

## ORIGINAL ARTICLE OPEN ACCESS

# Local Selection Shaped the Diversity of European Maize Landraces

Margarita Takou<sup>1</sup> | Kerstin Schulz<sup>1,2</sup> | Markus G. Stetter<sup>1,2</sup> 

<sup>1</sup>Institute for Plant Sciences, University of Cologne, Cologne, Germany | <sup>2</sup>Cluster of Excellence on Plant Sciences, University of Cologne, Cologne, Germany

**Correspondence:** Markus G. Stetter ([m.stetter@uni-koeln.de](mailto:m.stetter@uni-koeln.de))

**Received:** 2 July 2024 | **Revised:** 30 January 2025 | **Accepted:** 21 February 2025

**Handling Editor:** Sarah Turner-Hissong

**Funding:** This work was supported by Bundesministerium für Bildung und Forschung, 031B0195, project MAZE. Deutsche Forschungsgemeinschaft, EXC-2048/1–project ID 390686111.

**Keywords:** adaptation | European maize | genetic diversity | genetic load | population structure | range expansion

## ABSTRACT

The introduction of populations to novel environments can lead to a loss of genetic diversity and the accumulation of deleterious mutations due to selection and demographic changes. We investigate how the recent introduction of maize to Europe shaped the genetic diversity and differentiation of European traditional maize populations and quantify the impact of its recent range expansion and consecutive breeding on the accumulation of genetic load. We use genome-wide genetic markers of almost 2000 individuals from 38 landraces, 155 elite breeding lines, and a large set of doubled haploid lines derived from two landraces to find extensive population structure within European maize, with landraces being highly differentiated even over short geographic distances. Yet, diversity change does not follow the continuous pattern of range expansions. Landraces maintain high genetic diversity that is distinct between populations and does not decrease along the possible expansion routes. Signals of positive selection in European landraces that overlap with selection in Asian maize suggest convergent selection during maize introductions. At the same time, environmental factors partially explain genetic differences across Europe. Consistent with the maintenance of high diversity, we find no evidence of genetic load accumulating along the maize introduction route in European maize. However, modern breeding likely purged highly deleterious alleles but accumulated genetic load in elite germplasm. Our results reconstruct the history of maize in Europe and show that landraces have maintained high genetic diversity that could reduce genetic load in the European maize breeding pools.

## 1 | Introduction

Species distributions are the result of range expansion to new environments, such as the post-glacial colonisation of Northern Europe (Hewitt 2000). These dynamic population genetic processes have a strong influence on the genetic diversity of the expanding species (Excoffier et al. 2009), including a decrease in genetic diversity due to increased drift and the subsequent genetic differentiation among the newly established and core populations (Austerlitz et al. 1997; Excoffier

et al. 2009; Slatkin and Excoffier 2012; Wang et al. 2017; Takou et al. 2021). At the same time, deleterious mutations accumulate at the front of the expansion range, increasing genetic load (de Pedro et al. 2021; González-Martínez et al. 2017; Peischl et al. 2013). As locally adapted populations move to novel environments (Colautti and Barrett 2013; Savolainen et al. 2013), the adaptive potential of populations at the range edges can be compromised (Excoffier et al. 2009). Rapid range expansions are expected to become more frequent as climate change alters the original environments, which will

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Molecular Ecology* published by John Wiley & Sons Ltd.

push species to new suitable conditions outside their original range (Waldvogel et al. 2020). Crops have spread rapidly around the globe and likely experienced the effects of range expansion (Huang et al. 2022). They could have encountered novel selective pressures caused by the different climate, soil compositions, and ecological dynamics of the new regions, as well as the human-driven selection for agricultural traits (Purugganan and Fuller 2009; Purugganan 2019). This interplay between natural selection and human-driven intervention ultimately resulted in the appearance of local traditional crop varieties (Huang et al. 2022).

One of the most important crops worldwide is maize, which has shown great adaptability to locations outside its initial range, with one of the broadest cultivated ranges of all crops today (Tenaillon and Charcosset 2011). Maize domestication started approximately 9000 years ago in a small region in Mexico, where the wild grass species teosinte (*Zea mays* ssp *parviglumis* and *Zea mays* ssp *mexicana*) have given rise to the crop we know today (Tenaillon and Charcosset 2011; Yang, Xu, et al. 2017; Yang et al. 2023). Over time, maize spread throughout the Americas, facilitated by human movement (Kistler et al. 2018). The early dispersion process was characterised by two routes: one northward to the USA and Canada, and another extending southward to South America and the Caribbean coast (Tenaillon and Charcosset 2011; Kistler et al. 2018). This human-mediated range expansion has led to decreased genetic diversity, introgression, and local adaptation across the colonisation route (Wang et al. 2017; Hufford et al. 2012; Arca et al. 2023), which creates the opportunity to study range expansion and the dynamics of selection and decreasing genetic diversity. Hence, crops such as maize might have experienced similar range expansion forces as predicted for natural populations.

European maize landraces offer a compelling system to explore the recent introduction and expansion at potential range edges. While extensive research has been done on the evolution, domestication and traditional use of primary American maize landraces (Hufford et al. 2012; Yang et al. 2023; Wang et al. 2017), the secondary range expansion and development of their European counterparts have been little investigated (Revilla Temiño et al. 2003; Tenaillon and Charcosset 2011). Unlike the gradual expansion outside of the domestication area, maize's arrival in Europe is characterised by an abrupt introduction to a novel environment around 500 years ago (Janick and Caneva 2005; Brandenburg et al. 2017). Moreover, only a relatively small number of plants were initially brought from the Americas, likely resulting in a strong bottleneck (Tenaillon and Charcosset 2011). After its first introduction into Europe from the Caribbean through Spain, there is evidence for further diffusion routes via France and Italy to the rest of Europe, as well as a secondary introduction from North America to Central Europe (Brandenburg et al. 2017; Leng et al. 1962; Rebourg et al. 2003). Additionally, the varying local European environments likely required local adaptation, as has been observed in Mexican maize populations (Tittes et al. 2023). Historical accounts state that by the middle of the 16th century, maize fields were present across all of Europe (Finan 1948; Janick and Caneva 2005; Brandolini and Brandolini 2009), making European maize landraces an invaluable resource for understanding the rapid introduction to new environments.

Here, we study the genetic diversity across Europe and how it was potentially shaped by the multiple introductions of maize to the continent. Using a comprehensive dataset of traditional varieties spanning the range of European maize, we were able to show that European landraces are highly diverse but genetically differentiated. We detect two major clusters of landraces within Europe that might depict their introduction history. In addition, we identify local clusters of similarity between traditional varieties, likely reconstructing historical trade routes. Selection scans in two landrace doubled haploid (DH) line libraries with almost 1000 individuals reveal genomic regions that were likely under selection in individual landraces and suggest potential adaptation signatures within Europe. However, European maize does not show signatures of range-edge decreased diversity and accumulated genetic load. Together, our results give insights into the history of the rise of European maize in the last centuries.

## 2 | Materials and Methods

### 2.1 | Dataset Composition

We selected a sample set to represent the history of maize in Europe. Two primary categories of maize samples were used to form the panel: European landraces and elite lines. The European landraces, which represent authentic European populations (Villa et al. 2005; Casañas et al. 2017), were selected with the aim of capturing a high geographical range. The second category, elite lines, representing modern hybrid pools (Reif et al. 2005) were included as points of comparison with European landraces. Furthermore, doubled haploid lines (DH lines), which are completely homozygous lines from three European landraces (Maqbool et al. 2020), were also included.

All data used in the present study are publicly available (Unterseer et al. 2016; Mayer et al. 2017, 2020). In brief, all individuals used were genotyped at high density with the 600k Affymetrix Axiom Maize Array, which is comprised of 616,201 variants, of which 6759 represent insertions/deletions (Unterseer et al. 2014, 2016). Specifically, 941 DH lines derived from 2 European landraces were taken from Mayer et al. (2020), which were divided as follows: 501 DH lines of Kemater Landmais Gelb (KL; Austria) and 409 of Petkuser Ferdinand Rot (PE; Germany). 952 individuals from 35 European maize landraces from Mayer et al. (2017). Lastly, 3 landraces from Unterseer et al. (2016) and 155 elite lines were added from Unterseer et al. (2016). For the following analysis, we clustered the populations in two groups, based on population clustering (Table S1). The final panel had in total 2954 individuals (Table S1).

### 2.2 | Data Preparation

The merging of the three data sets (Unterseer et al. 2016; Mayer et al. 2017, 2020) was performed using custom python scripts, and all data sets were formatted in the HapMap format using Tassel 5.0 GUI (command: `./start_tassel.pl -Xmx4g`) (Bradbury et al. 2007). Post conversion, the IUPAC nucleotide codes in the converted files were replaced with the corresponding nucleotide bases, as stated in the user guide. For merging, the Unterseer et al. (2016) data sets were updated to AGPv4 maize reference

genome. All insertions/deletions in the datasets were removed, which resulted in 6759 and 6752 markers being removed from the Mayer and Unterseer data sets, respectively. For quality control, we checked if the alleles were correctly merged based on the individual marker information provided by the manufacturer. Additionally, in order to filter out possible sequencing errors and sites with low genotyping quality, we used the Affymetrix Quality classification, and only markers that had a quality of 'PolyHighResolution' 'MonoHighResolution' or 'NoMinorHom' across all datasets were kept, resulting in 419,477 SNPs. As none of the individuals had missing information in more than 0.8% of their sites and the majority of the sites had <1% missing information after the above filtering steps, we did not apply any other filtering criteria based on missingness. Subsequently, the dataset tables were merged and adjusted to fit the HapMap format, enabling conversion via TASSEL. Lastly, the Hapmap file was converted back into a common VCF file using TASSEL Bradbury et al. (2007). The resulting VCF file was then used for the subsequent analysis, each time filtered as indicated within the respective sections.

