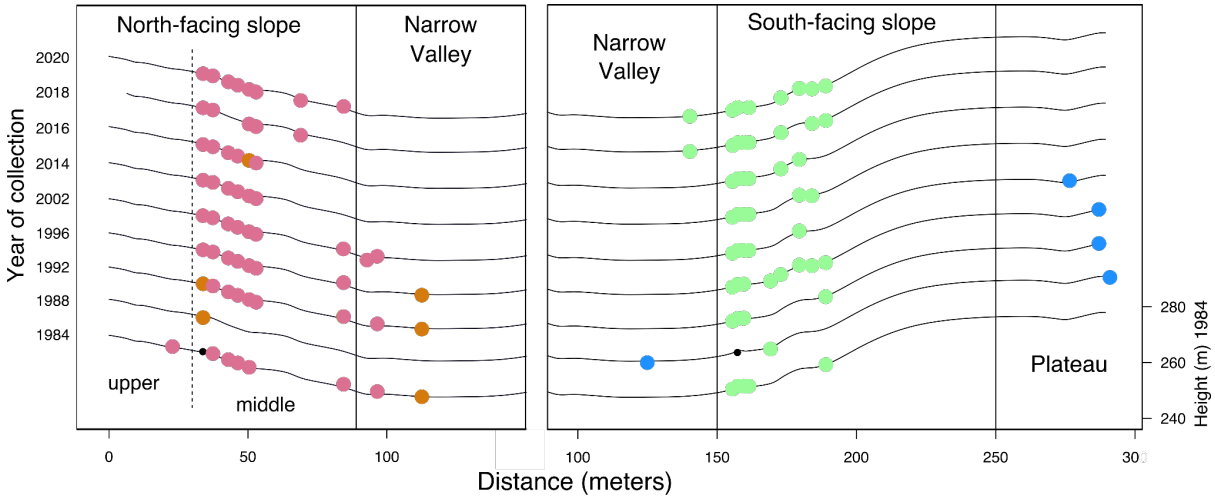


Figure S2. Outcrossing in Ammiad. (A) Outcrossing over time in Ammiad. The red continuous line represents the estimated average decrease of 0.06% every year, with confidence intervals shown as dashed lines. The highest rate of outcrossing is estimated for the middle north-facing slope, consistently with the identification of two highly heterozygous plants in this habitat from 1984 and 1988 (A17 and A20, respectively). Individual yearly estimates are shown as black dots. (B) Parental DGGs of 1984A17 (left panel) and 1988A47 (right panel) detected using SNPMatch. F1 plants are marked as a black dot.

