ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Repeated migration, interbreeding and bottlenecking shaped the phylogeography of the selfing grass *Brachypodium stacei*

Miguel Campos[1,2] | Ernesto Pérez-Collazos[1,2] | Antonio Díaz-Pérez[1,3,4] | Diana López-Alvarez[1,5] | Ali Oumouloud[6] | Luis A. J. Mur[7] | John P. Vogel[8] | Pilar Catalán[1,2]

[1]Departamento de Ciencias Agrarias y del Medio Natural, Escuela Politécnica Superior de Huesca, Universidad de Zaragoza, Huesca, Spain

[2]Grupo de Bioquímica, Biofísica y Biología Computacional (BIFI, UNIZAR), Unidad Asociada al CSIC, Zaragoza, Spain

[3]GESPLAN S.A. C, Las Palmas de Gran Canaria, Spain

[4]Instituto de Genética, Facultad de Agronomía, Universidad Central de Venezuela, Maracay, Venezuela

[5]Facultad de Ciencias Agropecuarias, Departamento de Ciencias Biológicas, Universidad Nacional de Colombia, Palmira, Colombia

[6]Institute Agronomique et Vétérinaire Hassan II, Agadir, Morocco

[7]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, UK

[8]Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA

Correspondence
Pilar Catalán, Departamento de Ciencias Agrarias y del Medio Natural, Escuela Politécnica Superior de Huesca, Universidad de Zaragoza, C/ Carretera de Cuarte Km 1, E-22071 Huesca, Spain.
Email: pcatalan@unizar.es

Funding information
Spanish Ministry of Science and Innovation, Grant/Award Number: TED2021-131073B-I00 and PDC2022-133712-I00; Gobierno de Aragon, Grant/Award Number: A01-23R

Handling Editor: Ana Caicedo

Abstract

Brachypodium stacei is the most ancestral lineage in the genus *Brachypodium*, a model system for grass functional genomics. *B. stacei* shows striking and sometimes contradictory biological and evolutionary features, including a high selfing rate yet extensive admixture, an ancient Miocene origin yet with recent evolutionary radiation, and adaptation to different dry climate conditions in its narrow distribution range. Therefore, it constitutes an ideal system to study these life history traits. We studied the phylogeography of 17 native circum-Mediterranean *B. stacei* populations (39 individuals) using genome-wide RADseq SNP data and complete plastome sequences. Nuclear SNP data revealed the existence of six distinct genetic clusters, low levels of intra-population genetic diversity and high selfing rates, albeit with signatures of admixture. Coalescence-based dating analysis detected a recent split between crown lineages in the Late Quaternary. Plastome sequences showed incongruent evolutionary relationships with those recovered by the nuclear data, suggesting interbreeding and chloroplast capture events between genetically distant populations. Demographic and population dispersal coalescent models identified an ancestral origin of *B. stacei* in the western-central Mediterranean islands, followed by an early colonization of the Canary Islands and two independent colonization events of the eastern Mediterranean region through long-distance dispersal and bottleneck events as the most likely evolutionary history. Climate niche data identified three arid niches of *B. stacei* in the southern Mediterranean region. Our findings indicate that the phylogeography of *B. stacei* populations was shaped by recent radiations, frequent extinctions, long-distance dispersal events, occasional interbreeding, and adaptation to local climates.

KEYWORDS
Brachypodium model species, Mediterranean and Macaronesian aridic niches, phylogeography, plastome capture, recent radiation, repeated dispersals and extinctions

1 | INTRODUCTION

The Mediterranean region is a hotspot of biodiversity (Médail & Diadema, 2009; Mittermeier et al., 2004) and has been used as a natural laboratory to study the phylogeography of many plants (Thompson, 2007). Several Mediterranean plant lineages emerged after the divergence of some proto-Mediterranean ancestors in the Oligocene–Miocene and further radiated after the last glacial and postglacial phases of the Quaternary (Feliner, 2011; Vargas et al., 2018). Although the onset of the Mediterranean climate has been dated to the early Pliocene (Suc, 1984), the radiations of many Mediterranean plants occurred earlier during the Miocene (Thompson, 2007; Vargas et al., 2018). While some of the ancestral diploid species have survived until the present (Vargas et al., 2020), many of them experienced hybridization and wide genome duplication events while others went extinct during glacial times (Kadereit & Abbott, 2021; Mandáková & Lysak, 2018). Biogeographical models and dated transcriptome-based phylogenies suggest that *Brachypodium* originated in the Holarctic region in the Mid–Late Miocene (Díaz-Pérez et al., 2018). This suggests that the Mediterranean region was colonized by an ancestral *Brachypodium* species with a karyotype ($x = 10$), similar to that of extant diploid *B. stacei*, in the Late Miocene (Sancho et al., 2022). Analysing the phylogeography of relict Miocene survivors, such as *B. stacei*, can provide insight into the evolutionary framework of the radiations experienced by many plants in the Mediterranean region over the last several million years.

The phylogeographical patterns of Mediterranean plants are influenced by their life cycle and reproductive strategies (Feliner, 2011; Thompson, 2007). The aridification of the circum-Mediterranean region at the end of the Miocene likely favoured the development of annual self-fertile species, since in stressful environments selfing is reproductively more efficient than outbreeding (Solbrig, 1976). Due to inbreeding, selfing species have fewer haplytypes than outcrossing species within populations (Holsinger, 2000). Thus, selfing species tend to have lower intrapopulation diversity and greater interpopulation differentiation than outcrossing species due to the increased genetic drift experienced by selfers (Charlesworth, 2003). Despite the genetic drawbacks imposed by inbreeding, selfers usually adapt to their ecological niches by shedding deleterious or maladaptive genetic load through selection (Greer et al., 2023; Roessler et al., 2019; Szövényi et al., 2014; Zeitler et al., 2023). Selfing annual grasses can spread and colonize new areas within a few generations because population establishment can occur after a single individual is dispersed to a new location (Mairal et al., 2023). Selfing could alter population demography and ecology by reducing effective population size and therefore genetic variation (Barrett et al., 2014; Charlesworth, 2003). Evidence suggests that other effects like selective sweeps and more rapid selection in selfing species could accelerate the loss of genetic diversity in selfing populations (Barrett et al., 2014). However, some of the genetic effects caused by selfing can be counteracted by occasional outcrossing, since complete selfing is extremely rare in nature (Wright et al., 2013).

Brachypodium stacei is one of the three circum-Mediterranean annual species in the otherwise perennial genus *Brachypodium* that serves as a model system for grass biology, physiology and functional genomics (Scholthof et al., 2018). *B. stacei* shows notable biological and evolutionary characteristics, such as a high selfing rate with signatures of genomic admixture (Mu, Li, Yang, Breiman, Lou, et al., 2023), an old Miocene origin (Sancho et al., 2022) with recent radiation in some of its populations (Mu, Li, Yang, Breiman, Lou, et al., 2023; Shiposha et al., 2016), and a rapid adaptation to the variable arid climate conditions of the Mediterranean (Lopez-Alvarez et al., 2015; Mu, Li, Yang, Breiman, Lou, et al., 2023). The diploid *B. stacei* ($2n = 2x = 20$, $x = 10$) is, together with the diploid *B. distachyon* ($2n = 2x = 10$, $x = 5$), a progenitor species of the allotetraploid *B. hybridum* ($2n = 4x = 30$, $x = 10 + 5$) (Catalán et al., 2012). This trio of annual Mediterranean species has been selected as a model to investigate recurrent origins and functional consequences of allopolyploidy (Hasterok et al., 2022; Mu, Li, Yang, Breiman, Yang, et al., 2023). Population genetic studies have shown that all three annual taxa are highly selfing species (Mu, Li, Yang, Breiman, Lou, et al., 2023; Shiposha et al., 2016, 2020; Vogel et al., 2009). Despite this fact, population genomics of the closest relative *B. distachyon* has identified several cases of chloroplast capture events between divergent populations (Gordon et al., 2020; Sancho et al., 2018). This suggests that long distance dispersal (LDD) and recurrent crossings can increase gene flow in this predominantly selfing species. Other studies, however, have postulated that migration without interbreeding can explain the existence of mixed populations of *B. distachyon* formed by individuals of different origins that grow together in the same environment (Stritt et al., 2022). The occurrence of mixed populations harbouring individuals from distinct evolutionary lineages was previously recorded for *B. distachyon* and *B. hybridum* (Gordon et al., 2020; Shiposha et al., 2020; Vogel et al., 2009). Furthermore, sympatric populations of two or even the three annual *Brachypodium* species are not rare in some western Mediterranean sites (Catalan, López-Álvarez, Díaz-Pérez, et al., 2016; Scholthof et al., 2018). This evidence suggests that frequent dispersal and local adaptive success may quickly reshape the genotypic composition of the annual *Brachypodium* populations.

A randomized GBS study of over eighteen hundred genotypes of the three annual *Brachypodium* species in its native Mediterranean region recovered an extremely low percentage of *B. stacei* accessions (3%), compared to those of its congeners, supporting its general rarity (Wilson et al., 2019). Environmental niche model analysis also detected the smallest niche range and smallest niche breadth of the three taxa for *B. stacei* (Lopez-Alvarez et al., 2015). Although considerable genomic resources have been generated for *B. distachyon*, a functional model system for monocots (Gordon et al., 2017; Hasterok et al., 2022; Minadakis et al., 2023; Scholthof et al., 2018; Stritt et al., 2022; Tyler et al., 2016), and for *B. hybridum*, a model species for allopolyploids (Gordon et al., 2020; Mu, Li, Yang, Breiman, Yang, et al., 2023; Scarlett et al., 2022), comparative genomic studies have only recently started for *B. stacei* (Mu, Li, Yang, Breiman, Lou, et al., 2023).

In this study, we investigated the evolutionary history of the poorly studied *B. stacei* in its native circum-Mediterranean region, aiming to unravel the underlying mechanisms that have driven the genetic composition and distribution of its populations. Using an integration of genome-wide nuclear RADseq SNP data and whole-plastome sequence data, we estimated the impact of LDD, potential interbreeding, structure and admixture on genotypic diversity. We tested alternative demographic models and reconstructed time-measured phylogenies to explore the temporal spans of population divergences and dispersals, and the potential effects of bottlenecks that contributed to shape the genetic constitution of populations. We also conducted environmental niche modelling (ENM) to uncover the current and past environmental niches to determine the potential range shifts in the adaptive success of populations to their niches. Our complementary phylogeographic and ecological studies have shed light on the evolutionary trajectory of the phylogeographically unexplored selfing species *B. stacei* and may provide insight about general mechanisms applicable to other plants.

2 | MATERIALS AND METHODS

2.1 | Plant material

A total of 17 *Brachypodium stacei* populations, distributed across its native circum-Mediterranean region, were selected for the phylogeographic study (Figure 1, Table S1). The sampling included 39 individuals of *B. stacei* (one to three individuals per population) plus two individuals of its close species *B. distachyon*, used as outgroups (Table S1). The taxonomic identity of individuals was confirmed through DAPI-stained chromosome counting ($2n=20$), discarding previously misidentified samples from SW Asia (Iran; Lopez-Alvarez et al., 2015). To avoid potential inter-population hybridizations during sample processing, seeds collected from natural populations and germplasm banks (Table S1) were grown in standard greenhouse conditions (16-h light period at 20°C) at different institutions [High Polytechnic School of Huesca – University of Zaragoza (Spain), Aberystwyth University (UK), Joint Genome Institute (USA)] in different years and for different genome sequencing experiments, following the procedures indicated in Catalán et al. (2012). Seven additional *B. stacei* populations retrieved from reliable records (López-Álvarez et al., 2017) were added to the initial sampling and used for environmental niche reconstruction (totalling 24 populations) (Figure 1, Table S1).

