The genetic structure of a wild wheat population has remained associated with microhabitats over 36 years

Tal Dahan-Meir[1], Thomas James Ellis[2], Fabrizio Mafessoni[1], Hanan Sela[3,4], 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.

[*]Co-corresponding authors: [avi.levy@weizmann.ac.il; magnus.nordborg@gmi.oeaw.ac.at]

Abstract

Wild progenitors of major crops can help us understand domestication, and may also provide the genetic resources needed for ensuring food security in the face of climate change. We examined the genetic structure of a wild emmer wheat population, sampled over 36 years while both temperature and CO_2 concentration increased significantly. The genotypes of 832 individuals revealed high genetic diversity over scales of tens of meters and were clustered spatially into ecological microhabitats. This pattern was remarkably stable over time. Simulations indicate that neutral processes alone are unlikely to fully explain the spatial and temporal stability of the population. These results are consistent with a role for local adaptation in shaping the fine-scale structure of plant populations, which is relevant for *in-situ* conservation strategies of biodiversity in nature.

Introduction

Climate change is a looming threat to plant biodiversity and food security[1–3]. This increases the importance of wild progenitors of crops because their genetic diversity is an immense reservoir of biotic and abiotic stress-resistance genes[4]. Long-term studies of progenitor species in their natural ecological niches can provide a better understanding of the likely impacts of climate change, and inform strategies for conservation of such resources *in-situ* as well as in gene banks[5].

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). During this period, local ambient temperatures increased by ~1.5 degrees Celsius (Figure S1) and atmospheric CO_2 concentration increased by more than 20% (Figure S2)[6]. This population was previously shown to be highly diverse in comparison with other populations, suggesting that it is located at the centre of distribution of the species[7,8]. Wild emmer wheat is a self-pollinating, allotetraploid annual cereal (genome BBAA) that was domesticated about 10,000 years ago, during the Neolithic period[9–11]. It is the direct 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). The population has been sampled at nine time points between 1984 and 2020. At each time point, seeds were sampled from plants at 100 marked locations along four transects traversing seven ecologically distinct microhabitats[12] (Figure 1A, Supplementary Table 1, Supplementary Movie 1). We genotyped 832 individuals using a genotyping-by-sequencing (GBS) protocol[13], resulting in 5196 single-nucleotide-polymorphism (SNP) markers spread throughout the genome (Figure S3). Earlier spatio-temporal studies of wild plants involved shorter periods, as well as less comprehensive and less precise sampling[14–18]. Our dataset provides a unique resource for assessing changes in genetic diversity of a population in its natural habitat over a long-term experiment.

Results

As expected for a highly selfing species[19], the population is composed of genetically distinct groups of nearly identical, highly homozygous individuals clustered in space. Pairwise genetic identity follows a multimodal distribution that is characterized by pairs of plants that are either identical at almost all SNPs, or identical at 70% to 90% of SNPs (Figure 1B). We unambiguously assigned all plants to distinct genotype groups (DGGs) of individuals that were identical at >98% of markers. 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 numbers of DGGs per year varied between 32 (2014) to 49 (1988; Supplementary Table 2). Rarefaction curves indicate that the sampling scheme was sufficient to detect almost all of the DGGs present close to the transects (Figure S4). These results provide a detailed picture of the genetic structure of the population through time and space.

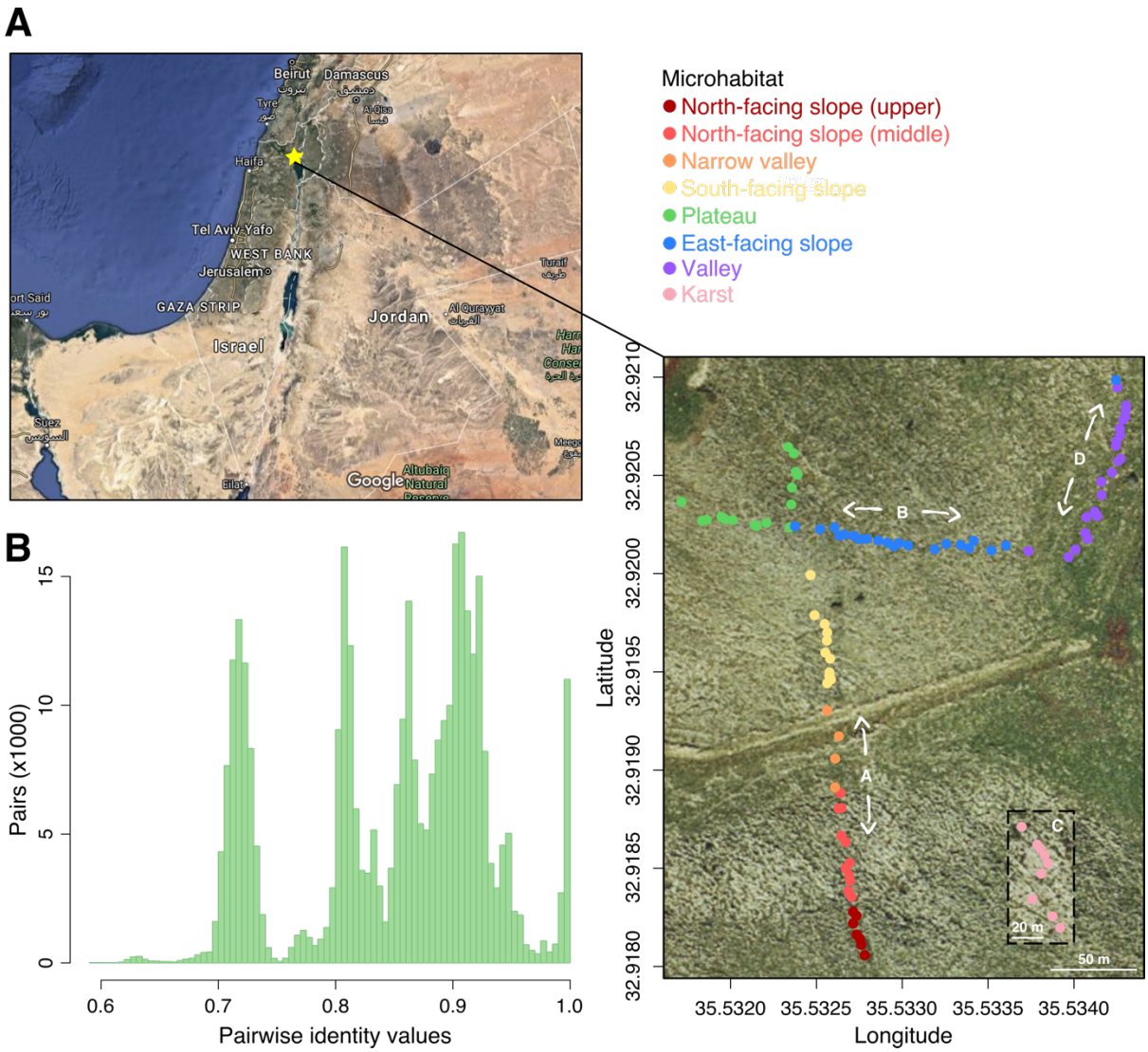

Figure 1. Spatial structure of the Ammiad population. (A) Location of Ammiad near the lake of Galilee (star, top left and geographic coordinates, bottom right). Sampling points of transects A, B, C, D collected in the years 1984, 1988, 1992, 1996, 2002, 2014, 2016, 2018 and 2020. Sampling points are shown on-site for transects A, B, D, but transect C, shown in the dashed rectangle, is located ~500 meters north-west of A, B, D. Sampling points colors correspond to the different microhabitats labelled in the legend. Image of map on the right is from gov.map.co.il, map on the left from Google Maps. (B) Distribution of pairwise identity values between the plants in the collection.

The spatio-temporal distribution of DGGs shows that plants of the same DGG tend to cluster together in both space and time (Figure 2 and Figure S5). This clustering was also found in a principal component analysis (Figure S6). Within years, the probability of belonging to the same DGG decays rapidly with distance (Figure 3A), however, there is little evidence for isolation by distance beyond this (Figure S7, S8). Overall, we find a striking concordance between DGGs and microhabitats, which were defined based on topographic and floristic data without knowledge of the genetic structure of the population[20] (Figure 2). For example, in transect A there is a sharp transition between DGGs of the upper and middle north-facing slope (Figure 2A). Another example is the genotype at collection point A59 in a unique location under the shade of a *Ziziphus spina-christi* tree (Figure 2A). In contrast, in transect D, where there is negligible microhabitat differentiation, plants of different DGGs were much more mixed (Figure 2D). The spatial pattern was highly stable over the course of the experiment. In contrast to the rapid spatial decay, the probability of belonging to the same DGG decays only slowly over time (Figure 3). For example, 33% of plants sampled in 1984 belonged to the same DGG as plants sampled at the same position in 2020. In conclusion, our results show a strong spatial structure, associated with microhabitat, which has remained stable for at least 36 years despite changes in climatic conditions.

