Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE OPEN ACCESS

Population Genomics of a Cosmopolitan Weed Provides Insights Into Its Local Adaptation and Recent Demographic History

Marilia Souza Lucas[1,2] | Christoph Rosche[1,2] | Isabell Hensen[1,2] | Stefan G. Michalski[3] | Dávid U. Nagy[4] | Diana Gamba[5] | Renske E. Onstein[2,6] | Karime Abidkulova[7] | Mohammad Al-Gharaibeh[8] | Ali A. Al-Namazi[9] | Sirwan Babaei[10] | Fernando Bastida[11] | Caio Brunharo[5] | Aurélien Caillon[12] | Ragan M. Callaway[13] | Sam St Clair[14] | Darron A. Cullen[15] | Ryan Donnelly[16] | Alyson Ennis[17] | David J. Ensing[18] | Özkan Eren[19] | Andrey Erst[20] | Rita Filep[21] | Luke Flory[22] | Lauren J. Frazee[23] | Yusufjon Gafforov[24,25] | Fidji Gendron[26] | Zigmantas Gudžinskas[27] | Christopher Matt Guilliams[28] | Avni Hajdari[29] | Jian-Hua Hao[30] | Erin Haramoto[31] | Anna A. Ivashchenko[32] | Li Jun[33,34] | Nolan Kane[17] | Matthew A. Kaproth[35] | Hesham Kassem[36] | Abdelmajid Khabbach[37] | Damase P. Khasa[38] | Matthew Koski[39] | Maria Kozhevnikova[40] | Nikos Krigas[41] | Denis Krivenko[42] | Jesse R. Lasky[5] | Ylva Lekberg[43] | Mohamed Libiad[44] | Vanessa Lozano[45] | Wenbo Luo[46] | Trobjon Makhkamov[47] | Elizabete Marchante[48] | Chandra Moffat[18] | Abigail Moore[49] | Albert Muldashev[50] | Anush Nersesyan[51,52] | Pål l Axel Olsson[53] | Adrian Oprea[54] | Robert Pal[55] | Astghik Papikyan[51,52] | Christopher Proctor[56] | Vadim Prokhorov[40] | Satu Ramula[57] | Emily Rauschert[58] | Bryan Reatini[59] | Christian Rixen[60,61] | Scott Rush[62] | Gemma Rutten[63] | Mariem Saadani[1] | Ioulietta Samartza[41] | Julian Selke[1,64] | Stepan Senator[65] | Manzoor A. Shah[66] | Min Sheng[67] | Jamal Sheriff[68] | Baasanmunkh Shukherdorj[69] | Peter Sikkema[70] | Tatiana Silaeva[71] | Satoshi N. Suzuki[72] | Arpad E. Thoma[1,2] | Baoliang Tian[73] | Valeriy Tokhtar[74] | Sabrina Träger[1,2] | Tomonori Tsunoda[75] | Orzimat Turginov[76] | Kathryn G. Turner[77] | Tatyana Vakhlamova[78] | Yulia Vinogradova[66] | Viktoria Wagner[79] | Lydia Westberg[80] | Sa Xiao[81] | Christopher D. Barratt[2,6,82] | Walter Durka[2,3]

Correspondence: Marilia Souza Lucas (marilia_souza2@hotmail.com)

Received: 24 October 2025 | **Revised:** 20 February 2026 | **Accepted:** 16 April 2026

Keywords: *Erigeron canadensis* | genetic structure | invasive species | population genomics | rapid evolution

ABSTRACT

Invasive plant species present a growing ecological and economic challenge, and often adapt rapidly to their novel environments through complex demographic and evolutionary processes. Invasion genomics offers powerful tools to disentangle these processes, but most studies rely on geographically narrow sampling across native and non-native ranges. *Erigeron canadensis* is a cosmopolitan weed native to North America that has successfully invaded diverse climates worldwide. We used double-digest restriction site-associated DNA sequencing (ddRADseq) to investigate the genetic diversity and structure of 280 *E. canadensis* populations across the Northern Hemisphere. We found that native and non-native populations maintained comparable genetic diversity. Population structure analyses revealed four genetic clusters that were mainly differentiated along latitudinal and aridity gradients. However, one cluster was strongly overrepresented in the non-native compared to the native range. In the native range, genetic differentiation was shaped by spatial and environmental gradients, while in non-native regions human-mediated dispersal and repeated introductions disrupted environmentally driven genetic structure. Migration network analyses revealed limited intercontinental connectivity and a possible role of long-distance dispersal in within-range expansions. Genomic offset analyses showed that genotype-environment mismatches in non-native

Christopher D. Baratt and Walter Durka shared last authorship.

For affiliations refer to page 13.

populations associated with reduced growth and reproduction. Together, our results indicate that the invasion dynamics of *E. canadensis* were driven by multiple introductions, population admixture, and lineage sorting, while some genotypes contributed disproportionately to the spread of this invader. The presence of apparent maladaptation suggests that even long-established invaders may still be evolving in response to their novel environment, raising concerns about potential future expansions.

1 | Introduction

Invasive plants, including many globally widespread weeds, are typically strong colonizers and capable of rapid evolutionary adaptation to novel environments (Turner et al. 2014; Clements and Jones 2021). The adaptability of an introduced species depends on the genetic variation present in the native source populations from which the invasive propagules originate (Encinas-Viso et al. 2022). The use of genome-wide markers provides valuable insights into the evolutionary history and genomic diversity of both native and invasive populations, thereby advancing our understanding of the spatial and environmental determinants of population genetic structure in invaders (McGaughran et al. 2024). Such information helps explain how selection and gene flow interact to shape invasion dynamics (Neve et al. 2009), offering important implications for management strategies (Montgomery et al. 2024; Ahmed et al. 2025). Applications include identifying source populations for targeted control (Muirhead et al. 2008), monitoring the spread of herbicide resistance alleles (Kalsing et al. 2024), and prioritizing regions for biosecurity interventions based on genetic connectivity (McGaughran et al. 2024). Despite progress in sequencing reference genomes, however, comprehensive population genomic data remain scarce for most invasive plant species (Matheson and McGaughran 2022). Moreover, many population genetic studies rely on few markers or geographically restricted sampling, which limits our ability to infer large-scale genomic patterns or their environmental correlates (Lucas et al. 2024).

A critical first step in mitigating the impact of invasive species is to reconstruct their invasion history to understand how non-native populations evolve genotypes that facilitate invasive success (Le Roux 2021; Hess et al. 2022). Reconstructing invasion history involves identifying the native source regions of non-native gene pools and assessing shifts in their genetic structure due to founder effects (Prentis et al. 2008; Estoup and Guillemaud 2010; Jiang et al. 2023). Beyond initial introductions, this also requires tracking spatial expansion routes and investigating potential genetic adaptations to novel environments (Novak and Mack 2001; Sherpa and Després 2021). Such analyses are crucial because changes in genetic composition are fundamentally shaped by demographic processes such as expansion, contraction, and migration, which together determine how populations respond to environmental gradients over time (Bellard et al. 2018).

Founder effects are common in plant invasions, which can reduce effective population sizes and therewith limit the potential of invaders to spread, especially during their initial colonization (Bock et al. 2016; Rosche et al. 2017; Hernández-Espinosa et al. 2022). Some non-native species can overcome these limitations through reproductive assurance by self-fertilization (selfing), allowing them to establish and spread successfully even in isolated conditions (Dlugosch and Parker 2008; Allendorf et al. 2012; Voss et al. 2012). High levels of selfing promote homozygosity, which may facilitate the purging of potentially harmful deleterious mutations over time (Barrett et al. 2008). As such, self-compatibility can play a crucial role in the ability of invasive plants to rapidly establish and reproduce in new environments (Traveset and Richardson 2006; Petanidou et al. 2012; Uesugi et al. 2020). Moreover, occasional outcrossing among self-compatible populations can introduce new genetic combinations, which can increase fitness due to hybrid vigour, promoting the spread of admixed genotypes (Cropano et al. 2021). This dynamic interplay between self-compatibility and occasional genetic mixing may enable invasive plants to maintain sufficient genetic diversity to thrive and adapt in diverse and changing environments (Pyšek et al. 2023). However, it remains unclear whether multiple introductions and population admixture combined with subsequent local adaptation are common in highly selfing weeds that have spread globally (Facon et al. 2008; van Boheemen et al. 2017).

In addition to intrinsic plant traits such as their reproductive system, the genomic composition of established invasive populations is affected by extrinsic factors (Schierenbeck 2017; Camus et al. 2024; Hodgins et al. 2025). For example, geography and climate influence the genetic structure by imposing selective pressures across native and non-native ranges (Colautti and Lau 2016; Chen et al. 2023). Genetic differentiation from native gene pools, particularly when driven by deterministic and rapid adaptation, can significantly contribute to the success of invasive species (Colautti and Lau 2016; Marchini et al. 2019; Stankowski et al. 2019). In selfing species, alleles that are beneficial under novel conditions may become rapidly fixed in non-native gene pools, as selfing minimizes the dilution of advantageous mutations that might otherwise occur through outcrossing (Rosche et al. 2019). Also, many highly selfing weeds possess small, gene-dense genomes (Wright et al. 2008), increasing the frequency of coding regions across their genomes. These streamlined genomes may facilitate rapid evolutionary responses because a higher proportion of coding regions increases the chance that mutations affect functional loci (Peng et al. 2014). However, the overall frequency of mutations also depends on mutation rate, population size, and generation time (Wang and Obbard 2023).

Colonization and subsequent adaptation under variable environmental conditions largely depends on how well introduced propagules fit into their new environment (Colautti and Lau 2016). Genomic approaches can help identify loci under selection, track the spread of adaptive alleles, and predict future invasions (Martin et al. 2019; Leon et al. 2021). Moreover, linking the genome structure to individual performance is particularly relevant for developing weed management strategies in the face of global change (MacLaren et al. 2020). In

this context, genomic offset (i.e., the genetic distance between predicted and observed genotypes; Gain et al. 2023) provides a framework to identify genotype-environment mismatches under shifting climates (Chen et al. 2022; Gain et al. 2023; Rellstab and Keller 2025), which can help prioritize management actions in populations at risk of maladaptation (Lind and Lotterhos 2025). However, this approach can be also applied for comparing genotype-environment (mis)matches across native and non-native ranges (Camus et al. 2024; Gamba et al. 2025). Such assessments may detect potential local (mal)adaptation and therewith help anticipate potential future spread of certain invasive genotypes (Camus et al. 2024). Efforts to evaluate these processes are often constrained by limited geographic sampling, which makes it difficult to determine whether non-native populations are genetically suited to the environments they colonize (Muirhead et al. 2008; Hamilton et al. 2016; Gamba et al. 2025).

Here we study the population genomics of the widespread invasive plant *Erigeron canadensis* L. (Canadian horseweed). *E. canadensis* (syn. *Conyza canadensis* (L.) Cronquist) is an annual herb with a high selfing rate (96%) and prolific seed output (Bhowmik and Bekech 1993), which suppresses many crop species in the Northern Hemisphere (Weaver 2001). Previous studies have shown that aridity strongly influences the phenotype and ecological interactions of *E. canadensis* (Sheng et al. 2022; Nagy et al. 2024), suggesting it may also shape patterns of genetic differentiation. Despite the agronomical and ecological importance of this weed, prior population genetic assessments relied on few microsatellite markers and a limited number of populations, underestimating variation within and between its native and non-native ranges (Okada et al. 2013; Rosche et al. 2019; Bhattacharya et al. 2022). This limitation restricts our ability to identify native source regions and key drivers of gene flow (Lucas et al. 2024).

To address these limitations, we sampled and genotyped 280 widely distributed populations across the Northern Hemisphere. We assessed the genomic structure and gene flow patterns among these populations to study how the interplay between environmental gradients and non-adaptive demographic processes has shaped the genetic structure of *E. canadensis* across its native and non-native ranges. Our study was directed by the following research questions: (i) what are the dominant patterns of population structure, and how do they differ between native and non-native ranges? (ii) How have migration and human-mediated dispersal shaped gene flow and population connectivity across the ranges of *E. canadensis*? (iii) Which environmental variables drive genetic differentiation, and can these be used to predict adaptation in non-native populations? By addressing these questions, we aim to shed light on the evolutionary and ecological mechanisms that have facilitated the global invasion success of *E. canadensis*.

2 | Materials and Methods

2.1 | Study Species

The cosmopolitan Asteraceae *E. canadensis* is native to North America, and has become a notorious and wide-ranging invader across the Northern Hemisphere (Shah et al. 2014). The first record in the non-native range was in 1646 in Germany (Kowarik 2023). *Erigeron canadensis* exhibits several weedy traits, including a high selfing rate (Davis et al. 2010), abundant seed production (Weaver 2001), rapid growth (Yan et al. 2020), and the ability to thrive in disturbed environments (Prieur-Richard et al. 2000). These weedy traits make *E. canadensis* a successful invader, an economically significant agricultural weed, and, as such, an important model species for weed science in general (Laforest et al. 2020). It is also recognized as the first eudicot to develop resistance to glyphosate (VanGessel 2001). The diploid genome is small (426 Mbp/1C), yet comprises a comparatively large number of genes (> 45,000; Laforest et al. 2020). *E. canadensis* represents an exceptional model for studying how selfing, demographic processes, and environmental filtering interact to shape genome-wide variation during biological invasions.

2.2 | Plant Material and DNA Extraction

Erigeron canadensis seeds were sampled within the iCONNECT network (Lucas et al. 2024), a large collaborative and interdisciplinary research network investigating mechanisms driving rapid evolution in this cosmopolitan weed through coordinated field sampling and common garden experiments across native and non-native ranges. For this study, we sampled 103 native and 177 non-native populations, widely distributed across the Northern Hemisphere (Figure S1, Table S1). We assigned the 280 populations into 19 regions (8 native and 11 non-native; see Table S1), considering geographic proximity and natural barriers. To be included in our sampling, populations had to have at least 10 *E. canadensis* individuals and cover at least a 10-m transect. While population sizes varied, we restricted our sampling to a maximum distance of 30 m to have comparable sampling conditions. We recorded field plant performance according to Sheng et al. (2022). We randomly selected five *E. canadensis* plants per population (> 2 m apart). To estimate individual plant performance, we harvested the aboveground biomass of the five *E. canadensis* plants and counted the number of flower heads (capitula). To quantify habitat productivity, we collected the aboveground biomass in five plots of a 0.5 × 0.5-m area where the upper right corner of the quadrate was placed next to (but excluding) each of the five focal plants. All aboveground biomass was dried at 65°C for 3 days and weighed.