2.2 | DNA extraction, RADseq library preparation and generation of nuclear and plastome genome data

Total genomic DNA was obtained from leaf tissue using a slightly modified Wizard® Genomic DNA Purification Kit protocol from Promega (including the addition of proteinase K (20mg/mL)). For nuclear genome sequencing RADseq libraries were prepared using
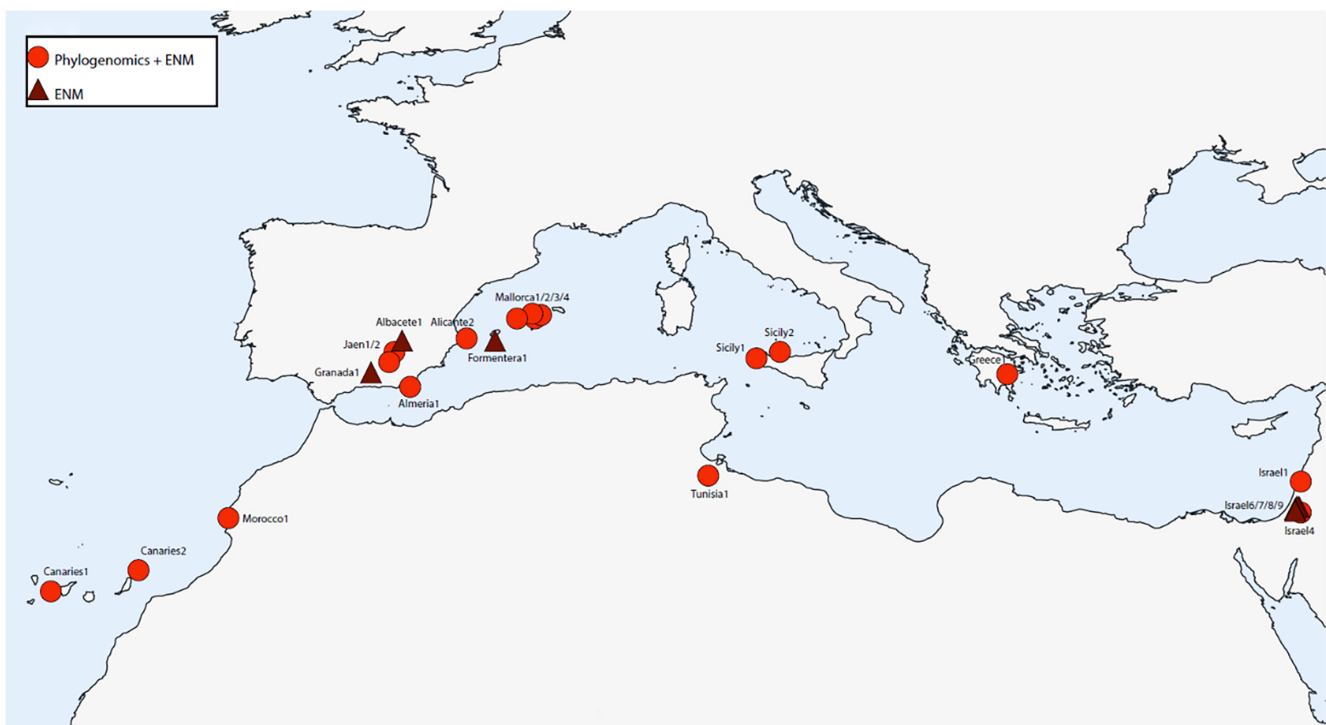

FIGURE 1 Geographical distribution of the studied *Brachypodium stacei* populations in its native circum-Mediterranean region. In circles, populations used for phylogeographic analysis with RAD-Seq data and, in triangles, additional populations used in environmental niche modelling analysis.

the protocol of Baird et al. (2008) with modifications from Peterson et al. (2012). For each sample, 0.1 μg of genomic DNA was digested with PstI (New England Biolabs, UK) and ligated to P1 barcode adaptors (25 nM). Ligated DNA was sheared using an ultrasonic sonicator for 15 cycles of 30 s on, followed by 30 s off; the size range was selected using the LabChip instrument for a range of 250 ± 10% bp. After end-repair and A-tailing of the size-selected DNA fragments, they were ligated to P′ adapters (2 nM) and PCR amplified for six independent reactions of 25 μL, using the following cycling parameters: 98°C for 30 s followed by 15 cycles of 98°C for 10 s, 65°C for 60 s and 75°C for 30 s, and a final cycle of 72°C for 5 min. After PCR, the library was cleaned using Ampure beads. This cleaning procedure was also used for all subsequent reactions. Final RAD libraries were quantified using the Qubit® dsDNA BR Assey and Qubit® dsDNA HS Assey kits (Life Technologies Corporation, Caribad, CA) and were pooled in equimolar proportions for sequencing in a single lane on an Illumina Hiseq2000 platform at the Institute of Biological, Environmental and Rural Sciences (IBERS) Genomics Core Facility at Aberystwyth University.

Single-end read lengths were de-multiplexed and filtered to obtain a SNP matrix using Ipyrad (Eaton & Overcast, 2020). We employed the reference approach using the *Brachypodium stacei* reference genome to filter and map the SNPs (ABR114; https://genome.jgi.doe.gov/portal/BrastaABR114_FD/BrastaABR114_FD.download.html). To mitigate the effects of linkage disequilibrium and refine the data set for subsequent genetic and evolutionary analyses, we filtered out plastid loci, monomorphic alleles, alleles with frequency <0.005, and alleles with a missing rate >0.2, with subsequent pruning for LD using a slide window of 10,000 bp and a threshold of 0.2 in SNPRelate in R (Mijangos et al., 2022; Sievert, 2020). Ipyrad recovered a total output of 281,273 SNPs, which were reduced to a final data set of 72,450 SNPs after all successive filtering steps (Table S2A). To explore the proportional distribution of selected SNPs on the *Brachypodium stacei* chromosomes we used SRplot (Tang et al., 2023), dividing each chromosome into multiple bins and calculating the number of SNPs located within each bin. All the SNPs were similarly distributed and densely mapped along the 10 chromosomes of *B. stacei* (Figure S1). The final data set of 72,450 filtered SNPs was used for all subsequent population genetic, phylogenetic and demographic analyses.

Chloroplast (plastome) DNA data were obtained from whole genome resequencing data. Total DNA was isolated using CTAB protocol with modifications (Doyle & Doyle, 1987). Sequence data for libraries were generated using Illumina technology. An Illumina library was constructed and sequenced using the Illumina NovaSeq S4 platform (2x151) at the DOE Joint Genome Institute (JGI) and duplicate reads were removed based on paired sequence matching using Clumpify in BBMap (Bushnell, 2014). BBDuk v.38.96 from BBMap was used to trim reads containing adapter sequence and G homopolymers of size ≥5 at the ends of the reads, and quality trim reads where quality was <6. BBDuk was used to remove reads that contained one or more 'N' bases, had an average read quality score <6, or had a minimum length ≤49 bp or 33% of the total read length. Whole plastome sequences (LSC, SSC, IRA and IRB) were assembled with Novoplasty 4.3.3 (Dierckxsens et al., 2017) (Table S2B) using the reference plastome of *B. stacei* ABR114 as seed (accession LT558589).

2.3 | Population genomics and phylogeographic analyses

Maximum likelihood (ML) phylogenomic trees were constructed for the unlinked nuclear SNP data set and the whole plastome sequence data set using IQtree (Nguyen et al., 2015), imposing the best substitution model selected by the program for each separate data set and 1000 ultrafast bootstrap replicates. A Bayesian genomic structure model-based analysis of nuclear data was performed to infer population structure and spatial ranges of *B. stacei* by estimating individual ancestry with ADMIXTURE v.1.3.0 (Alexander et al., 2009). This software alternately updates allele frequency and ancestry fraction parameters by using a block relaxation approach for multilocus SNP genotype data sets. One to 10 hypothetical genetic groups (K = 1–10) were evaluated with 10 iterations for each K. The hypothetical genetic group with the lowest cross-validation error was selected as the best K (Alexander et al., 2009). In addition, we also employed StructureSelector (Li & Liu, 2018), a method based on six alternative statistics (medmedk, medmeak, maxmedk, maxmeak, Ln Pr(X|K) and ΔK), to determine the optimal number of hypothetical genetic groups. This tool integrates Bayesian model-based clustering techniques (Raj et al., 2014), including variational inference from large SNP data sets and estimators for addressing uneven sampling (Puechmaille, 2016; Raj et al., 2014). Admixture results for each K value were plotted using the Pophelper package in R (Francis, 2017) and the best K-based genetic groups were plotted alongside the nuclear ML phylogenomic tree using the iTOL online tool (Letunic & Bork, 2021). A principal coordinates analysis (PCoA) was computed from a pairwise SNP distance matrix generated by the dartR and Plotly packages in R v 4.0.5 (Mijangos et al., 2022; Sievert, 2020) to reveal relationships between individuals.

Population genetic diversity was calculated using VCFtools (Danecek et al., 2011) and Arlequin v3.5.2.2 (Excoffier & Lischer, 2010). We estimated the observed (H_o) and the expected heterozygosity (H_e) values and the inbreeding coefficient (F_{is}) values at population, geographical zone and genomic (Admixture) group levels. Furthermore, the selfing rate (s) was estimated as $s = 2F_{is}/(1 + F_{is})$ (Ritland, 1990).

To test the significance of genetic variability within and among populations and between groups, we performed analyses of molecular variance (AMOVA) with Arlequin v.3.5.2.2. We computed a standard AMOVA, and two different hierarchical AMOVAs for (i) 10 sampled geographical groups [Sicily, Spanish regions (Mallorca, Jaen, Almeria, Alicante and Canary Islands), Greece, Israel, Morocco and Tunisia] and (ii) six genomic groups retrieved from ADMIXTURE [Canaries, M + S (Mallorca + Sicily), Greece, Israel, Jaen and WestMed (West Mediterranean: Alicante + Almeria + Morocco + Tunisia)] (see Section 3). We performed the AMOVAs with 10,000 permutations to quantify the variance among groups.

We assessed the putative occurrence of gene flow between populations using the formula $F_{st} \approx 1/(4N_m + 1)$ as an indirect measurement of the migration rate between populations with Arlequin v3.5.2.2. Additionally, we also calculated the extent of gene flow using TreeMix and custom scripts (Milanesi et al., 2017; Pickrell & Pritchard, 2012; Zecca et al., 2020, https://github.com/carolindahms/TreeMix), inferring the history of population splits and admixture events from allele frequency data, choosing the optimum number of migrations based on multiple linear models. We tested from 0 to 10 migration events using a bootstrap of 1000 replicates and 10 repetitions for each migration event.

As the chloroplast genome is considered non-recombinant in plants (but see Sancho et al., 2018 for putative heteroplasmy and recombination of the *B. distachyon* plastomes), we employed DAPC (discriminant analysis of principal components) to infer plastome clusters in *B. stacei* using the adegenet package in R (Jombart, 2008; Jombart & Collins, 2015). DAPC identifies genetic groups based on non-redundant data sets by combining principal component analysis (PCA) and linear discriminant analysis (LDA). The program employs Ward's hierarchical clustering algorithm and Akaike information criterion (AIC) to select the optimal number of clusters.

To obtain a time-measured phylogeny of *B. stacei* we used two different molecules and approaches. First, we applied a coalescent-based method, accounting for potential ILS, employing the nuclear SNP data matrix and the SNAPP package (Bryant et al., 2012) implemented in BEAST2 (Bouckaert et al., 2014). A Ruby script (snapp_prep.rb) was used to prepare the input file for SNAPP (https://github.com/mmatschiner/snapp_prep), feeding it with the SNP data matrix, a table linking species IDs and specimen IDs, and a file specifying age constraints (see Supplementary Information and data files at https://github.com/Bioflora/Bstacei_Phylogeography). We merged individuals into populations and used *B. distachyon* as an outgroup to match the plastome samples, and imposed a normal distribution for a secondary age constrain on the *Brachypodium* crown node (mean = 11.6 Ma, SD = 1.0) following Gordon et al. (2020). We also imposed OneOnX priors on the clock rate and the speciation rate (lambda), a uniform prior on the theta parameter with bounds from 0 to 10,000, and a strict clock rate, assuming negligible branch rate variation between closely related samples. The analysis was conducted using 2,500,000 MCMC generations with a sampling frequency of 500 generations, discarding the first 10% of steps as burn-in. TRACER v.1.6 (http://beast.bio.ed.ac.uk/Tracer) was used to verify the adequacy of parameters, showing an effective sample size >200 for all parameters. A maximum clade credibility tree was computed after discarding 10% of the respective saved trees as burn-in. Second, we applied a Bayesian dating analysis using the whole plastome sequence data matrix, considered as a single heritable unit, and estimated the nodal divergence times of the lineages using BEAST2 (Bouckaert et al., 2014). We imposed normal distributions for secondary calibrations for the crown nodes of Pooideae (mean = 33.2 Ma, SD = 9.52) and Brachypodium (mean = 11.6 Ma, SD = 1.0), following Sancho et al. (2018), a uniform prior for

nucleotide frequencies, a gamma prior for substitution rates, a lognormal distribution for branch rates of a relaxed clock model, and an exponential prior for the Yule birth rate speciation model. We ran 500,000,000 MCMC generations, with a sampling frequency of 1000 generations, and a 10% burn-in.