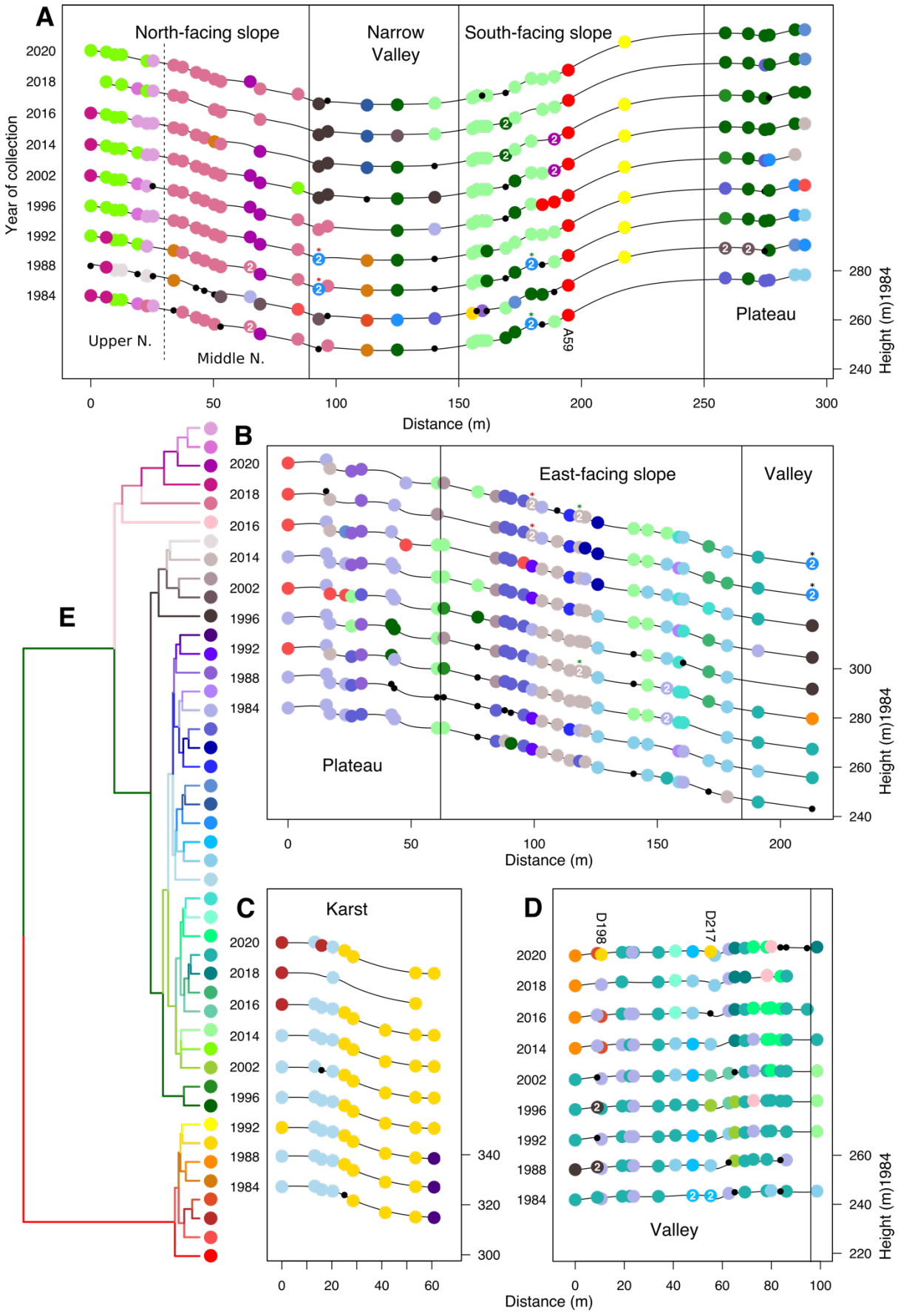

Figure 2. Spatio-temporal distribution of genotypes. (A-D) Distribution of samples colored by distinct genotype group (DGG) in transects A, B, C and D. Y-axes indicate year of collection (left) and altitude for 1984 (right). Singleton samples are marked in small black dots, while doubleton groups are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

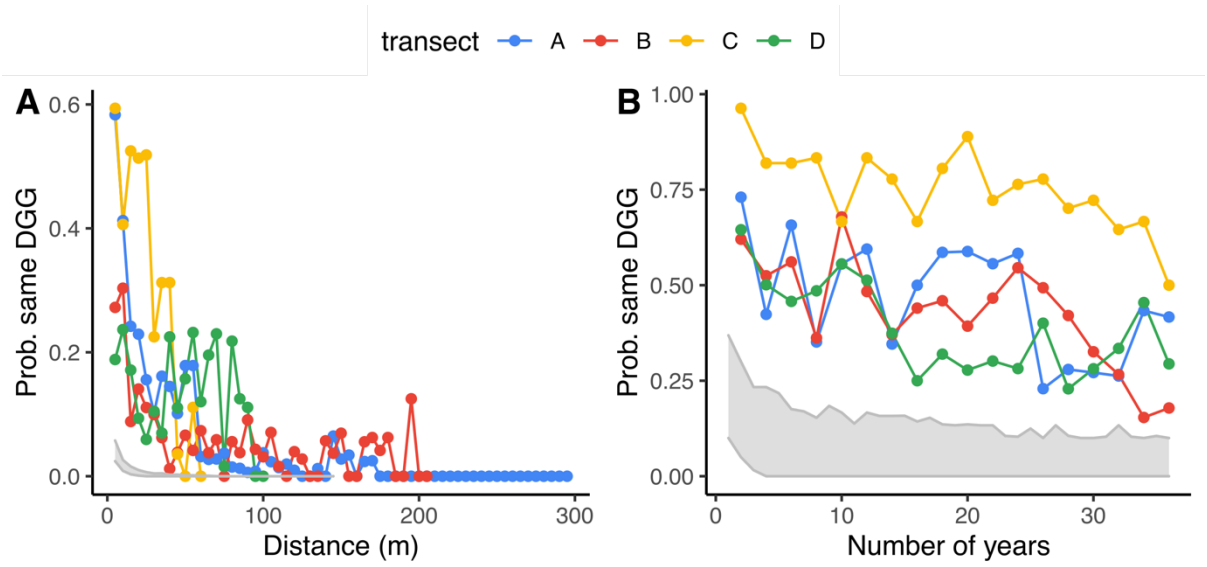

Figure 3. Spatial clustering and temporal stability at Ammiad. Points show the probability of observing genotypes of the same DGG at (A) pairs of sampling points in the same year at increasing distances along each transect and (B) the same sampling point between pairs of years. Grey regions indicate 95% confidence intervals over populations simulated under conditions closest to empirical estimates (3 plants/m^2, mean seed dispersal of 1m, 0.5% outcrossing per generation, 30% seeds germinating in the second year after dispersal). In A, points show the average for sampling points in bins of 5m.

The observed patterns cannot be explained by simple chance. Across transects, pairs of the same DGG were significantly more likely to be found in the same microhabitat in the observed

dataset than in 1000 permuted datasets where the order of microhabitats was randomly shuffled (Figure S9), or even when we control for distance and spatial autocorrelation, which could be caused by limited dispersal, as shown via partial Mantel tests and Moran-spectral-randomization Mantel statistics (MSR-Mantel) (Figure S10). Similarly, population differentiation measured by F_{st} was consistently higher between microhabitats than within, even when taking distance into account (Figure S11), and genetic distances are strongly correlated with microhabitat. Furthermore, an analysis of molecular variance (AMOVA) showed that 28% of the genetic variation between samples is explained by differences between microhabitats (Supplementary Table 3). In order to further test the hypothesis that the correspondence between the population structure of Ammiad and microhabitats is shaped by processes other than limited dispersal, we implemented an explicit demographic model that also incorporates self-fertilization and limited seed dispersal. The question is whether these neutral demographic processes - on their own - could generate the strong spatial clustering and temporal stability observed at Ammiad.

To investigate this, we performed individual-based simulations of populations of plants in a homogeneous environment, evolving by seed dispersal and outcrossing only. We ran simulations using different values of outcrossing rate, plant density, seed dormancy and mean seed dispersal distance that approximate a range of biologically realistic values. Empirical values of outcrossing rates in wild wheat have been estimated to be approximately 0.5%[21] with up to 30% of germinating seedlings coming from the seed bank[22]. Average plant density in the Ammiad population was measured in 1984 as 4 plants/m^2 [23] and again in 2020 for transect A as 3.8 plants/m^2 (Figure S12, Supplementary Table 4). We lack direct estimates of seed dispersal, but average gene flow, a composite of seed and pollen dispersal, was estimated to be 1.25 m per year[24]. Seed dispersal occurs through a combination of gravity, ants, birds, grazing mammals and rodents, enabling both short and long-distance dispersal. For example, the

appearance of a DDG prominent in transect C (and marked by yellow dot), in transect D in 2020 (locations D198, D217), more than 700 meters away, demonstrates that long distance movement occasionally occurs.

In simulations with the parameters we consider most likely (1 m mean seed dispersal, 0.5% outcrossing, density of 3 plants/m^2 and 30% seed dormancy) we find that neutral demographic processes alone could not account for the observed spatial clustering and temporal stability (Figure 3), in line with other studies[25]. However, the simulation could produce results consistent with the data if we assumed extremely low plant densities (≤1 plant/m^2) or rates of gene flow (outcrossing <0.5% when seed dispersal ≤ 1m) (Figure S13-S16). Without better estimates of these parameters or direct fitness estimates from reciprocal transplant experiments, it is not possible to definitely reject a purely neutral explanation for the observed patterns. That said, the correspondence between DGGs and microhabitats is strong (Figure 2 and Figure S9-S11). Thus, the simulations indicate that it would be difficult to generate the observed patterns of stability over 36 years through neutral demographic processes alone.

Discussion

Studies of natural populations rarely have fine-scale geographic resolution, and almost never have temporal resolution beyond the duration of a PhD[26]. Our study, spanning over 36 years of dense sampling, thus provides a unique dataset that could improve our understanding of gene flow and local adaptation.

Our analysis revealed that the collection of Ammiad population varies more in space than in time. We found unexpectedly strong long-term genetic stability, demonstrating that gene flow can be highly limited in a wild cereal despite its relatively wide distribution throughout the

fertile crescent. Remarkably, the stable genetic clusters observed in Ammiad matched ecological microhabitats previously identified at the site, suggesting that natural selection, in addition to limited dispersal, contributed to shaping population structure. The possibility that local adaptation occurs at such a fine spatial scale is particularly intriguing, and invites experimental confirmation.

These 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 urbanization, 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 progenitors, partially protected from anthropogenic activities, does not necessarily require large areas and is a feasible goal that is critical for biodiversity conservation.

In conclusion, we provide evidence for very strong genetic stability in a wild grass, most likely due to a combination of local adaptation and limited dispersal. This strong stability is remarkable given that, during the period of this study, temperatures and CO_2 levels increased significantly (Figure S1, S2). Yet only rather minor changes occurred in the population genetic

structure, reflecting the ability of wild plants to deal with changing environments and pointing to the robustness of wild emmer wheat. Such robustness further demonstrates the potential utility of wild relatives of crops as gene pools for breeding[27].

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). An R package, *simmiad*, is provided to perform the individual-based simulations (DOI:10.5281/zenodo.4762083).

Acknowledgements

We thank T. Azulay, T. Grinevich, C. Melamed-Bessudo, O. Gross, R. Epstein, O. Rudich, Z. Meir and members from the Institute for Cereal Crops Improvement at Tel-Aviv University for plant material collection and J. Müller for the drone movie. We also thank 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.

Funding

This work was supported by the Yeda-Sela Center 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.

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 and genetic analyses with input from T.J.E., F.M., M.N. and A.A.L. T.J.E. performed the simulations and microhabitat sorting analyses with input from T.D-M., M.N. and A.A.L. F.M. carried out the population differentiation analysis and the

MSR-Mantel tests. 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.

Ethics declaration

The authors declare no competing interests.