Given the global scope of this study, within-population sampling was necessarily limited (three individuals per population), which may reduce resolution for fine-scale population-level processes such as rare alleles or fine-scale local genetic structure. However, this design is appropriate for detecting broad geographic and regional patterns of population structure and gene flow (Dlugosch and Parker 2008). The seeds from these three randomly distributed individuals were collected, and one offspring per mother plant was grown in a greenhouse at Montana Technological University (USA) for genetic analyses. Fresh leaves were collected from one offspring per mother plant, dried in a freeze dryer, and conserved in silica gel until DNA extraction. Genomic DNA was extracted from the dry macerated leaves using the peqGOLD Plant DNA Mini Kit (Avantor). DNA was quantified using a

Qubit 2.0 fluorometer (Thermo Fisher Scientific), and DNA concentration was increased by evaporation if necessary to meet sequencing platform requirements.

2.3 | Genomic Data

We used a modified version of the ddRADseq protocol of Peterson et al. (2012) to construct reduced-representation libraries of genomic DNA for the genotyping of single nucleotide polymorphisms (SNPs) as described in Durka et al. (2025). In short, approximately 200 ng of DNA from each sample was digested with EcoRI-HF and MspI using Cutsmart buffer (NEB). Custom-barcoded adapters were then ligated to restriction fragments. Groups of 96 samples, each with a unique barcode, were pooled, purified with the Promega-Wizard Gel and PCR Clean-up Kit, and size-selected using the BluePippin (Sage Science) electrophoresis platform to select target fragments of 350 to 450 bp in length. Unique external indices were added to each pool through PCR amplification. The PCR products were purified with the Promega Clean-up kit and AMPure XP beads. Fragment sizes were verified on an Agilent 2100 Bioanalyzer. Libraries were then pooled in equimolar quantities and sequenced using the Illumina NovaSeq 6000 Sequencing System with 10% PhiX.

2.4 | SNP Genotyping

Demultiplexing of the raw reads was performed using process_radtags from the Stacks pipeline (Catchen et al. 2013). Further read processing (quality filtering, de novo assembly, construction of reference sequence, read mapping and SNP calling) was performed using dDocent (version 2.7.8; Puritz et al. 2014). The parameter configuration used default values (Table S2), except for Clustering_Similarity%, which was increased to 0.9 (Figure S2). Subsequently, we filtered loci and individuals following O'Leary et al. (2018). First, we used vcfallelicprimitives from the library vcflib (version 1.0.0; Garrison et al. 2021) and vcftools (version 0.1.16) to remove indels. We then applied population-level filters, keeping only biallelic SNPs with a minimum allele count (mac) of 3 and maximum missingness across individuals (max_missing) of 50%. Individual-level genotype filters included minimum genotype read depth (minDP) of 3 and minimum mean sequence quality (minQ) of 30. To exclude paralogous SNPs, we filtered putatively fixed heterozygous SNPs, by calculating the *p*-value for heterozygosity excess with vcftools and filtering SNPs with P-HET_EXCESS < 1e-15. We skipped individuals with > 75% missing values (imiss). Subsequently, using *vcffilter* from vcflib, we filtered SNPs according to allele balance (AB > 0.2 and AB < 0.8 | AB < 0.01 | AB > 0.99), strandedness (SAF/SAR > 100 and SRF/SRR > 100 | SAR/SAF > 100 and SRR/SRF > 100), mapping quality ratio of the two alleles (MQM/MQMR > 0.9 and MQM/MQMR < 1.05), and properly pairing alleles (PAIRED > 0.05 and PAIREDR > 0.05 and PAIREDR/PAIRED < 1.75 and PAIREDR/PAIRED > 0.25 | PAIRED < 0.05 and PAIREDR < 0.05). We then filtered SNPs to maximum missingness of 34% (max_missing 0.66), the minimum minor allele frequency (maf) of 0.05, minimum mean read depth (min-meanDP) of 10, and a maximum mean depth (max-meanDP) of 1000. Thus, from a total of 801,773 SNPs in 41,149 loci, 11,501 SNPs in 5080 loci were obtained of which we retained only one single SNP per locus. The genotypic data were analysed using the R-package dartR (version 2.9.7; Gruber et al. 2018). Initially, the VCF file was converted into a genlight object, suitable for subsequent analyses.

2.5 | Spatial Genetic Structure and Diversity

To investigate whether the transcontinental colonization has influenced genetic diversity of *E. canadensis*, we estimated genetic diversity per population and the average per region as expected and observed heterozygosity (H_e and H_o) and allelic richness (A_R) using the R-packages adegenet (version 2.1.10; Jombart 2008), hierfstat (version 0.5–11; Goudet and Jombart 2015), and poppr (version 2.9.6; Kamvar et al. 2014). Similarly, the inbreeding coefficient (F_{IS}; Wright 1965) was estimated for each region. We then tested whether these estimates differed between native and invaded regions, using independent *t*-tests.

To examine how genetic diversity is partitioned across different spatial scales and whether population structure differs between native and non-native populations, we performed analyses of molecular variance (AMOVA) using poppr. We used range (native vs. non-native), region, and population as hierarchical levels and quantified percentage variance and Phi statistics, analogous to *F*-statistics, at each hierarchical level.

For further investigations of the population genetic structure, we performed cluster analyses using ADMIXTURE (version 1.3.0; Alexander et al. 2009), testing *K* values ranging from 2 to 50. Cross-validation (CV) errors were obtained to determine the most likely *K* value. Pie charts of mean ancestry coefficients per population (i.e., each unique sampling locality) were then plotted on a distribution map.

We used conStruct (version 1.0.4; Bradburd et al. 2018) to compare the population structure between the spatially naïve approach of ADMIXTURE to a spatially aware approach that incorporates isolation by distance. Based on individual genotypes, we constructed a population level dataset of allele frequencies. Missing values were imputed by mean values of the next higher spatial level of region, range and total. We ran conStruct with spatial and nonspatial models with numbers of groups between $K = 1$ and $K = 7$. We ran a cross-validation to assess predictive accuracy with eight independent runs with a 1000 iteration chain length.

To explore the role of geographic and environmental factors in shaping genetic differentiation, we investigated isolation by distance (IBD) and isolation by environment (IBE) through partial Mantel tests, examining the relationships between genetic distance (calculated as pairwise $F_{ST}/[1 - F_{ST}]$, using the R package StAMPP [version 1.6.3; Pembleton et al. 2013]), spatial distance ($\log_e$-transformed), and climatic distance (Euclidean distance of PC1 [temperature seasonality] scores of 19 BioClim variables; see below). This analysis was conducted to determine whether genetic differentiation is primarily driven by geographic distance (neutral processes)

or environmental differences (adaptive processes), using permutation-based correlation tests in the R package vegan (version 2.6–4; Oksanen et al. 2022).

Climatic conditions were obtained from WorldClim 2.1 (Fick and Hijmans 2017) at a 30-arc sec resolution (~0.86 km^2 per grid cell at the equator), using the raster R-package (version 3.6–30; Hijmans et al. 2015). We extracted the 19 standard bioclimatic variables (BIO1–BIO19; Table S3), which describe temperature, precipitation, and seasonality patterns, and applied a principal component analysis (PCA) to reduce collinearity and summarize broad-scale environmental gradients. The first principal component (PC1; 36% of variance) primarily captured temperature seasonality and continentality, showing positive correlations with temperature seasonality (BIO4) and annual temperature range (BIO7), and negative correlations with mean temperature of the coldest quarter (BIO11) and minimum temperature of the coldest month (BIO6). The second principal component (PC2; 25% of variance) captured moisture- and warm-season climatic variation, including several precipitation-related variables (Figure S9). We retained PC1 for downstream analyses because it represented the dominant, spatially structured macroclimatic gradient across the Northern Hemisphere and minimized multicollinearity among predictors, while precipitation-driven water limitation was explicitly captured using climatic water deficit (CWD).

To identify patterns of differentiation between native and non-native populations and to infer potential colonization pathways, a Neighbour-Joining (NJ) tree was constructed based on a Jukes–Cantor distance matrix. The resulting tree was then optimized using a maximum likelihood approach under the General Time Reversible model, implemented in the R-package phangorn (version 2.11.1; Schliep 2011). The final tree was visualized in a circular layout using the R-package ggtree (version 3.6.2; Yu et al. 2017).

To investigate phylogenetic relationships among individuals, Nei's genetic distances were calculated using the StAMPP package. The resulting distance matrix was exported in Nexus format and analysed in SplitsTree4 (version 6.4.12; Huson and Bryant 2006) to construct a Neighbour-Net network (Bryant and Moulton 2004). The network output was visualized using the R-packages phangorn and ape (version 5.6–2; Paradis et al. 2004).

2.6 | Recent Gene Flow and Historical Demography

To investigate recent migration and genetic connectivity in *E. canadensis* across its global range, we analysed migration networks among regions using the divMigrate function in the R-package diveRsity (version 1.9.90; Keenan et al. 2013). This approach estimates relative migration rates from allele frequency differentiation, with values scaled as N_m (effective number of migrants per generation) to represent the effective number of migrants per generation inferred from genetic data. Higher N_m-values indicate substantial gene flow and genetic homogenization among populations, whereas lower values suggest restricted gene flow and stronger differentiation. To visualize these patterns, we filtered the migration matrix for $N_m > 0.5$ and mapped directional migration among regions. This threshold highlights stronger gene flow and is consistent with previous applications of the method, typically representing recent migration (Sundqvist et al. 2016).

Spatial genetic structure in *E. canadensis* was characterized using FEEMS (Fast Estimation of Effective Migration Surfaces; version 1.0.5; Marcus et al. 2021). Genotypes were imported from PLINK files, with missing values mean-imputed and loci scaled prior to analysis. A discrete global grid spanning the sampling extent was generated using the dggridR package (v3.1.1; Barnes and Sahr 2017). The spatial graph was constructed by linking sample locations to grid nodes within an outer boundary polygon defining the modelled geographic extent.

2.7 | Genotype–Environment Associations

To predict environmental and human factors associated with genetic differentiation, we performed a distance-based Redundancy Analysis (dbRDA) using the capscale function in the R-package vegan. We used the Bray–Curtis dissimilarity of allele frequencies as the response matrix and environmental predictors as constraints Candidate environmental variables were identified a priori based on previous studies demonstrating their biological and ecological relevance for *E. canadensis*. Collinearity among candidate predictors was assessed using pairwise Pearson correlations, and only variables with relatively low correlations ($r < 0.7$) were retained. A single dbRDA model including the retained predictors was fitted; no forward or stepwise selection procedure was applied. Finally, the chosen predictors included the first principal component (PC1[temperature seasonality]) of the PCA on climatic conditions obtained from WorldClim (see above) to account for broad macroclimatic gradients. Latitude was also included given its direct influence on annual sunlight duration patterns, which can influence species distribution, genetic differentiation and phenotypic variation (i.e., latitudinal clines). Because previous studies have shown that aridity strongly influences the phenotypic expression and ecological interactions of *E. canadensis* (Sheng et al. 2022; Nagy et al. 2024), we further incorporated the climate water deficit (CWD). This estimate describes the amount of water that is missing in a system for unrestricted transpiration and is calculated as the difference between potential evapotranspiration and the actual evapotranspiration (Nagy et al. 2024). CWD data for each population were obtained using the TerraClimate dataset (Abatzoglou et al. 2018) at a spatial resolution of 1/24° (~4 km). Because anthropogenic activity can be a main determinant of spatio-temporal invasion routes and the local abundance of invaders (Rosche et al. 2025), human footprint values were also included in our dbRDA. These values were obtained for each population from the Human Footprint raster dataset (Venter et al. 2018) at a 1-km^2 resolution. The footprint values quantify human impacts based on nine anthropogenic stressors, such as urbanization and deforestation, on a scale from 0 (minimal impact) to 50 (maximal impact).

The dbRDA was obtained as most parsimonious model based on forward selection which is a common approach to identify determinants of genetic structure (e.g., Rosche et al. 2018). More specifically, we used the *ordistep* function of the R package *vegan* which builds on a permutation approach that successively adds predictors to a model if they significantly improve the model fit ($p < 005$; see Borcard et al. 2011).

2.8 | Genotype-Environment Mismatches in Non-Native Populations

The dbRDA approach was further applied to assess whether non-native populations experience genotype–environment mismatches relative to their native counterparts, comparable to genomic offset statistics (Lachmuth et al. 2023). To this end, we first performed a dbRDA on native populations, using their genotypic data as response variables and the previously selected abiotic factors as predictors. The fitted dbRDA model was then used to predict genotypic values for non-native populations. Euclidean distances between the predicted and the observed genotypic values in the non-native range were subsequently calculated. These distances represent the degree of genotype–environment mismatch, or maladaptation, in non-native populations, under the assumption that native populations exhibit stronger adaptation to their local environments due to their longer evolutionary history (Gamba et al. 2025). To evaluate whether the observed mismatch deviated from expectations under random environmental associations, we conducted a permutation test with 1000 iterations.

To understand if these genomic offset values have biological relevance, it is important to test whether they correspond with field performance (Lotterhos 2024; Camus et al. 2024). To address this, we ran linear mixed-effects models using the $\log_e$-transformed biomass and the $\log_e$-transformed number of capitula of the five individuals per population as response variables. As explanatory variables, we used the genotype-environment mismatch (genomic offset) of the non-native populations. To control for variation in the general habitat productivity across populations, we also included $\log_e$-transformed community biomass as a covariate. Population nested within region were included as random factors in the models.