Demographic model selection analysis that takes into account the potential existence of incomplete lineage sorting (ILD) was performed using coalescence-based Approximate Bayesian Computation and supervised machine learning methods implemented in DIYABC-RF (Random Forest) v.1.1.1 (Collin et al., 2021). Through this approach we statistically tested alternative scenarios for the dispersal and demographic history of the *B. stacei* populations. We defined six main genomic groups of *B. stacei* identified by the Admixture analysis and according to geographical range [(i) Mallorca+Sicily (M+S): Mallorca1, 2, 3, 4, Sicily1, 2; (ii) Canaries: Canaries1, 2; (iii) Ibero-Maghrebian group (WestMed): Alicante2, Almeria1, Morocco1, Tunisia1; (iv) Jaen: Jaen1, 2; (v) Israel: Israel1, 4; and (vi) Greece: Greece1] and assessed six competing scenarios for the dispersal and demographic history of the *B. stacei* groups based on our previous population genomic results [(1) West-to-East independent dispersals, including M+S→EastMed (Greece), M+S→Canaries, M+S→EastMed (Israel) and M+S→WestMed(+Jaen) dispersals; (2) West–to-East dispersal and vicariance; (3) East–to-West dispersal and vicariance; (4) West-Canarian split plus one West-to-East (M+S→Israel(+Greece) early dispersal; (5) West-Canarian split plus two independent West-to-East (M+S→Greece, M+S→Israel) early dispersals; and (6) West-Canarian split plus one West-to-East (WestMed(+Jaen)→Israel (+Greece) recent dispersal]. For all scenarios, training sets were generated using 10,000 simulations per model. Due to the lack of ancestral divergence time information, we set the prior distribution of tree divergence times between the oldest and most recent splits of the coalescent-dated tree and the effective population sizes equal for each scenario. The most likely scenario was determined using the RF module of DIYABC-RF, computing a forest of 10,000 random trees and using the median with a 90% credibility interval as a point estimate. The model with the highest classification vote and the highest posterior probability was considered the most suitable for the target data set among the compared models.

After selecting the best dispersal scenario, we conducted a second DIYABC-RF analysis, imposing a new set of priors for alternative bottlenecking events (0–5 bottlenecks) on the optimal dispersal model, testing all the possible combinations. Global and local error rates were evaluated using RF regression, and these values were used to calculate the posterior probability of the selected scenario (Pudlo et al., 2016). After selecting the optimal number of bottlenecks in the best dispersal scenario, we estimated parameter values by running 10,000 simulations for the final selected model to infer time and populations size parameters.

2.4 | Environmental niche modelling analysis

We performed an ENM analysis of *B. stacei* to test the potential shifts of climatic niche ranges in the last glacial phases and how it

could have affected the genetic and phylogeographical patterns of its populations adapted to different Mediterranean regions with contrasting climatic features. We used occurrence data from a total of 24 circum-Mediterranean populations of *B. stacei*, increasing the sampling from the previous study of Lopez-Alvarez et al. (2015) with new populations (Figure 1, Table S1), and excluding previously misidentified samples.

The ENM analysis was conducted through the maximum entropy algorithm implemented in R package Maxnet v.0.1.4 (Phillips et al., 2021). Nineteen bioclimatic variables from current climate were extracted from the Bioclim dataset in a GIS-based raster format at 2.5 min resolution using WorldClim 1.4. We performed species distribution modelling, projection of realized niches to maps, and analysis of niche evolution employing 12 R packages (Devtools v.2.4.5; Dplyrv.1.0.10; Ecospatv.3.4; ENMTools v.1.0.7; HHv.3.1-49; Maptools v.1.1-6; Maxnet v.0.1.4; Raster v.3.6-14; Rgdal v.1.6-4; Rgeosv.0.6-1; Vegan v.2.6-4; Dismov.1.3-9). To evaluate the potential effect of the Quaternary glacial and interglacial cycles on the expansions or contractions of the *B. stacei* niche, we also constructed ENMs for the climatic conditions of the Last Interglacial (LIG, ~120,000–140,000 ya; Otto-Bliesner et al., 2006), Last Glacial Maximum (LGM, ~22,000 ya) and the Mid Holocene (MH, ~6000 ya), using palaeoclimatic data from the 'Community Climate System Model' at 2.5 min resolution (CCSM; Gent et al., 2011) and assuming that the species' climate-driven biological requirements were the same in the past as today. Niche breadth for each time window, and Schoener's *D* statistic were calculated to assess niche breath and overlap (Cardillo & Warren, 2016; Warren et al., 2010). The most suitable climatic variables were selected for the ENM, discarding those highly correlated according to the variance inflation factor. To reduce sampling bias, we overlaid 100 independent Maxnet models and their climate data (bootstrap analysis). For each modelling run, 75% of the records were used as training data and 25% as test data. We evaluated model performance based on the area under the curve (AUC) value, which discriminates a species model from a random one (>0.9 good support, <0.7 poor support; Lawson et al., 2014). Species distribution models at each time interval were exported in HDR format and mapped with QGISv.3.22 (QGIS Development Team, 2023). In addition, a multivariate ordination PCA was carried out with R v.4.0.5 using current bioclimatic variables obtained from WorldClim to detect potential bioclimatic groups within geographical ranges of the *B. stacei* populations. The PCA was conducted with a Pearson correlation matrix and the visualization of the output was done through 2D plotting.

3 | RESULTS

3.1 | Late Pleistocene divergence of eight main lineages and six genomic groups of *Brachypodium stacei*

The ML phylogenomic tree based on 72,450 unlinked nuclear SNP data recovered eight major lineages that showed bootstrap support (BS) values greater than 86% in all branches except the most recently evolved lineages that showed moderate BS values (Figure 2a). The optimal population genomic clustering retrieved six genomic groups (ADMIXTURE lowest cross-validation errors for $K=6$ (0.572); Figure 2a, Figure S2a; Table S3), revealing a high degree of genetic structure among populations of *B. stacei*. The *B. stacei* ML tree showed successive divergences from the oldest population lineage of western Mediterranean Balearic island Mallorca4, followed by those of Mallorca2, and Mallorca1, the eastern Mediterranean population of Greece1, the sister western-central Mediterranean islands Mallorca3 and Sicily (Sicily1/Sicily2) populations, the Canary Islands (Canaries1/Canaries2) populations, the sister eastern Mediterranean populations of Israel (Israel1/Israel4), and the recently evolved western Mediterranean populations of the Ibero-Maghrebian region (Almeria1, Alicante2, Morocco1 and Tunisia1) and Jaen (Jaen1/Jaen2) (Figure 2a). The six genomic groups corresponded to populations from Mallorca + Sicily (M + S), Canaries, Greece, Israel, Spain-Maghreb (Alicante + Morocco + Tunisia) and Jaen (Figure 2a). Individuals of the Almeria population group showed signatures of admixture between its sister Ibero-Maghrebian lineage (45%), the close Jaen lineage (8.5%), and the less related and geographically distant Greece lineage (46.5%). A few individuals from other clusters also showed genomic admixture (Canaries2: 12% (Jaen); Mallorca4-1: 17% (Jaen) (Figure 2a).

The same subdivision of populations was evident in the PCoA (Figure 2b). The main PCoA axis 1 (PC1, 9.96% of accumulated variance) separated the Canarian and Jaen groups at, respectively, its negative and positive ends, and the second axis (PC2, 8.08%) Canaries and Jaen at the positive end and Greece at the negative end, from the remaining groups that clustered in the centre. The last groups separated along the third axis (PC3, 7.6%), showing close spatial affinities for, respectively, the Mallorca-Sicily (M + S), Ibero-Maghrebian and Israel groups, as retrieved in the phylogeny. The Almeria individuals also clustered with the Ibero-Maghrebian group in the PCoA plot (Figure 2b).

The coalescent-dated nuclear phylogeny of *B. stacei* inferred a very recent origin for its crown node (0.250 Ma) and young ages for the subsequent splits, which spanned the Late Pleistocene and the Holocene (0.233–0.003 Ma) (Figure S3a). This topology was congruent with that of the ML tree for some lineages but not for others. The main discordances between the two trees concerned the divergent splits of the Mallorcan, Canaries, Greece versus Israel, and Tunisia versus Ibero-Mahgrebian lineages in the coalescent-based tree compared to the ML tree (Figure 2a, Figure S3a).

3.2 | High selfing rates and low genetic diversity of the *B. stacei* populations

The estimated values of genetic diversity, measured by observed heterozygosity (H_o), were overall low in all the *B. stacei* populations, ranging from 0.005 (Canaries2) to 0.133 (Mallorca1) and with a mean of 0.041 (Table 1). The inbreeding coefficient values were high and positive, indicating that all populations and population
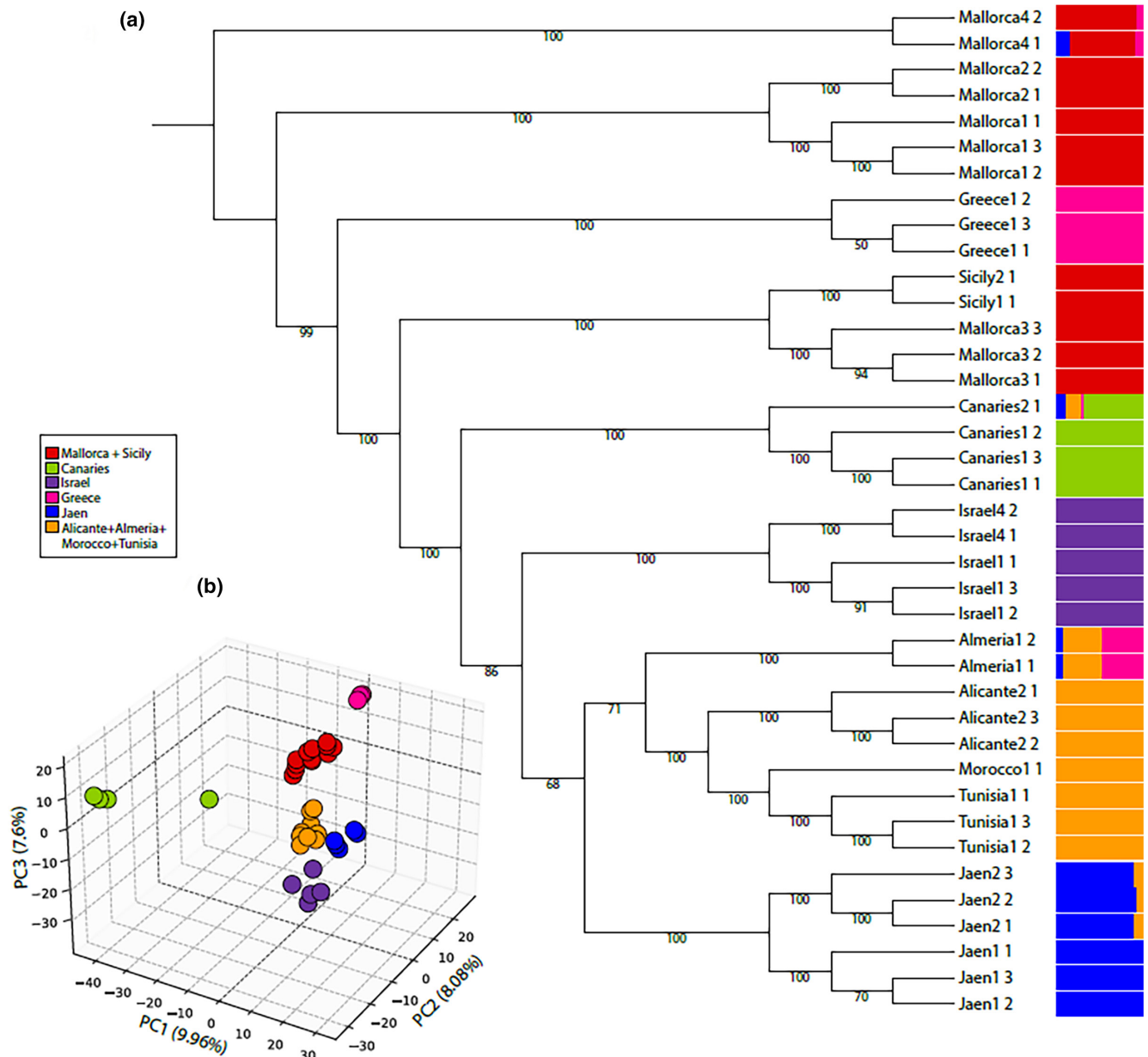
FIGURE 2 Phylogenomic tree and genome structure analysis of *Brachypodium stacei* populations based on nuclear SNPs. (a) Maximum likelihood phylogenetic tree inferred from 72,450 SNPs showing eight main lineages and bootstrap values. The plot of coloured bars displays percentages of membership of samples to six optimal Bayesian genomic groups [Mallorca + Sicily (red), Canaries (green), Greece (pink), Israel (purple), Jaen (blue), and Ibero-Maghrebian (orange)]. (b) Tridimensional plot of a principal coordinates analysis (PCoA) of the *B. stacei* population samples (colour codes correspond to the six genomic groups of Figure a). PC1 (9.96%), PC2 (8.08%), and PC3 (7.6%) axes explained 25.64% of the variance.