Methods

Sample collection and DNA extraction

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

Genotyping by sequencing (GBS) and variant calling

Plant genotyping was done according to the Genotyping-By-Sequencing method[13,28]. Samples were sequenced using Illumina NextSeq 550 mid-output 150 base-pairs single-end kits. A reference sample, Zavitan accession[10], was added to each run, undergoing the same plate GBS library preparation and sequencing. Samples were divided by barcodes using a python 2.7 code. Mapping to the WEW_v2.0 genome assembly[29] was done using *bwa-mem*[30]. Conversion to binary alignment map (BAM) format, filtering for uniquely mapped reads with minimum QUAL score of 30, sorting and indexing was done using SAMtools[31]. Variant calling was done in parallel running on all 845 (832 Ammiad samples and 13 Zavitan controls) samples using GATK HaplotypeCaller 3.8[32]. Filtering for quality (>30), read depth (>6), genotype quality (>15), maximum missing sites (10%) and no-indels was done using VCFtools[33]. Filtering for up to 10% heterozygous plants per polymorphic position was done using TASSEL5[34]. Filtered VCF was used for all further analyses.

Pairwise genetic identity

Identity was measured using R code which calculates the pairwise identity values between all samples, divided by the numbers of comparisons. The pipeline also includes the division of the samples to distinct genotype groups (DGGs), done by setting a threshold of 98.1% identity, which is the lowest threshold where no sample is assigned to two or more DGGs. and the construction of the dendrogram describing the relatedness between the DGGs. Hierarchical clustering was done in R by UPGMA. The 13 Zavitan reference samples added to each run were grouped together with 99.4% identity. Heatmap of samples from 1984 was done using *pheatmap* (https://cran.r-project.org/web/packages/pheatmap/index.html) with default hierarchical clustering.

Temperature and CO2 concentration data

Data of daily minimum and maximum temperature from 1984 to 2020 were obtained from Israel Meteorological Service at the Israel government portal (https://ims.data.gov.il/he/ims/2). Data of weekly atmospheric CO2 concentrations were obtained from The Scripps Institution of Oceanography UC San Diego[6].

Rarefaction, principal component and identity by distance analyses

Rarefaction analysis was done using R, sampling backwards by the number of samples and different genotypes, per year, 100 times and adjusting a polynomial to each sampling.

Principal component analysis was done on the .vcf file of the entire Ammiad collection using *SNPRelate* R Package[35]. Identity by distance was correlated for each transect separately. Pairwise Euclidean distances was calculated using the *distHaversine* function of *Geosphere* R package (https://cran.rproject.org/web/packages/geosphere/index.html). Mantel test with 1000

permutations was used to calculate correlation between the genetic identity and the Euclidean distance matrices, using *mantel.test* in *Ape* R package[36].

Individual-based simulations

To investigate whether neutral demographic processes could account for the observed spatial clustering and temporal stability we performed individual-based simulations of populations of plants evolving via seed dispersal and outcrossing only. Briefly, we allow populations to evolve in a 2-dimensional continuous space, then draw a transect through the population and sample plants close to 30 evenly spaced sampling points in the last 36 generations to replicate the sampling procedure in the empirical dataset. An R package, *simmiad*, is provided to perform these simulations (DOI:10.5281/zenodo.4762083).

Populations begin with N plants of a single genotype distributed randomly through a homogeneous habitat with 1, 2, 3 or 5 plants/m^2. At each generation, N plants are sampled with replication to produce N seeds. Wild emmer wheat has a two-year seed bank, so from generation 3 onwards, seeds are drawn from the previous generation with probability p, and from two generations previous with probability 1-p, for values of p of 0.5, 0.7, 0.9 or 1 .Since we are only concerned with tracking unique genotypes, and information about seed dispersal is in any case lacking model outcrossing as analogous to mutation; seeds are randomly assigned as outcrossed with probabilities 0.25%, 0.5%, 1%, or 2%, in which case it is assigned a new unique genotype. Offspring disperse from the mother in a random direction at a distance drawn from an exponential distribution with a mean of 0.5, 0.75, 1 or 2m. To eliminate edge effects, the habitat is modelled as a torus, such that if dispersal takes a seed over the edge of the habitat it is reflected back to the opposite side. We allowed populations to evolve for 2000 generations to ensure populations had reached equilibrium. We then drew a transect of 30 sampling points

5m apart through the middle of the population, and sampled the plant closest to each sampling point in each of the final 36 generations of the simulations, replicating the sampling procedure in the field. For each combination of plant densities, outcrossing rates, seed bank parameter *p*, and seed dispersal distances we performed 100 replicate simulations.

Microhabitat sorting and analysis

We aimed to test the hypothesis that the population structure of Ammiad wild wheat reflects that of microhabitats. We tested this using two different statistics and permutation schemes. First, we tested whether plants in the same DGG tend to be confined to a single microhabitat more than would be expected by chance using a procedure analogous to a permutation test. We concatenated microhabitat and DGG labels across transects, and calculated how often pairs of plants of the same DGG were located in the same microhabitat. We then shuffled the order of microhabitats, whilst keeping labels of the same microhabitat together, and recalculated how many pairs of plants of the same DGG were located in the same microhabitat after shuffling. We repeated this shuffling procedure 1000 times, and compared the observed number of matching pairs within the same microhabitat to the distribution of shuffled datasets. This procedure randomizes the relationship between the locations of DGGs and microhabitat boundaries, but maintains the overall spatial structure of DGGs and microhabitats.

Second, we tested whether population differentiation, measured by F_{st}, is higher when measured between plants occurring in different microhabitats than in the same microhabitat. To this goal we subdivided plants into demes with a maximum distance between plants of 30m. We then computed the pairwise Fst between demes using the R package *pegas*[36]. To correct for the effects of isolation by distance and spatial structure we stratified comparisons by distance, measured as the distance between the average locations of plants in the two demes.

We computed an analysis of molecular variance using the R package *pegas*[37], using plants sampled across all years and restricting the analysis only to genotypes present in Ammiad. We included two factors, microhabitat and year of collection, with the latter nested within the former.

To account for spatial autocorrelation we first performed a partial Mantel test using the function *multi.mantel* from the R package *phytools*. Specifically, we considered as a response variable the pairwise genetic distances used to define DGGs. We tested for the effect of habitat, represented as a distance matrix with value "0" for plants found in the same microhabitat and "1" for plants found in two different microhabitats. In addition, we included in the test a matrix of time distances between the sampling of two plants (distance measured in years) and spatial distance, measured in metres and obtained from latitude and longitude using the function *earth.dist* from the R package fossil. We detected a significant effect of space and microhabitat (p-values <0.005 for both matrices) and a non-significant effect of time (p-value 0.863).

Note that Mantel and partial Mantel tests have severely been criticized for their inaccuracy in presence of high spatial autocorrelation [38,39]. To account for these known biases, we used the approach proposed by[39], which leverages a Moran spectral randomization procedure to generate random samples which maintain the same autocorrelation properties of the original dataset. To this goal, we used the R packages *sdpep* to process the spatial matrices and the *msr* function of the package *adespatial*. The distances above were transformed into Euclidean distances with function *cailliez* from the R package *adegenet*. Using this approach, we tested the effect of habitat on both DGGs and pairwise genetic distances, using 1000 randomizations. In both cases the effect of microhabitat was significant (p-value<0.05).

Supplementary Figures

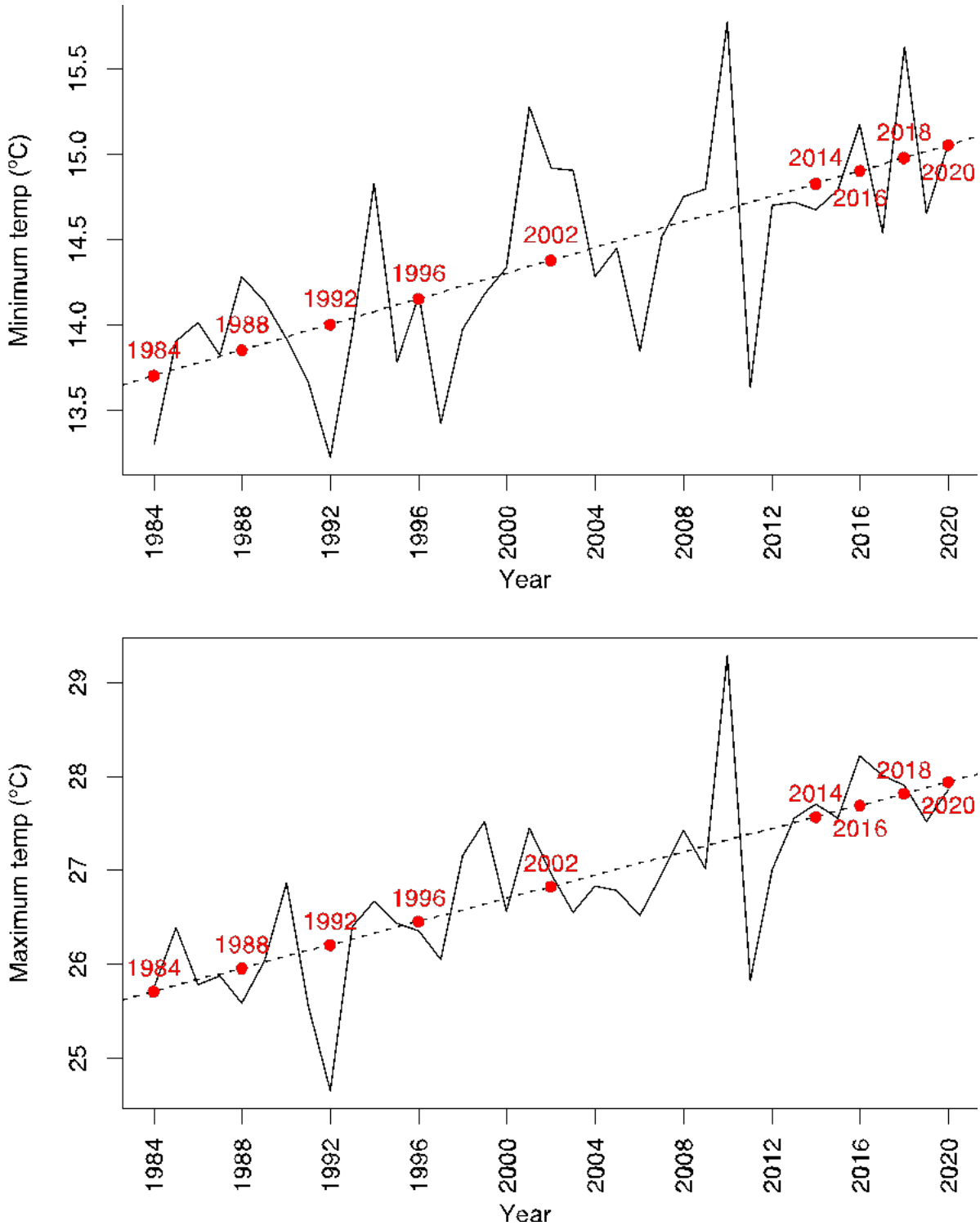

Figure S1. **Average yearly minimum (A) and maximum (B) temperature measured in Ayelet Haschahar weather station from 1984 to 2020**. Collection years are marked in red on the linear model regression line.