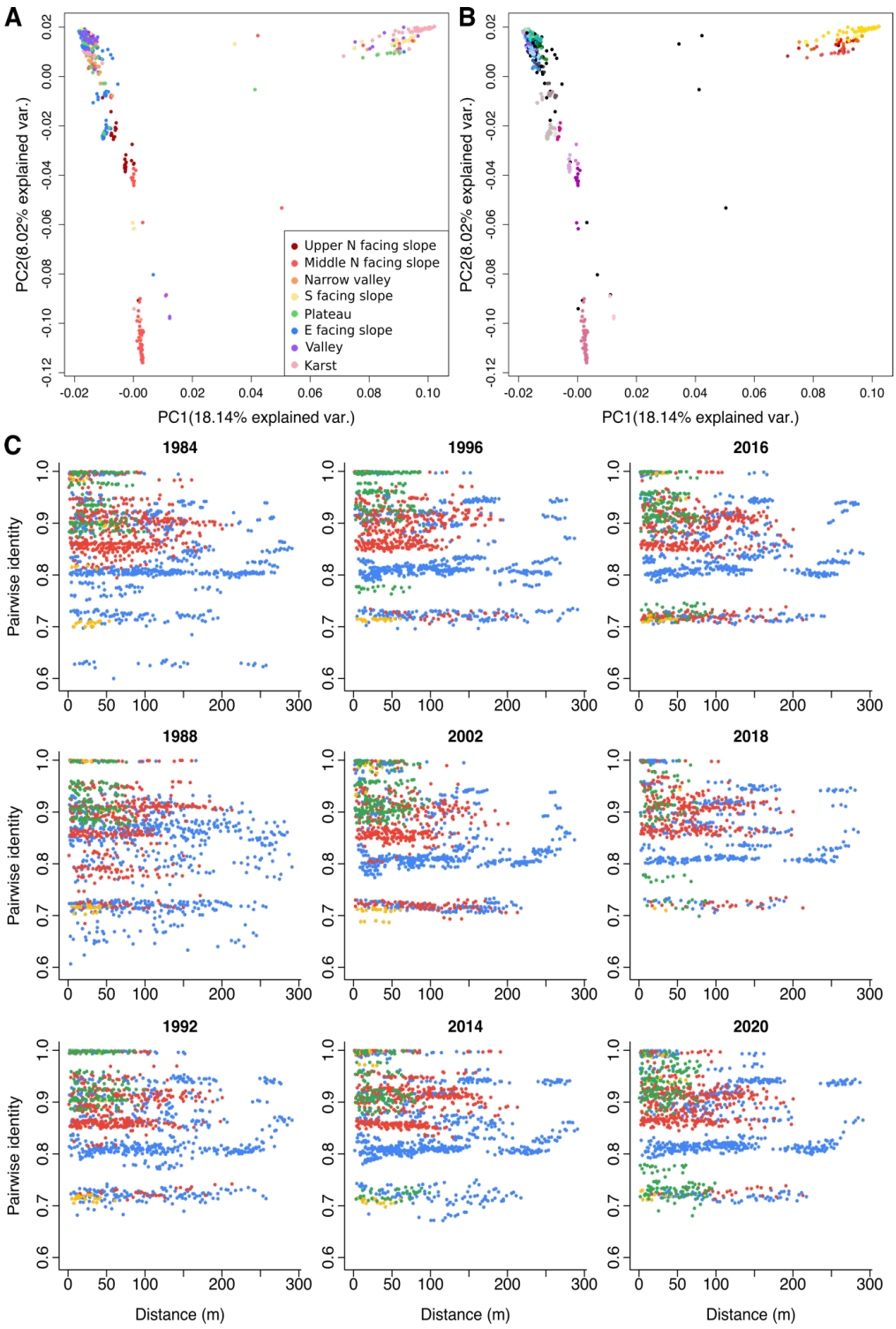


Figure S3. Principal Component Analysis (PCA) and isolation by distance of Ammiad collection. (A) PCA coloured by microhabitat. (B) PCA coloured by DGGs, black dots represent singleton DGGs. (B) Scatter plots of pairwise genetic identity and Euclidean distance (metres). Subplots show years of collection, and colours represent the different transects (A=blue, B=red, C=yellow, D=green).