3 | Results

3.1 | Spatial Genetic Structure and Diversity

The final data set comprised 640 *E. canadensis* individuals from 280 populations and 19 regions. We called a total of 5080 SNPs with 10% missing data. *E. canadensis* showed low population-level diversity and high individual inbreeding (Table 1). Observed heterozygosity (H_o) ranged from 0.05 to 0.10 for native regions, and from 0.05 to 0.09 for non-native regions. Similarly, expected heterozygosity (H_e) was consistently low, with values between 0.12 and 0.16 for the native regions, and from 0.10 to 0.15 across non-native regions. The inbreeding coefficient (F_{IS}) was relatively high across all regions, ranging from 0.44 to 0.50, suggesting a degree of inbreeding. The allelic richness (A_R) had values between 1.05 and 1.10. Notably, there were no significant differences (based on *t*-tests) between native and non-native populations in any of these population genetic metrics, suggesting that *E. canadensis* maintained a similar genetic diversity and breeding system in both its native and non-native ranges.

The overall AMOVA analysis showed that 8.2% of genomic variation was explained by differences between native and non-native ranges, whereas 44% occurred among populations within ranges (Table S4a). This indicated only moderate differentiation between ranges but very strong differentiation among populations. Considering the 19 defined regions, 21.9% of the variation was explained by differences among regions, while about half (51.6%) still occurred among populations (Table S4c). When native and non-native ranges were analysed separately, 47.3% and 48.9% of the variation, respectively, occurred among populations, pointing to similar population structure in both ranges (Table S4b).

For the population structure analysis in ADMIXTURE, the cross-validation error approach suggested $K = 4$ as the optimal number of clusters (Figure S3). However, the plot did not exhibit a clear elbow, which suggested a more gradual improvement in model fit with increasing K rather than a strong, singular optimal value. As K increased, more genetic clusters emerged, mostly geographically structured, revealing finer population differentiation (Figure S4). Note that $K = 4$ as the optimal number of clusters was also supported by the phylogenetic analysis presented below.

The spatial distribution of the four genetic clusters exhibited distinct patterns within and between the ranges. In the native range (Figure 1A), the yellow cluster was more prevalent in the southern and eastern parts, while the blue cluster was more common in northern and central populations. The green cluster appeared in substantial proportions in the western Mediterranean part of the native range. In the non-native range (Figure 1B), this green cluster occurred in the eastern Mediterranean and Near East. The blue cluster was distributed in the northern non-native range and extended into Middle and Central Asia, whereas the yellow cluster was dominant in the southern part of Eastern Asia. Across both ranges, the distributions of these three clusters largely corresponded to geographic regions characterized by distinct environments: the green cluster was predominantly found in areas with dry climates, the blue cluster was most common at higher latitudes, and the yellow cluster occurred primarily in oceanic regions. However, the orange cluster showed a differentiated distribution across ranges. It was largely absent from the native range, but in the non-native range, the orange cluster was dominant at intermediate latitudes from Western Europe to Western Asia. Moreover, the presence of multiple admixture components across populations suggested a complex evolutionary history and gene flow dynamics within and across native and non-native ranges.

The Neighbour-Net network (Figure 1C) revealed no clear genetic differentiation between native and non-native populations. It showed four clusters of populations that largely correspond to the four admixture clusters. The green cluster was the most distinct, positioned farthest from the others in the network. In contrast, the yellow, blue, and orange clusters were more closely

TABLE 1 | Characterization of genetic variability of *Erigeron canadensis* for regions within ranges (mean over populations ± SD). Region codes are shown in parentheses as consecutive numbers following the region names.

Range	Regions	N	H_o	H_e	F_{IS}	A_R
Native	South-Western North America (1)	34	0.09 ± 0.07	0.15 ± 0.10	0.45 ± 0.04	1.09 ± 0.07
	Pacific Coast (2)	19	0.06 ± 0.04	0.12 ± 0.06	0.46 ± 0.03	1.06 ± 0.04
	Intermountain US (3)	31	0.08 ± 0.07	0.15 ± 0.10	0.44 ± 0.05	1.08 ± 0.07
	Central US (4)	50	0.10 ± 0.08	0.16 ± 0.11	0.44 ± 0.06	1.10 ± 0.08
	North US Prairies (5)	14	0.09 ± 0.09	0.15 ± 0.11	0.44 ± 0.06	1.09 ± 0.09
	Great Lakes (6)	30	0.08 ± 0.07	0.14 ± 0.10	0.45 ± 0.05	1.08 ± 0.07
	South-Eastern US (7)	22	0.07 ± 0.05	0.12 ± 0.07	0.46 ± 0.03	1.07 ± 0.05
	East Coast (8)	36	0.07 ± 0.06	0.12 ± 0.08	0.46 ± 0.04	1.07 ± 0.06
Non-native	Western Mediterranean (9)	42	0.07 ± 0.06	0.13 ± 0.09	0.46 ± 0.03	1.07 ± 0.06
	Western Europe (10)	11	0.08 ± 0.05	0.12 ± 0.08	0.45 ± 0.03	1.08 ± 0.05
	Central Europe (11)	63	0.06 ± 0.04	0.12 ± 0.06	0.46 ± 0.02	1.06 ± 0.04
	Baltic Europe (12)	39	0.05 ± 0.02	0.10 ± 0.03	0.50 ± 0.04	1.06 ± 0.02
	Eastern Mediterranean (13)	43	0.07 ± 0.06	0.12 ± 0.09	0.46 ± 0.04	1.07 ± 0.06
	Armenia-Iran (14)	13	0.09 ± 0.06	0.15 ± 0.09	0.44 ± 0.04	1.09 ± 0.06
	Middle Asia (15)	39	0.08 ± 0.07	0.13 ± 0.10	0.45 ± 0.05	1.08 ± 0.07
	Eastern Europe (16)	53	0.08 ± 0.07	0.15 ± 0.10	0.45 ± 0.05	1.08 ± 0.07
	Siberia (17)	17	0.08 ± 0.08	0.14 ± 0.10	0.45 ± 0.06	1.08 ± 0.08
	Far East Asia (18)	55	0.06 ± 0.04	0.12 ± 0.06	0.46 ± 0.02	1.06 ± 0.04
	China (19)	29	0.05 ± 0.02	0.10 ± 0.03	0.47 ± 0.01	1.05 ± 0.02
	t-statistic, p		$t = 0.87$, $p = 0.39$	$t = 0.62$, $p = 0.53$	$t = 1.17$, $p = 0.25$	$t = 0.98$, $p = 0.33$

Abbreviations: A_R, allelic richness; F_{IS}, inbreeding coefficient; H_e, expected heterozygosity; H_o, observed heterozygosity; N, number of individuals per population; SD, standard deviation.

related. Additionally, several intermediate populations occupied positions between clusters and often exhibited mixed admixture proportions, suggesting gene flow or admixture among them. This pattern did not support a simple divergence scenario in which non-native populations evolved directly from a specific native origin. Instead, it pointed to a more complex, reticulate evolutionary history in the native range followed by multiple introduction events and genetic mixing. Notably, the orange cluster, which was mostly found in non-native populations, formed a distinct branch, highlighting a unique genetic component. Consistent with these findings, the circular NJ tree (Figure S7) also showed no distinct split between native and non-native populations but a separated branch of non-native populations that belonged to the orange cluster.

Partial Mantel tests showed that when controlling for climatic distance, genetic distance was correlated with geographic distance in both ranges (native: $r = 0.091$, $p = 0.035$; non-native: $r = 0.287$, $p = 0.001$, Figure 2), indicating isolation by distance (IBD). We also checked for potential non-linearity of these relationships. Across the full spatial range, genetic distance increased monotonically with geographic distance, with no evidence for non-linearity or distance thresholds (Figure S8). In contrast, when controlling for geographic distance, genetic distance was significantly correlated with climatic distance in the native range ($r = 0.147$, $p = 0.001$), but not in the non-native range ($r = 0.014$, $p = 0.355$). No consistent relationship between genetic and climatic distance was evident in the non-native range at either short or long geographic distances (Figure S8).

The spatially-aware population structure analysis using conStruct corroborated the strong signal of isolation by distance in our data. Cross-validation showed that spatial models consistently outperformed nonspatial models across all K values (Figure S5), confirming that IBD is a strong driver of genetic differentiation. However, spatial conStruct models assume approximately continuous gene flow, which is biologically unrealistic for our global dataset that includes major oceanic barriers separating native and non-native ranges. Thus, global conStruct models are not appropriate in this case. The nonspatial $K = 4$ model closely matched our ADMIXTURE results (Figure S6), providing independent support for four discrete genetic clusters despite higher overall admixture levels.

3.2 | Recent Gene Flow and Historical Demography

The relative migration network (Figure 3, Table S5) illustrated recent migration intensities and directions across the Northern

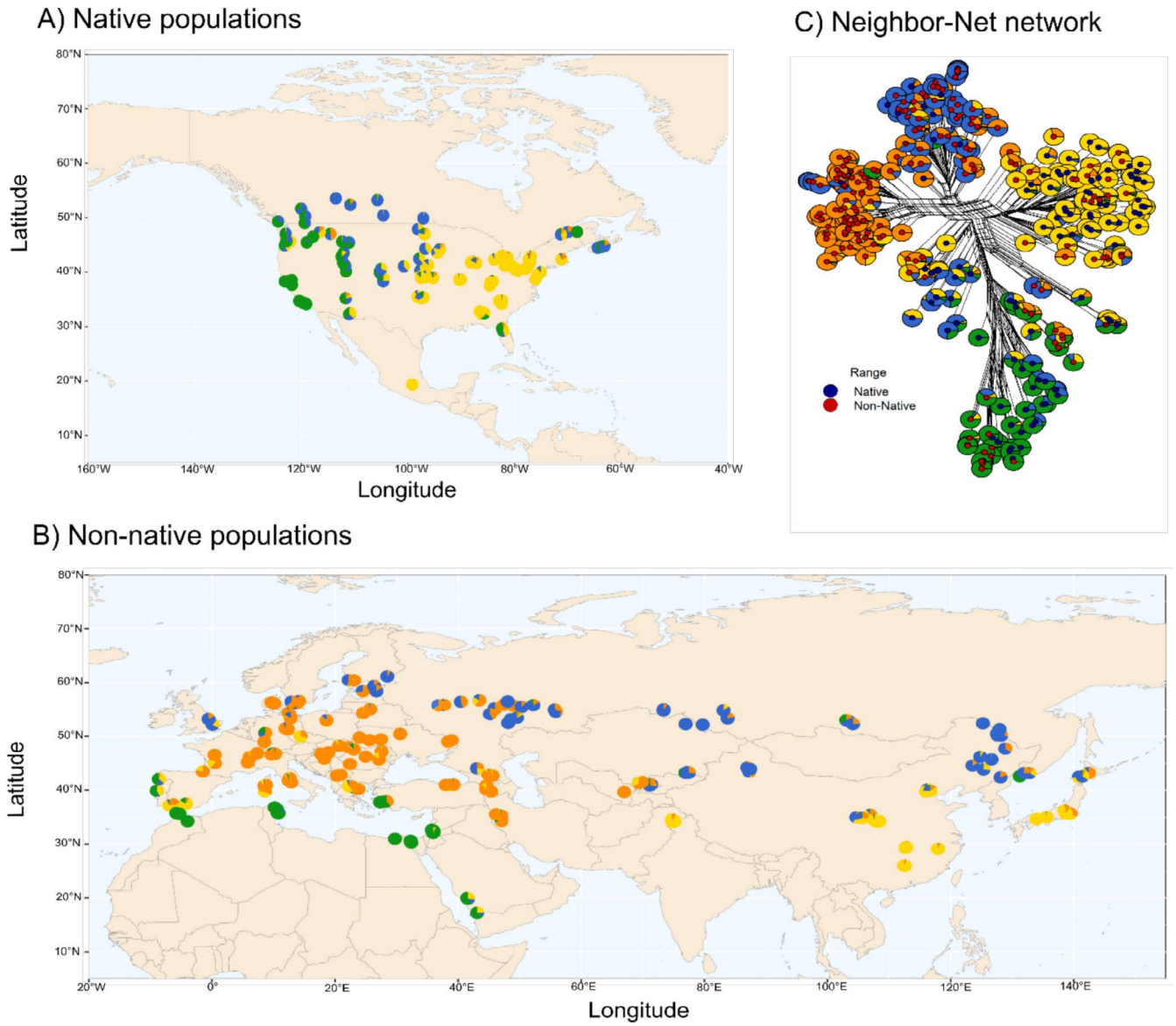

FIGURE 1 | Geographic distribution of *Erigeron canadensis* of (A) native and (B) non-native populations with admixture proportions of $K = 4$. Each pie chart represents the proportion of four colour-coded genetic clusters. Individual ancestry coefficients are shown in Figure S4. (C) Neighbour-Net network of *E. canadensis* based on Nei's distances. Tips represent the individuals, with edge lengths indicating genetic distances. Inner dot colours show range affiliation (native in blue and non-native in red), and outer ring pie charts depict admixture proportions.

Hemisphere. In North America, migration intensity was predominantly low, with no clear medium-high intracontinental gene flow, indicating reduced connectivity within the native range. Intercontinental migration was more prevalent, however, with signals of high migration between native regions in Central (4) and Eastern US (7) and non-native regions in Europe (16) to Siberia (17), possibly facilitated by human transport, prevailing westerlies, or the polar stream. Within Europe, medium-intensity migration was observed between the Baltics (12) and Eastern Europe (16), and from the Western (9) to the Eastern Mediterranean (13), reflecting a lack of major dispersal barriers and geographic proximity. A strong signal of recent migration was indicated from Central Europe (11) to Central US (4) likely reflecting that the orange cluster is present in low frequency in the native range and expanded only after transport to Europe, which is thus interpreted as the source region. The strongest migration signal was detected in the non-native range between Armenia-Iran (14) and Far East Asia (18) to Eastern Europe (16). Within Asia, medium levels of migration were detectable from China (19) to Middle Asia (15). Taken together, a complex mixture of natural and potentially human-mediated dispersal was evident.