### 2.3 | Population Structure and Genetic Diversity Analysis

To investigate the underlying patterns of variation in this maize dataset, a Principal Component Analysis (PCA) was conducted using the scikit allele package (Miles et al. 2023) for python v3.7. All individuals of the 35 landraces from Mayer et al. (2017), the 3 unique landrace lines from Unterseer et al. (2016) AL, FL, and SO, as well as the elite maize lines grouped into dent and flint populations were used for this analysis. We sub-sampled the dataset to exclude the DH lines and reduce the number of individuals to 1160 individuals divided into 40 populations, including the dent and flint elite lines, for the PCA. In order to do the PCA, using the scikit allele package, singletons were removed from the sub-sampled dataset and linkage disequilibrium (LD) filtering/pruning was done using 5 iterations, a window size of 500 bp, and a step size of 200 bp. The PCA was then performed using the function *allele.pca()* (Miles et al. 2023) and plotted using the matplotlib package (Hunter 2007). We further investigated the clustering of the data by using the software STRUCTURE (Pritchard et al. 2000). We ran the analysis for up to  $K=40$  and used the cross-validation error (cv) to decide the optimal number of clusters.

Using scikit allele package (Miles et al. 2023), all pairwise  $F_{ST}$  values between all the 35 European landrace population pairs from Mayer et al. (2017) were estimated based on segregating SNPs. We calculated Hudson's  $F_{ST}$  using the *allele.blockwise\_hudson\_fst()* function with a block size of 100,000 variants (Smaragdov and Kudinov 2020; Bhatia et al. 2013; Hudson et al. 1992). Additionally, we estimated Slatkin's linearized  $F_{ST}$ , as shown in (Bay et al. 2018). To visualise the population differentiation patterns, the calculated  $F_{ST}$  values were plotted with a hierarchical clustered heatmap, using the package ComplexHeatmap in R v4.1.2 (R Core Team 2021; Gu et al. 2016).

We estimated mean expected heterozygosity ( $H_{EXP}$ ) per population for all landraces to investigate the distribution of genetic diversity and potential patterns of diversity among European

landraces across the diverse European landscape. We calculated the expected heterozygosity using the *allele.heterozygosity\_expected()* function and estimated the mean using the *numpy.nan-mean()* function (Harris et al. 2020). The geographical map including the  $H_{EXP}$  values was created using the plotly package (Inc.P.T 2015). The sample locations (longitude and latitude values from Mayer et al. (2017)) were used as proxies for each population's distribution. We estimated the haversine geographic distance in kilometres of each landrace to the two populations that are approximately the closest to the expected entry points in Europe; the population Tuy (TU; 42.04N, -8.64W) marked the entry point in Spain and the population Barisis (BA; 49.57N, 3.328E) in Central Europe. We estimated all the distances with the function *dism* of the R package geosphere (Hijmans 2022). For the populations Altreier (AL; Italy), Fleimstal (FL; Italy) and Sornay (SO; France), whose coordinate information were missing, we used the best approximation of their locations. The inbreeding coefficient of each individual was estimated using *vcftools* (Danecek et al. 2011).

### 2.4 | Spatial Gene-Flow Analysis of Population Genetic Data and Selection Scans

The Fast Estimation of Effective Migration Surfaces (FEEMS, version 1.0.0) analysis was conducted to investigate non-homogeneous gene flow between the European landraces (Marcus et al. 2021). For this analysis, we used the 35 populations published in Mayer et al. 2017, as for the 3 additional landraces in Unterseer et al. (2016) we did not have accurate geographic information for them, leaving 834 individuals for 35 landrace populations. Filtering was performed on the final dataset and included the following steps: Variants were filtered to only include SNPs using the command *bcftools view -v snps*, and only biallelic SNPs were selected with *bcftools view -m2 -M2*. As no SNPs with more than 20% missing data were present, we did not remove any sites based on this filter. Monomorphic sites were removed with *bcftools* (Li 2011) and the input files were prepared with PLINK2 (Chang et al. 2015). We used the FEEMS imputation feature for any present missing data. Specifically, we used scikit's function *SimpleImputer* and the imputation strategy of 'mean'. Additionally, the script was run with the options *scale\_snps=True* and *translated=False* and the provided smaller grid size file *grid\_100.shp*. A custom smoothing regularisation parameter ( $\lambda=100$ ) was estimated from a cross-validation procedure using their provided cross-validation.ipynb.

### 2.5 | Selection Scans

We searched for signals of selection across the DH lines generated for Petkuser Ferdinand Rot (PE DH; Germany) and Kemater Landmais Gelb (KL DH; Austria) as those could be regarded as naturally phased datasets with approximately 500 individuals each. We used *selscan* (Szpiech and Hernandez 2014; Szpiech 2024) to estimate the integrated haplotype score (iHS). For each SNP, we estimated an empirical  $p$ -value by dividing its rank of the absolute iHS value by the number of SNPs. We classified SNPs as being under selection when their  $p$ -value was below 0.05 and their absolute iHS value was > 2. Then, we extracted either the upstream or

downstream genes that were the closest to the selected SNPs. We tested for Gene Ontology enrichment using the online database agriGO v2.2 (Tian et al. 2017).

## 2.6 | Climate Characterisation and Climate Redundancy Analysis

We used historical bioclimatic variables bio1 through bio19 to characterise the population's environment (Fick and Hijmans 2017). For this analysis, we grouped the populations based on east and west separation into clusters, and we compared the distributions of the environments with a Wilcoxon's rank test. Moreover, we correlated the bioclimatic variables with the estimated population genetic parameters using Spearman's  $\rho$  correlation coefficient.

We determined how genetic and environmental factors interact, by performing the multivariate ordination technique, Redundancy Analysis (RDA) (Capblancq and Forester 2021). We inferred the population structure of the landraces by extracting the first two PCs of a PCA on nearly neutral sites. Near neutral sites were identified as all the 4-fold sites using degenotate from <https://github.com/harvardinformatics/degenotate/tree/main>. We selected for the climatic variables that fully explain the genetic structure by a forward model building procedure, by using the function `rda()` from the R package `vegan` (Oksanen et al. 2022). After we identified which bioclimatic variables were part of the model, we tested for the drivers of genetic variation by comparing four different models. A full model that included the selected bioclimatic variables, the first two PCs of neutral sites, latitudinal and longitudinal information as fixed effects; a purely climatic model that incorporated the bioclimatic variables as fixed effects and the geographical and population structure information as 'condition'; a pure neutral population structure model, which incorporated the PCs as fixed and the rest of the parameters as conditional and finally, a geographic model, which had the longitudinal and latitudinal data as fixed effects and the rest as conditional parameters. We selected for the model that best described the dataset based on their  $R^2$  values via the function `RsquareAdj()` from the `vegan` package. All results were plotted with `ggplot2` (Wickham 2016).

For each pair of populations we calculated their environmental distance, as described in (Bay et al. 2018). We first scaled and centered each bioclimatic variable and subsequently calculated the pairwise Euclidean distances between each landrace pair.

## 2.7 | Genomic Load Estimations

We estimated genetic load per individual of the landrace populations and elite pools to estimate the accumulation of deleterious mutations. We used the Genomic Evolutionary Rate Profiling (GERP) scores from Kistler et al. (2018). To estimate genetic load, for each individual belonging to one of the landraces or elite pools, we scored each site based on the number of derived alleles they have and then summed up the scores across all sites. We excluded 100 individuals that had more than 1.69% of missing sites (95th quantile of the distribution) in order to exclude potential differences in load due to missing information. Furthermore, we

estimated the fixed and segregating load within the dent elite and flint breeding pool, as well as for the dent and flint landraces. Finally, we counted the number of highly deleterious alleles, which we defined as sites with a GERP score larger than 5, within each sample. We compared the genetic load distributions between the different groups with a Kolmogorov–Smirnov test.

## 3 | Results

### 3.1 | Genetic Diversity Remains Constant Across European Maize Landraces

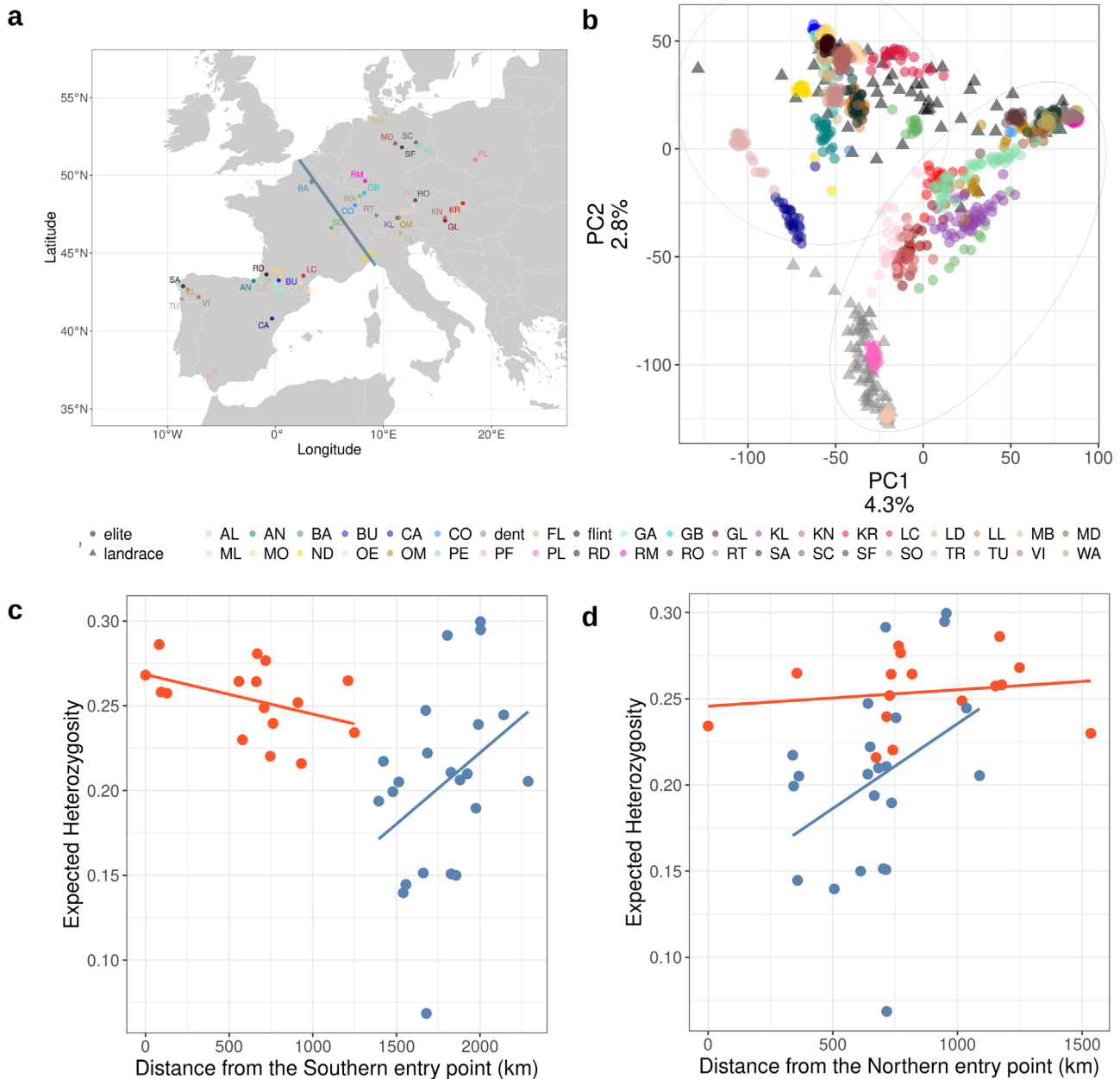
Maize was likely introduced multiple times to Europe through different routes (Brandenburg et al. 2017). This introduction history and the consecutive expansion should be reflected in the population structure of European traditional varieties. We assessed the population structure within European maize of 1253 individuals, which were divided into 38 traditional populations and modern elite inbreds of the two heterotic pools, flint and dent, representing a diverse sampling of the European maize (Figure 1a). We performed a PCA based on 491,477 genome-wide SNP markers. The PCA showed the integrity of landrace populations, as individuals from the same landrace grouped together (Figure 1b). The first two principal components, explaining 7.1% of the variation, mostly separated the populations into a Western and Eastern cluster.