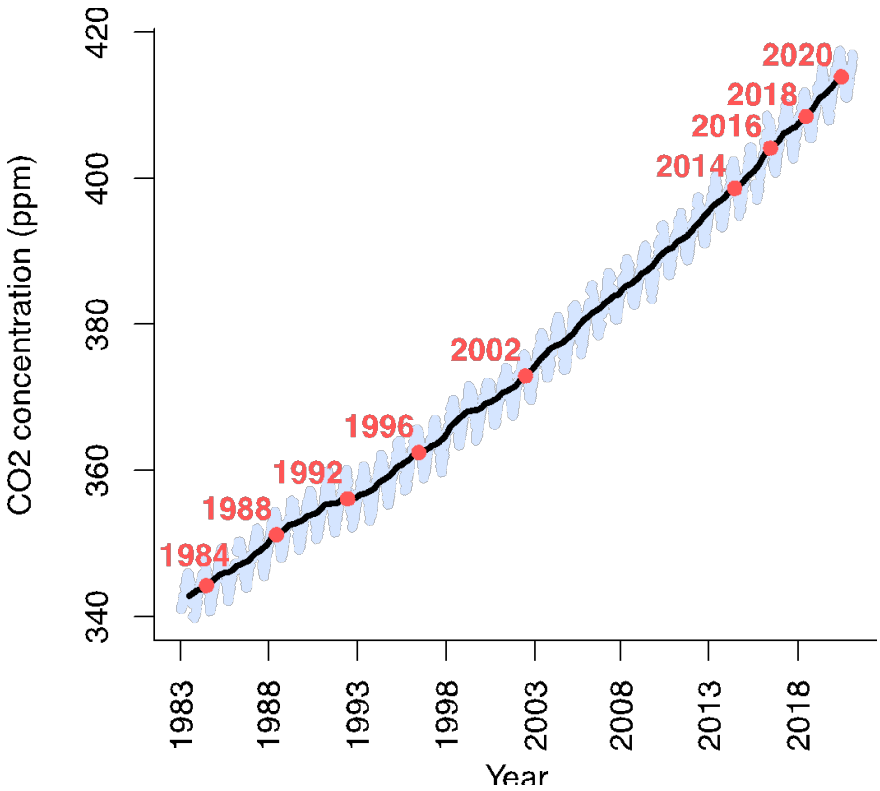

Figure S2. **CO_2 concentration 1983 to 2021**. Data obtained from the Mauna Loa record, Scripps Institution of Oceanography, UC San Diego (*1*).

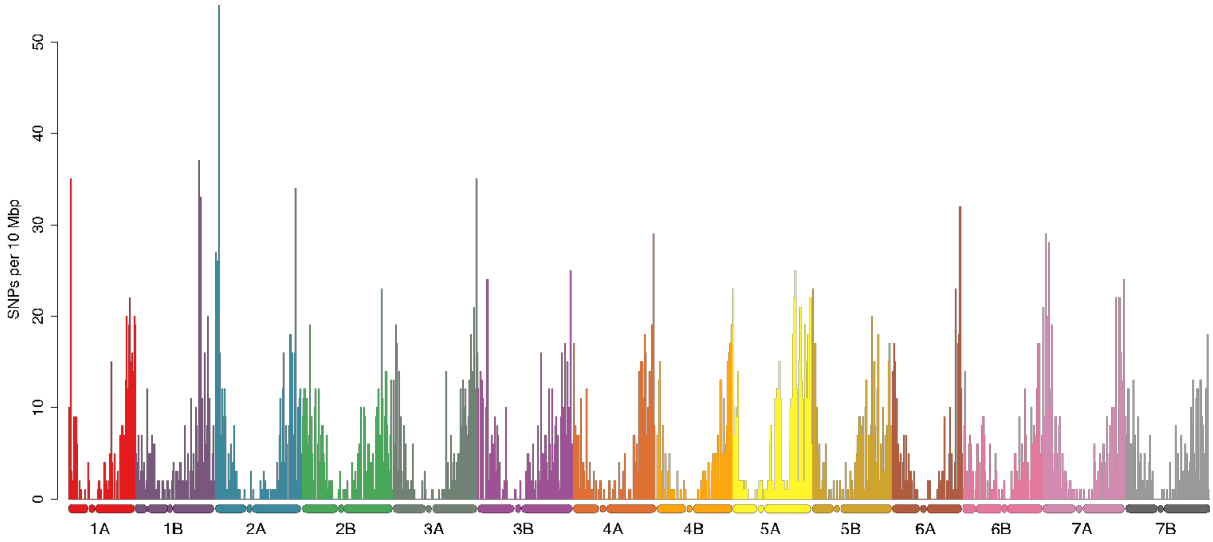

Figure S3. **SNP density of 5196 markers along the wild emmer wheat genome.**

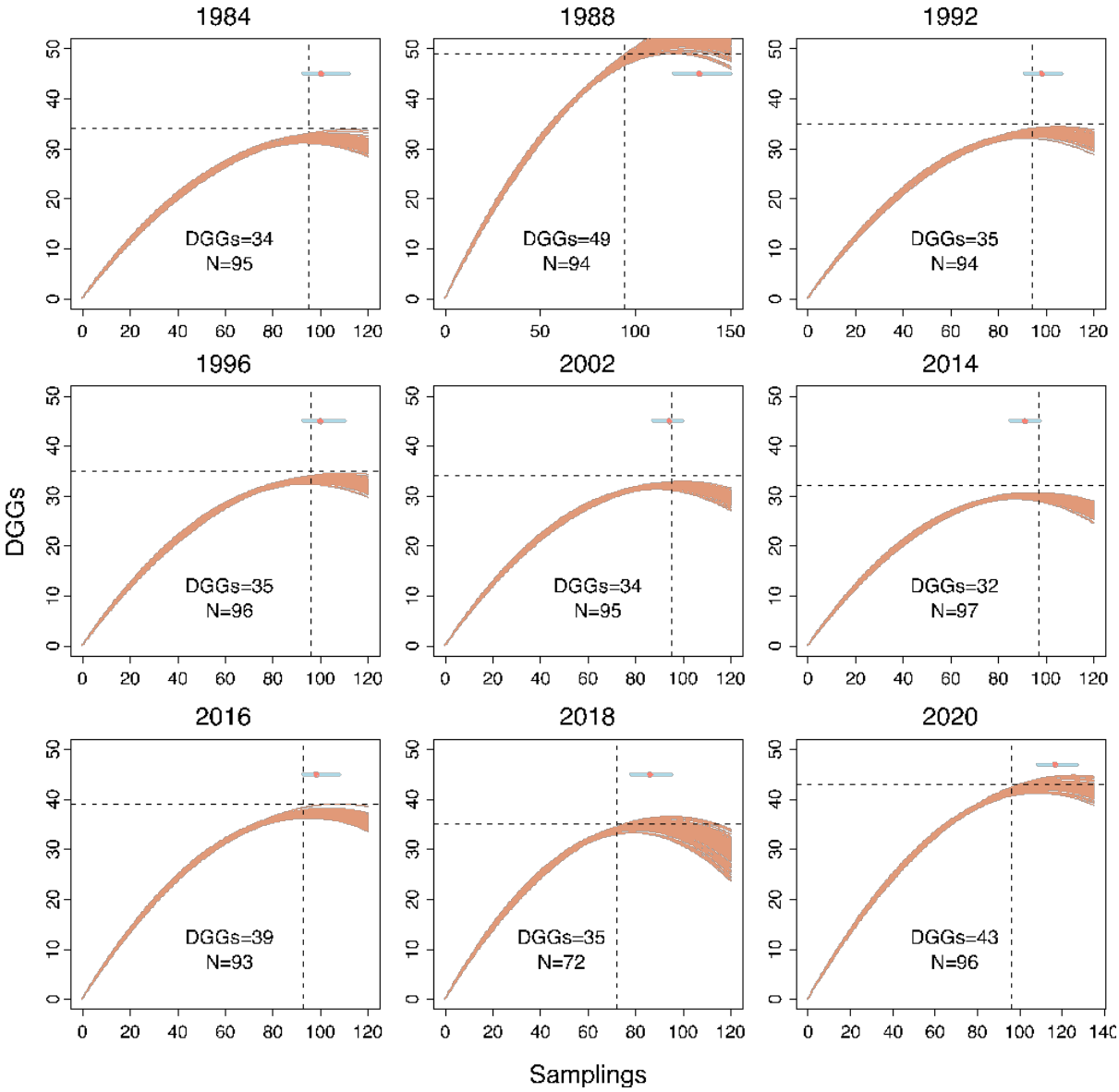

Figure S4. **Rarefaction curves.** Adjusted 2nd degree polynomials of 100x backwards sample of unique genotype number and number of samples collected per year. Dashed horizontal line represents the observed number of unique genotypes per year, dashed vertical lines represent number of samples per year. Light blue line and red dot represent range and mean of inflection points over replicate subsamples respectively.