To further visualize spatial patterns in gene flow across the global range of *E. canadensis*, we applied FEEMS to estimate effective migration surfaces (Figure 4). The FEEMS analysis revealed pronounced spatial variation in effective migration rates, with regions of elevated connectivity interspersed with areas of reduced migration. In particular, extensive zones of low effective migration were observed between Europe and Africa, within parts of North America and across central Asia, indicating regions of migration resistance, whereas higher effective

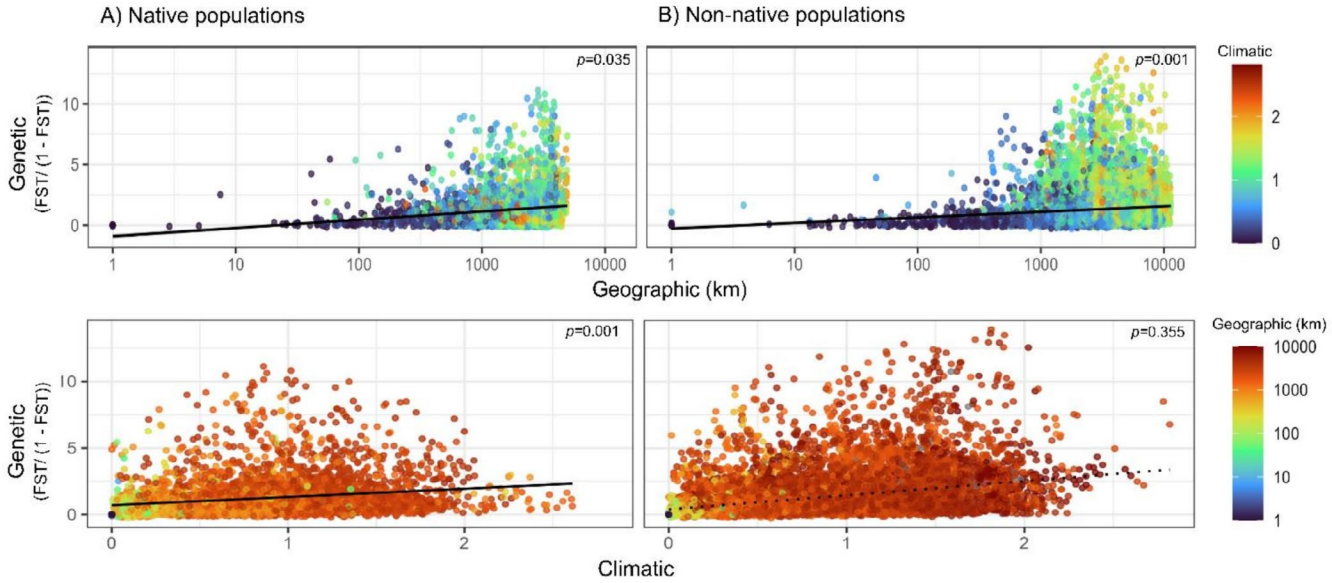

FIGURE 2 | Isolation by distance and isolation-by-environment correlating genetic distance ($F_{ST}/[1-F_{ST}]$) with geographic and environmental distance for native (A) and non-native (B) ranges. The upper panels show the relationship between geographic distance (km, log-scaled) and genetic distance, with the colour gradient representing climatic distance. The lower panels depict the relationship between climatic distance and genetic distance, with the colour gradient representing the geographic distance (km, log-scaled). The black lines in each plot represent the fitted regression lines. Scale dependence of the linear relationships of isolation by distance and isolation-by-environment are explored the Supporting Information (Figure S8).

migration characterized several regions in Europe and western Eurasia. These spatial patterns are consistent with isolation by distance processes and complement the directional estimates obtained with divMigrate by providing a continuous, landscape-level representation of gene flow and migration barriers across both native and non-native ranges.

3.3 | Genotype–Environment Associations

The dbRDA indicated that latitude, climatic water deficit, and the first and second principal components of the bioclimatic variables (PC1 and PC2 Bioclim) were the main environmental gradients associated with genetic structure in *E. canadensis*. In contrast, human footprint exhibited a relatively short vector and a weaker association with genetic structure, suggesting negligible influence on genomic differentiation at the spatial scale assessed. The first two constrained axes explained 55.3% (CAP1) and 32.5% (CAP2) of the fitted genetic variation, corresponding to 12.5% and 7.3% of the total genetic variation, respectively (Figure 5).

When the distribution of Admixture clusters is overlaid onto the dbRDA plot, it provided additional information on how genetic clusters are distributed along these environmental gradients. Specifically, populations belonging to the blue cluster were found at high latitudes (see Table S6 for population-level site scores and environmental values). The orange cluster appeared to occupy a narrow environmental space and was positioned toward positive values of PC1_BioClim, reflecting association with seasonal climate. However, the vectors for latitude and PC1 Bioclim showed similar orientations, reflecting strong latitudinal gradients in temperature-related climatic variables. Populations belonging to the green genetic cluster were positioned in the direction of increasing climatic water deficit, suggesting selection in more arid environments. The yellow cluster was positioned at higher CAP2 values and showed a stronger association with PC2 Bioclim.

3.4 | Genome-Environment Mismatches in Non-Native Populations

The genomic offset analysis indicated potential genetic mismatches to local environmental conditions in certain non-native regions (Table S7). The highest offset values were found in Western, Baltic, and Central Europe. In Middle Asia and Armenia-Iran, moderate genetic offset values were observed, indicating some degree of pre-adaptation. China, in contrast, showed relatively low offset values. To provide context for the magnitude of genomic offset values observed in the non-native range, we also summarized genomic offset values for native populations using the same framework, allowing direct comparison across native and non-native regions (Tables S7 and S8).

Populations with higher offset values consistently showed reduced growth and reproductive output in the field. Specifically, aboveground biomass declined with increasing genomic offset ($\chi^2 = 5.29$, $p = 0.021$), corresponding to a 13.42% decrease in biomass per unit increase in genomic offset (95% CI: 2.11%–23.43%; Figure 6A; Table S9). Likewise, the number of capitula decreased with increasing genomic offset ($\chi^2 = 5.84$, $p = 0.016$), corresponding to a 13.78% decrease per unit increase in genomic offset (95% CI: 2.76%–23.54%; Figures 6B and S11; Table S9). Together, these results indicate that higher genomic offset values are associated with reduced field performance and are consistent with maladaptation to local environmental conditions.

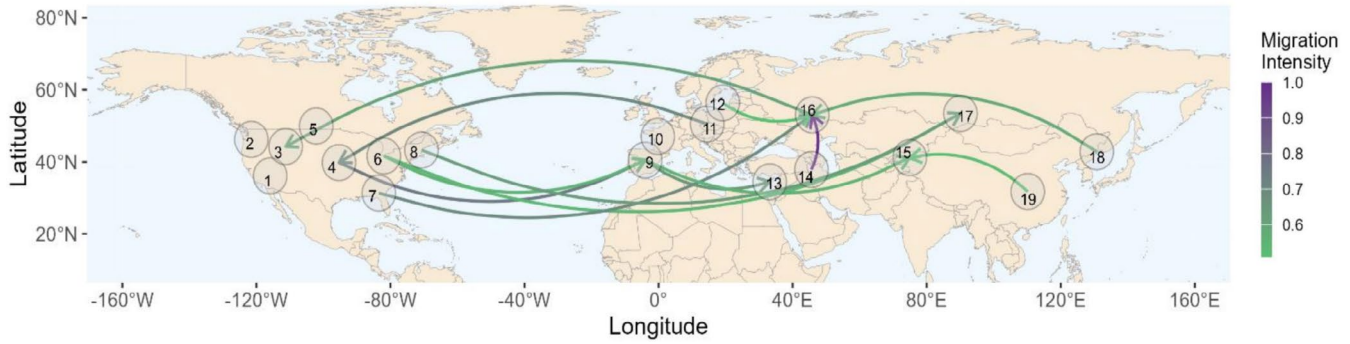

FIGURE 3 | Relative migration network map showing migration intensities across native and non-native regions (1 to 19, see Table 1) with N_m (effective number of migrants per generation) > 0.5. Circles correspond to the centroid of each region. Arrow colours indicate migration strength, with dark purple representing high intensity and green representing intermediate intensity.

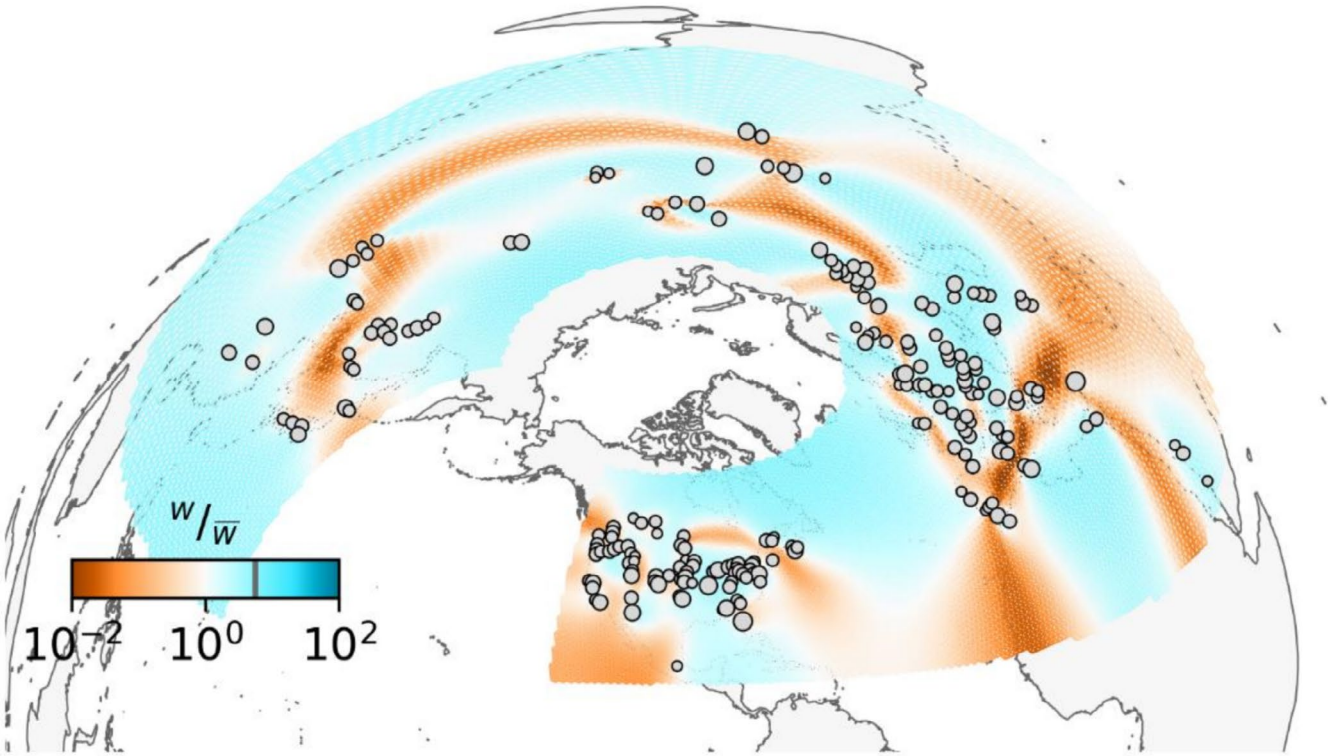

FIGURE 4 | Spatial patterns of effective migration in *Erigeron canadensis* inferred using FEEMS (Fast Estimation of Effective Migration Surfaces). Colours indicate relative effective migration rates across the modelled geographic extent (blue = higher migration; orange = reduced migration/increased drift). Grey circles denote sampled populations, with circle size proportional to sample size. The spatial graph was constructed within an outer boundary polygon defining the sampled geographic extent; areas outside this extent were not included in model fitting. Apparent discontinuities (e.g., across the Pacific) reflect the chosen projection and graph boundary rather than inferred biological barriers.

4 | Discussion

4.1 | Spatial Genetic Structure and Diversity

We found similar levels of genetic diversity and inbreeding across the native and non-native ranges, challenging expectations that founder effects and limited dispersal reduce variation after introduction (Dlugosch and Parker 2008). However, our results supported previous work showing no significant differences in microsatellite diversity between native and non-native *E. canadensis* populations (Rosche et al. 2019; Bhattacharya et al. 2022), which is consistent with the species' predominantly selfing breeding system. High selfing rates suggest a reproductive strategy that supports population survival and reproduction even in small or isolated populations (Pannell 2015), while founder effects do not necessarily increase inbreeding rates because genomes are already largely homozygous (Voss et al. 2012; Rosche et al. 2019). Species with high propagule pressure, like *E. canadensis*, have colonization advantages by producing large numbers of wind-dispersed seeds that promoted rapid range expansion (Shields et al. 2006; Wilson et al. 2009; Liu et al. 2018). Such pressure also increased the likelihood of pre-adapted genotypes establishing, and multiple introductions of independent lineages may also have led to repeated admixture events, despite generally low outcrossing frequencies (Bossdorf et al. 2005; Li et al. 2019; Greve and Pertierra 2022).