groups had excess of homozygotes (Table 1). Consistent with previous studies of this selfing species (Mu, Li, Yang, Breiman, Lou, et al., 2023; Shiposha et al., 2016), the estimated selfing rates were high for most populations and genomic groups, ranging from the lowest values in Sicily (0.35, Sicily1) to the highest in Canaries (0.976, Canaries2), and with an overall mean value of 0.777 (Table 1). The standard and hierarchical AMOVAs further corroborated the low percentages of variation within the population and the large genomic differentiation of the populations (Table S4); the geographical and genomic groups tested showed

considerably higher percentages of variation among populations than among groups (Table S4).

The estimated population divergences and migration rates showed contrasting differences between the western-central Mediterranean islands (Mallorca and Sicily) and the rest of the groups (Table 2). The levels of pairwise population differentiation (F_{st} values) were lowest for Mallorca and Sicily and highest for Greece and the other regions (especially with respect to the geographically distant Ibero-Maghrebian populations: Alicante, Almeria, Morocco and Tunisia) (Table 2a). Consequently, Mallorca exhibited the highest rate

TABLE 1 Genetic diversity of the studied populations, geographical and genomic groups of *Brachypodium stacei*. Estimated values of observed and expected heterozygosity (H_o, H_e), inbreeding coefficient (F_{is}) and selfing rate (s). The geographical groups were established according to 10 main geographical distribution zones of the populations and the genomic groups correspond to those of the best $K=6$ retrieved by ADMIXTURE (see text and Figure 2).

Population	H_o	H_e	F_{is}	s
Alicante2	0.017	0.105	0.836	0.906
Almeria1	0.011	0.106	0.894	0.944
Canaries1	0.063	0.106	0.409	0.578
Canaries2	0.005	0.106	0.953	0.976
Greece1	0.007	0.106	0.934	0.966
Israel1	0.036	0.106	0.661	0.733
Israel4	0.075	0.106	0.654	0.752
Jaen1	0.007	0.106	0.932	0.964
Jaen2	0.012	0.106	0.883	0.936
Mallorca1	0.133	0.105	0.387	0.534
Mallorca2	0.039	0.106	0.630	0.746
Mallorca3	0.050	0.106	0.526	0.675
Mallorca4	0.015	0.105	0.854	0.920
Morocco1	0.006	0.106	0.943	0.970
Sicily1	0.083	0.106	0.218	0.359
Sicily2	0.078	0.106	0.269	0.424
Tunisia1	0.053	0.106	0.496	0.661
Geographic zone				
Alicante	0.017	0.105	0.836	0.906
Almeria	0.011	0.106	0.894	0.944
Canaries	0.048	0.106	0.545	0.677
Greece	0.007	0.106	0.934	0.966
Israel	0.051	0.106	0.658	0.740
Jaen	0.010	0.106	0.907	0.950
Mallorca	0.066	0.106	0.570	0.696
Morocco	0.006	0.106	0.943	0.970
Sicily	0.080	0.106	0.244	0.391
Tunisia	0.053	0.106	0.496	0.661
Genomic group				
Canaries	0.048	0.106	0.545	0.677
Greece	0.007	0.106	0.934	0.966
Israel	0.051	0.106	0.658	0.740
Jaen	0.010	0.106	0.907	0.950
M+S	0.068	0.106	0.516	0.645
WMed	0.027	0.106	0.748	0.840
Overall				
B. stacei	0.041	0.106	0.683	0.777

of migrants with the other population regions (between 0.970 and 3.471 migrants per generation), especially with Sicily, while the lowest rates were observed between Greece and the rest (0.036–0.390)

(Table 2b). According to our N_m estimates, populations from Mallorca showed gene flow with almost all Mediterranean populations, and Sicily with the Canaries and Israel ($N_m > 1$ values), while gene flow tended to be low in all other pairwise comparisons (Table 2b). Complementing these findings, the TreeMix analysis identified four migration events based on the Piecewise and Bent cable linear models, assuming that migration rates may change over time (Sicily to the Canaries, Almeria to Jaen, Morocco to Almeria and Morocco to Tunisia) (Figure S4). The other two models tested, which assume constant migration rates over time (Simple exponential and non-linear least squares models) detected three migrations that were included within the previous set. The two approaches indicated that migration may have occurred between western Mediterranean mainland and island populations, especially between spatially close populations, and from the western Mediterranean region to the Canaries and the eastern Mediterranean, but not vice versa (Table 2b, Figure S4).

3.3 | Plastome phylogenomics and evidence of interbreeding in *B. stacei*

The 15 assembled *B. stacei* plastomes were, in general, very similar to each other (Table S2b). They ranged between 136,333 and 136,279 bp and showed the same structure and gene content characteristics as the reference plastome of *B. stacei* (ABR114) (Sancho et al., 2018). They contained 76 protein-coding genes, 38 tRNAs, four rRNAs, five pseudogenes and two hypothetical ORF, plus a ~1161 bp insertion between *psa*I and *rbc*L in the LSC region (including a *rpl*23 pseudogene) and a deletion of a *rps*19 copy between *psb*A and *trn*H in the IRb repeat with respect to the reference plastome of *B. distachyon* (Bd21) (Sancho et al., 2018). Only 138 polymorphic positions were found in the multiple sequence plastome alignment. The ML plastome tree showed a robust topology for major divergent lineages (BS > 92%) but with very short branches for the recent splits (Figure 3a). DAPC-based plastome genetic structure grouped the *B. stacei* plastomes into five genetic clusters, according to the Ward's best grouping and AIC statistics (Figure 3a, Figure S5).

The DAPC groups were compatible with the divergences of the plastome phylogeny lineages. In the plastome tree the crown node splits into two main clades, one containing the sister Canaries1 (DAPC group 1) and Jaen2/Almeria1 (DAPC group 2) lineages, and the other showing the successive splits of the Mallorca1/Sicily2 (DAPC group 3), the ((Canaries2/Israel), (Mallorca3/Tunisia)) (DAPC group 4), and the (Morocco1, (Mallorca4, (Mallorca2, (Sicily1, (Greece1/Alicante2))))) (DAPC group 5) lineages (Figure 3a, Figure S5). The plastome-based statistical parsimony haplotypic network (Figure 3b) showed a highly congruent topology with that of the plastome tree. The haplotypic network revealed that DAPC 1 and 2 haplotypes separated well from the rest and accumulated a high number of mutations in their respective stem and terminal branches, while DACP 3- and 5-type haplotypes had few shared and unique mutations and DACP 4-type haplotypes separated into the same two subgroups as the phylogenetic tree (Figure 3).

TABLE 2 (a) Pairwise divergence between the studied *Brachypodium stacei* populations based on fixation index (F_{st}) distances and (b) inferred rate of migrants per generation (N_m) between those pairs of populations.

(a)

F_{st}	Mallorca	Sicily	Canaries	Greece	Israel	Jaen	Alicante	Almeria	Morocco	Tunisia
Mallorca	–	–	–	–	–	–	–	–	–	–
Sicily	0.126	–	–	–	–	–	–	–	–	–
Canaries	0.279	0.325	–	–	–	–	–	–	–	–
Greece	0.340	0.564	0.591	–	–	–	–	–	–	–
Israel	0.244	0.327	0.402	0.562	–	–	–	–	–	–
Jaen	0.263	0.414	0.466	0.648	0.428	–	–	–	–	–
Alicante	0.244	0.396	0.457	0.732	0.410	0.473	–	–	–	–
Almeria	0.241	0.376	0.464	0.822	0.438	0.523	0.566	–	–	–
Morocco	0.332	0.435	0.499	0.933	0.501	0.618	0.625	0.753	–	–
Tunisia	0.282	0.376	0.448	0.697	0.429	0.500	0.465	0.542	0.573	–

(b)

No. migrants (nm)	Mallorca	Sicily	Canaries	Greece	Israel	Jaen	Alicante	Almeria	Morocco	Tunisia
Mallorca	–	–	–	–	–	–	–	–	–	–
Sicily	3.471	–	–	–	–	–	–	–	–	–
Canaries	1.293	1.038	–	–	–	–	–	–	–	–
Greece	0.970	0.386	0.346	–	–	–	–	–	–	–
Israel	1.553	1.029	0.742	0.390	–	–	–	–	–	–
Jaen	1.401	0.707	0.574	0.272	0.667	–	–	–	–	–
Alicante	1.546	0.762	0.595	0.183	0.719	0.557	–	–	–	–
Almeria	1.578	0.830	0.577	0.108	0.641	0.457	0.384	–	–	–
Morocco	1.004	0.650	0.502	0.036	0.497	0.308	0.300	0.164	–	–
Tunisia	1.272	0.829	0.615	0.218	0.665	0.501	0.576	0.423	0.373	–

However, the matrilineal plastome tree recovered a distinct evolutionary scenario for the divergence of the *B. stacei* populations than that reconstructed from the nuclear genome data (Figure 2a, Figure S3a). Furthermore, Bayesian-based age estimates of lineages from the plastome tree (Figure S3b) differed substantially from those from the nuclear tree (Figure S3a). In the plastome tree, the inferred age for the crown ancestor of the *B. stacei* clade dated back to the Late Miocene (5.921 Ma) and subsequent splits spanned from the early Pliocene (4.361 Ma) to the Late Pleistocene (0.446 Ma) (Figure S3b). The posterior estimate was earlier than the inferred age for the MRCA of *B. stacei* in the nuclear coalescent tree (0.256 Ma; Figure S3a). The large disagreement between the two dated trees and the dating analyses was likely a consequence of the high conservatism of the plastome data set and its low mutation rate, reflecting older estimates for all clades, in contrast to rapid coalescence of nuclear SNP alleles and the youngest estimates for all nodes of the tree (Figure S3a,b). Therefore, the nuclear tree coalescent dating analysis was used to interpret *B. stacei* phylogeography. We also detected major topological discordances between the plastome tree and the nuclear ML and coalescent trees (Figure S3a–c). Specifically, they corresponded to the different resolutions of the Canaries (1 and 2), Sicily (1 and 2), Mallorca (1 and 2) and Ibero-Mahgrebian (Almeria1, Alicante2, Morocco1 and Tunisia1)

population lineages, which were nested in divergent positions in the plastome tree but resolved as monophyletic in the nuclear ML tree or with distinct topological positions in the coalescent nuclear tree (Figures 2a and 3a, Figure S3a–c). The discordances between the plastome and nuclear topologies suggest the occurrence of frequent gene flow and interbreeding between largely divergent and geographically distant *B. stacei* populations.