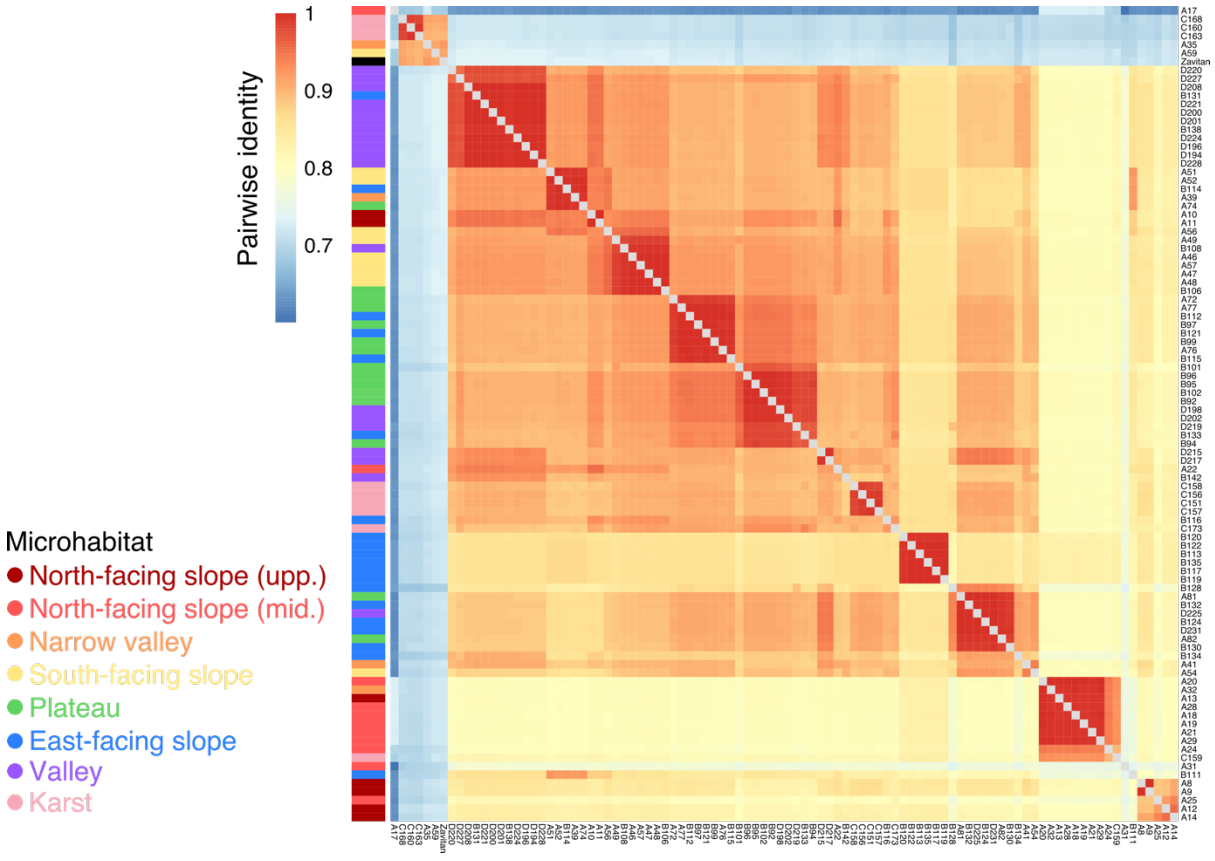

Figure S5. **Population structure of Ammiad in the first year of collection (1984).** Heatmap colors represent identity value for each pairwise comparison, annotation colors (on the left) correspond to the list of microhabitats.

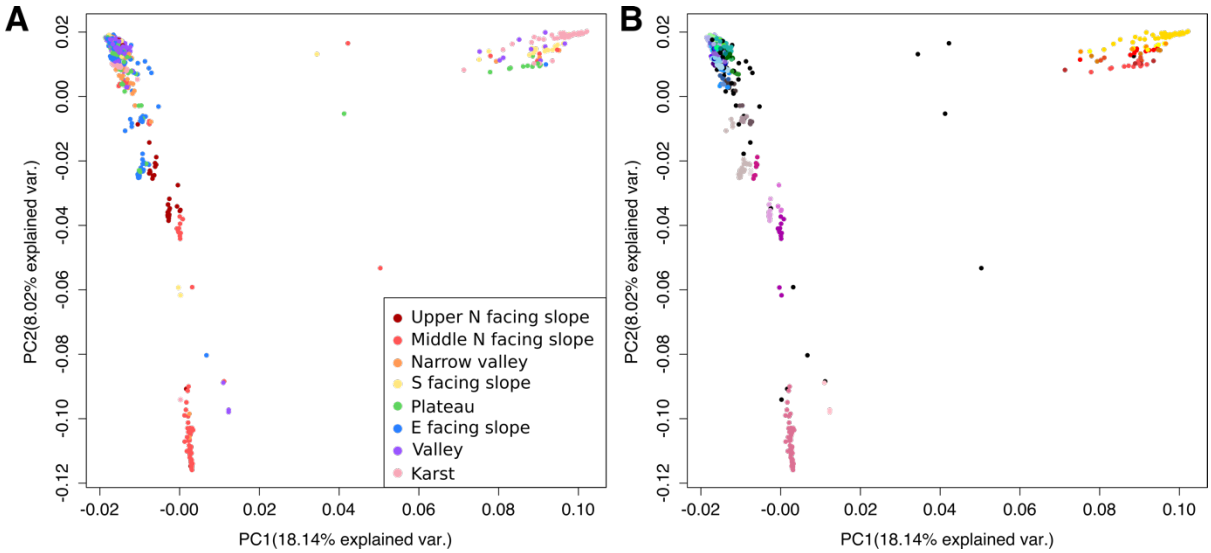

Figure S6. **Principal Component Analysis of Ammiad collection.** (A) Colored by microhabitat. (B) Colored by DGGs, black dots represent singleton DGGs, which are not related to each other.

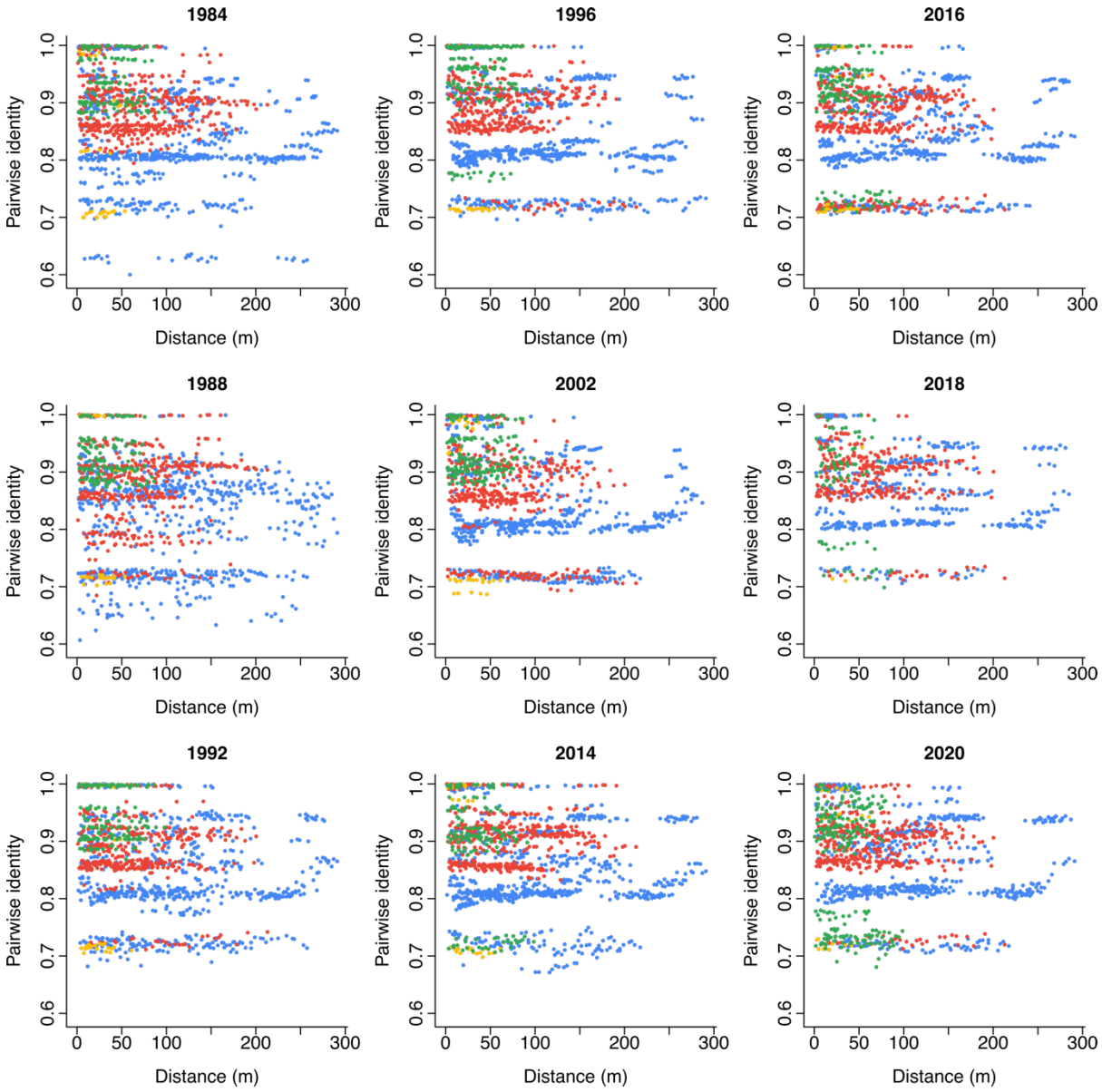

Figure S7. **No evidence for isolation by distance in Ammiad collection.** Scatter plots of pairwise genetic identity and Euclidean distance (meters). Subplots show years of collection, and colors represent the different transects (A=blue, B=red, C=yellow, D=green).

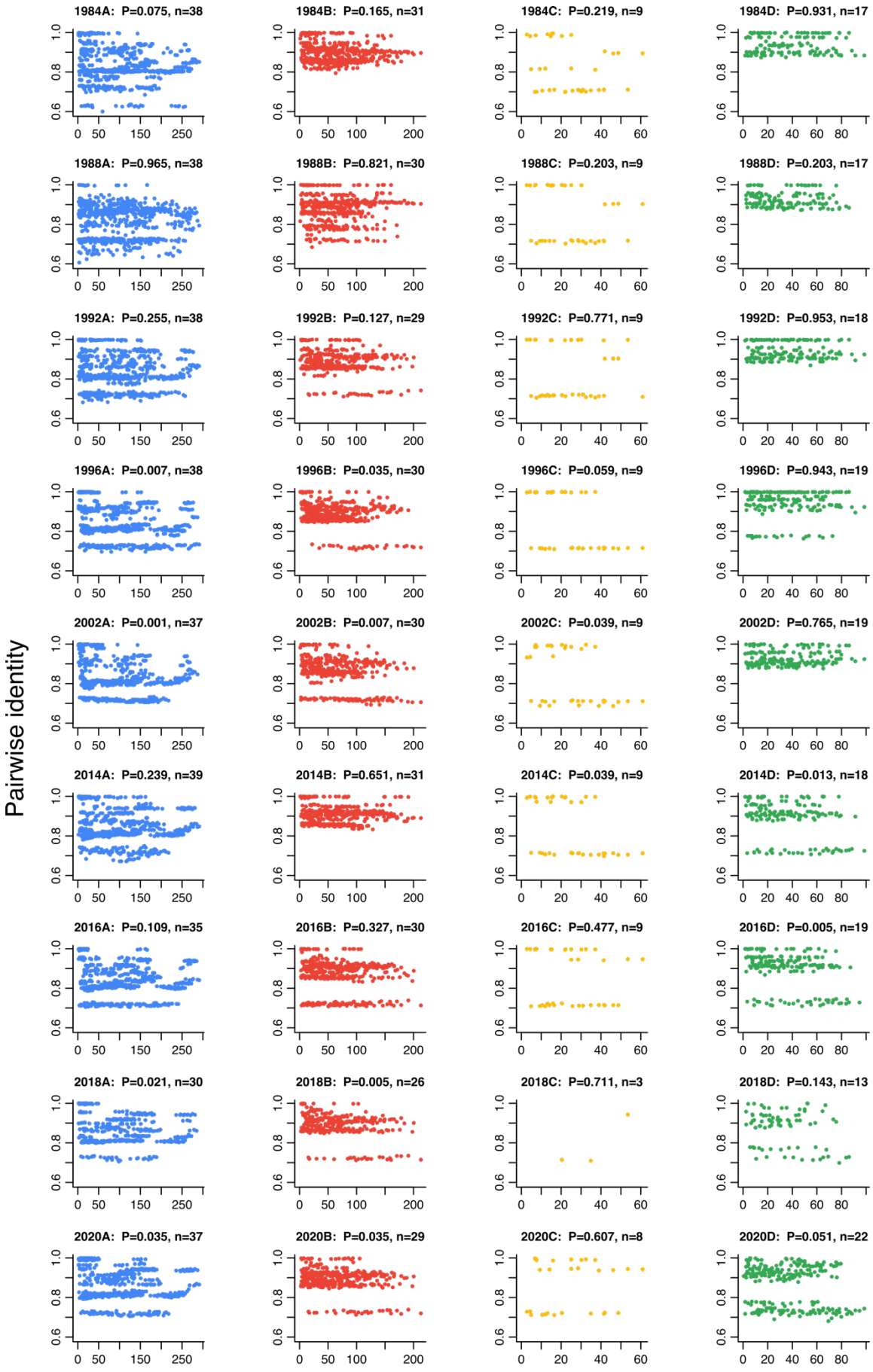

Figure S8. **Correlation and significance of isolation by distance in Ammiad collection.** Pairwise genetic identity by Euclidean distance (meters). Subplots show years of collection (rows) and transects (columns). The title of each plot indicates its number of samples and p-value from a Mantel test with 1000 permutations.