Furthermore, PC2, which explained 2.8% of variation, clearly separated flint (Europe) and dent (US and Europe) maize. It has been shown before that the flint and dent pools are genetically differentiated, despite sharing haplotypes (Brown and Anderson 1947; Unterseer et al. 2016; Haberer et al. 2020). This reflects the long divergent germplasm pools of the two heterotic groups Brown and Anderson 1947; Unterseer et al. 2016. Moreover, there are 2 landrace populations, Polnischer Landmais (PL; Poland) and Altreier (AL; Italy), which clustered with the dent instead of the flint heterotic group. The existence of multiple well-defined populations within the European maize landraces was also supported by a STRUCTURE analysis of the individuals (Figure S1). The STRUCTURE supported the presence of 35 ( $cv=0.38024$ ) separate populations (Mayer et al. 2017), with varying levels of shared ancestry between the populations.

We estimated genetic diversity using expected heterozygosity ( $H_{Exp}$ ) per site within each population (Table S1). The median  $H_{Exp}$  was 0.232 and mean  $H_{Exp}$  was 0.224 across all populations (Figure S2a). The highest  $H_{Exp}$  was observed within the landrace Gleisdorfer (GL; Austria) and lowest for Fleimetal (FL; Italy), with  $H_{Exp}$  0.299 and 0.068, respectively. The elite breeding lines had the highest inbreeding coefficients (median  $F_{IS}=0.992$ ), but a few landraces also had high inbreeding coefficients (Figure S2c). Surprisingly, the populations in Eastern Europe showed higher variation in inbreeding than in the West (Figure 1c).

In line with the theory of range expansion, we looked for patterns of decreased  $H_{Exp}$  across the geographic distance from potential introduction points to study the impact of maize range expansion within Europe (Figure 1c,d). We estimated the Haversine geographical distance between the populations closest to the potential entry points of maize in Europe and the rest of the populations. We





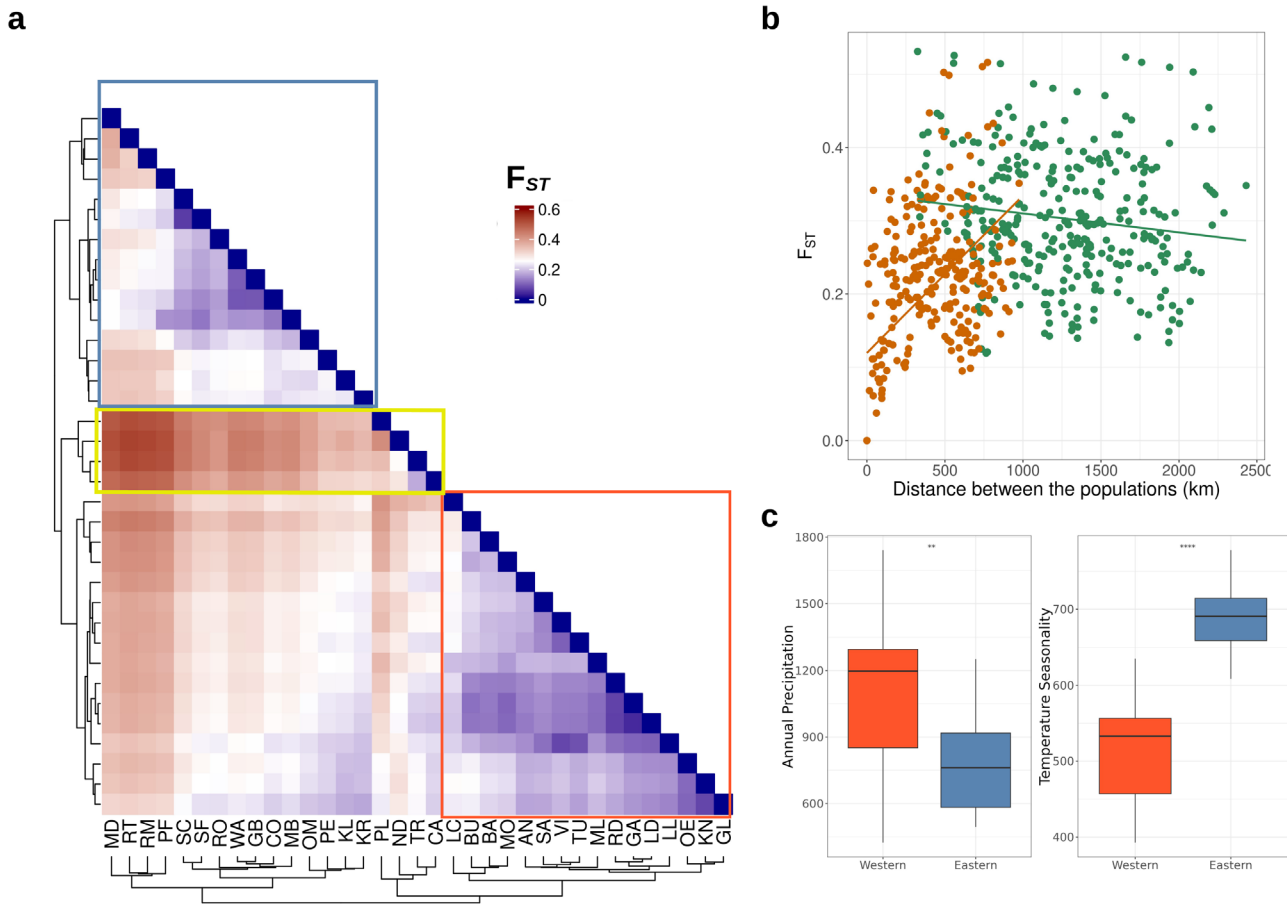
**FIGURE 1** | Population diversity of European maize. (a) Geographic distribution of the landraces in Europe. Blue line denotes approximate geographic position which separates the populations into two clusters, the Eastern and Western European clusters. (b) First two PCs of genomic PCA on 38 landraces and 2 modern elite pools. Colour of dots corresponds to population, shape to the type of population; landrace (circle) or elite breeding line (triangle). The two major clusters identified are marked by red (Western Cluster) and blue (Eastern Cluster). (c, d) Mean expected heterozygosity is plotted against the distance from the potential introduction points in Europe via (c) the South and (d) North Europe. Populations are coloured based on cluster membership; Eastern (blue) or Western (red) cluster.

used the population Tuy (TU) in Spain as the closest population to the entry for introduction via the Southern route and the coordinates of the population Barisis (BA; France) as the closest population to the second entry point in Central Europe, as described in Tenaillon and Charcosset (2011). The  $H_{Exp}$  was significantly negatively correlated with the geographic distance from Tuy (Spain;  $p=0.024$ ,  $\rho=-0.366$ ), which might reflect a pattern of declining genetic diversity along the range expansion route. However, the opposite pattern was observed for the Northern route from Barisis (France;  $p=0.00099$ ,  $\rho=0.518$ ). There was a significant negative correlation ( $p=0.0191$ ,  $\rho=-0.568$ ) of  $H_{Exp}$  and distance within

the Western cluster of populations from the introduction point in the South. No equivalent correlation was detected within the Eastern cluster of populations, or with the introduction point close to Barisis.

### 3.2 | Population Structure and Expansion Across Europe

The calculated  $F_{ST}$  values, which ranged from 0 to 0.531, with the mean and median of 0.255 and 0.257, respectively, supported



**FIGURE 2** | Population clustering of European maize landraces. (a) Pairwise  $F_{ST}$  values between 35 European maize landraces. The Western (red box) and Eastern European (blue box) clusters, and 'dent' cluster (yellow). (b) The  $F_{ST}$  values between each pair of populations against the geographic distance between the populations. Pairs within cluster (either the Western or Eastern European) are coloured in orange, pairs belonging to different clusters are shown in green. (c) Bioclimatic of landrace geographic origin factors by population cluster (both  $p < 0.05$ ).

the presence of separate landrace populations (Figure 2a). The Polnischer Landmais (PL; Poland) population, which clustered with the dent elite lines in the PCA, had on average high  $F_{ST}$  values with the other populations (median of 0.352), corroborating the results of the PCA.

The two population clusters, each of which included landraces from the Western or from Eastern Europe, were also definable by the  $F_{ST}$  values. On average, the  $F_{ST}$  values were significantly lower (Kolmogorov–Smirnov test  $p = 4.663e-09$ ; Figure S3) between populations within the same cluster (mean of 0.259 and median of 0.2542) than between populations between the two clusters (mean of 0.303 and median of 0.295). A third smaller cluster of populations was observed within the Western cluster, which consists of Polnischer landmais (PL; Poland), Nostrano dell Isola (ND; Italy), Tremesino (TR; Spain) and Castellote (CA; Spain). Those correspond to landraces, which clustered with the dent elite lines in the PCA (Figure 1b).