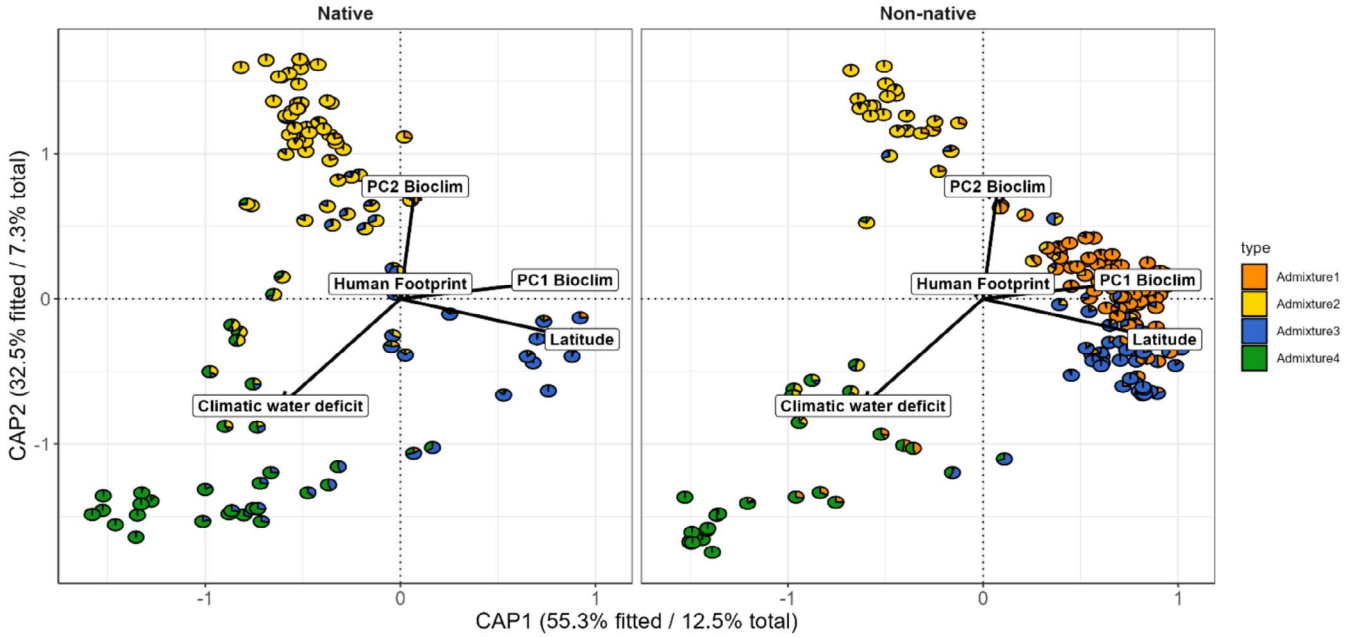

FIGURE 5 | Redundancy Analysis (RDA) plot showing the genomic relationship among populations and their relationship to environmental gradients (latitude, PC1 and PC2 of BioClim variables, climatic water deficit) and anthropogenic influence (human footprint). For clarity, the native and non-native populations are depicted in separate plots. Pie charts represent admixture proportions of the populations.

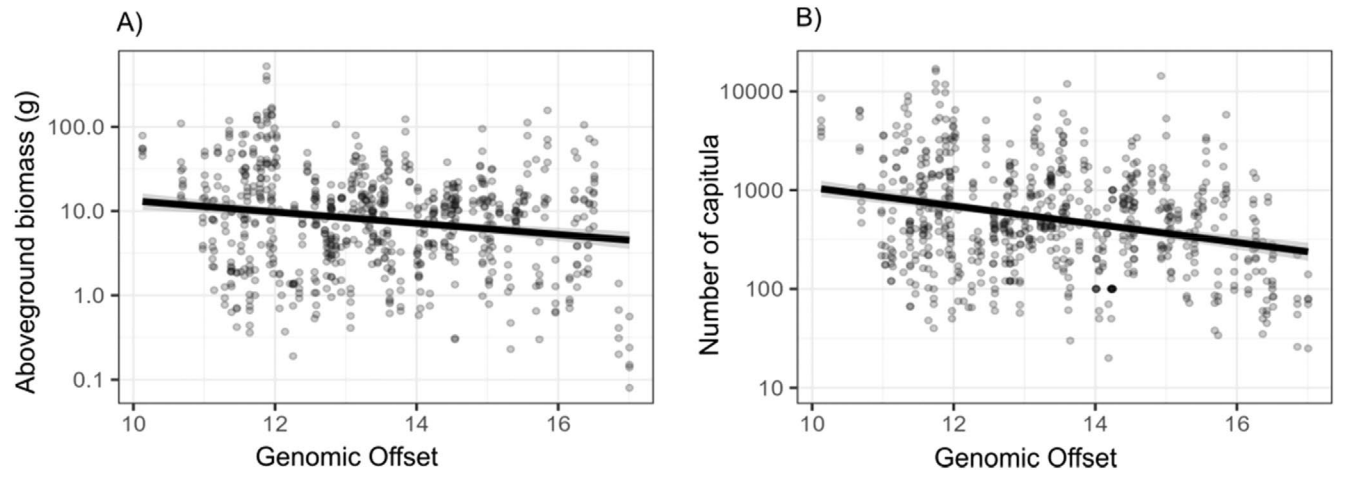

FIGURE 6 | Relationship between field performance and genomic offset in non-native *Erigeron canadensis* for (A) aboveground biomass (log-scaled) and (B) number of capitula (log-scaled) of the *E. canadensis* individuals in the field. Details of the linear mixed-effects models and the map of genomic offset can be found in the Supporting Information (Figure S11, Table S9).

Our admixture analysis indeed showed that most genetic clusters occurred in both native and non-native ranges, supporting multiple introductions and enhancing establishment through admixture and increased variation (Bhattacharya et al. 2022; van Boheemen and Hodgins 2020). The four genetic clusters identified in our analyses contrast with Rosche et al. (2019) who identified two climate-associated clusters in this species. Using genome-wide SNP data across more populations, we detected finer-scale substructures across the Northern Hemisphere. The distribution of three of these clusters likely reflect historical differentiation, shaped by both isolation by distance and isolation-by-environment, within the native range and subsequent establishment in environmentally similar regions of the non-native range (Smith et al. 2020; Hernández-Espinosa et al. 2022; Kreiner et al. 2022). The orange cluster was almost absent in the native range but widespread in Eurasia. The most parsimonious explanation is that it represents the introduction of a pre-existing, low-frequency lineage from North America (e.g., from prairies and/or the East Coast where non-admixed orange individuals occur), followed by demographic expansion and diversification in Europe. Given the long establishment history of *E. canadensis* in Europe (> 400 years; Kowarik 2023), stochastic introduction processes combined with subsequent expansion could plausibly generate this pattern.

An alternative, less parsimonious scenario is that the orange cluster arose after introduction (e.g., via divergence from other introduced lineages) and subsequently contributed to gene

flow back into North America. Signals of asymmetric gene flow detected in our analyses are compatible with this possibility, although such patterns could also reflect recent, human-mediated transport rather than historical invasion dynamics. Distinguishing between these scenarios would require time-explicit genomic data (e.g., herbarium material) and/or higher-resolution genomic analyses to estimate divergence times and the directionality of gene flow.

The observed population structure reflected both geographic and ecological factors, with differently pronounced relative effects of these drivers across native and non-native populations. Both ranges showed isolation by distance; yet the relationship was more strongly pronounced in non-native populations, likely due to recent introductions followed by step-wise range expansion (Shirk et al. 2014; Smith et al. 2020; Hernández-Espinosa et al. 2022). In the native range, climate also contributed to differentiation, highlighting the combined influence of demographic history and environmental heterogeneity (Bhattacharya et al. 2022). In contrast, the absence of isolation-by-environment in the non-native range may reflect the disruptive effects of human-mediated long-distance dispersal. Although the human footprint did not strongly predict genetic differentiation in the dbRDA analysis, anthropogenic transport of propagules can facilitate gene flow across environmental gradients, thereby homogenizing genetic variation and eroding climate-genetic associations without directly structuring genomic variation itself.

4.2 | Recent Gene Flow and Historical Demography

Although *E. canadensis* is globally distributed, our analyses detected only moderate levels of intercontinental migration between North America and Eurasia, likely due to the strong genetic drift within both ranges. Still, Eastern Europe, the Western Mediterranean, Middle Asia, and China appeared as major hubs of gene flow, likely driven by intensive agriculture and high anthropogenic pressure in these regions (Smith et al. 2020). Frequent trading and travel may have established efficient dispersal routes that promoted long-distance propagule movement (Hulme 2009; Wilson et al. 2009; Einfeldt et al. 2020).

The spatial patterns revealed by FEEMS analyses support a model of invasion driven by heterogeneous gene flow rather than uniform connectivity across the global range (Marcus et al. 2021; Smith et al. 2024). Areas of reduced effective migration likely correspond to regions where founder effects and genetic drift have been particularly strong, whereas zones of higher connectivity are broadly consistent with regions identified as major hubs of gene flow in our divMigrate analyses. This spatial heterogeneity is consistent with a predominantly selfing species in which local demographic processes dominate within regions, while human-mediated dispersal facilitates connectivity across larger spatial scales (Rosche et al. 2017; Smith et al. 2020; Hudson et al. 2022).

From these hubs, rapid expansion across Eurasia was likely facilitated by small wind-dispersed seeds, a short life cycle, prolific seed production, and ecological adaptability (Weaver 2001; Bajwa et al. 2016). These traits, along with its success in anthropogenic habitats such as agricultural fields, roadsides, and urban areas (Clements 2018; Clements and Jones 2021), facilitated its establishment and further proliferation across large parts of Eurasia. Historical intercontinental spread, predating the recent migration signals detected by our SNP analyses, may have involved rare long-distance dispersal events (Smith et al. 2020), potentially mediated by wind currents (Benvenuti 2007), bird movement (Martín-Vélez et al. 2021), or historical human-mediated introductions through colonial trade routes (Kowarik 2023). Together, our findings highlight the challenges of reconstructing invasion histories in selfing species with strong dispersal capacity, where repeated introductions and stochastic processes obscure phylogenetic signals. This complexity is consistent with recent reviews emphasizing how multiple introductions and ongoing gene flow hinder the resolution of invasion pathways (Hudson et al. 2022). Spatio-temporal explicit analyses using herbarium specimens may provide more nuanced information on the invasion history in this species (Rosche et al. 2025).

4.3 | Genotype–Environment Associations and Potential Mismatches in Some Non-Native Populations

Despite range expansion, non-native populations retained a genetic composition similar to native ones across several bioclimatic regions, suggesting that pre-existing variation supported establishment in novel environments. For example, in both ranges, the RDA showed that the green cluster was associated with climatic water deficit, while the blue cluster was widespread at higher latitudes. Despite this general genomic-environmental match across both ranges, the genomic offset analysis indicated potential environmental mismatches in some regions of the non-native range. Although high genomic offset values have often been linked to extinction risk in rare native species (e.g., Fitzpatrick et al. 2021), this logic does not necessarily apply to invasives (Camus et al. 2024; Chen et al. 2024). We found a negative association between offset and field performance, and thus, offset values appeared to reflect ecological constraints on plant performance, consistent with maladaptation. Similar results were reported in the invasive cheatgrass (*Bromus tectorum* L.), where populations with higher offsets occurred in areas of lower abundance (Gamba et al. 2025).

Note that the persistence of populations exhibiting high genomic offset alongside moderate field performance suggests that *E. canadensis* may partially buffer maladaptation through microhabitat selection or phenotypic plasticity (Nicotra et al. 2010; Moran and Alexander 2014). As a ruderal species (Grime 2006), it frequently colonizes disturbed and resource-rich patches that may locally mitigate climatic mismatch at finer spatial scales. However, our results indicate that such local-scale buffering does not eliminate fitness costs associated with broader genotype–environment mismatch (Lotterhos 2024). Instead, our findings reinforce the biological relevance of genomic offset estimates (Gain et al. 2023; Lind and Lotterhos 2025) and specifically support the use of genomic offset as a proxy for adaptive mismatch in non-native species. As such, high offset values in some regions suggested that the species had not reached equilibrium with local environmental conditions. In terms of invasiveness, our findings raise concern that *E. canadensis* could exert

stronger impacts on native vegetation in the future. This interpretation is supported by Nagy et al. (2024), who showed that *E. canadensis* is not yet fully adapted in its non-native range but already attains high abundance. Thus, even long-established populations may still be evolving in response to abiotic conditions, potentially explaining pronounced lag times in biological invasions (Durand et al. 2024).

5 | Conclusions

Together, our findings highlight the interplay between demographic history and environmental variation in shaping the genomic structure of *E. canadensis*. Comparable levels of genetic diversity across ranges likely reflect the species' selfing reproductive system rather than post-introduction retention, while admixture among native lineages indicates that multiple introductions contributed to the non-native gene pool. A single lineage was particularly dominant outside the native range, suggesting that it largely drove the species' expansion in Europe. Environmental variables, especially climatic water deficit and latitude, further structured genetic variation, and human activities may have facilitated ongoing evolutionary change by promoting long-distance dispersal and admixture across environments. By integrating population genomics with field performance, we show that maladapted populations still exist in the non-native range and that continued adaptation could facilitate further spread. Overall, this work advances our understanding of the mechanisms underlying plant invasions and underscores the joint roles of demographic and environmental drivers in predicting range expansion. Future research integrating reciprocal transplant experiments and functional genomics may help clarify the processes driving adaptation and dispersal in invasive *E. canadensis* populations.

Author Contributions

Marilia Souza Lucas, Christoph Rosche, Renske E. Onstein, Christopher D. Barratt, and Walter Durka designed the research. Marilia Souza Lucas performed the analyses and wrote the manuscript with input from Christoph Rosche, Isabell Hensen, Christopher D. Barratt, Walter Durka, and Renske E. Onstein. Stefan G. Michalski, Dávid U. Nagy, Diana Gamba, Christopher D. Barratt, Christoph Rosche, and Walter Durka contributed to data analysis. Karime Abidkulova, Mohammad Al-Ghareibeh, Ali A. Al-Namazi, Sirwan Babaei, Fernando Bastida, Caio Brunharo, Aurélien Caillon, Ragan M. Callaway, Sam St. Clair, Darron A. Cullen, Ryan Donnelly, Alyson Ennis, David J. Ensing, Özkan Eren, Andrey Erst, Rita Filep, Luke Flory, Lauren J. Frazee, Yusufjon Gafforov, Fidji Gendron, Zigmantas Gudžinskas, Christopher M. Guilliams, Avni Hajdari, Jian-Hua Hao, Erin Haramoto, Anna A. Ivashchenko, Li Jun, Nolan Kane, Matthew A. Kaproth, Hesham Kassem, Abdelmajid Khabbach, Damase P. Khasa, Matthew Koski, Maria Kozhevnikova, Nikos Krigas, Denis Krivenko, Jesse R. Lasky, Ylva Lekberg, Mohamed Libiad, Vanessa Lozano, Wenbo Luo, Trobjon Makhkamov, Elizabete Marchante, Chandra Moffat, Abigail Moore, Albert Muldashev, Anush Nersesyan, Pål Axel Olsson, Adrian Oprea, Robert Pal, Astghik Papikyan, Christopher Proctor, Vadim Prokhorov, Satu Ramula, Emily Rauschert, Bryan Reatini, Christian Rixen, Scott Rush, Gemma Rutten, Mariem Saadani, Ioulietta Samartza, Julian Selke, Stepan Senator, Manzoor A. Shah, Min Sheng, Jamal Sheriff, Baasanmunkh Shukherdorj, Peter Sikkema, Tatiana Silaeva, Satoshi N. Suzuki, Baoliang Tian, Valeriy Tokhtar, Sabrina Träger, Tomonori Tsunoda, Orzimat Turginov, Kathryn G. Turner, Tatyana Vakhlamova, Yulia Vinogradova, Viktoria Wagner, Lydia Westberg, and Sa Xiao contributed to this study by collecting seed material and associated field data, including measurements of above-ground biomass and reproductive output. These authors further contributed through participation in project workshops and meetings over multiple years, where preliminary results and future research directions were discussed, and by providing feedback and comments on the final version of the manuscript. All authors read and approved the final manuscript.