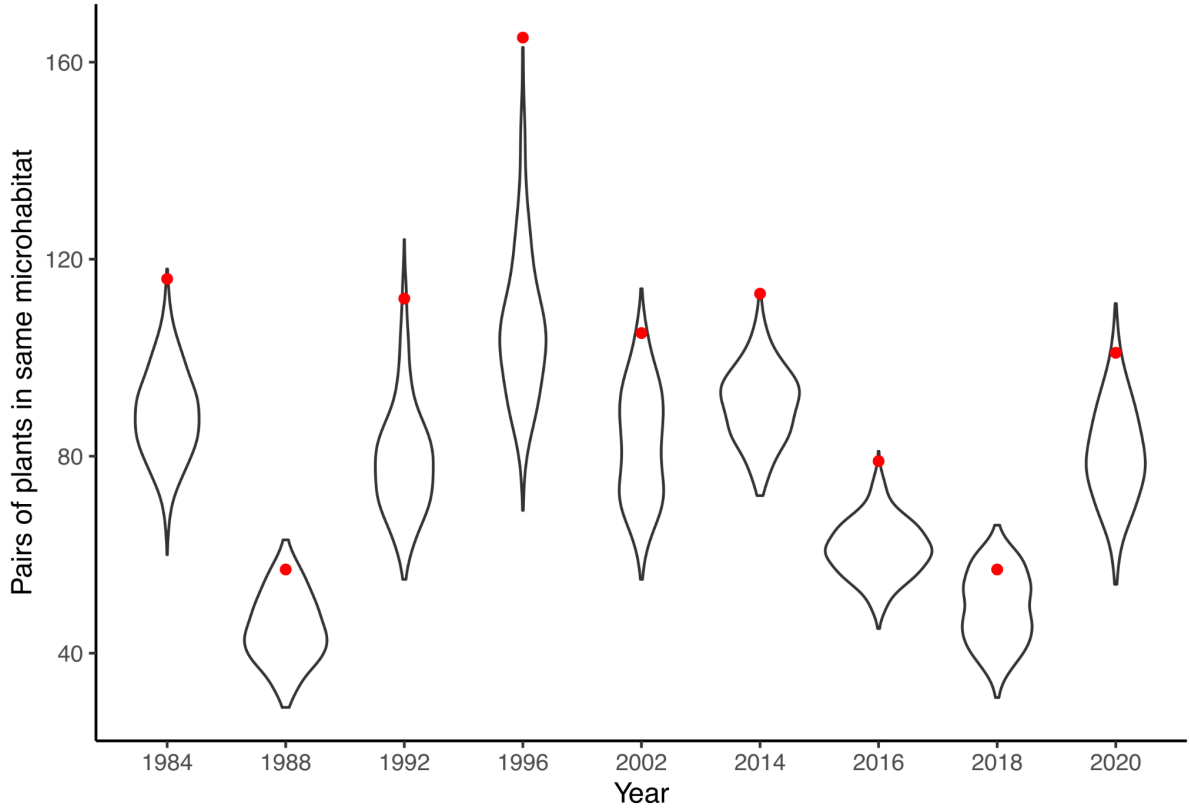

Figure S9. **Sorting of plants of the same DGG by microhabitat comparing observed values (red) to values obtained by permuting habitat labels (violins).**

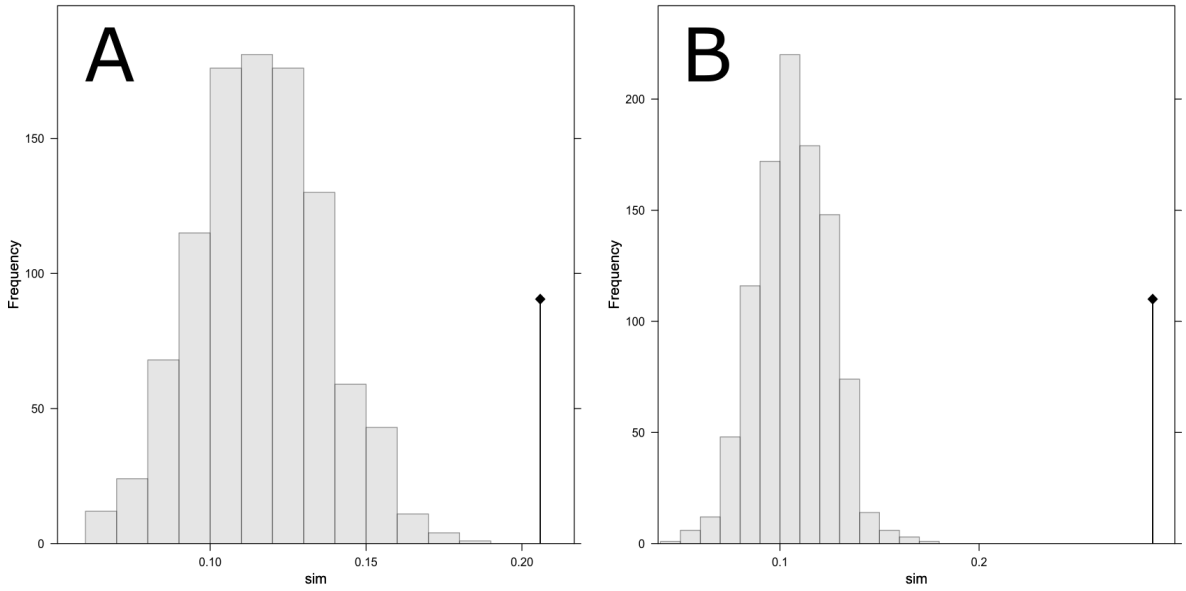

Figure S10. **Histograms of 1000 spatially constrained randomizations of Mantel statistics (MSR-Mantel).** Observed values (black dots) versus randomizations for (A) pairwise genetic distances (p-value < 0.001) and (B) DGGs (p-value < 0.001).

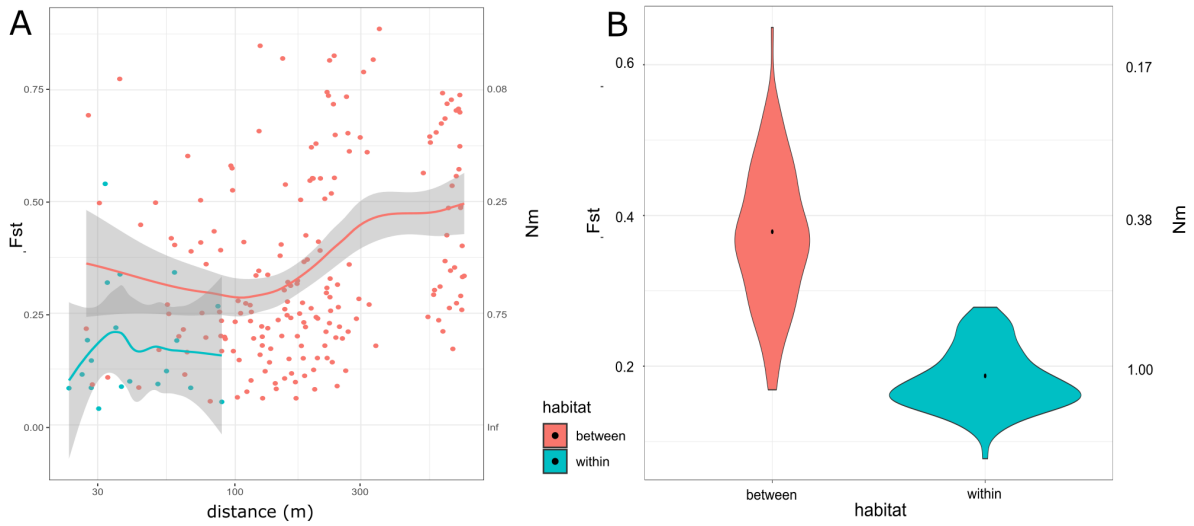

Figure S11. **Genetic differentiation between and within habitats, accounting for spatial distance.** Plants within each habitat were grouped in 30m demes. Pairwise F_{st} was calculated among all demes, within and between habitats. A) Pairwise F_{st} plotted along distance, measured as the average distance of plants in the two demes. Credibility intervals of a loess fit with degree 1 are shown in grey. B) Pairwise F_{st} between demes with average distance <50m. Average observed values are shown as a black dot. Uncertainty was estimated with 1000 bootstraps of the parcels considered for the comparisons. Average within-habitat Fst was higher than between-habitat F_{st} in 3% of the bootstraps. For both A and B, Nm values (i.e. the theoretical number of migrants between parcels in an infinite island model) calculated are shown on the right axis.