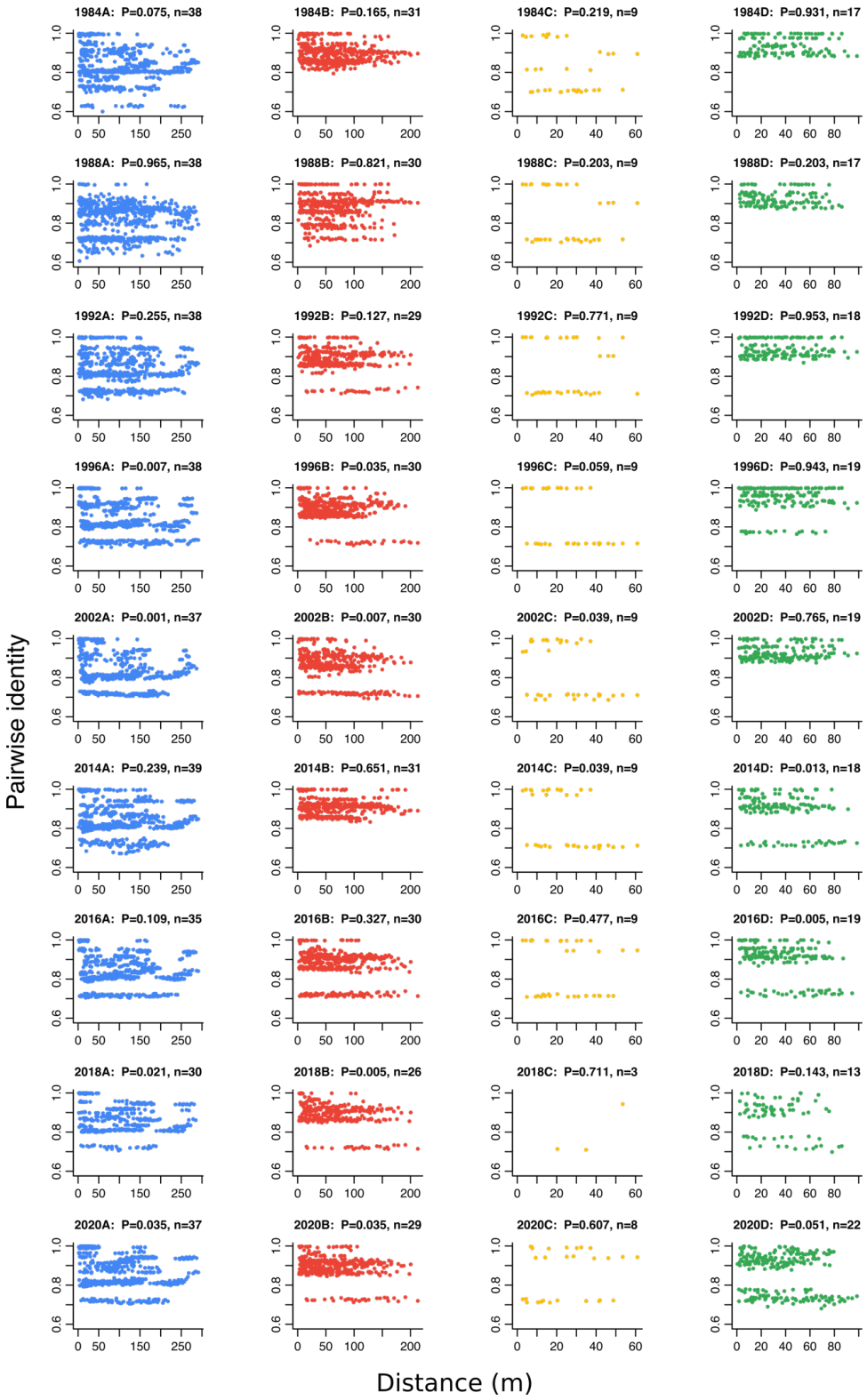


Figure S4. Correlation and significance of isolation by distance in Ammiad collection. Pairwise genetic identity by Euclidean distance (metres). Subplots show years of collection (rows) and transects (columns). The title of each plot indicates its number of samples and p-value from a Mantel test with 1000 permutations.