We summarised historical bioclimatic variables (Fick and Hijmans 2017) from the origin of landraces, according to the two major clusters we identified in both the PCA and  $F_{ST}$  analysis. The Western cluster was characterised by significantly higher (Wilcoxon's  $p = 1.6e-05$ ) annual precipitation and overall higher mean temperature in the coldest quarter and minimum

temperatures in the coldest month, with  $p$ -values of  $9e-10$  and  $1e-06$ , respectively (Figure 2c). Meanwhile, the Eastern cluster had a significantly ( $p = 2.6e-10$ ) greater temperature seasonality than the Western cluster. In general, out of the 19 bioclimatic variables, 13 were significantly different between the two population clusters, when using a  $p$ -value cut-off of 0.001 (Figure 2c; Figure S3). Since the population subdivision corresponded well to both climatic and genetic differences, we used this clustering for the following analyses in this study.

The genetic differentiation between populations increased with the distance between them ( $p < 2.2e-16$ ,  $\rho = 0.409$ ; Figure 2b). The same result was reached when we used Slatkin's linearized  $F_{ST}$  ( $p < 2.2e-16$ ,  $\rho = 0.409$ ; Figure S4b). The pattern of isolation by distance was also evident when taking the environmental distance between the populations instead of the physical distance into account. The correlation between  $F_{ST}$  values and environmental distance was  $\rho = 0.282$  ( $p = 4.882e-13$ ; Figure S4c), with the results being the same when we took into account Slatkin's linearized  $F_{ST}$ . This general pattern of correlation with the geographic distance was stronger within the Western European group ( $p = 7.14e-16$ ,  $\rho = 0.678$ ), than the Eastern European cluster of populations ( $p = 1.179e-07$ ,  $\rho = 0.367$ ). This pattern was also observed when the environmental distance was correlated within each

cluster, with the Eastern and Western clusters of populations having positive correlations of 0.346 ( $p = 2.5e-10$ ) and 0.471 ( $p < 2.2e-16$ ), even though the median  $F_{ST}$  value between the Western European populations was 0.167. In contrast, the Eastern European populations had a more defined population structure, with median  $F_{ST}$  of 0.2391. On the contrary, the correlation of  $F_{ST}$  with the environmental distance was higher within each cluster than between clusters. Between the populations of the same cluster the correlation with the environmental distance was  $\rho = 0.290$  ( $p = 9.015e-08$ ), while when the populations belonged in different clusters the  $F_{ST}$  values had a significant negative correlation with the environmental distance ( $p = 0.0016$ ,  $\rho = -0.179$ ; Figure S4c).

The strongest genetic differentiation was observed between the landrace populations of Rheintaler St. Gallen (RT; Switzerland) and Nostrano dell Isola in Italy (ND;  $F_{ST} = 0.531$ ), Rheintaler Monsheim (RM; Germany) and Nostrano dell Isola (ND;  $F_{ST} = 0.525$ ), as well as between Rheintaler St. Gallen (RT; Switzerland) and Tremesino in Spain (TR;  $F_{ST} = 0.523$ ) despite their geographic proximity to each other (approximately 324 and 559 km between RT/ND and RM/ND, respectively). These examples demonstrate the complex local history of maize in Europe.

To further investigate the geographic patterns of diversity, we used the FEEMS analysis, which estimates effective migration surfaces ( $\lambda = 100$ ; cv error = 0.126; Figure S4d). The effective migration rate was higher than neutral ( $\log_{10}(w) > 1$ ) within the Iberian peninsula, with a barrier to migration present along the Pyrenees ( $\log_{10}(w) < 1$ ), reducing gene flow between the Iberian peninsula and Central Europe. Within the Central European cluster of populations, there were surfaces with lower effective migration ( $\log_{10}(w) < 1$ ). Those areas overlapped with the locations of the Rheintaler Monsheim (RM; Germany), Nostrano dell Isola (ND; Italy) and Rheintaler St. Gallen (RT; Switzerland), which also had high  $F_{ST}$  values between them (Figure 2c).

### 3.3 | Evidence for Selection Within Two European Maize Landraces

The observed population structure within Europe could have been the result of selection pressures in the new environment. We scanned the genome of two doubled haploid (DH) line libraries derived from the Petkuser Ferdinand Rot (PE; Germany) and Kemater Landmais Gelb (KL; Austria) landraces (Mayer et al. 2020) for signals of relatively recent positive selection. The iHS statistic measures the relative extended haplotype homozygosity between alleles and indicates signatures of recent selection. Within both Petkuser Ferdinand Rot (PE; Germany) DH and Kemater Landmais Gelb (KL; Austria) DH libraries, we detected SNPs with elevated absolute iHS values (Figure 3).

In total, we found 266 and 103 significant markers for Kemater Landmais Gelb (KL; Austria) DH and Petkuser Ferdinand Rot (PE; Germany) DH respectively ( $p < 0.05$ ; Table S2). In both DH populations the signatures of selection were distributed along all chromosomes. We identified the closest unique genes

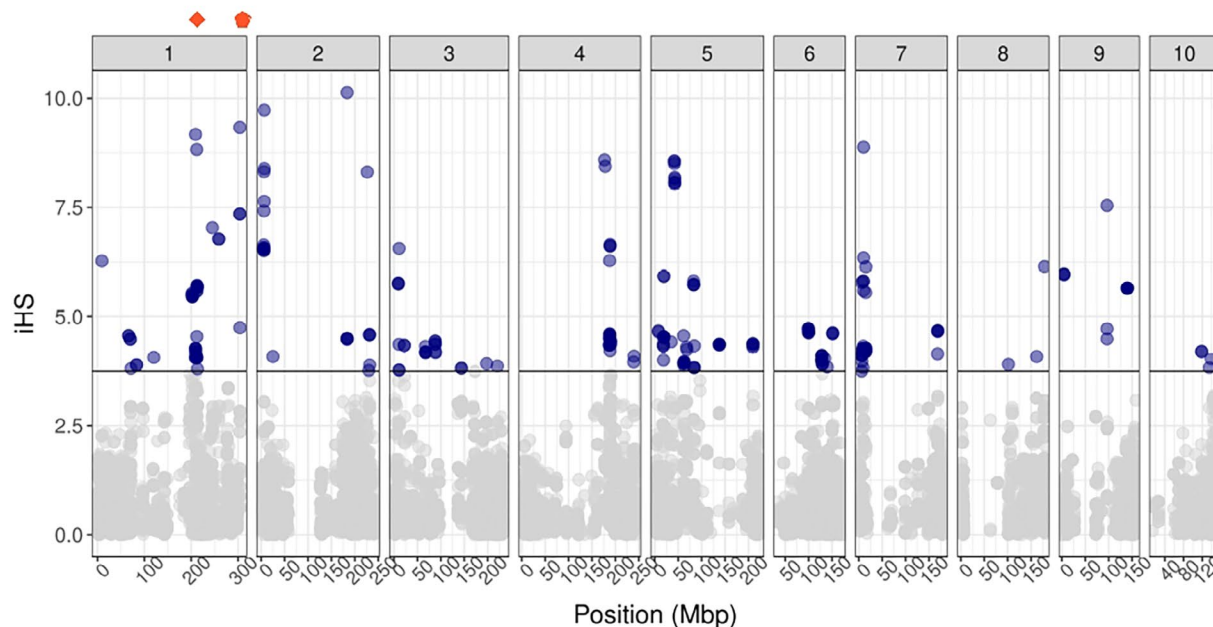
to the selected SNPs, and specifically, we found 140 genes for Kemater Landmais Gelb (KL; Austria). These were significantly enriched ( $p < 0.05$ ) for two Gene Ontology functions 'lipid metabolic process' and 'cellular lipid metabolic process' (Figure S5; Table S2). Moreover, the markers on chromosome 1 overlapped with previously identified selective sweep regions related to maize breeding in the US and China (Huang et al. 2022). Only 48 unique genes were located in close distance upstream or downstream from the identified selection signals in the Petkuser Landmais Rot (PE; Germany) DH, but we did not find overlapping regions under selection in the two populations. Those genes were significantly enriched for the 18 GOs ( $p < 0.05$ ), including as the top 3 the functions 'regulation of nitrogen compound metabolic process', 'regulation of cellular process' and 'regulation of biological process' (Figure S5; Table S3). Hence, this might reflect selection based on the local environment rather than the introduction to Europe or the general breeding process.