3.4 | Colonization and demographic scenarios of *B. stacei*

For all DIYABC-RF analyses, the ABC PCA and the random forest LDA projections of simulated data and error metrics fit the observed data, indicating that the model choices were reliable (Figures S6 and S7). Summary statistics from the DIYABC-RF pipeline are shown in Appendix S1. Our first DIYABC-RF analysis supported model 5 as the dispersal scenario with the largest number of votes and the highest posterior probability (Figure 4a, Table S5a). The most likely scenario suggests an ancestral origin of *B. stacei* in the western-central Mediterranean islands, followed by an early colonization of the oceanic Canary Islands, and subsequent independent dispersals
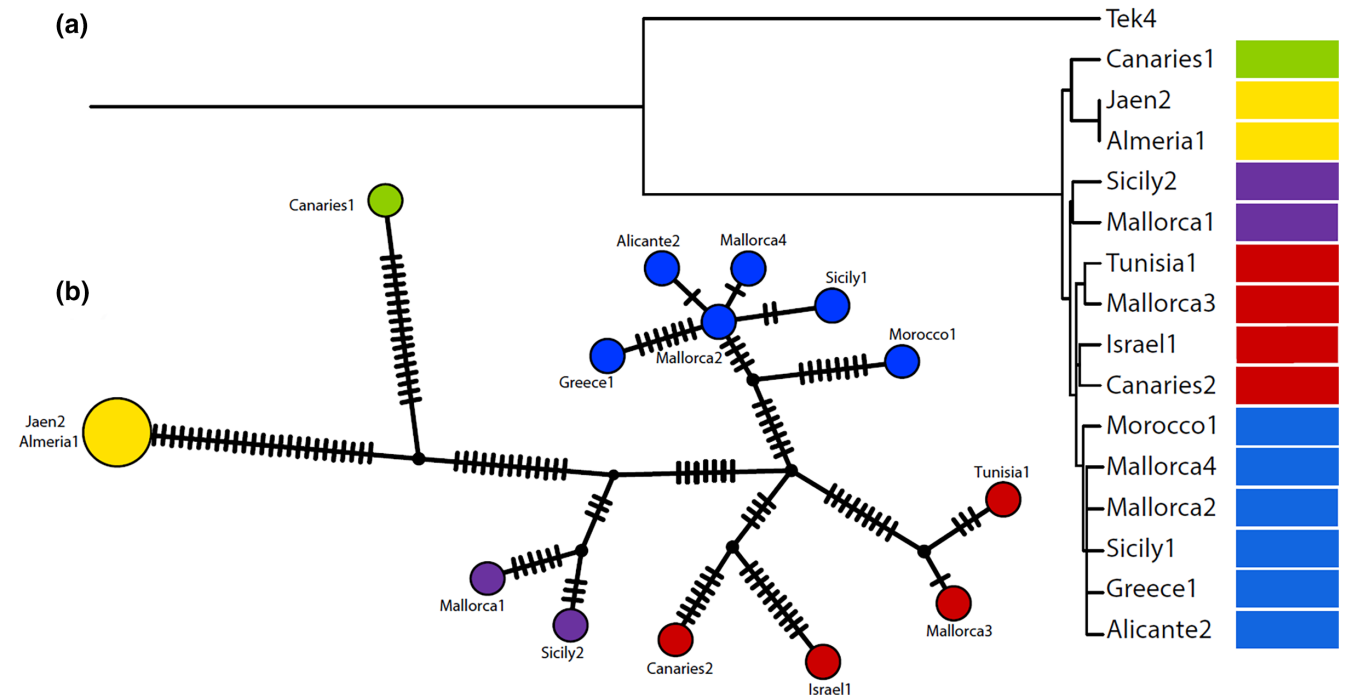
FIGURE 3 Evolution of plastomes of representative samples of the *Brachypodium stacei* populations studied. (a) Maximum likelihood phylogenomic tree based on whole plastome sequences. (b) Statistical parsimony haplotype network showing plastome mutations on branches. Coloured bars and circles correspond to the best five plastome genetic groups retrieved by DAPC.
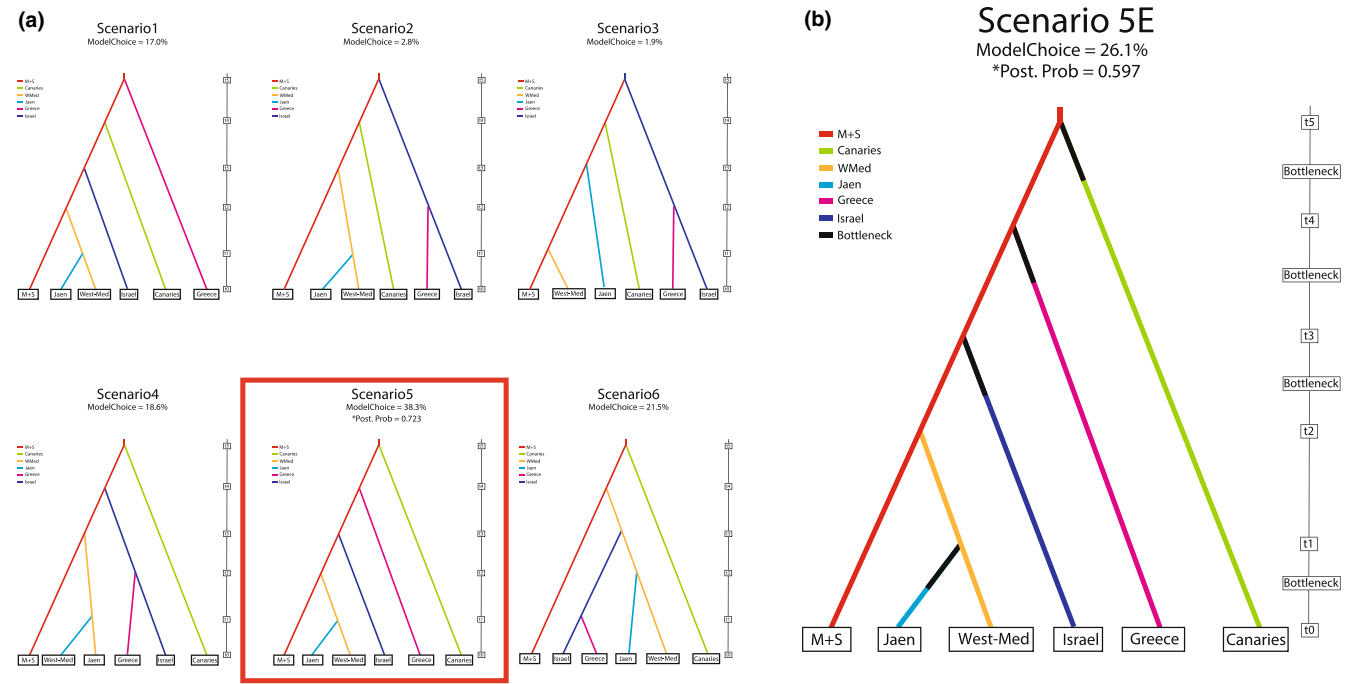

FIGURE 4 Graphical representation and posterior probabilities of (a) the six compared scenarios on the alternative dispersal routes and demographic models of the *Brachypodium stacei* population groups from the first DIYABC-RF analysis. Scenario 5 showed the highest posterior probability (red square). Population groups: M + S (Mallorca1, 2, 3 and 4 + Sicily1 and 2; red), Canary Islands (Canaries1 and 2; green), Ibero-Maghrebian (Alicante2, Morocco1 and Tunisia1; yellow), Jaen (Jaen 1 and 2; blue), Israel (Israel 1 and 4; navy) and Greece (Greece1; pink). The right scale shows the states for the time parameter (t5 > t4 > t3 > t2 > t1). (b) Best model selected after the second DIYABC-RF analysis for optimal bottleneck events (see Figure S8). Bottlenecks are indicated as black portions of branch lineages.

to the eastern Mediterranean ranges of Greece and Israel, and then to the continental Ibero-Maghrebian region, from which there was subsequent dispersal to Jaen (Figure 4a). Our second refined DIYABC-RF analysis, which imposed different bottlenecking events in the previous scenario, selected as best model the one in which all colonization experienced plausible bottlenecks except the WestMed group (Figure 4b, Figure S8, Table S5b). Our optimal model estimated the occurrence of the early dispersal from M + S to Canaries 89,852 (±37,266) generations (years) ago, and to the eastern Mediterranean areas of Greece and Israel 46,915 (±17,259) and 30,440 (±7209) ya respectively. The Ibero-Maghrebian region and Jaen were colonized, in turn, 28,489 (±6931) and 27,055 (± 6173) ya (Figure 4b, Table S5c). These demographic and temporal estimates for the different spatial dispersals are in general agreement with the young ages inferred from nuclear tree coalescent dating analysis (Figure S3a).

3.5 | Climatic niche models of *B. stacei*

Eighteen bioclimatic traits were used to define three main bioclimatic groups containing lines from the 24 analysed populations (Figure 5a, Table S6). The populations of these bioclimatic groups separated in the environmental space defined by the first two PCA axes

(Figure 5b). The 2D PCA plot differentiated an eastern and southern Mediterranean group (EastSouthMed: Israel, Greece and Tunisia) from a western and central Mediterranean group (WestCentralMed: Mallorca, S Spain and Sicily) along the positive and negative extremes of the PCA1 axis (explaining 40.44% of the variance) and a third Macaronesian group (Canaries and Atlantic Morocco) along the negative extreme of the PCA2 axis (26.61%) (Figure 5b). Seven of the 18 discriminant variables differentiated Macaronesia from WestCentralMed and EastSouthMed, four WestCentralMed from the others, and two EastSouthMed for the others (Table S6). These analyses indicate that the climatic conditions within the narrow distribution range of *B. stacei* populations vary considerably among these three main zones.

The current climatic niche of *B. stacei*, constructed with five uncorrelated variables (Figure S9), showed a high AUC value (0.98) and covered all sampled localities (Figure 5a), supporting its high predictive value. Our niche model improved on previous models (Lopez-Alvarez et al., 2015) by more accurately projecting the distribution range of *B. stacei* in the southern Mediterranean and Macaronesian aridic belt. Niche occupancy probabilities were higher in the lowland and coastal areas of the western circum-Mediterranean region, the Balearic, Tyrrhenian and Aegean islands, and several high-latitude locations in the eastern Mediterranean than in the northern Mediterranean mesic belt (Figure 5a).
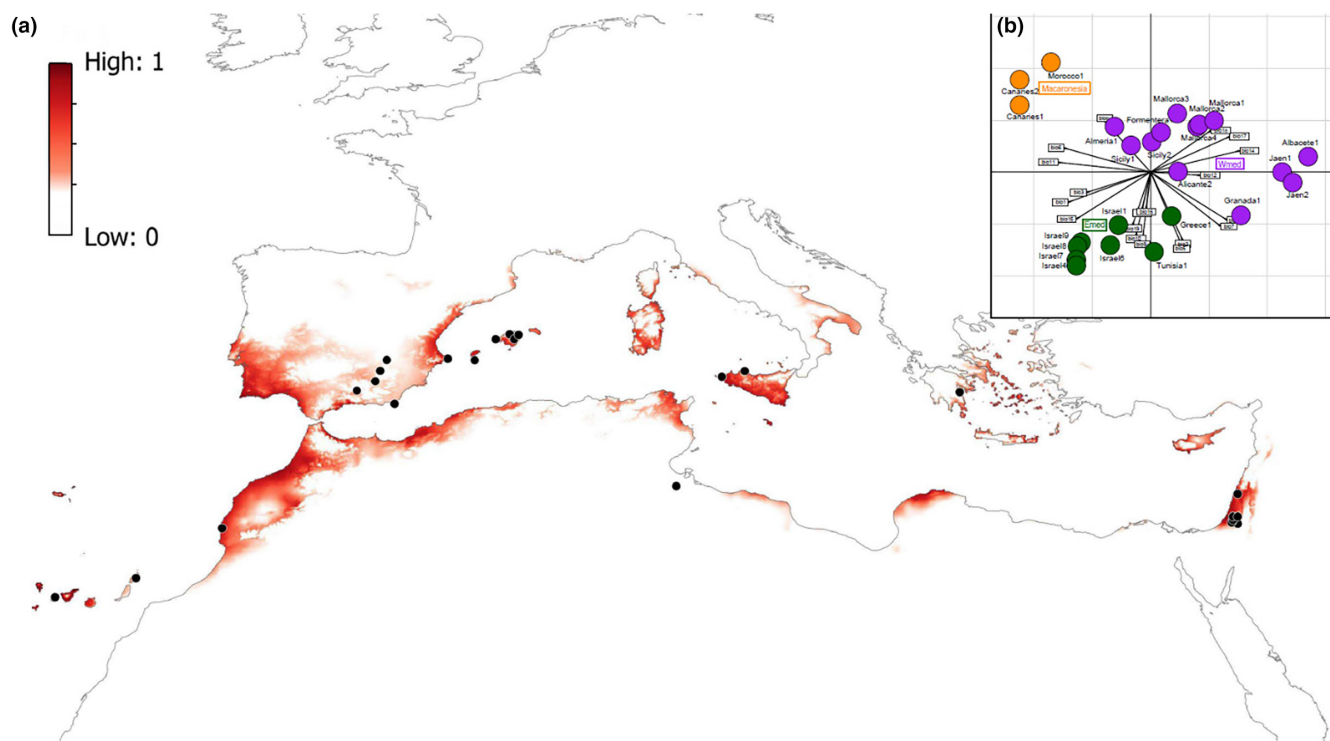