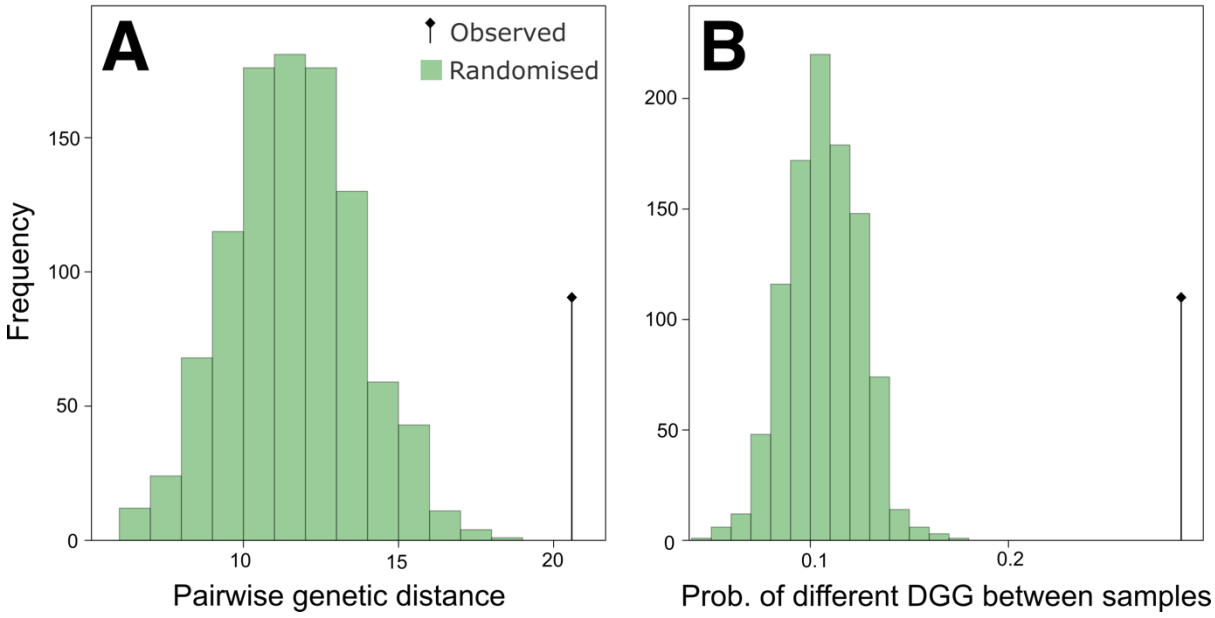


Figure S5. Moran Spectral Randomization for Mantel test to test the effect of habitat on genetic structure accounting for spatial autocorrelation. The histograms represent 1000 spatially constrained randomizations of Mantel statistics (MSR-Mantel). Each randomization was constrained to have a spatial autocorrelation comparable to the observed data, controlling for dispersal and other random factors. For each randomization the effect of habitat on pairwise genetic distances (A) and the probability of two samples to belong to the same DGG (B) was measured, and compared to the observed (black line and dot), in order to, test for the role of habitats in determining the spatial autocorrelation in genetic similarity in Ammiad. For both A and B, we observed a significant role of habitat in shaping the autocorrelation and the genetic distances in Ammiad, both with p-value <0.01.