### 3.4 | Climatic Gradients Along the Expansion Routes Could Explain Part of the Observed Genetic Variance

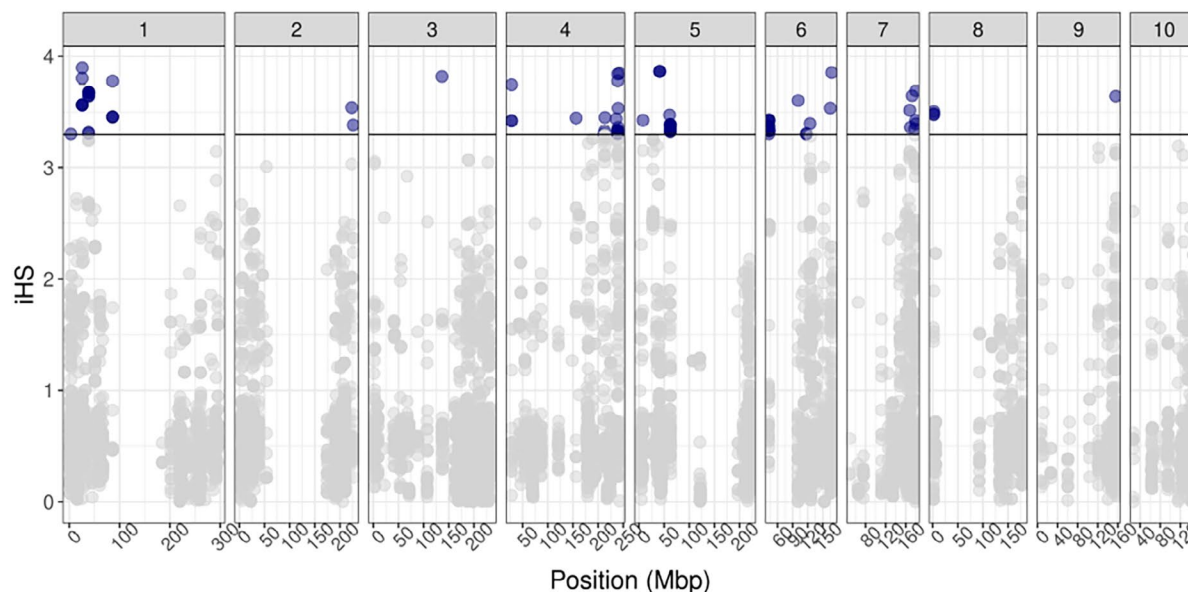
We further investigated the link between the local environment and genetic diversity within European maize landraces using a redundancy analysis (Capblancq and Forester 2021). We used the historical bioclimatic factors for this purpose, latitude and longitude, and corrected for population structure by using neutral variation. We identified three bioclimatic factors, which explained the largest variance. The first bioclimatic variable was 'mean temperature of the driest quarter' (adjusted  $R^2 = 0.132$ ;  $p = 0.002$ ), then followed the 'precipitation seasonality' (adjusted  $R^2 = 0.159$ ;  $p = 0.028$ ) and 'precipitation of coldest quarter' (adjusted  $R^2 = 0.176$ ). Interestingly, two of those bioclimatic variables were also significantly different between the Eastern and Western population clusters. The 'mean temperature of driest quarter' and the 'precipitation of the coldest quarter' were significantly higher in the environment of the Western cluster than the Eastern one, with Wilcoxon test  $p$ -values of  $3.6e-05$  and  $4.8e-05$ , respectively.

When we correlated each population's value for the three bioclimatic variables with the distance to the two different hypothetical introduction points of maize in Europe, we were able to identify gradients of climatic change across the expansion routes. Along the introduction route starting in Spain, the bioclimatic variables 'mean temperature of driest quarter' and the 'precipitation of the coldest quarter' were significantly negatively correlated ( $p < 2.2e-16$ ), with  $\rho$  values of  $-0.76$  and  $-0.74$ , respectively. On the contrary, the bioclimatic variable "precipitation seasonality" had a significant positive correlation of  $\rho = 0.476$  ( $p = 0.0027$ ) with the distance from the introduction point within the Eastern Cluster in Barisis (BA; France). However, the RDA model that had the best fit on the dataset, based on the adjusted  $R^2$  value of 0.39 ( $p = 0.001$ ), was the full model, which incorporates information about the latitude, longitude and population structure besides the three bioclimatic variables (Figure S6). Therefore, we inferred that the climatic conditions along the expansion routes partly contributed to the genetic diversity across the European landscape.

KL



PE



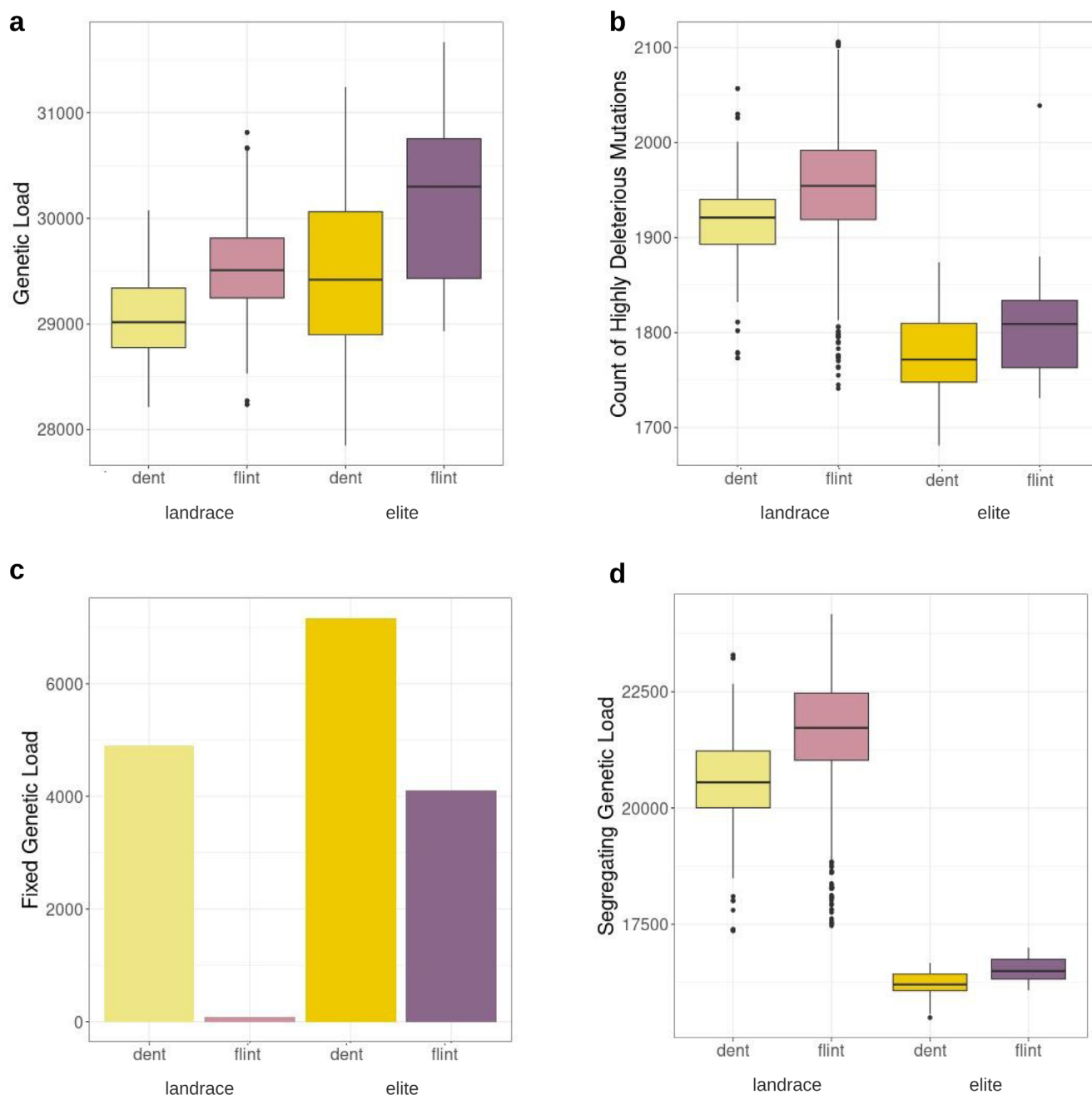
**FIGURE 3** | Recent selective sweeps in two European maize double haploid (DH) libraries. The absolute iHS values across the genome are plotted for the Kemater Landmais Gelb (KL; Austria) DH (top) and Petkuser Landmais Rot (PE; Germany) DH (bottom) populations. In blue are highlighted the positions that are significant for being under positive selection (empirical  $p$  value cut-off of 0.05 on the ranked iHS values). In the Kemater Landmais Gelb population the two locations are marked that have been previously identified as under selection in Chinese (diamond) or US (pentagon) breeding lines.

### 3.5 | Change of Genetic Load Along the Introduction Routes and Due to Breeding

Range expansion has been previously associated with the accumulation of deleterious mutations and overall genetic load (de Pedro et al. 2021; Peischl et al. 2013). We estimated the genetic load per population, including the elite breeding pools, using the GERP scores from Kistler et al. (2018). Estimates of genetic load per population differed significantly ( $p < 2.2e-16$ ), with the population Rheintaler Monsheim (RM; Germany) having the highest median genetic load (sum of GERP=29,881.23), while

the population with the least accumulated genetic load (sum of GERP=29,578.20) was Bugard (BU; France) (Figure S7). We further tested for correlation between accumulated genetic load and the distance from the two hypothetical introduction points in Southern and Central Europe and found significant correlations of  $\rho = -0.16$  ( $p = 1.18e-11$ ) and  $\rho = -0.156$  ( $p = 3.845e-11$ ), respectively. The accumulated genetic load also had a significant ( $p < 2.2e-16$ ) negative correlation with the bioclimatic factor ‘mean temperature of the driest quarter’ of  $\rho = -0.214$ . However, there was no positive correlation with the other two bioclimatic factors identified as important for partitioning the dataset’s variance.





**FIGURE 4** | Purging segregating deleterious genetic load within the elite breeding lines. (a) Total genetic load estimated as the sum of the GERP scores for all sites within an individual for dent landraces, flint landraces, flint elite lines, and dent elite lines. (b) Number of highly deleterious alleles. Count of derived alleles at sites with GERP score larger than 5. (c) Fixed genetic load for each group, given as the sum of the GERP scores at fixed derived alleles within each group of maize lines. (d) Segregating genetic load per individual for each of the four groups as the sum of GERP scores of the segregating sites.

Furthermore, we tested the differences in genetic load between landrace and elite groups, as inbreeding, drift and selection during breeding might have changed genetic load. We grouped both the landraces and the elite lines based on their germplasm origin of either dent or flint, as identified in the PCA (Figure 4a; Table S1). Genetic load was significantly higher in both elite pools than their counterpart landrace group, with  $p$ -values of 0.0486 and 0.01346 for flint and dent germplasm, respectively. The accumulated genetic load was also different between the dent landraces and the flint elite

germplasm lines ( $p = 0.0162$ ). Those differences were not observed for highly deleterious mutations. Both elite pools had a significantly ( $p < 2.2 \times 10^{-16}$ ) lower number of highly deleterious mutations than the landraces (Figure 4b), suggesting potential purging during breeding. In contrast, fixed load was significantly higher ( $p < 2.2 \times 10^{-16}$ ) in the elite lines than the landraces (Figure 4c), while the segregating load was significantly higher ( $p < 2.2 \times 10^{-16}$ ) in the landraces (Figure 4d). Those differences were observed irrespective of whether the lines belonged to the flint or dent germplasm.