Affiliations

[1]Martin Luther University Halle-Wittenberg, Halle, Germany | [2]German Centre for Integrative Biodiversity Research, Halle, Germany | [3]Helmholtz Center for Environmental Research, Leipzig, Germany | [4]Goethe-University, Frankfurt, Germany | [5]Pennsylvania State University, University Park, Pennsylvania, USA | [6]Naturalis Biodiversity Center, Leiden, the Netherlands | [7]Al-Farabi Kazakh National University, Almaty, Kazakhstan | [8]Jordan University of Science and Technology, Irbid, Jordan | [9]National Center for Vegetation Cover, Riyadh, Saudi Arabia | [10]University of Kurdistan, Sanandaj, Iran | [11]University of Huelva, Huelva, Spain | [12]Conservatoire Botanique National Sud-Atlantique, Audenge, France | [13]University of Montana, Missoula, Montana, USA | [14]Brigham Young University, Provo, Utah, USA | [15]University of Hull, Hull, UK | [16]Kansas State University, Manhattan, Kansas, USA | [17]University of Colorado, Denver, Colorado, USA | [18]Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada | [19]Aydın Adnan Menderes University, Aydin, Turkey | [20]Russian Academy of Sciences, Moscow, Russia | [21]University of Pécs, Pécs, Hungary | [22]University of Florida, Gainesville, Florida, USA | [23]PharmaLex, Conshohocken, Pennsylvania, USA | [24]New Uzbekistan University, Tashkent, Uzbekistan | [25]National University of Uzbekistan, Tashkent, Uzbekistan | [26]First Nations University of Canada, Regina, Saskatchewan, Canada | [27]Nature Research Centre, Vilnius, Lithuania | [28]Santa Barbara Botanic Garden, Santa Barbara, California, USA | [29]University of Prishtina, Prishtine, Kosovo | [30]Changshu Institute of Technology, Changshu, China | [31]University of Kentucky, Lexington, Kentucky, USA | [32]Institute of Zoology Republic of Kazakhstan, Almaty, Kazakhstan | [33]Chinese Academy of Sciences, Beijing, China | [34]Aksu National Oasis Agroecosystem Observation and Research Station, Aksu, China | [35]Minnesota State University, Mankato, Minnesota, USA | [36]Suez Canal University, Ismailia, Egypt | [37]Sidi Mohamed Ben Abdallah University, Fez, Morocco | [38]Université Laval, Quebec City, Quebec, Canada | [39]Clemson University, Clemson, South Carolina, USA | [40]Kazan Federal University, Kazan, Russia | [41]Hellenic Agricultural Organization, Demeter, Greece | [42]Siberian Institute of Plant Physiology and Biochemistry, Irkutsk, Russia | [43]MPG Ranch Missoula, Missoula, Montana, USA | [44]Abdelmalek Essaadi University, Tetouan, Morocco | [45]University of Sassari, Sassari, Italy | [46]Northeast Normal University, Changchun, China | [47]Tashkent State Agrarian University, Tashkent, Uzbekistan | [48]University of Coimbra, Coimbra, Portugal | [49]University of Oklahoma, Norman, Oklahoma, USA | [50]Federal Scientific Centre of the Russian Academy of Sciences, Moscow, Russia | [51]National Academy of Sciences of the Republic of Armenia, Yerevan, Armenia | [52]Nature Heritage NGO, Yerevan, Armenia | [53]Lund University, Lund, Sweden | [54]Alexandru Ioan Cuza University, Iași, Romania | [55]Montana Technological University, Butte, Montana, USA | [56]University of Nebraska, Lincoln, Nebraska, USA | [57]University of Turku, Turku, Finland | [58]Cleveland State University, Cleveland, Ohio, USA | [59]United States Forest Service, Washington, DC, USA | [60]WSL Institute for Snow and Avalanche Research, Davos Dorf, Switzerland | [61]Climate Change Extremes and Natural Hazards in Alpine Regions Research Centre, Davos Dorf, Switzerland | [62]Mississippi State University, Starkville, Mississippi, USA | [63]University of Bern, Bern, Switzerland | [64]University of Regensburg, Regensburg, Germany | [65]Tsitsin Main Botanical Garden, Moscow, Russia | [66]University of Kashmir, Srinagar, Jammu and Kashmir, India | [67]Northwest A&F University, Xianyang, China | [68]University of Illinois, Champaign, Illinois, USA | [69]Changwon National University, Changwon, South

Korea | [70]University of Guelph Ridgetown Campus, Ridgetown, Ontario, Canada | [71]Ogarev Mordovia State University, Saransk, Russia | [72]Hokkaido University, Hokkaido, Japan | [73]Henan University, Henan, China | [74]Belgorod State University, Belgorod, Russia | [75]Fukui Prefectural University, Fukui, Japan | [76]National University of Uzbekistan Named After Mirzo Ulugbek, Tashkent, Uzbekistan | [77]Idaho State University, Pocatello, Idaho, USA | [78]Toraighyrov University, Pavlodar, Kazakhstan | [79]University of Alberta, Edmonton, Alberta, Canada | [80]University of Kansas, Lawrence, Kansas, USA | [81]Lanzhou University, Lanzhou, China | [82]Wageningen University & Research, Wageningen, the Netherlands

Acknowledgements

We thank all collaborators who contributed to field sampling across the species' native and non-native ranges. We are also grateful to the technical staff and student assistants for their invaluable help with greenhouse work and laboratory processing. This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and by the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (see funding information). Open Access funding enabled and organized by Projekt DEAL.

Funding

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; Project number: RO 6418/1-1) and iDiv Flexpool funding (grant number W47038118). The German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig is funded by the DFG (FZT 118, 202548816).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Demultiplexed individual raw sequence data are available under project accession number PRJEB91610, sample accession numbers ERS25137122–ERS25137761 and ERS25147679–ERS25147684, experiment accession numbers ERX14597121–ERX14597760 and ERX14603920–ERX14603925, and run accession numbers ERR15191494–ERR15192133 and ERR15198279–ERR15198284 at the European Nucleotide Archive (https://www.ebi.ac.uk/ena). The variant call sets (VCF genotype files), environmental data and R code used in this study are available on Dryad: Dataset https://doi.org/10.5061/dryad.8w9ghx40k.

References

Abatzoglou, J. T., S. Z. Dobrowski, S. A. Parks, and K. C. Hegewisch. 2018. "TerraClimate, a High-Resolution Global Dataset of Monthly Climate and Climatic Water Balance From 1958-2015." *Scientific Data* 5, no. 1: 1–12.

Ahmed, D. A., R. Sousa, A. Bortolus, et al. 2025. "Parallels and Discrepancies Between Non-Native Species Introductions and Human Migration." *Biological Reviews* 100, no. 3: 1365–1395.

Alexander, D. H., J. Novembre, and K. Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19, no. 9: 1655–1664.

Allendorf, F. W., G. H. Luikart, and S. N. Aitken. 2012. *Conservation and the Genetics of Populations*. John Wiley & Sons.

Bajwa, A. A., S. Sadia, H. H. Ali, K. Jabran, A. M. Peerzada, and B. S. Chauhan. 2016. "Biology and Management of Two Important *Conyza* Weeds: A Global Review." *Environmental Science and Pollution Research* 23: 24694–24710.

Barnes, R., and K. Sahr. 2017. "dggridR: Discrete Global Grids for R. R Package." https://github.com/r-barnes/dggridR/.

Barrett, S. C., R. I. Colautti, and C. G. Eckert. 2008. "Plant Reproductive Systems and Evolution During Biological Invasion." *Molecular Ecology* 17, no. 1: 373–383.

Bellard, C., J. Jeschke, B. Leroy, and G. Mace. 2018. "Insights From Modeling Studies on How Climate Change Affects Invasive Alien Species Geography." *Ecology and Evolution* 8: 5688–5700.

Benvenuti, S. 2007. "Weed Seed Movement and Dispersal Strategies in the Agricultural Environment." *Weed Biology and Management* 7, no. 3: 141–157.

Bhattacharya, S., F. Hernández, M. F. Alves, et al. 2022. "Genetic Diversity and Population Structure of Invasive and Native Populations of *Erigeron canadensis L*." *Journal of Plant Ecology* 15, no. 4: 864–876.

Bhowmik, P. C., and M. M. Bekech. 1993. "Horseweed (*Conyza canadensis*) Seed Production, Emergence, and Distribution in No-Tillage and Conventional-Tillage Corn (*Zea mays*)." *Agronomy Trends in Agricultural Science* 1: 67–71.

Bock, D. G., C. Caseys, R. D. Cousens, et al. 2016. "What We Still Don't Know About Invasion Genetics." *Invasion Genetics: The Baker and Stebbins Legacy*: 346–370.

Borcard, D., F. Gillet, and P. Legendre. 2011. *Numerical Ecologywith R*. Springer.

Bossdorf, O., H. Auge, L. Lafuma, W. E. Rogers, E. Siemann, and D. Prati. 2005. "Phenotypic and Genetic Differentiation Between Native and Introduced Plant Populations." *Oecologia* 144, no. 1: 1–11.

Bradburd, G. S., G. M. Coop, and P. L. Ralph. 2018. "Inferring Continuous and Discrete Population Genetic Structure Across Space." *Genetics* 210, no. 1: 33–52.

Bryant, D., and V. Moulton. 2004. "Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks." *Molecular Biology and Evolution* 21, no. 2: 255–265.

Camus, L., M. Gautier, and S. Boitard. 2024. "Predicting Species Invasiveness With Genomic Data: Is Genomic Offset Related to Establishment Probability?" *Evolutionary Applications* 17, no. 6: e13709.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. "Stacks: An Analysis Tool Set for Population Genomics." *Molecular Ecology* 22, no. 11: 3124–3140.

Chen, X., W. Liu, Y. Y. Zhang, and Y. Zhang. 2023. "Altered Trait Covariances Between Invasive and Native Ranges of a Global Plant Invader." *Functional Ecology* 37, no. 5: 1280–1290.

Chen, Y., Y. Gao, X. Huang, S. Li, Z. Zhang, and A. Zhan. 2024. "Incorporating Adaptive Genomic Variation Into Predictive Models for Invasion Risk Assessment." *Environmental Science and Ecotechnology* 18: 100299.

Chen, Y., Z. Jiang, P. Fan, et al. 2022. "The Combination of Genomic Offset and Niche Modelling Provides Insights Into Climate Change-Driven Vulnerability." *Nature Communications* 13, no. 1: 4821.

Clements, D. R. 2018. "Invasive Weed Species and Their Effects." In *Integrated Weed Management for Sustainable Agriculture*, 65–88. Burleigh Dodds Science Publishing.

Clements, D. R., and V. L. Jones. 2021. "Rapid Evolution of Invasive Weeds Under Climate Change: Present Evidence and Future Research Needs." *Frontiers in Agronomy* 3: 664034.

Colautti, R. I., and J. A. Lau. 2016. "Contemporary Evolution During Invasion: Evidence for Differentiation, Natural Selection, and Local Adaptation." *Invasion Genetics: The Baker and Stebbins Legacy*: 101–121.

Cropano, C., I. Place, C. Manzanares, et al. 2021. "Characterization and Practical Use of Self-Compatibility in Outcrossing Grass Species." *Annals of Botany* 127: 84–852.

Davis, V. M., G. R. Kruger, S. G. Hallett, P. J. Tranel, and W. G. Johnson. 2010. "Heritability of Glyphosate Resistance in Indiana Horseweed (*Conyza canadensis*) Populations." *Weed Science* 58, no. 1: 30–38.

Dlugosch, K. M., and I. M. Parker. 2008. "Founding Events in Species Invasions: Genetic Variation, Adaptive Evolution, and the Role of Multiple Introductions." *Molecular Ecology* 17, no. 1: 431–449.

Durand, K., S. Yainna, and K. Nam. 2024. "Population Genomics Unravels a Lag Phase During the Global Fall Armyworm Invasion." *Communications Biology* 7, no. 1: 957.

Durka, W., S. G. Michalski, J. Höfner, et al. 2025. "Assessment of Genetic Diversity Among Seed Transfer Zones for Multiple Grassland Plant Species Across Germany." *Basic and Applied Ecology* 84: 50–60.

Einfeldt, A. L., L. K. Jesson, and J. A. Addison. 2020. "Historical Human Activities Reshape Evolutionary Trajectories Across Both Native and Introduced Ranges." *Ecology and Evolution* 10, no. 13: 6579–6592.

Encinas-Viso, F., L. Morin, S. Raghu, N. Knerr, C. Roux, and L. Broadhurst. 2022. "Population Genomics Reveal Multiple Introductions and Admixture of *Sonchus oleraceus* in Australia." *Diversity and Distributions* 28, no. 9: 1951–1965.