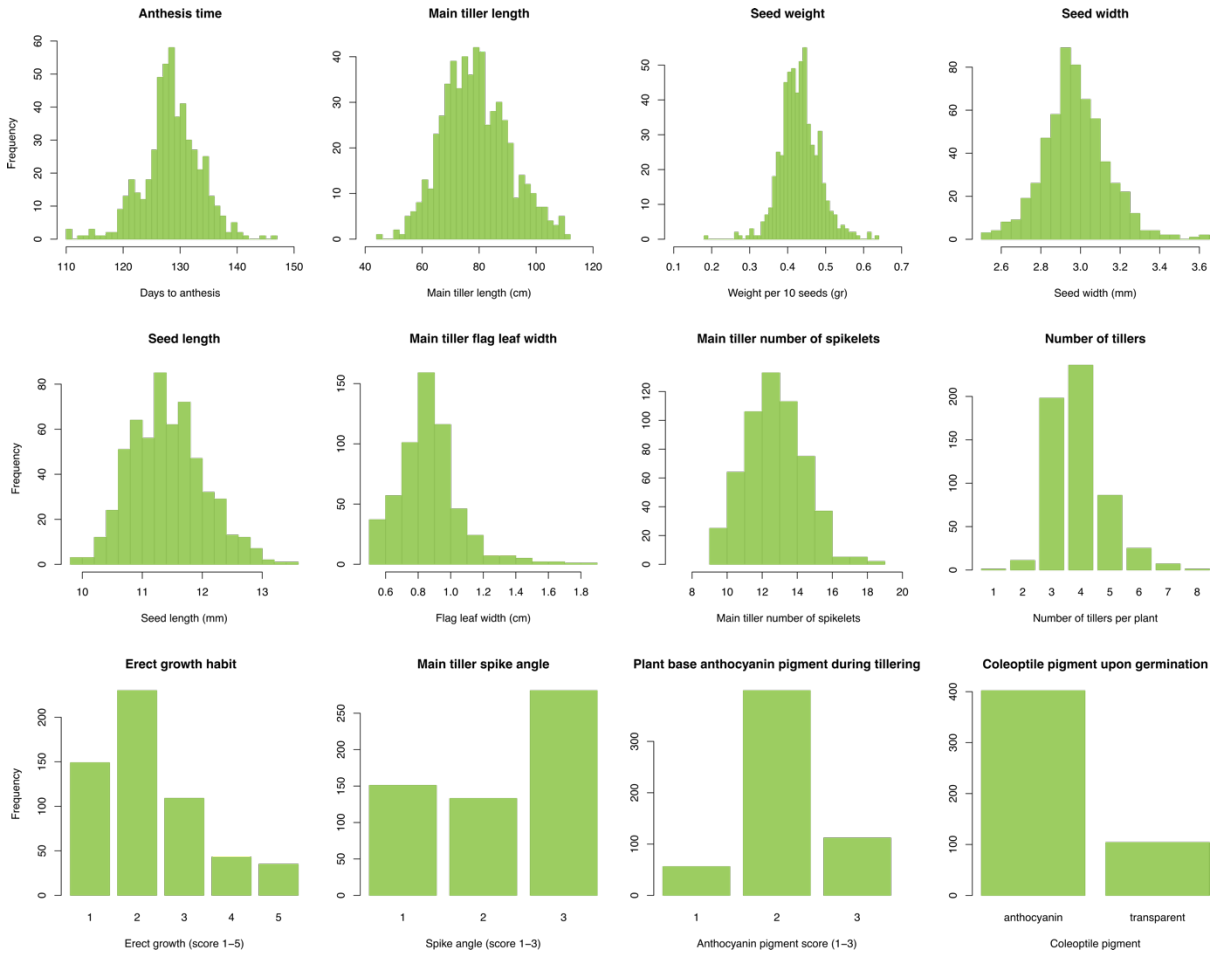


Figure S6. Distribution of phenotypes measured in the common-garden experiment.