## 4 | Discussion

### 4.1 | Genetic Diversity Is Shaped by Local Factors Rather Than the Introduction Route

Maize has been introduced to Europe within the last 500 years, and its range expanded rapidly throughout the European continent (Janick and Caneva 2005). Evidence exists for two introductory routes to Europe: one from the Caribbean via Southern Europe and one to Central Europe with material from North America (Tenaillon and Charcosset 2011; Brandenburg et al. 2017). Along this expansion route, local openly pollinating traditional varieties have formed and have been kept and traded by local farmers for centuries. Indeed, the population structure and relatedness between samples of these traditional varieties trace the history of maize in Europe. We used a large sample of European maize landraces to explore how the introduction and expansion of maize in Europe has shaped genetic structure and diversity. The wide geographic sampling and representation of each landrace population enabled us to track the impact of colonisation of novel environments on a more detailed scale. The analysis of genetic diversity within European landrace populations showed three major clusters across populations (Figure 1). One cluster, which includes dent elite lines, shows that 4 landraces are most likely derived from North American dent populations, while all other landraces clustered with flint elite lines, confirming their flint ancestry and their involvement in the creation of European maize hybrid pools (Mayer et al. 2020; Tenaillon and Charcosset 2011; Brandenburg et al. 2017). The four landraces have been previously noted for their difference in genetic composition from the rest of the European landraces (Mayer et al. 2017). The further separation into two groups within the flint cluster across Europe might represent the two introduction routes of maize into Europe (Tenaillon and Charcosset 2011).

Range expansion can have a strong influence on the genetic diversity of the expanding species, specifically by decreasing genetic diversity and increasing genetic differentiation between the newly established and core populations (Excoffier et al. 2009). Despite the pattern of introduction routes and ancestry within Europe, we did not observe a pattern of decreasing genetic diversity along the potential introduction routes. We used the potential introduction locations of maize from Brandenburg et al. (2017) to investigate the impact of expansion along the separate routes, but did not find a pattern that would be expected for the colonisation by a natural species (Zimmermann et al. 2014; Smith et al. 2020), which indicates the maintenance of genetic diversity across the continent. On the contrary, genetic diversity is maintained in similar levels across the studied maize populations, suggesting a stronger influence of local maintenance and potential adaptation. This is reflected in the phenotypic level; some of the European maize landraces have been documented to show differences in early vigour, plant height and tillering (Hölker et al. 2019), as well as flowering (Balconi et al. 2024). However, part of this pattern can be explained by the presence of two distinct germplasm pools. Within Europe, the observed pattern is different from what has been observed within the Americas, where a decrease in genetic diversity has been documented

(Huang et al. 2022). American maize landraces have followed an expansion route from Central Mexico to North and South America, allowing tracking of changes in genetic composition along the route (Tenaillon and Charcosset 2011; Huang et al. 2022; Yang et al. 2023). In Europe, however, the pattern is less clear and suggests long-range dispersal and a more rapid spread through the continent (Figures 1 and 2).

In potato and European bean, admixture and different admixture histories in the introduced range have been suggested as reason for maintenance of diversity in introduced ranges but often little population structure is observed within a smaller geographic region, such as in the Nordic countries (Ortiz et al. 2023; Gutaker et al. 2019; Ames and Spooner 2008; Bellucci et al. 2023). Admixture between Central and Southern European maize populations might be prevalent (Bradbury et al. 2007; Mayer et al. 2017; Unterseer et al. 2016), however, we find that not only there is population structure dividing the landraces into multiple geographic distinct populations, but also an evident pattern of isolation by distance. We identify smaller local genetic clusters between landraces, including one along the Rhine and into northern Germany that consisted of four populations (Figure 2). The  $F_{ST}$  values between those populations and their neighbouring ones are up to 0.5. This sharp distinction likely represents historical trade routes and political and linguistic borders in the area during introduction of maize to Europe. Tenaillon and Charcosset (2011) described that Caribbean maize was introduced to Southern Europe by Columbus in the 15th Century and from there it was introduced to the Vatican and Italy.

### 4.2 | Genomic Regions Selected During Maize Introduction Due to Adaptation to Local Environments

We used two very large populations of DH lines derived from divergent landraces to infer signals of recent selection in the landraces. The advantage of such a population is their natural phasing through the DH process, which removes problems with computational phasing in highly heterozygous outcrossing maize. We were able to detect signals of positive selection within the Central European Petkuser Ferdinand Rot (PE; Germany) landrace and Kemater Landmais Rot (KL; Austria) landrace. We found 53 and 3 potentially selected loci in each population but no overlap in selection signals, confirming the importance of local selection in the formation of these local landraces. In addition, we found a partial overlap of detected selection signals in one of the two populations with selective sweeps related to breeding efforts in American and Asian breeding maize lines (Huang et al. 2022), suggesting convergent selection targets in different regions of the world. Additional selection signals might be linked to adaptation to the novel environments and subsequent breeding targets that maize encountered after its arrival in Europe. Compared to the Caribbean and Mexican environments, which range from tropical highlands to wet and dry lowlands, the temperate European environment offered novel challenges to adapt to (Bellon et al. 2011; Brandenburg et al. 2017). Even though we have only tested selection in two populations, the traces of selection corroborate with the results of the population structure within European maize. Local adaptation can be achieved by

even small changes in allele frequencies within populations (Le Corre and Kremer 2012), especially for polygenic traits such as the ones observed here. Regions of high genetic differentiation have been found to harbour genetic variation linked to adaptation to local climate in various plant species (De La Torre et al. 2019) and it can occur rapidly during range expansion (Colautti and Barrett 2013), even if there is extensive gene flow (Hämälä and Savolainen 2019).

The local environment can have profound effects on genetic diversity. Patterns of maintained genetic diversity and population structure in the short evolutionary times of European maize might be related to local environmental adaptation, but they have also been connected to human activity in other crops (Bellucci et al. 2023). For instance, human activity has facilitated the geographic spread and genetic differentiation of rice (Courtois et al. 2012). Within Southern Europe, landrace variability and genetic structure of tomato landraces, which were introduced from South America at a similar time as maize, are related to the fruit type (Corrado et al. 2014; García-Martínez et al. 2013), a highly coveted breeding trait.

We document three historical bioclimatic factors, ‘mean temperature of the driest quarter’, ‘precipitation seasonality’ and ‘precipitation of coldest quarter’, which significantly explain part of the genetic variance. This indicates that at least part of the genetic variation is due to adaptation to the more diverse environments of the Iberian peninsula and Central Europe. The idea is reinforced, both, by the significant differences of the two European clusters and by the gradual change of the significant bioclimatic factors along the possible introduction routes in Europe. Historically, crops such as maize and potato were introduced in Europe during the 15th to 17th centuries to solve problems created by famines, which took place due to cold and dry years (Ljungqvist et al. 2024). The three historical bioclimatic factors could be related to the tendency of farmers to select traits that would improve the survivability of maize plants under those climatic conditions. This practice has been proposed as a solution for crop improvement under current and future climate change scenarios in crops such as pearl millet (Rhoné et al. 2020), and even maize (Tamang et al. 2024).

However, we cannot exclude the possibility that some traits were selected as reaction to local micro-environments. For instance, within the double haploid population of Kemater Landmais (KL; Austria), genes close to the markers under selection are enriched for the function ‘lipid metabolic process’ and ‘cellular lipid metabolic process’. Tropical maize lines have adapted to colder temperatures via introgression of alleles related to lipid metabolism (Barnes et al. 2022). The Petkuser Ferdinand Rot (PE; Germany) population showed enrichment for functions such as ‘obsolete positive regulation of nitrogen compound metabolic process’. Adaptations to low nitrogen in sub-Saharan African maize has been documented (Worku et al. 2007).

### 4.3 | Local Purging of Genetic Load Due to Inbreeding and Selection

Finally, we estimated the genetic load within the European landrace populations. In maize, deleterious mutations accumulate

within complex phenotypes of interest, even though, overall, they are maintained at low frequencies (Mezmouk and Ross-Ibarra 2014). However, the accumulation of genetic load during range expansion has been observed in the Americas (Wang et al. 2017). Here, we observed that genetic load within the European landrace populations is variable but does not follow a strong pattern of accumulation in any cluster. Population subdivision, similar to the one observed here, could maintain the genetic load in similar levels (Glémin et al. 2003), though the observed migration between the different landraces could have also helped to alleviate the accumulation of load across the expansion route (Theodorou and Couvet 2006). Additionally, the introduction of maize in Europe has been relatively recent, and therefore the crop spread throughout the European continent relatively fast. This fast range expansion could have limited genetic load accumulation (Gilbert et al. 2018), leading to the absence of a correlation between load and geographic distance from the possible introduction points in Europe. On the other hand, adaptation to climatic parameters and environmental gradient can promote the accumulation of genetic load (Gilbert et al. 2017; Fiscus et al. 2024), maintaining differences between populations as we observe here. Indeed, we observe a significant correlation between one of the bioclimatic factors (mean temperature of the driest quarter) and the accumulation of genetic load. This bioclimatic parameter is significantly associated with differences between the two population clusters and has a positive association with longitude. Together, they add support to the partial impact of the environment to accumulation of load, but the correlation between location and environment might obscure the causal reason for load accumulation.

When we compared the landraces to the elite breeding lines in our populations, we found that the number of highly deleterious mutations decreased in the elite lines, while the fixed load was significantly higher (Figure 4b,c). These results suggest purging of highly deleterious alleles but fixation of mildly deleterious load, potentially due to inbreeding and decreased population size during selective breeding. For example, in *Arabidopsis thaliana*, selfing, in contrast to outcrossing, has led to a decrease in the highly deleterious load (Zeitler et al. 2020). Inbreeding can effectively remove the deleterious mutations, particularly when they are recessive, making it more effective than selection alone (Glémin 2003), which might explain the higher load in landrace populations. Purging of highly deleterious mutations has been shown to decrease load in maize compared to teosinte, reducing inbreeding depression in maize compared to teosinte (Samayoa et al. 2021). Moreover, the need for inbred fitness in breeding programs would further favour purging during breeding and can lead to the removal of especially lethal mutations by inbreeding (Wang et al. 1999). Yet, the high genetic diversity observed within the landraces might enable masking of deleterious alleles but leads to the higher segregating genetic load. Masking of fixed mildly deleterious alleles in hybrids might contribute to heterosis in European maize hybrids (Yang, Mezmouk, et al. 2017), but further reduction of fixed genetic load might enable further selective gain.