FIGURE 5 Environmental niche model analysis of *Brachypodium stacei* populations in its native circum-Mediterranean region. (a) Map projection of current species distribution niche based on five uncorrelated climate variables (see Figure S9) (lower 10% niche probability threshold trimmed). (b) Two-dimensional principal component analysis (PCA) plot of current climate variables of *B. stacei* showing the separate clusters of three main bioclimatic groups of populations (West Central Mediterranean (WestCentralMed, violet); East South Mediterranean (EastSouthMed, green); Macaronesian, orange). The PCA1 (40.44%) and PCA2 (26.61%) axes explained 67.05% of the total variance.

TABLE 3 *Brachypodium stacei* climate niche breadth calculated using Levin's B1 (inverse concentration) for four frame times and total area (in km²), and climate niche overlap between current and past times calculated with Schoener's *D* and the total overlapping area (%).

Niche breadth	Levin's B1	km²
Current time	0.7695	1,069,890
Mid Holocene	0.7688	1,148,848
Last Glacial Maximum	0.7708	1,150,993
Last Inter Glacial	0.7774	1,421,020
Niche overlap	Schoener's *D*	%
Current – MH	0.8602	87.26
Current – LGM	0.5656	52.83
Current – LIG	0.6895	85.16

Projection of current *B. stacei* occurrence data to past LIG, LGM and MH climate envelopes showed considerable range shifts and a general trend of interglacial (LIG, MH and present) niche contractions and glacial (LGM) niche expansion (Table 3) but with continuous niche occupancy in the southern Mediterranean belt since the LIG (Figure 5a, Figure S10). All past climate niche models also had high predictive power (AUC = 0.980). Niche breath values (Levin's B) were similar for all temporal climatic models; however, Schoener's *D* overlaps of the current niche were lowest with the LGM niche (Table 3).

4 | DISCUSSION

4.1 | Frequent extinctions, recent radiations and long-distance-dispersals shaped the evolutionary history of the selfing species *B. stacei*

Our study has uncovered the phylogeography of the poorly investigated model grass *Brachypodium stacei* throughout its native circum-Mediterranean region (Figures 1–4). In contrast to the large geographical distribution and clearly diverging patterns of the different eastern and western Mediterranean linages of its close congener *B. distachyon* (Gordon et al., 2017; 2020; Minadakis et al., 2023; Scholthof et al., 2018; Stritt et al., 2022; Tyler et al., 2016), *B. stacei* populations are relatively rare (Figure 1; Wilson et al., 2019), showing a restricted distribution in the southern Mediterranean belt (Figures 1 and 5a), and have radiated from the common ancestor a few thousand years ago (Figure S3a). The genetic diversity of *B. stacei* populations is low, although comparable to that of some *B. distachyon* populations (Marques et al., 2017; Vogel et al., 2009). Most importantly, for similar inferred radiation ages (*B. distachyon* 0.107 Ma; Stritt et al., 2022; *B. stacei* 0.256 Ma, Figure S3a), the divergence rate of *B. distachyon* lineages was greater than of *B. stacei* lineages, probably favoured by its reduced $x = 5$ karyotype (Hasterok et al., 2022), since chromosome fusions may be associated with adaptive advantages, such as shortening DNA replication, cell cycle and

meiosis, and reducing recombination between locally adapted alleles (Guerrero & Kirkpatrick, 2014). It could have facilitated the rapid expansion of *B. distachyon* in the northern Mediterranean region, compared to the slow and limited dispersal of *B. stacei* in the south (Figure 5a, Figure S10; Lopez-Alvarez et al., 2015). Genome-wide evolutionary studied have shown that the allotetraploid *B. hybridum* has recurrently and independently originated at least three times in different locations in the western and eastern Mediterranean (Mu, Li, Yang, Breiman, Yang, et al., 2023). The few available population-level studies of *B. hybridum* indicate that its genetic diversity is relatively low, but higher for some populations than that of the diploids *B. distachyon* and *B. stacei* (Shiposta et al., 2020). Although *B. hybridum* shows no evidence of subgenome dominance (Mu, Li, Yang, Breiman, Yang, et al., 2023; Scarlett et al., 2022), the co-occurrence of the two parental subgenomes in the allotetraploid and their multiple origins from different parental genotypes have likely increased its genetic substrate for higher genetic and phenotypic diversity and, therefore, for faster and greater colonization success than that of its diploid progenitor species, inside and outside the Mediterranean region (Mu, Li, Yang, Breiman, Lou, et al., 2023; Mu, Li, Yang, Breiman, Yang, et al., 2023; Shiposha et al., 2020; Wilson et al., 2019).

The remarkable differences in divergent ages between the stem node of *B. stacei* (~12–11 Ma), probably the oldest lineage of the genus (Díaz-Pérez et al., 2018; Sancho et al., 2022; Scholthof et al., 2018), and its young crown node (0.250 Ma), support the probably occurrence of multiple local extinction events throughout its evolutionary history from the Miocene to the present. This hypothesis, emanating from a previous biogeographical reconstruction of the species (Díaz-Pérez et al., 2018), has been reinforced here by our optimal DIYABC-RF evolutionary model (Figure 4), which predicts the possible existence of strong bottlenecks in almost every past dispersal event (Figure 4, Table S5). Although it is generally accepted that the high turnover of extinctions and radiations of most Mediterranean annual plants occurred during the late Neogene–Pleistocene, as a consequence of different biological and ecological events (Feliner, 2011; Thompson, 2007), the contrasting differences in low frequency and limited range distribution of *B. stacei* populations compared to the abundance of those of its more recently evolved annual Mediterranean congeners *B. distachyon* and *B. hybridum* (Catalan, López-Álvarez, Díaz-Pérez, 2016; Lopez-Alvarez et al., 2015; Minadakis et al., 2023; Wilson et al., 2019), points towards a niche displacement. It is likely that *B. distachyon*, with a more successfully adaptation, colonized more mesic habitats of the northern Mediterranean belt, where *B. stacei* does not currently thrive. Then, the recently evolved and ecologically aggressive allotetraploid *B. hybridum* also expanded into the aridic habitats of the southern Mediterranean belt where it coexists with its progenitor species *B. stacei* in some places (Lopez-Alvarez et al., 2015). Currently, the survival of *B. stacei* is mostly confined to warm and shady habitats, where the plant grows protected by xerophytic shrubs or trees, which are not colonized by the heliophilous *B. hybridum* (Catalán, López-Álvarez, Bellosta, 2016; Mu, Li, Yang, Breiman, Yang, et al., 2023).

The recent radiation of *B. stacei* explains the low diversification rates of its main lineages and, consequently, some of the topological differences observed in the nuclear phylogenetic reconstructions based on ML methods versus coalescent approaches (Figure 2a, Figure S3a). However, both trees coincide in the greater ancestry of the populations of Mallorca+Sicily, the separate divergence of the eastern Mediterranean lineage, and the very recent split of the Ibero-Maghrebian group. These reconstructions explain a multiple colonization scenario for *B. stacei* populations that expanded rapidly across the Mediterranean in very short time periods (Figure S3a) and successfully drove in disjunct and isolated warm localities during the Late Pleistocene (Figure 5a, Figure S10). The high population divergences and low migration rates observed for most population pairs (Table 2, Figure S4) further reinforce this hypothesis. The best demographic and dispersal model also support the notion that *B. stacei* populations have experienced genetic bottlenecks and geographical isolation. The persistence of these isolated habitats across several geological periods, including the Mid Holocene, Last Glacial Maximum and Last Interglacial, emphasizes the importance of long-distance dispersal as a prevailing colonization mechanism, rather than habitat fragmentation (Figures 4 and 5a, Figure S10). Colonization scenarios through LDD in large areas of the circum-Mediterranean basin have also been proposed for the annual congeners *B. distachyon* and *B. hybridum*, as well as for the invasion of other continents by *B. hybridum* (Catalan, López-Álvarez, Díaz-Pérez, 2016; Minadakis et al., 2023; Scholthof et al., 2018; Shiposha et al., 2016, 2020; Wilson et al., 2019). Our dispersal model summarizes a congruent phylogeographic pattern that places ancestral populations of *B. stacei* in the western-central Mediterranean region and its continental islands, followed by an early colonization of the oceanic Canary Islands, a spread in the eastern Mediterranean, and more recent expansion in the Ibero-Mahgrebian continental region (Figure 4). It is consistent with the high sources of migrants inferred from the Mallorca, Sicily and WestMed populations to other populations (Figure 4b, Figure S4, Table 2b). This colonization scenario is coherent with previous phylogeographic findings of *B. stacei* in the western Mediterranean region (Shiposha et al., 2016) and with the recent Holocene origin estimated for its Evolution Canyon AS (African Slope) and ES (European Slope) populations in Israel (Mu, Li, Yang, Breiman, Lou, et al., 2023).

4.2 | Prevalent selfing causes low heterozygosity, but occasional interbreeding explains admixture and plastome capture in *B. stacei*

The colonization success of annual plants is enhanced by selfing, which ensures rapid establishment of the population within a few generations (Gaut et al., 2011; Mairal et al., 2023). The low values of genetic diversity and, consequently, the high rates of selfing and the high values of the inbreeding coefficient detected in the studied populations and genomic groups of *B. stacei* (mean $H_o=0.041$, $s=0.777$ and $F_{is}=0.683$) indicate that it is a predominantly selfing species.

However, it varies from extremely inbred (Canaries2, $s=0.976$) to less inbred (Sicily1, $s=0.359$; Sicily2, $s=0.424$; Mallorca1, $s=0.534$) (Table 1). Some of the most genetically variable and less-inbred populations are those of Mallorca and Sicily, which were revealed as the most ancestral populations (Figure 2a), while the less genetically variable and most inbred populations are the geographically disjunct Morocco ($s=0.970$) and Greece ($s=0.966$) populations that likely resulted from LDD founder events (Figures 1 and 4; Table 1). Our findings are in agreement with other studies that have also detected high but variable levels of inbreeding in western Mediterranean *B. stacei* populations using SSR data ($s=1$–0.555; Shiposha et al., 2016) and high levels in eastern Mediterranean populations using whole genome sequencing data ($s>0.92$; Mu, Li, Yang, Breiman, Lou, et al., 2023). These results parallel those recorded for the extreme selfer congener *B. distachyon* in most of its Mediterranean populations (Marques et al., 2017; Stritt et al., 2022; Vogel et al., 2009). The analogous mating system between the two congeners is in accord with similarities in floral morphology and floral structure in *B. stacei* and *B. distachyon*, as both species have relatively small cleistogamous or cleistogamous-like florets (Catalán, López-Álvarez, Bellosta, 2016). However, *B. stacei* can display exerted anthers, which could facilitate occasional outcrossing, while pollination in *B. distachyon* usually occurs in closed flowers (Dinh Thi et al., 2016), leading to extremely high levels of homozygosity (Vogel et al., 2009).