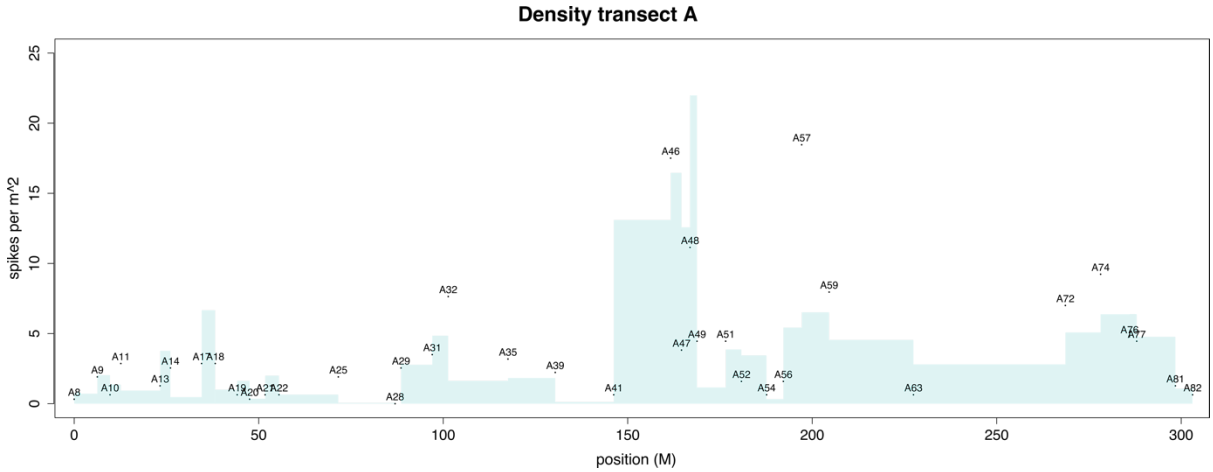

Figure S12. **Spike density in transect A.** Spike density (corresponding to density of fertile plants) was measured around and between sampling sites (Supplementary Table 4). Points represent the number of fertile spikes per 1 m^2 around the sampling point. Light blue bars represent normalized (to area) number of spikes from one sampling point to the following.

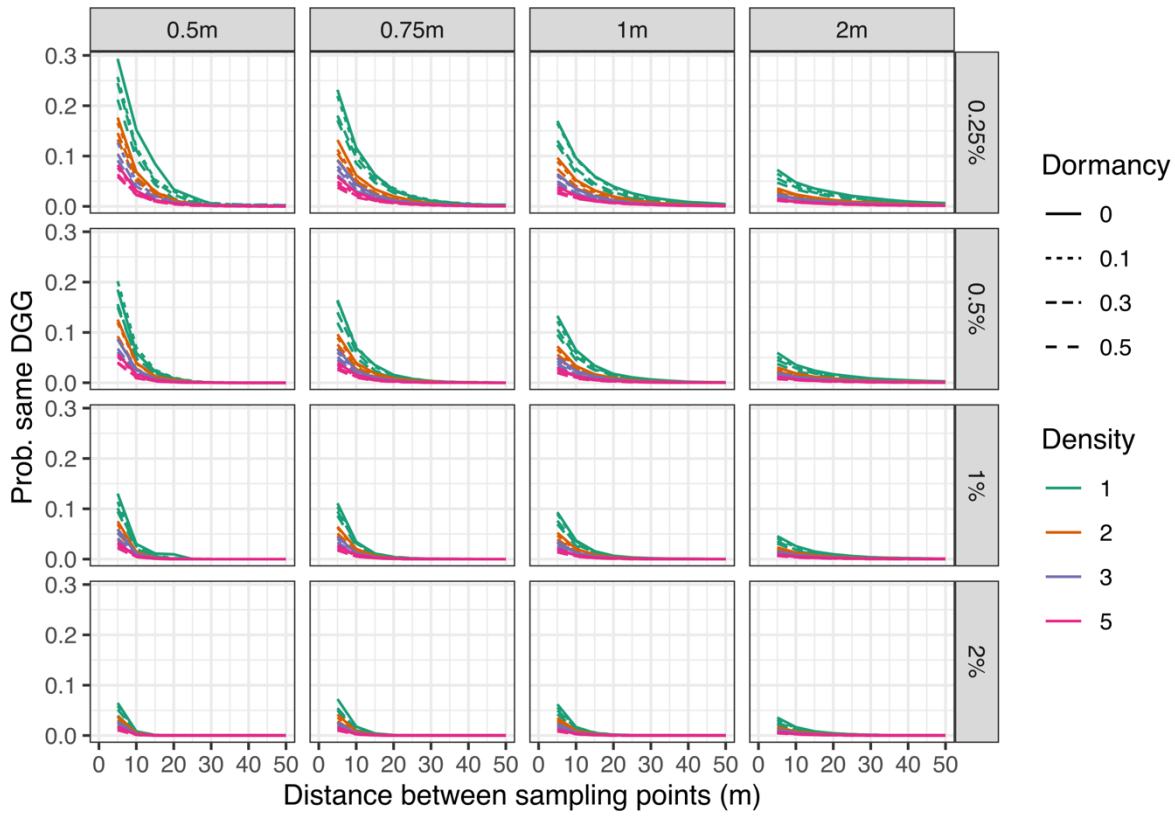

Figure S13. **Spatial clustering along a transect in simulated populations.** Subplots show the probability of observing plants of the same DGG at pairs of sampling points at increasing distances, averaged over 100 replicate simulations. Panels indicate increasing mean seed dispersal distances (left to right) and increasing outcrossing rates (top to bottom). 'Dormancy' indicates the proportion of seeds drawn from the seed bank in each generation, and 'Density' the number of plants per square meter.

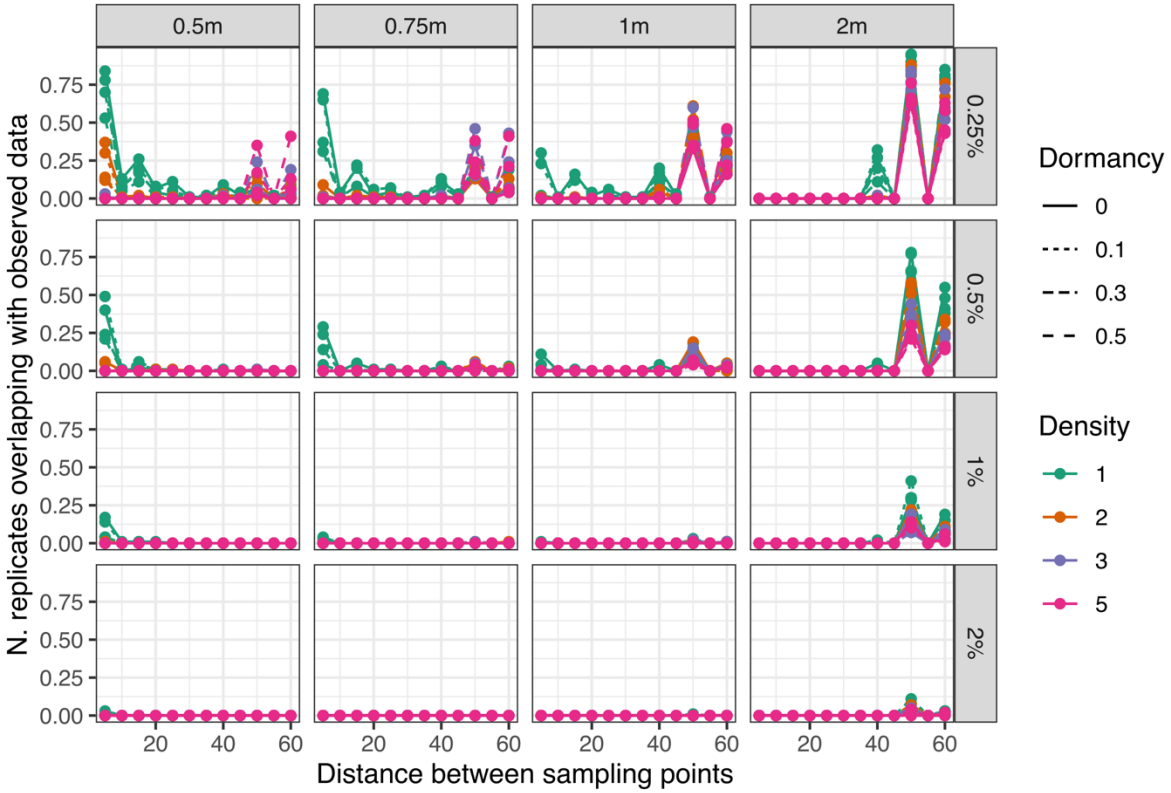

Figure S14. **Overlap in spatial clustering between observed data and populations simulated under neutral demography.** Plots show the proportion of replicate simulations where pairs sampling points at different distances were more likely to be occupied by plants of the same DGG than in any observed transect. Panels indicate increasing mean seed dispersal distances (left to right) and increasing outcrossing rates (top to bottom). 'Dormancy' indicates the proportion of seeds drawn from the seed bank in each generation, and 'Density' the number of plants per square meter.

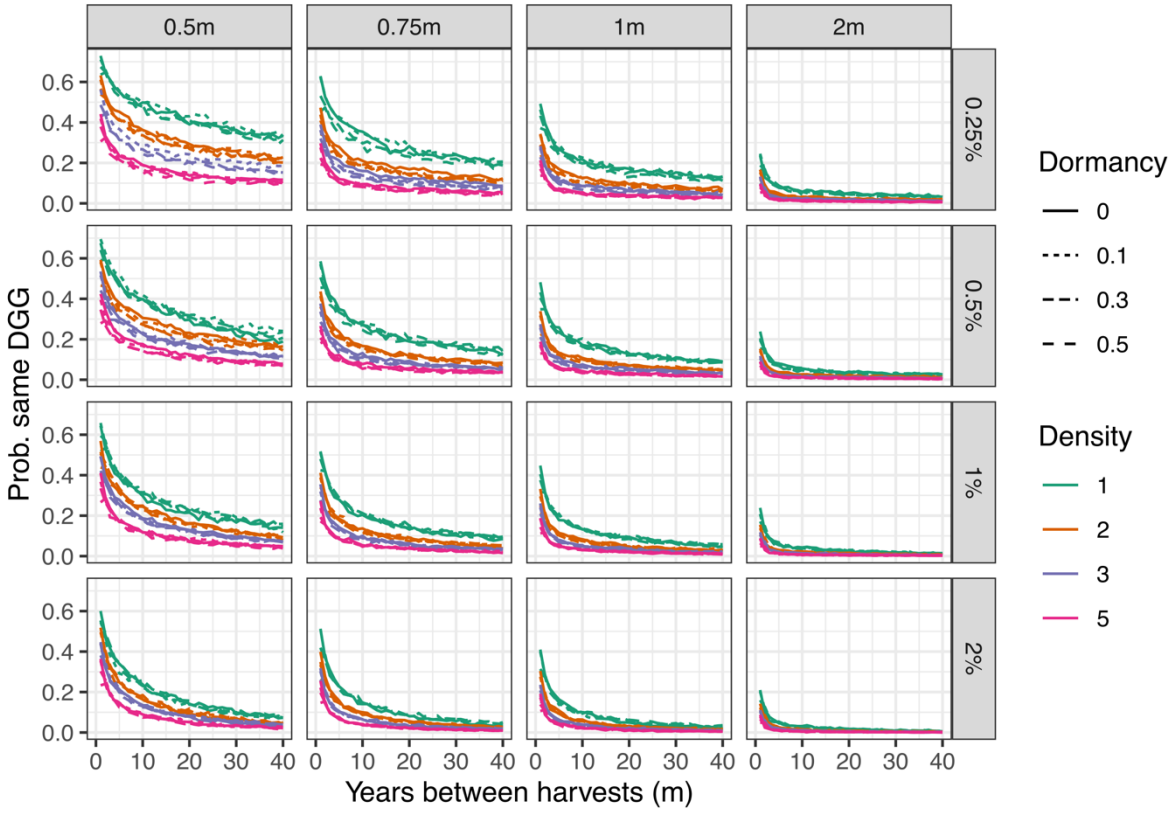

Figure S15. **Temporal stability in simulated populations.** Subplots show the probability of observing plants of the same DGG at a single sampling in different years, averaged over 100 replicate simulations. Panels indicate increasing mean seed dispersal distances (left to right) and increasing outcrossing rates (top to bottom). 'Dormancy' indicates the proportion of seeds drawn from the seed bank in each generation, and 'density' the number of plants per square meter.