Taken together, our results illustrate the complexity of the introduction of maize to Europe and the individual history of landraces within the continent and provide an example of local populations maintaining a broad genetic diversity. We

have expanded the knowledge on the genetic composition of European maize landraces by exploring how it has been shaped by possible introduction events and adaptation to novel environments. We show how genetic load accumulation during rapid range expansion and through the impact of breeding has shaped modern maize.

## Author Contributions

M.G.S. conceived the study. M.T. and K.S. conducted the analysis. M.T. prepared figures and tables. M.T., M.G.S. and K.S. wrote the manuscript. All authors discussed the results, edited and approved the manuscript.

## Acknowledgements

We would like to thank A. Singh for valuable feedback on the manuscript. We thank the Schoen lab for making their data publicly available and C.C. Schoen and M. Meyer for discussions about European maize landraces and their clustering. We acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2048/1—project ID 390686111 and TRR341/1-456082119 and financial support by the Federal Ministry of Education and Research (BMBF, Germany) under the Plant Breeding Research for the Bioeconomy initiative (funding ID: 031B0195, project MAZE) to M.G.S. Open Access funding enabled and organized by Projekt DEAL.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

All data used are publicly available from their respective publications of origin (Unterseer et al. 2016; Mayer et al. 2017, 2020). The merged dataset is available at [10.6084/m9.figshare.28219994](https://doi.org/10.6084/m9.figshare.28219994). All scripts used in the analysis are available under <https://github.com/cropevolution/europeanMaize.git>.

## References

- Ames, M., and D. M. Spooner. 2008. "Dna From Herbarium Specimens Settles a Controversy About Origins of the European Potato." *American Journal of Botany* 95: 252–257.
- Arca, M., B. Gouesnard, T. Mary-Huard, et al. 2023. "Genotyping of Dna Pools Identifies Untapped Landraces and Genomic Regions to Develop Next-Generation Varieties." *Plant Biotechnology Journal* 21: 1123–1139.
- Austerlitz, F., B. Jung-Muller, B. Godelle, and P.-H. Gouyon. 1997. "Evolution of Coalescence Times, Genetic Diversity and Structure During Colonization." *Theoretical Population Biology* 51: 148–164.
- Balconi, C., A. Galaretto, R. A. Malvar, et al. 2024. "Genetic and Phenotypic Evaluation of European Maize Landraces as a Tool for Conservation and Valorization of Agrobiodiversity." *Biology* 13: 454.
- Barnes, A. C., F. Rodríguez-Zapata, K. A. Juárez-Núñez, et al. 2022. "An Adaptive Teosinte Mexicana Introgression Modulates Phosphatidylcholine Levels and Is Associated With Maize Flowering Time." *Proceedings of the National Academy of Sciences* 119: e2100036119.
- Bay, R. A., R. J. Harrigan, V. L. Underwood, H. L. Gibbs, T. B. Smith, and K. Ruegg. 2018. "Genomic Signals of Selection Predict Climate-Driven Population Declines in a Migratory Bird." *Science* 359, no. 6371: 83–86. <https://doi.org/10.1126/science.aan4380>.
- Bellon, M. R., D. Hodson, and J. Hellin. 2011. "Assessing the Vulnerability of Traditional Maize Seed Systems in Mexico to Climate Change." *Proceedings of the National Academy of Sciences* 108: 13432–13437.

- Bellucci, E., A. Benazzo, C. Xu, et al. 2023. "Selection and Adaptive Introgression Guided the Complex Evolutionary History of the European Common Bean." *Nature Communications* 14: 1908.
- Bhatia, G., N. Patterson, S. Sankaraman, and A. L. Price. 2013. "Estimating and Interpreting Fst: The Impact of Rare Variants." *Genome Research* 23: 1514–1521.
- Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, and E. S. Buckler. 2007. "Tassel: Software for Association Mapping of Complex Traits in Diverse Samples." *Bioinformatics* 23: 2633–2635.
- Brandenburg, J.-T., T. Mary-Huard, G. Rigai, et al. 2017. "Independent Introductions and Admixtures Have Contributed to Adaptation of European Maize and Its American Counterparts." *PLoS Genetics* 13: e1006666.
- Brandolini, A., and A. Brandolini. 2009. "Maize Introduction, Evolution and Diffusion in Italy." *Maydica* 54: 233.
- Brown, W. L., and E. Anderson. 1947. "The Northern Flint Corns." *Annals of the Missouri Botanical Garden* 34: 1–29.
- Capblancq, T., and B. R. Forester. 2021. "Redundancy Analysis: A Swiss Army Knife for Landscape Genomics." *Methods in Ecology and Evolution* 12: 2298–2309.
- Casañas, F., J. Simó, J. Casals, and J. Prohens. 2017. "Toward an Evolved Concept of Landrace." *Frontiers in Plant Science* 8: 145.
- Chang, C. C., C. C. Chow, L. C. Tellier, et al. 2015. "Second-Generation Plink: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4: 7.
- Colautti, R. I., and S. C. Barrett. 2013. "Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant." *Science* 342: 364–366.
- Corrado, G., M. Caramante, P. Piffanelli, and R. Rao. 2014. "Genetic Diversity in Italian Tomato Landraces: Implications for the Development of a Core Collection." *Scientia Horticulturae* 168: 138–144.
- Le Corre, V., and A. Kremer. 2012. "The Genetic Differentiation at Quantitative Trait Loci Under Local Adaptation." *Molecular Ecology* 21: 1548–1566.
- Courtois, B., J. Frouin, R. Greco, et al. 2012. "Genetic Diversity and Population Structure in a European Collection of Rice." *Crop Science* 52: 1663–1675.
- Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and Vcftools." *Bioinformatics* 27: 2156–2158.
- De La Torre, A. R., B. Wilhite, and D. B. Neale. 2019. "Environmental Genome-Wide Association Reveals Climate Adaptation Is Shaped by Subtle to Moderate Allele Frequency Shifts in Loblolly Pine." *Genome Biology and Evolution* 11: 2976–2989.
- de Pedro, M., M. Riba, S. C. González-Martínez, et al. 2021. "Demography, Genetic Diversity and Expansion Load in the Colonizing Species *Leontodon Longirostris* (Asteraceae) Throughout Its Native Range." *Molecular Ecology* 30: 1190–1205.
- Excoffier, L., M. Foll, and R. J. Petit. 2009. "Genetic Consequences of Range Expansions." *Annual Review of Ecology, Evolution, and Systematics* 40: 481–501.
- Fick, S. E., and R. J. Hijmans. 2017. "Worldclim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37: 4302–4315.
- Finan, J. J. 1948. "Maize in the Great Herbals." *Annals of the Missouri Botanical Garden* 35: 149–191.
- Fiscus, C. J., J. A. Aguirre-Liguori, G. R. Gaut, and B. S. Gaut. 2024. "Climate, Population Size, and Dispersal Influences Mutational Load Across the Landscape in *Vitis Arizonica*." *bioRxiv*.
- García-Martínez, S., G. Corrado, J. J. Ruiz, and R. Rao. 2013. "Diversity and Structure of a Sample of Traditional Italian and Spanish Tomato Accessions." *Genetic Resources and Crop Evolution* 60: 789–798.