Despite the predominant selfing nature of *B. stacei*, the genetic admixture found in some populations (Figure 2a), the rates of gene flow detected between populations (Table 2b, Figure S4), and the disparate phylogenetic patterns retrieved from the plastome tree and haplotype network with respect to those of the nuclear trees (Figures 2 and 3, Figure S3) support a relatively frequent occurrence of gene flow and interbreeding operating between isolated and largely diverging *B. stacei* lineages. The most notable cases correspond to the Canaries2, Sicily2, Mallorca 2 and 4, and Jaen and Tunisia plastotypes, which probably resulted from LDD and introgressions (Figures 2 and 3, Figure S3), while the origin of the Almeria-type plastotype, shared with Jaen2, agrees with the nuclear admixture scenario with the Jaen genomic group (Figure 2a, Figure S2a). Although some form of incomplete lineage sorting of plastome haplotypes and of nuclear genes may have affected the phylogenies (Jakob & Blattner, 2006), the mounting evidence reinforces occasional interbreeding as a likely phenomenon shaping the evolution of *B. stacei*. Chloroplast capture and introgression have also been detected among divergent and geographically distant populations of the close congener *B. distachyon* (Gordon et al., 2020; Sancho et al., 2018) pointing to the high frequency of LDD events in these annual Mediterranean plants, which probably favour secondary contacts and subsequent introgression and admixture. In selfing plants, gene flow is mainly facilitated through seed dispersal (Holsinger, 2000; Mu, Li, Yang, Breiman, Lou, et al., 2023); however, occasional interbreeding has explained the admixed patterns observed at the genome level among sympatrically divergent subpopulations of *B. stacei* (Mu, Li, Yang, Breiman, Lou, et al., 2023). The outcrossing capability of *B. stacei*

has been reaffirmed by the existence of three independent and bidirectional interspecific hybridization events with its congener *B. distachyon* in nature at different times that gave rise to the allotetraploid *B. hybridum* (Gordon et al., 2020; Mu, Li, Yang, Breiman, Yang, et al., 2023; Scarlett et al., 2022; Shiposha et al., 2020), and the artificial crossing that resulted in the synthetic allopolyploid (Dinh Thi et al., 2016). Furthermore, very successful intraspecific crosses have been carried out between divergent parental lines for the generation of mapping populations in *B. distachyon* and *B. stacei*, respectively (Gordon et al., 2020). All this evidence supports the feasibility of natural spontaneous natural crosses between diverging genotypes of the same species in both grasses. Moreover, we postulate that shared ancestry may also reflect genomic admixture, as observed in some *B. stacei* (Figure 2a) and *B. distachyon* (Gordon et al., 2017) populations. However, the strong drift effects of selfing drastically reduces heterozygosity within a few generations, resulting in highly inbred and isolated populations, as observed for *B. stacei* in spatially close but ecologically diverging sister populations in Israel (Mu, Li, Yang, Breiman, Lou, et al., 2023).

4.3 | Ecologically driven differentiation of *B. stacei* populations

Our *B. stacei* distribution models and climatic niche analyses have improved on previous ENMs for this species (Lopez-Alvarez et al., 2015), detecting three significantly different climatic niches for its western central Mediterranean (WestCentralMed), eastern southern Mediterranean (EastSouthMed) and Macaronesian populations (Figure 5b, Table S6). Our niche projections for past climatic envelopes evidenced that subsequent climate changes caused the loss of suitable habitats and changes in the distribution range of *B. stacei* under each new climatic condition (Figure 5a, Figure S10). The predicted niche contractions in the interglacials and niche expansion in the LGM could have caused local extinctions in unsuitable areas and bottlenecks in others during the LIG and the present (Figure 5a, Figure S10, Table 3), and probably in other Quaternary glacial–interglacial cycles that led to population shrinkage and eradication. These events would have contributed to increasing genetic drift and isolation of the recently evolved diverging lineages estimated in our nuclear time-measured phylogeny (Figure S3a) and the selected colonization scenario (Figure 4).

Our climate niche data also explain the rapid adaptive success of colonizing lineages to new environments. The significant differences observed for seven variables between the Macaronesian and the continental niches, and also for another six variables between the WestCentralMed and EeastSouthMed niches indicate seasonally warm conditions in the western Mediterranean, extremely dry and aridic conditions in the latitudinal range of the eastern Mediterranean, and relatively more humid conditions in the Atlantic African front and the Canary isles (Figure 5b, Table S6). Therefore, in addition to historical factors, such as potential LGM refugia

(Figure 5a, Figure S10), the genomic diversity of *B. stacei* in its native Mediterranean region could have been determined by the environment rather than geographical distance.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The RADseq and plastome Illumina paired-end sequencing data generated for this project are available at the European Nucleotide Archive, project numbers PRJEB67979 and PRJEB68148, respectively. Documents containing the filtered SNPs and input and output data related to the evolutionary and ENM analyses performed in this study are available on Github (https://github.com/Bioflora/Bstacei_Phylogeography).

ORCID

Miguel Campos https://orcid.org/0000-0002-6555-3576
Antonio Díaz-Pérez https://orcid.org/0000-0001-8517-1404
Diana López-Alvarez https://orcid.org/0000-0002-7734-8481
Luis A. J. Mur https://orcid.org/0000-0002-0961-9817
John P. Vogel https://orcid.org/0000-0003-1786-2689
Pilar Catalán https://orcid.org/0000-0001-7793-5259

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*(10), e3376.

Barrett, S. C. H., Arunkumar, R., & Wright, S. I. (2014). The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1648), 20130344. https://doi.org/10.1098/rstb.2013.0344

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*(8), 1917–1932.

Bushnell, B. (2014). *BBMap: A fast, accurate, splice-aware aligner*. Lawrence Berkeley National Lab (LBNL).

Cardillo, M., & Warren, D. L. (2016). Analysing patterns of spatial and niche overlap among species at multiple resolutions. *Global Ecology and Biogeography*, *25*(8), 951–963. https://doi.org/10.1111/geb.12455

Catalán, P., López-Álvarez, D., Bellosta, C., & Villar, L. (2016). Updated taxonomic descriptions, iconography, and habitat preferences of *Brachypodium distachyon*, *B. stacei*, and *B. hybridum* (Poaceae). *Anales Jardin Botánico Madrid*, *73*(1), e028. https://doi.org/10.3989/ajbm.2428

Catalan, P., López-Álvarez, D., Díaz-Pérez, A., Sancho, R., & López-Herránz, M. L. (2016). Phylogeny and evolution of the genus *Brachypodium*. In P. J. Vogel (Ed.), *Genetics and genomics of Brachypodium* (pp. 9–38). Springer. https://doi.org/10.1007/7397_2015_17

Catalán, P., Müller, J., Hasterok, R., Jenkins, G., Mur, L. A. J., Langdon, T., Betekhtin, A., Siwinska, D., Pimentel, M., & López-Alvarez, D. (2012). Evolution and taxonomic split of the model grass *Brachypodium distachyon*. *Annals of Botany*, *109*(2), 385–405. https://doi.org/10.1093/aob/mcr294

Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1434), 1051–1070. https://doi.org/10.1098/rstb.2003.1296

Collin, F., Durif, G., Raynal, L., Lombaert, E., Gautier, M., Vitalis, R., Marin, J., & Estoup, A. (2021). Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC random forest. *Molecular Ecology Resources*, *21*(8), 2598–2613. https://doi.org/10.1111/1755-0998.13413

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Díaz-Pérez, A., López-Álvarez, D., Sancho, R., & Catalán, P. (2018). Reconstructing the origins and the biogeography of species' genomes in the highly reticulate allopolyploid-rich model grass genus *Brachypodium* using minimum evolution, coalescence and maximum likelihood approaches. *Molecular Phylogenetics and Evolution*, *127*, 256–271. https://doi.org/10.1016/j.ympev.2018.06.003

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *45*(4), e18. https://doi.org/10.1093/nar/gkw955

Dinh Thi, V. H., Coriton, O., Le Clainche, I., Arnaud, D., Gordon, S. P., Linc, G., Catalan, P., Hasterok, R., Vogel, J. P., Jahier, J., & Chalhoub, B. (2016). Recreating stable *Brachypodium hybridum* allotetraploids by uniting the divergent genomes of *B. distachyon* and *B. stacei*. *PLoS One*, *11*(12), e0167171. https://doi.org/10.1371/journal.pone.0167171

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*(1), 11–15.

Eaton, D. A. R., & Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*(8), 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and windows. *Molecular Ecology Resources*, *10*(3), 564–567.

Feliner, G. N. (2011). Southern European glacial refugia: A tale of tales. *Taxon*, *60*(2), 365–372. https://doi.org/10.1002/tax.602007

Francis, R. M. (2017). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27–32.

Gaut, B., Yang, L., Takuno, S., & Eguiarte, L. E. (2011). The patterns and causes of variation in plant nucleotide substitution rates. *Annual Review of Ecology, Evolution, and Systematics*, *42*(1), 245–266. https://doi.org/10.1146/annurev-ecolsys-102710-145119

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z. L., & Zhang, M. (2011). The community climate system model version 4. *Journal of Climate*, *24*(19), 4973–4991. https://doi.org/10.1175/2011JCLI4083.1

Gordon, S. P., Contreras-Moreira, B., Levy, J. J., Djamei, A., Czedik-Eysenberg, A., Tartaglio, V. S., Session, A., Martin, J., Cartwright, A., Katz, A., Singan, V. R., Goltsman, E., Barry, K., Dinh-Thi, V. H., Chalhoub, B., Diaz-Perez, A., Sancho, R., Lusinska, J., Wolny, E., … Vogel, J. P. (2020). Gradual polyploid genome evolution revealed by pan-genomic analysis of *Brachypodium hybridum* and its diploid progenitors. *Nature Communications*, *11*(1), 1–16. https://doi.org/10.1038/s41467-020-17302-5

Gordon, S. P., Contreras-Moreira, B., Woods, D. P., Des Marais, D. L., Burgess, D., Shu, S., Stritt, C., Roulin, A. C., Schackwitz, W., & Tyler, L. (2017). Extensive gene content variation in the *Brachypodium distachyon* pan-genome correlates with population structure. *Nature Communications*, *8*(1), 1–13.

Greer, S. U., Wright, S. I., & Eckert, C. G. (2023). Population bottleneck associated with but likely preceded the recent evolution of self-fertilization in a coastal dune plant. *Evolution; International Journal of Organic Evolution*, *77*(2), 454–466. https://doi.org/10.1093/evolut/qpac047

Guerrero, R. F., & Kirkpatrick, M. (2014). Local adaptation and the evolution of chromosome fusions. *Evolution*, *68*(10), 2747–2756. https://doi.org/10.1111/evo.12481

Hasterok, R., Catalan, P., Hazen, S. P., Roulin, A. C., Vogel, J. P., Wang, K., & Mur, L. A. J. (2022). Brachypodium: 20 years as a grass biology model system; the way forward? *Trends in Plant Science*, *27*(10), 1002–1016. https://doi.org/10.1016/j.tplants.2022.04.008

Holsinger, K. E. (2000). Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(13), 7037–7042. https://doi.org/10.1073/pnas.97.13.7037

Jakob, S. S., & Blattner, F. R. (2006). A chloroplast genealogy of *Hordeum* (Poaceae): Long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Molecular Biology and Evolution*, *23*(8), 1602–1612. https://doi.org/10.1093/molbev/msl018

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Jombart, T., & Collins, C. (2015). *A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0*. Imperial College London, MRC Centre for Outbreak Analysis and Modelling.