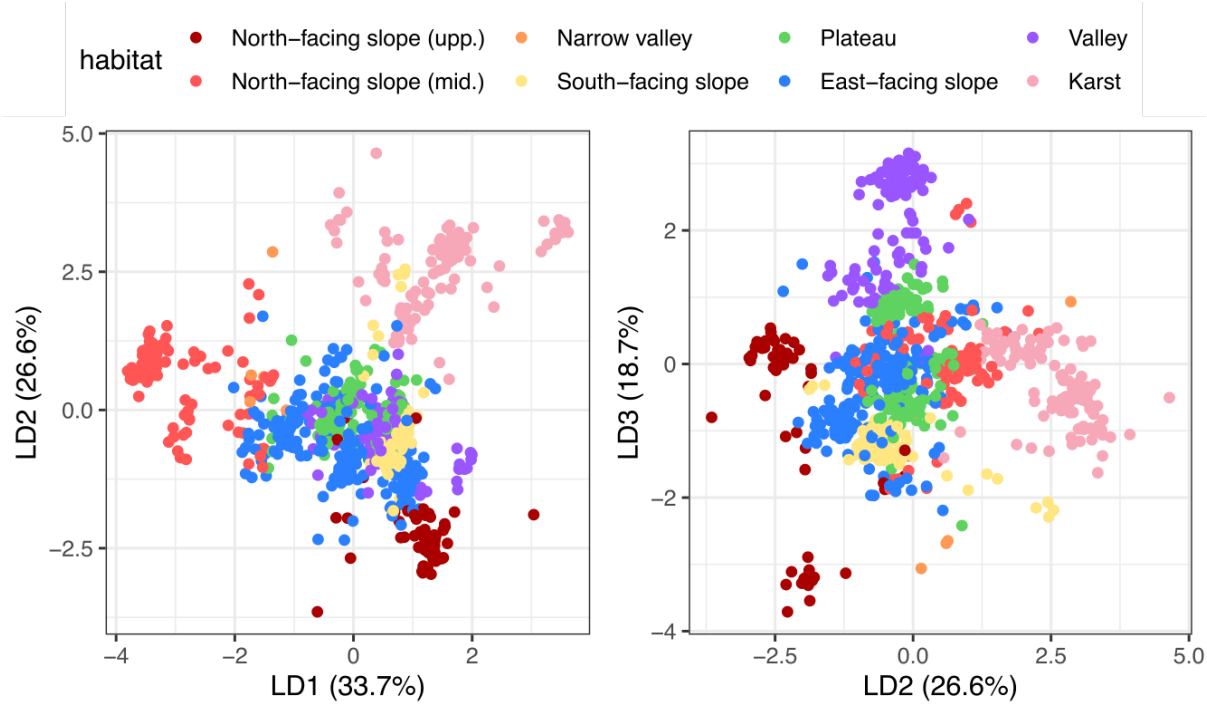


Figure S7. Phenotype-habitat linear discriminant analysis. Projection of the first three dimensions of the linear discriminant analysis of the posterior-mean predicted phenotypes of 801 samples from the wild wheat Ammiad population.

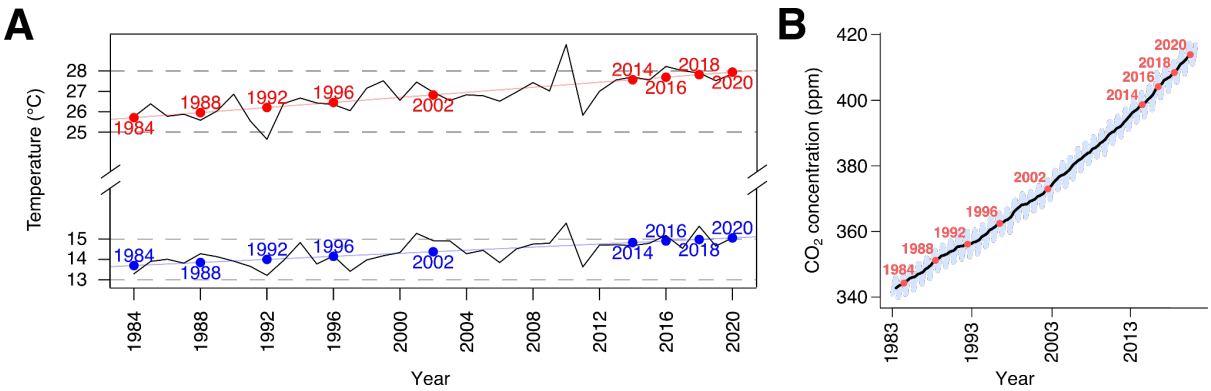


Figure S8: Climatic changes during the course of the Ammiad experiment. (A) Average yearly minimum (blue) and maximum (red) temperature measured in Ayelet Haschahar weather station from 1984 to 2020. Collection years are marked on the respective linear model regression line. (B) CO2 concentration 1983 to 2021. Data obtained from the Mauna Loa record, Scripps Institution of Oceanography, UC San Diego.

855 ## Supplementary Tables

856 **Table S1. Bootstrap-based ranges of F_{IS} and estimated outcrossing rates per habitat.**

	F_{IS}	F_{IS}	outcrossing	outcrossing
Habitat	2.5%	97.5%	2.5%	97.5%
Upper North facing slope	0.9074	0.9246	0.0392	0.0485
Middle north facing slope	0.8385	0.8654	0.0722	0.0878
Narrow valley	0.9548	0.9627	0.0190	0.0231
South facing slope	0.9274	0.9348	0.0337	0.0377
Plateau	0.9565	0.9640	0.0183	0.0222
East facing slope	0.9294	0.9405	0.0307	0.0366
Valley	0.9535	0.9614	0.0197	0.0238
Karst	0.9082	0.9317	0.0354	0.0481

857

858 **Table S2A. Analysis of molecular variance of genotypes sampled in Ammiad across all**
859 **years, subdivided by habitat, using only polymorphic sites within Ammiad.** Sum of
860 squared deviations (SSD) and mean square deviations (MSD), variance components (σ^2) and
861 corresponding p-values were computed using the R package *pegas*. The Φ statistic for
862 habitat was used as an estimate of the proportion of molecular variance explained by the
863 habitat structure.