- Gilbert, K. J., S. Peischl, and L. Excoffier. 2018. "Mutation Load Dynamics During Environmentally-Driven Range Shifts." *PLoS Genetics* 14: e1007450.
- Gilbert, K. J., N. P. Sharp, A. L. Angert, et al. 2017. "Local Adaptation Interacts With Expansion Load During Range Expansion: Maladaptation Reduces Expansion Load." *American Naturalist* 189: 368–380.
- Glémin, S. 2003. "How Are Deleterious Mutations Purged? Drift Versus Nonrandom Mating." *Evolution* 57: 2678–2687.
- Glémin, S., J. Ronfort, and T. Bataillon. 2003. "Patterns of Inbreeding Depression and Architecture of the Load in Subdivided Populations." *Genetics* 165: 2193–2212.
- González-Martínez, S. C., K. Ridout, and J. R. Pannell. 2017. "Range Expansion Compromises Adaptive Evolution in an Outcrossing Plant." *Current Biology* 27: 2544–2551.
- Gu, Z., R. Eils, and M. Schlesner. 2016. "Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data." *Bioinformatics* 32: 2847–2849.
- Gutaker, R. M., C. L. Weiß, D. Ellis, et al. 2019. "The Origins and Adaptation of European Potatoes Reconstructed From Historical Genomes." *Nature Ecology & Evolution* 3: 1093–1101.
- Haberer, G., N. Kamal, E. Bauer, et al. 2020. "European Maize Genomes Highlight Intraspecific Variation in Repeat and Gene Content." *Nature Genetics* 52: 950–957.
- Hämälä, T., and O. Savolainen. 2019. "Genomic Patterns of Local Adaptation Under Gene Flow in *Arabidopsis lyrata*." *Molecular Biology and Evolution* 36: 2557–2571.
- Harris, C. R., K. J. Millman, S. J. van der Walt, et al. 2020. "Array Programming With NumPy." *Nature* 585: 357–362.
- Hewitt, G. 2000. "The Genetic Legacy of the Quaternary Ice Ages." *Nature* 405: 907–913.
- Hijmans, R. J. 2022. "Geosphere: Spherical Trigonometry." R Package Version 1.5–18.
- Hölker, A. C., M. Mayer, T. Presterl, et al. 2019. "European Maize Landraces Made Accessible for Plant Breeding and Genome-Based Studies." *Theoretical and Applied Genetics* 132: 3333–3345.
- Huang, X., S. Huang, B. Han, and J. Li. 2022. "The Integrated Genomics of Crop Domestication and Breeding." *Cell* 185: 2828–2839.
- Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. "Estimation of Levels of Gene Flow From Dna Sequence Data." *Genetics* 132: 583–589.
- Hufford, M. B., E. Martínez-Meyer, B. S. Gaut, L. E. Eguiarte, and M. I. Tenaillon. 2012. "Inferences From the Historical Distribution of Wild and Domesticated Maize Provide Ecological and Evolutionary Insight." *PLoS One* 7: e47659.
- Hunter, J. D. 2007. "Matplotlib: A 2d Graphics Environment." *Computing in Science & Engineering* 9: 90–95.
- Inc., P. T. 2015. "Collaborative Data Science."
- Janick, J., and G. Caneva. 2005. "The First Images of Maize in Europe." *Maydica* 50: 71.
- Kistler, L., S. Y. Maezumi, J. de Gregorio Souza, et al. 2018. "Multiproxy Evidence Highlights a Complex Evolutionary Legacy of Maize in South America." *Science* 362: 1309–1313.
- Leng, E. R., A. Tavčar, and V. Trifunović. 1962. "Maize of Southeastern Europe and Its Potential Value in Breeding Programs Elsewhere." *Euphytica* 11: 263–272.
- Li, H. 2011. "A Statistical Framework for Snp Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation From Sequencing Data." *Bioinformatics* 27: 2987–2993.
- Ljungqvist, F. C., A. Seim, and D. Collet. 2024. "Famines in Medieval and Early Modern Europe—Connecting Climate and Society." *Wiley Interdisciplinary Reviews: Climate Change* 15: e859.
- Maqbool, M. A., A. Beshir, and E. S. Khokhar. 2020. "Doubled Haploids in Maize: Development, Deployment, and Challenges." *Crop Science* 60: 2815–2840.
- Marcus, J., W. Ha, R. F. Barber, and J. Novembre. 2021. "Fast and Flexible Estimation of Effective Migration Surfaces." *eLife* 10: e61927.
- Mayer, M., A. C. Hölker, E. González-Segovia, et al. 2020. "Discovery of Beneficial Haplotypes for Complex Traits in Maize Landraces." *Nature Communications* 11, no. 1: 1–10. <https://doi.org/10.1038/s41467-020-18683-3>.
- Mayer, M., S. Unterseer, E. Bauer, N. de Leon, B. Ordas, and C. C. Schön. 2017. "Is There an Optimum Level of Diversity in Utilization of Genetic Resources?" *Theoretical and Applied Genetics* 130: 2283–2295.
- Mezmouk, S., and J. Ross-Ibarra. 2014. "The Pattern and Distribution of Deleterious Mutations in Maize." *G3: Genes, Genomes, Genetics* 4: 163–171.
- Miles, A., pyup.io bot, M. R. P. Ralph, and J. Kelleher. 2023. "cggh/scikit-allel: v1.3.6."
- Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, and P. Legendre. 2022. "Vegan: Community Ecology Package. R package Version 2.6-4."
- Ortiz, R., F. Reslow, R. Vetukuri, M. R. García-Gil, P. Pérez-Rodríguez, and J. Crossa. 2023. "Inbreeding Effects on the Performance and Genomic Prediction for Polysomic Tetraploid Potato Offspring Grown at High Nordic Latitudes." *Genes* 14: 1302.
- Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. "On the Accumulation of Deleterious Mutations During Range Expansions." *Molecular Ecology* 22: 5972–5982.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155: 945–959.
- Purugganan, M. D. 2019. "Evolutionary Insights Into the Nature of Plant Domestication." *Current Biology* 29: R705–R714.
- Purugganan, M. D., and D. Q. Fuller. 2009. "The Nature of Selection During Plant Domestication." *Nature* 457: 843–848.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Rebourg, C., M. Chastanet, B. Gouesnard, C. Welcker, P. Dubreuil, and A. Charcosset. 2003. "Maize Introduction Into Europe: The History Reviewed in the Light of Molecular Data." *Theoretical and Applied Genetics* 106: 895–903.
- Reif, J. C., S. Hamrit, M. Heckenberger, et al. 2005. "Trends in Genetic Diversity Among European Maize Cultivars and Their Parental Components During the Past 50 Years." *Theoretical and Applied Genetics* 111: 838–845.
- Revilla Temiño, P., M. d. P. Soengas Fernández, M. E. Cartea González, R. A. Malvar Pintos, and A. Ordás Pérez. 2003. "Isozyme Variability Among European Maize Populations and the Introduction of Maize in Europe." *Maydica* 48: 141–152.
- Rhoné, B., D. Defrance, C. Berthouly-Salazar, et al. 2020. "Pearl Millet Genomic Vulnerability to Climate Change in West Africa Highlights the Need for Regional Collaboration." *Nature Communications* 11: 5274.
- Samayoa, L. F., B. A. Olukolu, C. J. Yang, et al. 2021. "Domestication Reshaped the Genetic Basis of Inbreeding Depression in a Maize Landrace Compared to Its Wild Relative, Teosinte." *PLoS Genetics* 17: e1009797.
- Savolainen, O., M. Lascoux, and J. Merilä. 2013. "Ecological Genomics of Local Adaptation." *Nature Reviews Genetics* 14: 807–820.

- Slatkin, M., and L. Excoffier. 2012. "Serial Founder Effects During Range Expansion: A Spatial Analog of Genetic Drift." *Genetics* 191: 171–181.
- Smaragdov, M., and A. Kudinov. 2020. "Assessing the Power of Principal Components and Wright's Fixation Index Analyzes Applied to Reveal the Genome-Wide Genetic Differences Between Herds of Holstein Cows." *BMC Genetics* 21: 1–15.
- Smith, A. L., T. R. Hodkinson, J. Villellas, et al. 2020. "Global Gene Flow Releases Invasive Plants From Environmental Constraints on Genetic Diversity." *Proceedings of the National Academy of Sciences* 117: 4218–4227.
- Szpiech, Z. A. 2024. "Selscan 2.0: Scanning for Sweeps in Unphased Data." *Bioinformatics* 40: btac006.
- Szpiech, Z. A., and R. D. Hernandez. 2014. "Selscan: An Efficient Multithreaded Program to Perform Ehh-Based Scans for Positive Selection." *Molecular Biology and Evolution* 31: 2824–2827.
- Takou, M., T. Hämälä, E. M. Koch, et al. 2021. "Maintenance of Adaptive Dynamics and no Detectable Load in a Range-Edge Outcrossing Plant Population." *Molecular Biology and Evolution* 38: 1820–1836.
- Tamang, A., M. W. Macharia, L. Caproni, et al. 2024. "Genomic, Climatic, and Cultural Diversity of Maize Landraces From the Himalayan Kingdom of Bhutan." *Plants, People, Planet* 6, no. 4: 965–978. <https://doi.org/10.1002/ppp3.10513>.
- Tenaillon, M. I., and A. Charcosset. 2011. "A European Perspective on Maize History." *Comptes Rendus Biologies* 334: 221–228.
- Theodorou, K., and D. Couvet. 2006. "Genetic Load in Subdivided Populations: Interactions Between the Migration Rate, the Size and the Number of Subpopulations." *Heredity* 96: 69–78.
- Tian, T., Y. Liu, H. Yan, et al. 2017. "Agrigo v2. 0: A Go Analysis Toolkit for the Agricultural Community, 2017 Update." *Nucleic Acids Research* 45: W122–W129.
- Tittes, S., A. Lorant, S. McGinty, et al. 2023. "Not So Local: The Population Genetics of Convergent Adaptation in Maize and Teosinte." *eLife* 12: PR92405.
- Unterseer, S., E. Bauer, G. Haberer, et al. 2014. "A Powerful Tool for Genome Analysis in Maize: Development and Evaluation of the High Density 600 k Snp Genotyping Array." *BMC Genomics* 15, no. 1: 1–15. <https://doi.org/10.1186/1471-2164-15-823>.
- Unterseer, S., S. D. Pophaly, R. Peis, et al. 2016. "A Comprehensive Study of the Genomic Differentiation Between Temperate Dent and Flint Maize." *Genome Biology* 17, no. 1: 1–14. <https://doi.org/10.1186/s13059-016-1009-x>.
- Villa, T. C. C., N. Maxted, M. Scholten, and B. Ford-Lloyd. 2005. "Defining and Identifying Crop Landraces." *Plant Genetic Resources* 3: 373–384.
- Waldvogel, A.-M., B. Feldmeyer, G. Rolshausen, et al. 2020. "Evolutionary Genomics Can Improve Prediction of Species' Responses to Climate Change." *Evolution Letters* 4: 4–18.
- Wang, J., W. G. Hill, D. Charlesworth, and B. Charlesworth. 1999. "Dynamics of Inbreeding Depression due to Deleterious Mutations in Small Populations: Mutation Parameters and Inbreeding Rate." *Genetics Research* 74: 165–178.
- Wang, L., T. M. Beissinger, A. Lorant, C. Ross-Ibarra, J. Ross-Ibarra, and M. B. Hufford. 2017. "The Interplay of Demography and Selection During Maize Domestication and Expansion." *Genome Biology* 18, no. 1: 1–13. <https://doi.org/10.1186/s13059-017-1346-4>.
- Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.
- Worku, M., M. Bänziger, G. S. a. Erley, D. Friesen, A. O. Diallo, and W. J. Horst. 2007. "Nitrogen Uptake and Utilization in Contrasting Nitrogen Efficient Tropical Maize Hybrids." *Crop Science* 47, no. 2: 519–528. <https://doi.org/10.2135/cropsci2005.05.0070>.
- Yang, J., S. Mezmouk, A. Baumgarten, et al. 2017. "Incomplete Dominance of Deleterious Alleles Contributes Substantially to Trait Variation and Heterosis in Maize." *PLoS Genetics* 13: e1007019.
- Yang, N., Y. Wang, X. Liu, et al. 2023. "Two Teosintes Made Modern Maize." *Science* 382, no. 6674: eadg8940. <https://doi.org/10.1126/science.adg8940>.
- Yang, N., X.-W. Xu, R.-R. Wang, et al. 2017. "Contributions of *zea mays* Subspecies Mexicana Haplotypes to Modern Maize." *Nature Communications* 8: 1874.
- Zeitler, L., J. Ross-Ibarra, and M. G. Stetter. 2020. "Selective Loss of Diversity in Doubled-Haploid Lines From European Maize Landraces." *G3: Genes, Genomes, Genetics* 10: 2497–2506.
- Zimmermann, H., P. Brandt, J. Fischer, and E. Welk. 2014. "The Human Release Hypothesis for Biological Invasions: Human Activity as a Determinant of the Abundance of Invasive Plant Species." *F1000Research* 3: 109. <https://doi.org/10.12688/f1000research.3740.2>.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.