Kadereit, J. W., & Abbott, R. J. (2021). Plant speciation in the quaternary. *Plant Ecology & Diversity*, *14*(3–4), 105–142. https://doi.org/10.1080/17550874.2021.2012849

Lawson, C. R., Hodgson, J. A., Wilson, R. J., & Richards, S. A. (2014). Prevalence, thresholds and the performance of presence–absence models. *Methods in Ecology and Evolution*, 5(1), 54–64. https://doi.org/10.1111/2041-210X.12123

Letunic, I., & Bork, P. (2021). Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293–W296. https://doi.org/10.1093/NAR/GKAB301

Li, Y., & Liu, J. (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources*, 18(1), 176–177.

Lopez-Alvarez, D., Manzaneda, A. J., Rey, P. J., Giraldo, P., Benavente, E., Allainguillaume, J., Mur, L., Caicedo, A. L., Hazen, S. P., Breiman, A., Ezrati, S., & Catalan, P. (2015). Environmental niche variation and evolutionary diversification of the *Brachypodium distachyon* grass complex species in their native circum-Mediterranean Range. *American Journal of Botany*, 102(7), 1073–1088. https://doi.org/10.3732/ajb.1500128

López-Álvarez, D., Zubair, H., Beckmann, M., Draper, J., & Catalán, P. (2017). Diversity and association of phenotypic and metabolomic traits in the close model grasses *Brachypodium distachyon*, *B.stacei* and *B. hybridum*. *Annals of Botany*, 119(4), 545–561. https://doi.org/10.1093/aob/mcw239

Mairal, M., García-, C., Le Roux, J. J., Chau, J. H., Jansen, B., Cang, V. V., Münzbergová, Z., Chown, S. L., & Shaw, J. D. (2023). Multiple introductions, polyploidy and mixed reproductive strategies are linked to genetic diversity and structure in the most widespread invasive plant across Southern Ocean archipelagos. *Molecular Ecology*, 32(4), 756–771. https://doi.org/10.1111/mec.16809

Mandáková, T., & Lysak, M. A. (2018). Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology*, 42, 55–65. https://doi.org/10.1016/j.pbi.2018.03.001

Marques, I., Shiposha, V., López-Alvarez, D., Manzaneda, A. J., Hernandez, P., Olonova, M., & Catalán, P. (2017). Environmental isolation explains Iberian genetic diversity in the highly homozygous model grass *Brachypodium distachyon*. *BMC Evolutionary Biology*, 17(1), 1–14.

Médail, F., & Diadema, K. (2009). Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, 36(7), 1333–1345. https://doi.org/10.1111/j.1365-2699.2008.02051.x

Mijangos, J. L., Gruber, B., Berry, O., Pacioni, C., & Georges, A. (2022). dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in Ecology and Evolution*, 13(10), 2150–2158. https://doi.org/10.1111/2041-210X.13918

Milanesi, M., Capomaccio, S., Vajana, E., Bomba, L., Garcia, J. F., Ajmone-Marsan, P., & Colli, L. (2017). BITE: An R package for biodiversity analyses. *BioRxiv*, 181610.

Minadakis, N., Williams, H., Horvath, R., Caković, D., Stritt, C., Thieme, M., Bourgeois, Y., & Roulin, A. C. (2023). The demographic history of the wild crop relative *Brachypodium distachyon* is shaped by distinct past and present ecological niches. *Peer Community Journal*, 3, e84. https://doi.org/10.24072/pcjournal.319

Mittermeier, R., Gil, P., Hoffman, M., Pilgrim, J., Brooks, T., Mittermeier, C. G., Lamoreux, J., Da Fonseca, G., & Saligmann, P. (2004). *Hotspots revisited Earth's biologically richest and most endangered terrestrial ecoregions*. Cemex.

Mu, W., Li, K., Yang, Y., Breiman, A., Lou, S., Yang, J., Wu, Y., Wu, S., Liu, J., Nevo, E., & Catalan, P. (2023). Scattered differentiation of unlinked loci across the genome underlines ecological divergence of the selfing grass *Brachypodium stacei*. *Proceedings of the National Academy of Sciences of the United States of America*, 120(45), e2304848120. https://doi.org/10.1073/pnas.2304848120

Mu, W., Li, K., Yang, Y., Breiman, A., Yang, J., Wu, Y., Zhu, M., Wang, S., Catalan, P., Nevo, E., & Liu, J. (2023). Subgenomic stability of progenitor genomes during repeated allotetraploid origins of the same grass *Brachypodium hybridum*. *Molecular Biology and Evolution*, 40(12), 1–20. https://doi.org/10.1093/molbev/msad259

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274. https://doi.org/10.1093/molbev/msu300

Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H., & Hu, A. (2006). Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science*, 311(5768), 1751–1753. https://doi.org/10.1126/SCIENCE.1120808

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, S. J., Dudík, M., & Schapire, R. E. (2021). Maxent software for modeling species niches and distributions (Version 3.4. 1). 2018, 5–21. http://biodiversityinformatics.amnh.org/open_source/maxent/accessed on

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859–866. https://doi.org/10.1093/bioinformatics/btv684

Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16(3), 608–627.

QGIS Development Team. (2023). QGIS Geographic Information System.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, 197(2), 573–589.

Ritland, K. (1990). Inferences about inbreeding depression based on changes of the inbreeding coefficient. *Evolution*, 44(5), 1230–1241.

Roessler, K., Muyle, A., Diez, C. M., Gaut, G. R. J., Bousios, A., Stitzer, M. C., Seymour, D. K., Doebley, J. F., Liu, Q., & Gaut, B. S. (2019). The genome-wide dynamics of purging during selfing in maize. *Nature plants*, 5(9), 980–990. https://doi.org/10.1038/s41477-019-0508-7

Sancho, R., Cantalapiedra, C. P., López-Alvarez, D., Gordon, S. P., Vogel, J. P., Catalán, P., & Contreras-Moreira, B. (2018). Comparative plastome genomics and phylogenomics of *Brachypodium*: Flowering time signatures, introgression and recombination in recently diverged ecotypes. *New Phytologist*, 218(4), 1631–1644. https://doi.org/10.1111/NPH.14926

Sancho, R., Inda, L. A., Díaz-Pérez, A., Des Marais, D. L., Gordon, S., Vogel, J. P., Lusinska, J., Hasterok, R., Contreras-Moreira, B., & Catalán, P. (2022). Tracking the ancestry of known and 'ghost' homeologous subgenomes in model grass *Brachypodium polyploids*. *Plant Journal*, 109(6), 1535–1558. https://doi.org/10.1111/tpj.15650

Scarlett, V. T., Lovell, J. T., Shao, M., Phillips, J., Shu, S., Goodstein, D. M., Jenkins, J., Grimwood, J., Barry, K., Chalhoub, B., Hasterok, R., Catala, P., Vogel, J. P., Lusinska, J., & Schmutz, J. (2022). Multiple origins, one evolutionary trajectory: Gradual evolution characterizes distinct lineages of allotetraploid *Brachypodium*. *Genetics*, 223(2), iyac146. https://doi.org/10.1093/genetics/iyac146

Scholthof, K. B. G., Irigoyen, S., Catalan, P., & Mandadi, K. K. (2018). Brachypodium: A monocot grass model genus for plant biology. *Plant Cell*, 30(8), 1673–1694. https://doi.org/10.1105/tpc.18.00083

Shiposha, V., Catalán, P., Olonova, M., & Marques, I. (2016). Genetic structure and diversity of the selfing model grass *Brachypodium stacei* (Poaceae) in Western Mediterranean: Out of the Iberian Peninsula and into the islands. *PeerJ*, 4, e2407. https://doi.org/10.7717/peerj.2407

Shiposha, V., Marques, I., López-Alvarez, D., Manzaneda, A. J., Hernandez, P., Olonova, M., & Catalán, P. (2020). Multiple founder events explain the genetic diversity and structure of the model allopolyploid grass *Brachypodium hybridum* in the Iberian Peninsula hotspot. *Annals of Botany*, *125*(4), 625–638. https://doi.org/10.1093/aob/mcz169

Sievert, C. (2020). *Interactive web-based data visualization with R, plotly, and shiny*. CRC Press.

Solbrig, O. T. (1976). On the relative advantages of cross-and self-fertilization. *Annals of the Missouri Botanical Garden*, *63*, 262–276.

Stritt, C., Gimmi, E. L., Wyler, M., Bakali, A. H., Skalska, A., Hasterok, R., Mur, L. A. J., Pecchioni, N., & Roulin, A. C. (2022). Migration without interbreeding: Evolutionary history of a highly selfing Mediterranean grass inferred from whole genomes. *Molecular Ecology*, *31*(1), 70–85. https://doi.org/10.1111/mec.16207

Suc, J.-P. (1984). Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature*, *307*(5950), 429–432. https://doi.org/10.1038/307429a0

Szövényi, P., Devos, N., Weston, D. J., Yang, X., Hock, Z., Shaw, J. A., Shimizu, K. K., McDaniel, S. F., & Wagner, A. (2014). Efficient purging of deleterious mutations in plants with haploid selfing. *Genome biology and evolution*, *6*(5), 1238–1252. https://doi.org/10.1093/gbe/evu099

Tang, D., Chen, M., Huang, X., Zhang, G., Zeng, L., Zhang, G., Wu, S., & Wang, Y. (2023). SRplot: A free online platform for data visualization and graphing. *PLoS One*, *18*(11), e0294236.

Thompson, J. D. (2007). *Plant evolution in the Mediterranean*. Oxford University Press. https://doi.org/10.1093/ACPROF:OSO/9780198515340.001.0001

Tyler, L., Lee, S. J., Young, N. D., DeIulio, G. A., Benavente, E., Reagon, M., Sysopha, J., Baldini, R. M., Troìa, A., Hazen, S. P., & Caicedo, A. L. (2016). Population structure in the model grass *Brachypodium distachyon* is highly correlated with flowering differences across broad geographic areas. *The Plant Genome*, *9*(2), 1–55. https://doi.org/10.3835/plantgenome2015.08.0074

Vargas, P., Fernández-Mazuecos, M., & Heleno, R. (2018). Phylogenetic evidence for a Miocene origin of Mediterranean lineages: Species diversity, reproductive traits and geographical isolation. *Plant Biology*, *20*, 157–165. https://doi.org/10.1111/plb.12626

Vargas, P., Jiménez-Mejías, P., & Fernández-Mazuecos, M. (2020). 'Endangered living fossils' (ELFs): Long-term survivors through periods of dramatic climate change. *Environmental and Experimental Botany*, *170*, 103892. https://doi.org/10.1016/j.envexpbot.2019.103892

Vogel, J. P., Tuna, M., Budak, H., Huo, N., Gu, Y. Q., & Steinwand, M. A. (2009). Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon*. *BMC Plant Biology*, *9*(1), 1–11. https://doi.org/10.1186/1471-2229-9-88/FIGURES/5

Warren, D. L., Glor, R. E., Turelli, M., Warren, D. L., Turelli, M., & Glor, Á. R. E. (2010). ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography*, *33*(3), 607–611. https://doi.org/10.1111/J.1600-0587.2009.06142.X

Wilson, P. B., Streich, J. C., Murray, K. D., Eichten, S. R., Cheng, R., Aitken, N. C., Spokas, K., Warthmann, N., Gordon, S. P., Accession, C., Vogel, J. P., & Borevitz, J. O. (2019). Global diversity of the *Brachypodium* species complex as a resource for genome-wide association studies demonstrated for agronomic traits in response to climate. *Genetics*, *211*(1), 317–331. https://doi.org/10.1534/genetics.118.301589

Wright, S. I., Kalisz, S., & Slotte, T. (2013). Evolutionary consequences of self-fertilization in plants. *Proceedings. Biological Sciences*, *280*(1760), 20130133. https://doi.org/10.1098/RSPB.2013.0133

Zecca, G., Labra, M., & Grassi, F. (2020). Untangling the evolution of American wild grapes: Admixed species and how to find them. *Frontiers in Plant Science*, *10*, 497733.

Zeitler, L., Parisod, C., & Gilbert, K. J. (2023). Purging due to self-fertilization does not prevent accumulation of expansion load. *PLoS Genetics*, *19*(9), e1010883. https://doi.org/10.1371/journal.pgen.1010883

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.