Component	**SSD**	**MSD**	**df**	**σ^2**	**p-value**	Φ
Habitat	0.9946	0.1421	7	0.0014	$<10^{-12}$	0.2833
Error	2.9429	0.0036	824	0.0036	/	/
Total	3.9374	0.0047	831	/	/	/

864

865 **Table S2B. Analysis of molecular variance of genotypes sampled in Ammiad across all**
866 **Habitats, subdivided by year, using only polymorphic sites within Ammiad.** Sum of
867 squared deviations (SSD) and mean square deviations (MSD), variance components (σ^2) and
868 corresponding p-values were computed using the R package *pegas*. The Φ statistic for year
869 was used as an estimate of the proportion of molecular variance explained by the temporal
870 structure.
871

Component	**SSD**	**MSD**	**df**	**σ^2**	**p-value**	Φ
Year	0.1197	0.0150	8	2.8618e-06	0.4226	0.00019
Error	12.1003	0.0147	823	1.4703e-02	/	/
Total	12.2200	0.0047	831	/	/	/

872

873 **Table S2C. Analysis of molecular variance of genotypes sampled in Ammiad, modelling**
874 **both Habitat and time, using only polymorphic sites within Ammiad.** Sum of squared
875 deviations (SSD) and mean square deviations (MSD), variance components (σ^2) and
876 corresponding p-values were computed using the R package *pegas*. Year was modelled as
877 nested within habitat. Note that the low variance explained by time is not due to this (Table
878 S2B). The Φ statistic for habitat and year was used as an estimate of the proportion of
879 molecular variance explained by the year and habitat structure.
880

Component	**SSD**	**MSD**	**df**	**σ^2**	**p-value**	**Φ**
Habitat	3.1348	0.4478	7	4.3370e-03	$<10^{-12}$	0.2831
Year	0.7146	0.0112	64	1.3254e-05	0.7902	0.0012
Error	8.3706	0.0110	824	1.1014e-02	/	/
Total	12.2200	0.0147	831	/	/	/

881

882 **Table S3. Linear discriminant analysis dimension values.**

883

	LD1	LD2	LD3	LD4	LD5	LD6	LD7
Anthesis date	-0.005	-0.528	-0.346	0.177	-0.198	-0.053	0.124
Coleoptile colour	-0.576	-0.524	-0.194	-0.525	0.285	-0.904	-0.779
Erect growth habit	-0.822	0.32	0.746	-0.985	0.188	0.336	-1.242
Germination date	0.168	0.398	-0.067	-0.111	0.533	-0.148	0.24
Main tiller flag leaf width	0.739	-0.27	-0.828	-0.49	-0.28	-0.355	0.21
Main tiller length	0.759	-0.497	-0.328	0.675	0.144	-1.084	0.159
Main tiller number of spikelets	-0.036	0.054	0.002	0.106	0.08	0.16	0.02
Number of tillers	0.056	-0.096	-0.052	-0.044	0.105	-0.012	0.231
Plant base pigment during tillering	0.617	-0.299	-0.052	0.058	-0.049	0.752	-0.062
Seed area	-1.083	-1.866	-2.63	0.055	-2.035	-1.184	0.422
Seed length	1.344	0.816	2.349	-0.173	1.634	1.42	-0.05
Seed width	1.113	1.303	1.233	1.001	1.606	1.104	-0.937
Main tiller spike angle	-0.214	-0.147	-0.8	0.138	0.553	0.562	0.362
Seed weight	-0.217	0.239	0.024	-0.68	-0.724	-0.234	0.704

884

885 ## Supplementary Files

886 **Movie S1 (separate file).** Drone video of the different transects and microhabitats in Ammiad
887 wild wheat reserve.

888 **Dataset S1 (separate file).** Dataset of Ammiad samples by genotype, coordinates and spike
889 density measured at transect A in 2020.

890 **Dataset S2 (separate file).** Dataset of plant and seed phenotypes measured in the common
891 garden experiment.