Estoup, A., and T. Guillemaud. 2010. "Reconstructing Routes of Invasion Using Genetic Data: Why, How and So What?" *Molecular Ecology* 19, no. 19: 4113–4130.

Facon, B., J. P. Pointier, P. Jarne, V. Sarda, and P. David. 2008. "High Genetic Variance in Life-History Strategies Within Invasive Populations by Way of Multiple Introductions." *Current Biology* 18, no. 5: 363–367.

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, no. 12: 4302–4315.

Fitzpatrick, M. C., V. E. Chhatre, R. Y. Soolanayakanahally, and S. R. Keller. 2021. "Experimental Support for Genomic Prediction of Climate Maladaptation Using the Machine Learning Approach Gradient Forests." *Molecular Ecology Resources* 21, no. 8: 2749–2765.

Gain, C., B. Rhoné, P. Cubry, et al. 2023. "A Quantitative Theory for Genomic Offset Statistics." *Molecular Biology and Evolution* 40, no. 6: msad140.

Gamba, D., M. L. Vahsen, T. M. Maxwell, et al. 2025. "Local Adaptation to Climate Facilitates a Global Invasion." BioRxiv [Preprint]: 2024.09.12.612725.

Garrison, E., Z. N. Kronenberg, E. T. Dawson, B. S. Pedersen, and P. Prins. 2021. "Vcflib and Tools for Processing the VCF Variant Call Format. bioRxiv, 2021-05."

Goudet, J., and T. Jombart. 2015. "Hierfstat: Estimation and Tests of Hierarchical F-Statistics." *R Package Version* 10: 04–22.

Greve, M., and L. Pertierra. 2022. "Opportunities for Studying Propagule Pressure Using Gene Flow Reveal Its Role in Accelerating Biological Invasions." *Molecular Ecology* 31: 1609–1611.

Grime, J. P. 2006. *Plant Strategies, Vegetation Processes, and Ecosystem Properties*. John Wiley & Sons.

Gruber, B., P. J. Unmack, O. F. Berry, and A. Georges. 2018. "Dartr: An R Package to Facilitate Analysis of SNP Data Generated From Reduced Representation Genome Sequencing." *Molecular Ecology Resources* 18, no. 3: 691–699.

Hamilton, J. A., M. Okada, T. Korves, and J. Schmitt. 2016. "The Role of Climate Adaptation in Colonization Success in *Arabidopsis thaliana*." *Invasion Genetics: The Baker and Stebbins Legacy* 24, no. 9: 2253–2263.

Hernández-Espinosa, R., J. González-Astorga, Y. Rico, and J. B. Gallego-Fernández. 2022. "Effect of Life-History Traits and Habitat Condition on Genetic Diversity Between Invasive and Native Plant Populations." *Diversity* 14, no. 12: 1025.

Hess, M. C., F. Mesléard, T. P. Young, B. de Freitas, N. Haveneers, and E. Buisson. 2022. "Altering Native Community Assembly History Influences the Performance of an Annual Invader." *Basic and Applied Ecology* 59: 70–81.

Hijmans, R. J., J. Van Etten, J. Cheng, et al. 2015. "Package 'raster'." *R Package* 734: 473.

Hodgins, K. A., P. Battlay, and D. G. Bock. 2025. "The Genomic Secrets of Invasive Plants." *New Phytologist* 245, no. 5: 1846–1863.

Hudson, J., S. D. Bourne, H. Seebens, M. A. Chapman, and M. Rius. 2022. "The Reconstruction of Invasion Histories With Genomic Data in Light of Differing Levels of Anthropogenic Transport." *Philosophical Transactions of the Royal Society B* 377, no. 1846: 20210023.

Hulme, P. E. 2009. "Trade, Transport and Trouble: Managing Invasive Species Pathways in an Era of Globalization." *Journal of Applied Ecology* 46, no. 1: 10–18.

Huson, D. H., and D. Bryant. 2006. "Application of Phylogenetic Networks in Evolutionary Studies." *Molecular Biology and Evolution* 23, no. 2: 254–267.

Jiang, H., Y. Zhang, W. Tu, G. Sun, N. Wu, and Y. Zhang. 2023. "The General Trends of Genetic Diversity Change in Alien Plants' Invasion." *Plants* 12, no. 14: 2690.

Jombart, T. 2008. "Adegenet: A R Package for the Multivariate Analysis of Genetic Markers." *Bioinformatic* 24: 1403–1405.

Kalsing, A., E. D. Velini, A. Merotto Jr., and C. A. Carbonari. 2024. "The Population Genomics of *Conyza* spp. in Soybean Macroregions Suggest the Spread of Herbicide Resistance Through Intraspecific and Interspecific Gene Flow." *Scientific Reports* 14, no. 1: 19536.

Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. "Poppr: An R Package for Genetic Analysis of Populations With Clonal, Partially Clonal, and/or Sexual Reproduction." *PeerJ* 2: e281.

Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, and P. A. Prodöhl. 2013. "diveRsity: An R Package for the Estimation and Exploration of Population Genetics Parameters and Their Associated Errors." *Methods in Ecology and Evolution* 4, no. 8: 782–788.

Kowarik, I. 2023. "Historical Evidence for Context-Dependent Assessment of *Erigeron canadensis* Invasions in an 18th-Century European Landscape." *NeoBiota* 89: 1–15.

Kreiner, J. M., S. M. Latorre, H. A. Burbano, et al. 2022. "Rapid Weed Adaptation and Range Expansion in Response to Agriculture Over the Past Two Centuries." *Science* 378, no. 6624: 1079–1085.

Lachmuth, S., M. Fitzpatrick, A. Prakash, S. Keller, and T. Capblancq. 2023. "Novel Genomic Offset Metrics Integrate Local Adaptation Into Habitat Suitability Forecasts and Inform Assisted Migration." *Ecological Monographs* 94, no. 1: e1593.

Laforest, M., S. L. Martin, K. Bisaillon, B. Soufiane, S. Meloche, and E. Page. 2020. "A Chromosome-Scale Draft Sequence of the Canada Fleabane Genome." *Pest Management Science* 76, no. 6: 2158–2169.

Le Roux, J. 2021. *The Evolutionary Ecology of Invasive Species*, 12–21. Academic Press.

Leon, R. G., J. C. Dunne, and F. Gould. 2021. "The Role of Population and Quantitative Genetics and Modern Sequencing Technologies to Understand Evolved Herbicide Resistance and Weed Fitness." *Pest Management Science* 77, no. 1: 12–21.

Li, F., M. van Kleunen, J. Li, et al. 2019. "Patterns of Genetic Variation Reflect Multiple Introductions and Pre-Admixture Sources of Common Ragweed (*Ambrosia artemisiifolia*) in China." *Biological Invasions* 21: 2191–2209.

Lind, B. M., and K. E. Lotterhos. 2025. "The Accuracy of Predicting Maladaptation to New Environments With Genomic Data." *Molecular Ecology Resources* 25, no. 4: e14008.

Liu, J., M. Qi, and J. Wang. 2018. "Long-Distance and Dynamic Seed Dispersal From Horseweed (*Conyza canadensis*)." *Écoscience* 25, no. 3: 271–285.

Lotterhos, K. 2024. "Interpretation Issues With "Genomic Vulnerability" Arise From Conceptual Issues in Local Adaptation and Maladaptation." *Evolution Letters* 8: 331–339.

Lucas, M. S., I. Hensen, C. D. Barratt, et al. 2024. "Re-Focusing Sampling, Design and Experimental Methods to Assess Rapid Evolution by Non-Native Plant Species." *Biological Invasions* 26, no. 5: 1327–1343.

MacLaren, C., J. Storkey, A. Menegat, H. Metcalfe, and K. Dehnen-Schmutz. 2020. "An Ecological Future for Weed Science to Sustain Crop Production and the Environment. A Review." *Agronomy for Sustainable Development* 40: 1–29.

Marchini, G. L., C. A. Maraist, and M. B. Cruzan. 2019. "Trait Divergence, Not Plasticity, Determines the Success of a Newly Invasive Plant." *Annals of Botany* 123, no. 4: 667–679.

Marcus, J. H., W. Ha, R. Foygel Barber, and J. Novembre. 2021. "Fast and Flexible Estimation of Effective Migration Surfaces." *eLife* 10: e61927.

Martin, S. L., J. S. Parent, M. Laforest, E. Page, J. M. Kreiner, and T. James. 2019. "Population Genomic Approaches for Weed Science." *Plants* 8, no. 9: 354.

Martín-Vélez, V., C. H. van Leeuwen, M. I. Sánchez, et al. 2021. "Spatial Patterns of Weed Dispersal by Wintering Gulls Within and Beyond an Agricultural Landscape." *Journal of Ecology* 109, no. 4: 1947–1958.

Matheson, P., and A. McGaughran. 2022. "Genomic Data Is Missing for Many Highly Invasive Species, Restricting Our Preparedness for Escalating Incursion Rates." *Scientific Reports* 12, no. 1: 13987.

McGaughran, A., M. K. Dhami, E. Parvizi, et al. 2024. "Genomic Tools in Biological Invasions: Current State and Future Frontiers." *Genome Biology and Evolution* 16, no. 1: evad230.

Montgomery, J., S. Morran, D. R. MacGregor, et al. 2024. "Current Status of Community Resources and Priorities for Weed Genomics Research." *Genome Biology* 25, no. 1: 139.

Moran, E. V., and J. M. Alexander. 2014. "Evolutionary Responses to Global Change: Lessons From Invasive Species." *Ecology Letters* 17, no. 5: 637–649.

Muirhead, J. R., D. K. Gray, D. W. Kelly, S. M. Ellis, D. D. Heath, and H. J. Macisaac. 2008. "Identifying the Source of Species Invasions: Sampling Intensity vs. Genetic Diversity." *Molecular Ecology* 17, no. 4: 1020–1035.

Nagy, D. U., A. E. Thoma, M. Al-Gharaibeh, et al. 2024. "Among-Population Variation in Drought Responses Is Consistent Across Life Stages but Not Between Native and Non-Native Ranges." *New Phytologist* 243, no. 3: 922–935.

Neve, P., M. Vila-Aiub, and F. Roux. 2009. "Evolutionary-Thinking in Agricultural Weed Management." *New Phytologist* 184, no. 4: 783–793.

Nicotra, A. B., O. K. Atkin, S. P. Bonser, et al. 2010. "Plant Phenotypic Plasticity in a Changing Climate." *Trends in Plant Science* 15, no. 12: 684–692.

Novak, S. J., and R. N. Mack. 2001. "Tracing Plant Introduction and Spread: Genetic Evidence From *Bromus tectorum* (Cheatgrass)." *Bioscience* 51, no. 2: 114–122.

Okada, M., B. D. Hanson, K. J. Hembree, et al. 2013. "Evolution and Spread of Glyphosate Resistance in *Conyza canadensis* in California." *Evolutionary Applications* 6, no. 5: 761–777.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, and D. McGlinn. 2022. "Vegan: Community Ecology Package." *R Package Version* 2: 2–3.

O'Leary, S. J., J. B. Puritz, S. C. Willis, C. M. Hollenbeck, and D. S. Portnoy. 2018. "These Aren't the Loci You're Looking for: Principles of Effective SNP Filtering for Molecular Ecologists." *Molecular Ecology* 27, no. 16: 3193–3206.

Pannell, J. R. 2015. "Evolution of the Mating System in Colonizing Plants." *Molecular Ecology* 24, no. 9: 2018–2037.

Paradis, E., J. Claude, and K. Strimmer. 2004. "APE: Analyses of Phylogenetics and Evolution in R Language." *Bioinformatics* 20, no. 2: 289–290.

Pembleton, L. W., N. O. Cogan, and J. W. Forster. 2013. "StAMPP: An R Package for Calculation of Genetic Differentiation and Structure of Mixed-Ploidy Level Populations." *Molecular Ecology Resources* 13, no. 5: 946–952.

Peng, Y., Z. Lai, T. Lane, et al. 2014. "De Novo Genome Assembly of the Economically Important Weed Horseweed Using Integrated Data From Multiple Sequencing Platforms." *Plant Physiology* 166, no. 3: 1241–1254.

Petanidou, T., R. Godfree, D. Song, A. Kantsa, Y. Dupont, and N. Waser. 2012. "Self-Compatibility and Plant Invasiveness: Comparing Species in Native and Invasive Ranges." *Perspectives in Plant Ecology, Evolution and Systematics* 14: 3–12.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. "Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species." *PLoS One* 7, no. 5: e37135.

Prentis, P. J., J. R. Wilson, E. E. Dormontt, D. M. Richardson, and A. J. Lowe. 2008. "Adaptive Evolution in Invasive Species." *Trends in Plant Science* 13, no. 6: 288–294.

Prieur-Richard, A. H., S. Lavorel, K. Grigulis, and A. Dos Santos. 2000. "Plant Community Diversity and Invasibility by Exotics: Invasion of Mediterranean Old Fields by *Conyza bonariensis* and *Conyza canadensis*." *Ecology Letters* 3, no. 5: 412–422.

Puritz, J. B., C. M. Hollenbeck, and J. R. Gold. 2014. "dDocent: A RADseq, Variant-Calling Pipeline Designed for Population Genomics of Non-Model Organisms." *PeerJ* 2: e431.

Pyšek, P., M. Lučanová, W. Dawson, et al. 2023. "Small Genome Size and Variation in Ploidy Levels Support the Naturalization of Vascular Plants but Constrain Their Invasive Spread." *New Phytologist* 239: 2389–2403.

Rellstab, C., and S. Keller. 2025. "Can we Use Genomic Data to Predict Maladaptation to Environmental Change?" *Molecular Ecology Resources* 25, no. 4: e14059.

Rosche, C., O. Broennimann, A. Novikov, et al. 2025. "Herbarium Specimens Reveal a Cryptic Invasion of Polyploid *Centaurea stoebe* in Europe." *New Phytologist* 245, no. 1: 392–405.

Rosche, C., S. Heinicke, I. Hensen, et al. 2018. "Spatio-Environmental Determinants of the Genetic Structure of Three Steppe Species in a Highly Fragmented Landscape." *Basic and Applied Ecology* 28: 48–59.