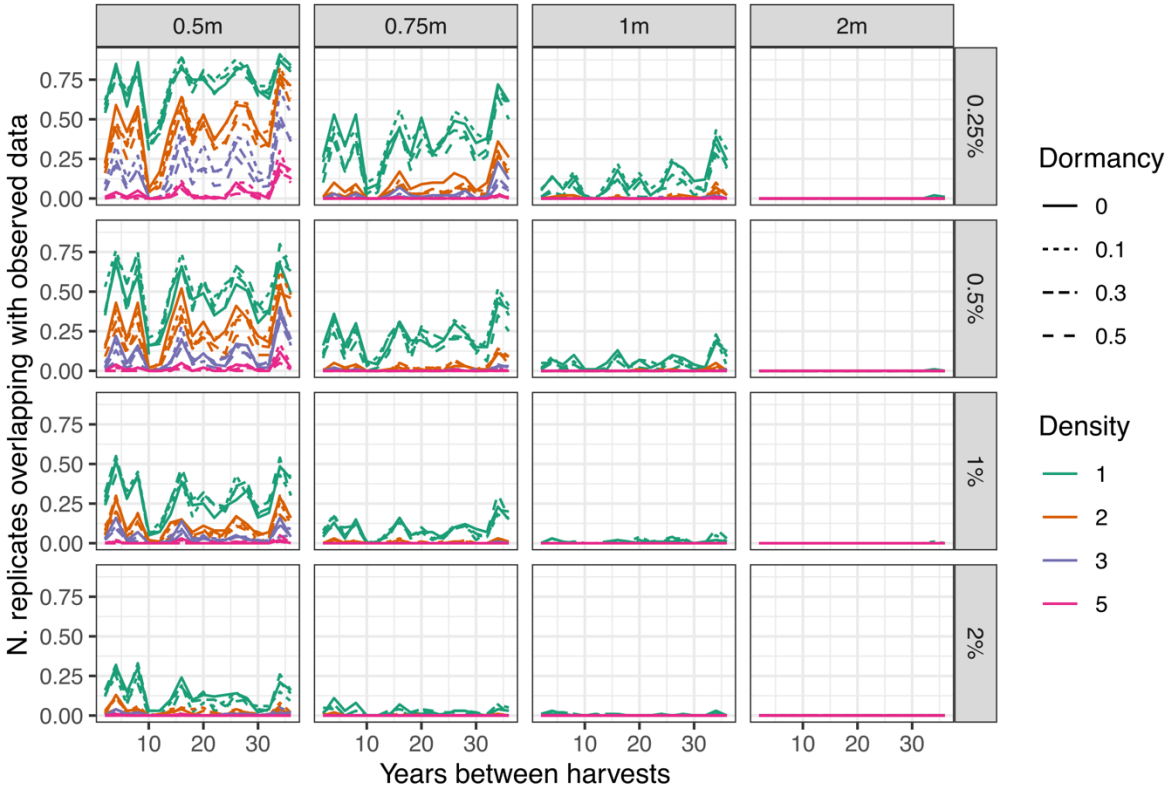

Figure S16. **Probabilities that observed temporal stability could be generated by neutral demography.** Plots show the proportion of replicate simulations where a single sampling point sampled in different harvest years was more likely to be occupied by plants of the same DGG than in any observed transect. Panels indicate increasing mean seed dispersal distances (left to right) and increasing outcrossing rates (top to bottom). 'Dormancy' indicates the proportion of seeds drawn from the seed bank in each generation, and 'Density' the number of plants per square meter.

Supplementary Tables

Supplementary Table 1. **Ammiad sampling points coordinates.**

Supplementary Table 2. **Dataset of Ammiad samples by genotype.**

Supplementary Table 3. **Analysis of molecular variance of genotypes sampled in Ammiad across all years, subdivided by habitat, using only polymorphic sites within Ammiad.** Sum of squared deviations (SSD) and mean square deviations (MSD), variance components ($\boldsymbol{\sigma^2}$) and corresponding p-values were computed using the R package *pegas*. The Φ statistic for habitat was used as an estimate of the proportion of molecular variance explained by the 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	/	/	/

Supplementary Table 4. **Spike density measured at transect A in 2020.**

References

1. Urban, M. C. Accelerating extinction risk from climate change. *Science* **348**, 571–573 (2015).
2. Hoffmann, A. A. & Sgró, C. M. Climate change and evolutionary adaptation. *Nature* **470**, 479–485 (2011).
3. Pecl, G. T. *et al.* Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* **355**, (2017).
4. Khoury, C. K. *et al.* Crop wild relatives of the United States require urgent conservation action. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33351–33357 (2021).
5. Wintle, B. A. *et al.* Global synthesis of conservation studies reveals the importance of small habitat patches for biodiversity. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 909–914 (2019).
6. Keeling, C. D. *et al.* Exchanges of Atmospheric CO2 and 13CO2 with the Terrestrial Biosphere and Oceans from 1978 to 2000. I. Global Aspects. *I. Glob. Asp. SIO Ref. Ser.* Scripps In, **88** (2001).
7. Levy, A. A. & Feldman, M. Ecogeographical distribution of HMW glutenin alleles in populations of the wild tetraploid wheat Triticum turgidum var. dicoccoides. *Theor. Appl. Genet.* **75**, 651–658 (1988).
8. Felsenburg, T., Levy, A. A., Galili, G. & Feldman, M. Polymorphism of high-molecular-weight glutenins in wild tetraploid wheat: spatial and temporal variation in a native site. *Isr. J. Bot.* **40**, 451–479 (1991).
9. Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R. & Martin, W. Genetics and geography of wild cereal domestication in the near east. *Nat. Rev. Genet.* **3**, 429–441 (2002).
10. Avni, R. *et al.* Wild emmer genome architecture and diversity elucidate wheat

evolution and domestication. *Science* **357**, 93–97 (2017).

11. Nave, M. *et al.* Wheat domestication in light of haplotype analyses of the Brittle rachis 1 genes (BTR1-A and BTR1-B). *Plant Sci.* **285**, 193–199 (2019).
12. Anikster, Y. & Noy-Meir, I. The wild-wheat field laboratory at Ammiad. *Isr. J. Bot.* **40**, 351–362 (1991).
13. Poland, J. A., Brown, P. J., Sorrells, M. E. & Jannink, J.-L. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. *PLoS One* **7**, e32253 (2012).
14. Fu, Y.-B. *et al.* Elevated mutation and selection in wild emmer wheat in response to 28 years of global warming. *Proc. Natl. Acad. Sci.* **116**, 20002–20008 (2019).
15. Franks, S. J., Sim, S. & Weis, A. E. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc. Natl. Acad. Sci.* **104**, 1278–1282 (2007).
16. Ozbek, O., Millet, E., Anikster, Y., Arslan, O. & Feldman, M. Spatio-temporal genetic variation in populations of wild emmer wheat, Triticum turgidum ssp. dicoccoides, as revealed by AFLP analysis. *Theor. Appl. Genet. 2007 1151* **115**, 19–26 (2007).
17. Volis, S., Anikster, Y., Olsvig-Whittaker, L. & Mendlinger, S. The influence of space in genetic-environmental relationships when environmental heterogeneity and seed dispersal occur at similar scale. *Am. Nat.* **163**, 312–327 (2004).
18. Horovitz, A. & Feldman, M. Evaluation of the wild-wheat study at Ammiad. *Isr. J. Bot.* **40**, 501–508 (1991).
19. Consortium, T. 1001 G. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166**, 481–491 (2016).
20. Noy-Meir, I., Agami, M., Cohen, E. & Anikster, Y. Floristic and ecological differentiation of habitats within a wild wheat population at Ammiad. *Isr. J. Bot.* **40**,

363–384 (1991).

21. Golenberg, E. M. Outcrossing rates and their relationship to phenology in Triticum dicoccoides. *Theor. Appl. Genet.* **75**, 937–944 (1988).

22. Horovitz, A. The soil seed bank in wild emmer. in *The Proceedings of International Symposium of In Situ Conservation of Plant Genetic Diversity* 185–188 (1998).

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

24. Golenberg, E. M. Estimation of gene flow and genetic neighborhood size by indirect methods in a selfing annual, Triticum dicoccoides. *Evolution* **41**, 1326–1334 (1987).

25. Nevo, E. Evolution in action across life at "Evolution Canyons", Israel. *Trends Evol. Biol.* **1**, e3–e3 (2009).

26. Siepielski, A. M., Dibattista, J. D. & Carlson, S. M. It's about time: The temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* **12**, 1261–1276 (2009).

27. Zsögön, A., Peres, L. E. P., Xiao, Y., Yan, J. & Fernie, A. R. Enhancing crop diversity for food security in the face of climate uncertainty. *Plant J.* (2021).

28. Elshire, R. J. *et al.* A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS One* **6**, e19379 (2011).

29. Zhu, T. *et al.* Improved Genome Sequence of Wild Emmer Wheat Zavitan with the Aid of Optical Maps. *G3 (Bethesda)*. **9**, 619–624 (2019).

30. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 1303.3997 (2013).

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

32. Poplin, R. *et al.* Scaling accurate genetic variant discovery to tens of thousands of

samples. *bioRxiv* 201178 (2018).

33. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

34. Bradbury, P. J. *et al.* TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).

35. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326 (2012).

36. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).

37. Paradis, E. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).

38. Legendre, P., Fortin, M. J. & Borcard, D. Should the Mantel test be used in spatial analysis? *Methods Ecol. Evol.* **6**, 1239–1247 (2015).

39. Crabot, J., Clappe, S., Dray, S. & Datry, T. Testing the Mantel statistic with a spatially-constrained permutation procedure. *Methods Ecol. Evol.* **10**, 532–540 (2019).