Rosche, C., I. Hensen, P. Mraz, W. Durka, M. Hartmann, and S. Lachmuth. 2017. "Invasion Success in Polyploids: The Role of Inbreeding in the Contrasting Colonization Abilities of Diploid Versus Tetraploid Populations of *Centaurea stoebe* Sl." *Journal of Ecology* 105, no. 2: 425–435.

Rosche, C., I. Hensen, A. Schaar, et al. 2019. "Climate Outweighs Native vs. Nonnative Range-Effects for Genetics and Common Garden Performance of a Cosmopolitan Weed." *Ecological Monographs* 89, no. 4: e01386.

Schierenbeck, K. A. 2017. "Population-Level Genetic Variation and Climate Change in a Biodiversity Hotspot." *Annals of Botany* 119, no. 2: 215–228.

Schliep, K. P. 2011. "Phangorn: Phylogenetic Analysis in R." *Bioinformatics* 27, no. 4: 592–593.

Shah, M. A., R. M. Callaway, T. Shah, et al. 2014. "*Conyza canadensis* Suppresses Plant Diversity in Its Nonnative Ranges but Not at Home: A Transcontinental Comparison." *New Phytologist* 202, no. 4: 1286–1296.

Sheng, M., C. Rosche, M. Al-Gharaibeh, et al. 2022. "Acquisition and Evolution of Enhanced Mutualism—An Underappreciated Mechanism for Invasive Success?" *ISME Journal* 16, no. 11: 2467–2478.

Sherpa, S., and L. Després. 2021. "The Evolutionary Dynamics of Biological Invasions: A Multi-Approach Perspective." *Evolutionary Applications* 14: 1463–1484.

Shields, E. J., J. T. Dauer, M. J. VanGessel, and G. Neumann. 2006. "Horseweed (*Conyza canadensis*) Seed Collected in the Planetary Boundary Layer." *Weed Science* 54, no. 6: 1063–1067.

Shirk, R. Y., J. L. Hamrick, C. Zhang, and S. Qiang. 2014. "Patterns of Genetic Diversity Reveal Multiple Introductions and Recurrent Founder Effects During Range Expansion in Invasive Populations of *Geranium carolinianum* (Geraniaceae)." *Heredity* 112, no. 5: 497–507.

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, no. 8: 4218–4227.

Smith, T. A., M. D. Martin, M. Nguyen, and T. C. Mendelson. 2024. "Estimation of Spatial Demographic Maps From Polymorphism Data Using a Neural Network." *Molecular Ecology Resources* 24, no. 1: e13920.

Stankowski, S., M. A. Chase, A. M. Fuiten, M. F. Rodrigues, P. L. Ralph, and M. A. Streisfeld. 2019. "Widespread Selection and Gene Flow Shape the Genomic Landscape During a Radiation of Monkeyflowers." *PLoS Biology* 17, no. 7: e3000391.

Sundqvist, L., K. Keenan, M. Zackrisson, P. Prodöhl, and D. Kleinhans. 2016. "Directional Genetic Differentiation and Relative Migration." *Ecology and Evolution* 6, no. 11: 3461–3475.

Traveset, A., and D. M. Richardson. 2006. "Biological Invasions as Disruptors of Plant Reproductive Mutualisms." *Trends in Ecology & Evolution* 21, no. 4: 208–216.

Turner, K. G., R. A. Hufbauer, and L. H. Rieseberg. 2014. "Rapid Evolution of an Invasive Weed." *New Phytologist* 202, no. 1: 309–321.

Uesugi, A., D. J. Baker, N. de Silva, K. Nurkowski, and K. A. Hodgins. 2020. "A Lack of Genetically Compatible Mates Constrains the Spread of an Invasive Weed." *New Phytologist* 226, no. 6: 1864–1872.

van Boheemen, L., and K. Hodgins. 2020. "Rapid Repeatable Phenotypic and Genomic Adaptation Following Multiple Introductions." *Molecular Ecology* 29: 4102–4117.

van Boheemen, L. A., E. Lombaert, K. A. Nurkowski, B. Gauffre, L. H. Rieseberg, and K. A. Hodgins. 2017. "Multiple Introductions, Admixture and Bridgehead Invasion Characterize the Introduction History of *Ambrosia artemisiifolia* in Europe and Australia." *Molecular Ecology* 26, no. 20: 5421–5434.

VanGessel, M. J. 2001. "Glyphosate-Resistant Horseweed From Delaware." *Weed Science* 49, no. 6: 703–705.

Venter, O., E. W. Sanderson, A. Magrach, et al. 2018. *Last of the Wild Project, Version 3 (LWP-3): 2009 Human Footprint, 2018 Release (Version 2018.00) [Data Set]*. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H46T0JQ4.

Voss, N., R. L. Eckstein, and W. Durka. 2012. "Range Expansion of a Selfing Polyploid Plant Despite Genetic Uniformity." *Annals of Botany* 110: 585–593.

Wang, Y., and D. J. Obbard. 2023. "Experimental Estimates of Germline Mutation Rate in Eukaryotes: A Phylogenetic Meta-Analysis." *Evolution Letters* 7, no. 4: 216–226.

Weaver, S. E. 2001. "The Biology of Canadian Weeds. 115. *Conyza canadensis*." *Canadian Journal of Plant Science* 81, no. 4: 867–875.

Wilson, J. R., E. E. Dormontt, P. J. Prentis, A. J. Lowe, and D. M. Richardson. 2009. "Something in the Way You Move: Dispersal Pathways Affect Invasion Success." *Trends in Ecology & Evolution* 24, no. 3: 136–144.

Wright, S. 1965. "The Interpretation of Population Structure by F-Statistics With Special Regard to Systems of Mating." *Evolution* 19, no. 3: 395–420.

Wright, S. I., R. W. Ness, J. P. Foxe, and S. C. Barrett. 2008. "Genomic Consequences of Outcrossing and Selfing in Plants." *International Journal of Plant Sciences* 169, no. 1: 105–118.

Yan, H., L. Feng, Y. Zhao, et al. 2020. "Predicting the Potential Distribution of an Invasive Species, *Erigeron canadensis* L., in China With a Maximum Entropy Model." *Global Ecology and Conservation* 21: e00822.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T. Y. Lam. 2017. "Ggtree: An R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data." *Methods in Ecology and Evolution* 8, no. 1: 28–36.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Information on *Erigeron canadensis* populations used in the analyses, including their range, country of origin, region, and geographical coordinates (longitude and latitude). **Table S2:** Parameters used during the dDocent SNP genotyping process for assembly, mapping, and SNP calling as given in the configuration file. **Table S3:** Bioclimatic variables from the WorldClim database used in environmental analyses. Temperature variables (BIO1–BIO11) represent thermal conditions and seasonality patterns, while precipitation variables (BIO12–BIO19) capture moisture availability and temporal distribution. These 19 variables were used to characterize environmental conditions across the native and non-native ranges of Erigeron canadensis. **Table S4a:** Hierarchical analyses of molecular variance of *Erigeron canadensis* populations. **Table S4b:** Analyses of molecular variance of native and non-native *Erigeron canadensis* populations. **Table S4c:** Hierarchical analyses of molecular variance of *Erigeron canadensis* regions. **Table S5:** Medium to high relative migration rates (values >0.5) among *Erigeron canadensis* populations, estimated using divMigrate. Each entry represents a directional gene flow estimate between a source region and an immigrant region. Values correspond to relative migration rates from 0 to 1, with higher values indicating stronger directional gene flow. **Table S6:** Population-level site scores from the distance-based redundancy analysis (CAP1 and CAP2) together with associated environmental variables. Values represent means across individuals per population and include latitude, climatic water deficit (CWD), human footprint (HF), and the first two principal components of bioclimatic variation (PC1_Bioclim, PC2_Bioclim). This table facilitates interpretation of cluster–environment relationships shown in Figure 5. **Table S7:** Mean genotype–environment mismatch (genomic offset) values for *Erigeron canadensis* populations grouped by geographic region non-native ranges. Values represent means across populations within each region. Higher genomic offset values indicate greater genotype–environment mismatch. To provide context for the magnitude of genomic offset values observed in the non-native range, we also summarized genomic offset values for native populations using the same framework, allowing direct comparison across native and non-native regions (Table S8). **Table S8:** Mean genotype–environment mismatch (genomic offset) values for *Erigeron canadensis* populations grouped by geographic region across native. Values represent means across populations within each region. Higher genomic offset values indicate greater genotype–environment mismatch. **Table S9:** Summary of linear mixed-effects models testing the effect of genomic offset on field performance in *Erigeron canadensis*. Two models were fitted with log-transformed aboveground biomass and number of capitula as response variables. Mean genomic offset and community productivity were included as fixed effects, and region and population (nested within region) were included as random effects. For each model, the chi-squared statistic ($\chi 2$, Type II test) and corresponding p-value for genomic offset are reported. **Figure S1:** Distribution of the 280 populations of *Erigeron canadensis* used in the work, (A) classified by range and (B) by region. **Figure S2:** Parameter optimization in dDocent using the script ReferenceOpt.sh (https://github.com/jpuritz/dDocent/raw/master/scripts/ReferenceOpt.sh). The plot shows the number of contigs obtained for different values of Clustering_Similarity% (0.80–0.98) and different values for k1 and k2. K1 is the minimum within individual

coverage level to include a read for assembly and k2 is the minimum number of individuals a read must be present in to include for assembly. Each line represents a unique k1 k2 pair. At clustering similarity percentage (%) = 0.9 the slope of the curve is suddenly increasing, indicating optimal performance at this similarity level. **Figure S3:** Fifty-fold cross-validation plot of the ADMIXTURE analysis of 640 samples of *Erigeron canadensis* for K-values from 2 to 50. The x-axis represents the number of clusters (K) in the model, and the y-axis represents the cross-validation score. The curve does not show the typical elbow with increasing scores at high values of K, indicating subtle population structures emerging at high levels of K. **Figure S4:** Bar plots depicting the ancestry coefficients of the ADMIXTURE analysis for K = 2 to K = 9 for all the *Erigeron canadensis* samples of different regions. Each vertical bar represents an individual sample. The height of the bar is proportional to the percentage of estimated ancestry assigned to each E. canadensis individual. The different colours correspond to the genetic clusters. **Figure S5:** Cross-validation of the conStruct models showing predictive accuracy (±95% CI) of the spatial and nonspatial analyses for K = 1–7 for the global range. **Figure S6:** Population-level ancestry proportions inferred using conStruct models. (A) Non-spatial conStruct models and (B) spatial conStruct models are shown for K = 2, 3, and 4 (rows). Pie charts represent admixture proportions at each population location. The optimal spatial model (K = 3, blue frame) and the optimal non-spatial model (K = 4, green frame) are highlighted. The non-spatial K = 4 model closely matches the ADMIXTURE results presented in the main text. **Figure S7:** Circular Neighbour-Joining (NJ) tree of *Erigeron canadensis* populations showing genetic relationships between native (blue) and non-native (red) populations. The tree reveals clear genetic differentiation, with native populations exhibiting extensive branching and region-specific clustering, reflecting long-term evolutionary history and natural gene flow. In contrast, non-native populations form tighter clusters with shorter branch lengths, suggesting multiple independent introduction events and restricted post-introduction gene flow. For example, similarities between non-native Mediterranean and certain native populations suggest potential source regions within North America. **Figure S8:** Scale dependence of isolation-by-distance and isolation-by-environment in native and non-native populations of *Erigeron canadensis*. (A) Genetic distance (FST/[1 − FST]) as a function of geographic distance (km, log-scaled) for native (left) and non-native (right) populations, with points coloured by climatic distance (see legend). Solid lines indicate LOESS smoothers (±95% confidence intervals) and dashed lines indicate linear fits. (B) Genetic distance as a function of climatic distance, with points coloured by geographic distance (km, log-scaled; see legend). Solid lines indicate LOESS smoothers (±95% confidence intervals) and dashed lines indicate linear fits. (C) Mean genetic distance (±SE) across geographic distance bins for native (left) and non-native (right) populations. (D) Mean genetic distance (±SE) across climatic distance bins for native (left) and non-native (right) populations. Together, these analyses show a monotonic increase in genetic differentiation with geographic distance and no evidence for scale-dependent isolation-by-environment in the non-native range. **Figure S9:** Principal component analysis (PCA) of 19 WorldClim bioclimatic variables (BIO1–BIO19; see Table S3), retrieved from WorldClim 2.1 (Fick and Hijmans 2017) at a 30-arc-second resolution (~0.86 km2 per grid cell at the equator), used to characterize macroclimatic conditions across the native and non-native ranges of *Erigeron canadensis*. The plot illustrates correlations among bioclimatic variables based on their loadings on the first two principal components. Arrows indicate the strength and direction of correlations. The first axis (PC1; 36.1% of variance) primarily reflects temperature seasonality and continentality, showing positive loadings for temperature seasonality (BIO4) and annual temperature range (BIO7) and negative loadings for minimum temperature of the coldest month (BIO6) and mean temperature of the coldest quarter (BIO11). The second axis (PC2; 25.2% of variance) is associated with moisture- and warm-season climate variation, including precipitation of the warmest quarter (BIO18), precipitation of the driest quarter (BIO14), and mean temperature of the warmest quarter (BIO17). **Figure S10:** Maps showing population-level mean values of the first (PC1; panel A) and second (PC2; panel B) principal components derived from a PCA of the 19 WorldClim bioclimatic variables (Figure S9). Points represent sampling populations with colours indicating PC scores. PC1 primarily reflects large-scale gradients in temperature seasonality and continentality, whereas PC2 captures variation associated with moisture and warm-season climatic conditions. **Figure S11:** Geographic distribution of genomic offset in non-native *Erigeron canadensis* populations. Darker points indicate higher offset values, suggesting greater genetic mismatch and potential maladaptation in certain regions. Surrounding histograms depict the null distributions of mean genomic offset for each region (based on 1000 permutations), with vertical red lines representing the observed regional means. Regions where the red line lies beyond the null distribution indicate cases where observed values exceed null expectations, suggesting potential signals of local adaptation or maladaptation.