












Genomic Footprints of Selfing, Introduction History, and Long-Distance Dispersal in an Invasive Alien Plant

Raúl Sánchez-García ^{1,*}, Andy J. Green ^{1,2}, María A. Ortiz ³, Cristina García ⁴, Francisco Hortas ⁵, Chevonne Reynolds ⁶, Jennifer Rowntree ⁷, Ester A. Serrão ⁸, Lina Tomasson⁹, Karin Tremetsberger ¹⁰, Casper H.A. van Leeuwen ¹¹, Joaquín Ortego ¹²

¹Departamento de Biología de la Conservación y Cambio Global, Estación Biológica de Doñana (EBD), CSIC, Seville, Spain

²Department of Natural Sciences, Manchester Metropolitan University, Manchester, UK

³Department of Plant Biology and Ecology, University of Sevilla, Seville, Spain

⁴Department of Biological Sciences, School of Life and Environmental Sciences, Royal Holloway University of London, Egham, Surrey, UK

⁵Department of Biology, Institute of Marine Research (INMAR), University of Cadiz, Puerto Real, Spain

⁶School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa

⁷School of Biological and Marine Sciences, University of Plymouth, Drake Circus, Plymouth, UK

⁸Centre of Marine Sciences, CCMAR, University of Algarve, Campus de Gambelas, Faro, Portugal

⁹National Coordinator for Aquatic Invasive Alien Species, Swedish Agency for Marine and Water Management, Gothenburg, Sweden

¹⁰Department of Ecosystem Management, Climate and Biodiversity, BOKU University, Institute of Botany, Vienna, Austria

¹¹Department of Ecology, Radboud Institute for Biological and Environmental Sciences, Radboud University, Nijmegen, the Netherlands

¹²Departamento de Ecología y Evolución, Estación Biológica de Doñana (EBD), CSIC, Seville, Spain

*Corresponding author: E-mails: raul.sangar98@gmail.com, raul.sanchez@ebd.csic.es

Accepted: May 15, 2026

Abstract

Biological invasions are natural experiments for studying the evolutionary and ecological processes underlying colonization success and range expansion. Using genome-wide data—generated via genotyping-by-sequencing (GBS) from 30 populations spanning Europe and South Africa—we investigated the colonization history and successful spread of the invasive buttonweed *Cotula coronopifolia*, an annual plant introduced into Europe from South Africa about 300 years ago. Our analyses identified three major lineages in Europe distributed across the continent, often co-occurring without evidence of admixture. Phylogenomic dating revealed that these lineages diverged > 2,000 years ago—well before the earliest European records—suggesting divergence within the native range and either multiple introductions or a single introduction with multiple lineages. Mating-system inference shows that reproduction occurs primarily via self-fertilization (~70% on average), although outcrossing predominates in some populations, revealing a facultative mating system. This high selfing rate has led to extremely low heterozygosity in most populations and a strong genetic structure. Genetic clustering also revealed admixed individuals resulting from rare inter-lineage outcrossing; comparisons of empirical and simulated data indicate that increased diversity after sporadic admixture

events decays rapidly under subsequent selfing. Both the introduction history and long-distance dispersal facilitated by waterbirds likely explain the scattered distribution of lineages across Europe. Altogether, these results illustrate Baker's "ideal weed" concept, highlighting the role of a flexible mating system in providing reproductive assurance during colonization and showing how predominant selfing can shape the genomic landscape of an invasive species.

Key words: buttonweed, *Cotula coronopifolia*, genomics of invasions, population genomics, seed dispersal, selfing.

Significance

Understanding how reproductive strategies influence colonization and spread of alien species is central to invasion biology. By combining population genomics and phylogenomic inference, this study provides key insights into the colonization history and successful invasion of the self-fertilizing plant *C. coronopifolia*, introduced from South Africa to Europe by the 18th century. We show that the species' invasion success relies on the introduction of multiple lineages and the predominance of selfing, which has drastically reduced genetic diversity yet contributed to reproductive assurance and spread across diverse habitats. Occasional outcrossing and long-distance dispersal by waterbirds or through horticultural transport have further shaped the species' genetic landscape. These findings are consistent with Baker's "ideal weed" concept, highlighting how self-compatibility and ecological generalism can help overcome genetic constraints during range expansion and providing a framework for understanding the evolutionary dynamics of selfing plant invasions.

Introduction

Biological invasions have long captured scientific attention, as they often pose a direct threat to biodiversity and the conservation of ecosystems (Simberloff et al. 2013; Pyšek et al. 2020). In addition to their ecological and socioeconomic impacts, invasions also serve as natural experiments in evolution, providing opportunities to study how species reach distant locations, spread and establish outside their native range, and adapt to new environments (Bock et al. 2015). Advances in next-generation sequencing (NGS) technologies have further expanded the scope of invasion genetics, allowing us to uncover the demographic and evolutionary processes underpinning invasion success across diverse taxa (Rius et al. 2015).

A common assumption in invasion biology is that high genetic variability is critical for adaptive success, as it provides the raw material for evolution and enables populations to respond to selective pressures in new environments (Dlugosch et al. 2015). However, recently established populations of invasive species often experience reductions in genetic diversity due to founder effects and demographic bottlenecks, creating an apparent contradiction between theoretical expectations and empirical observations (Schieber and Lachmuth 2017). This inconsistency is known as the "genetic paradox of invasion" (Allendorf and Lundquist 2003). Several non-exclusive mechanisms have been proposed to resolve this paradox, including phenotypic plasticity, multiple introductions leading to admixture, self-fertilization, clonal or asexual

reproduction, and hybridization (Liu et al. 2006; Roman and Darling 2007; Hao et al. 2011; Rius and Darling 2014; Williams et al. 2014). One scenario in which the genetic paradox is particularly relevant is described by Baker's law, which states that colonization success of remote regions is more likely in self-compatible organisms, as a single propagule would be capable of establishing a sexually reproducing colony (Baker 1955). While this strategy facilitates colonization, it may also constrain genetic diversity, creating a potential trade-off between short-term establishment success and long-term evolutionary potential (Pannell et al. 2015). In plants, the interplay between self-compatibility, clonal reproduction, and phenotypic plasticity can therefore strongly influence invasion outcomes (Rollins et al. 2013).

The buttonweed *Cotula coronopifolia* (Asteraceae) provides an excellent model system to explore these processes. This annual or short-lived perennial herb inhabits a variety of wet to seasonally wet habitats, ranging from freshwater to brackish wetlands. It exhibits dual reproductive strategies: asexual vegetative propagation via rooting nodes and sexual outcrossing or self-fertilization through hermaphroditic florets (van der Toorn 1980). Originally native to South Africa, *C. coronopifolia* has spread globally and is now established across Europe, the Americas, Australia, and other regions (van der Toorn 1980; Marfella et al. 2023) (Fig. 1). In Europe, the species was cultivated in an Edinburgh garden by 1683 (Stroh et al. 2023). In the wild, it was first recorded in Emden, Germany (~1739), followed by the

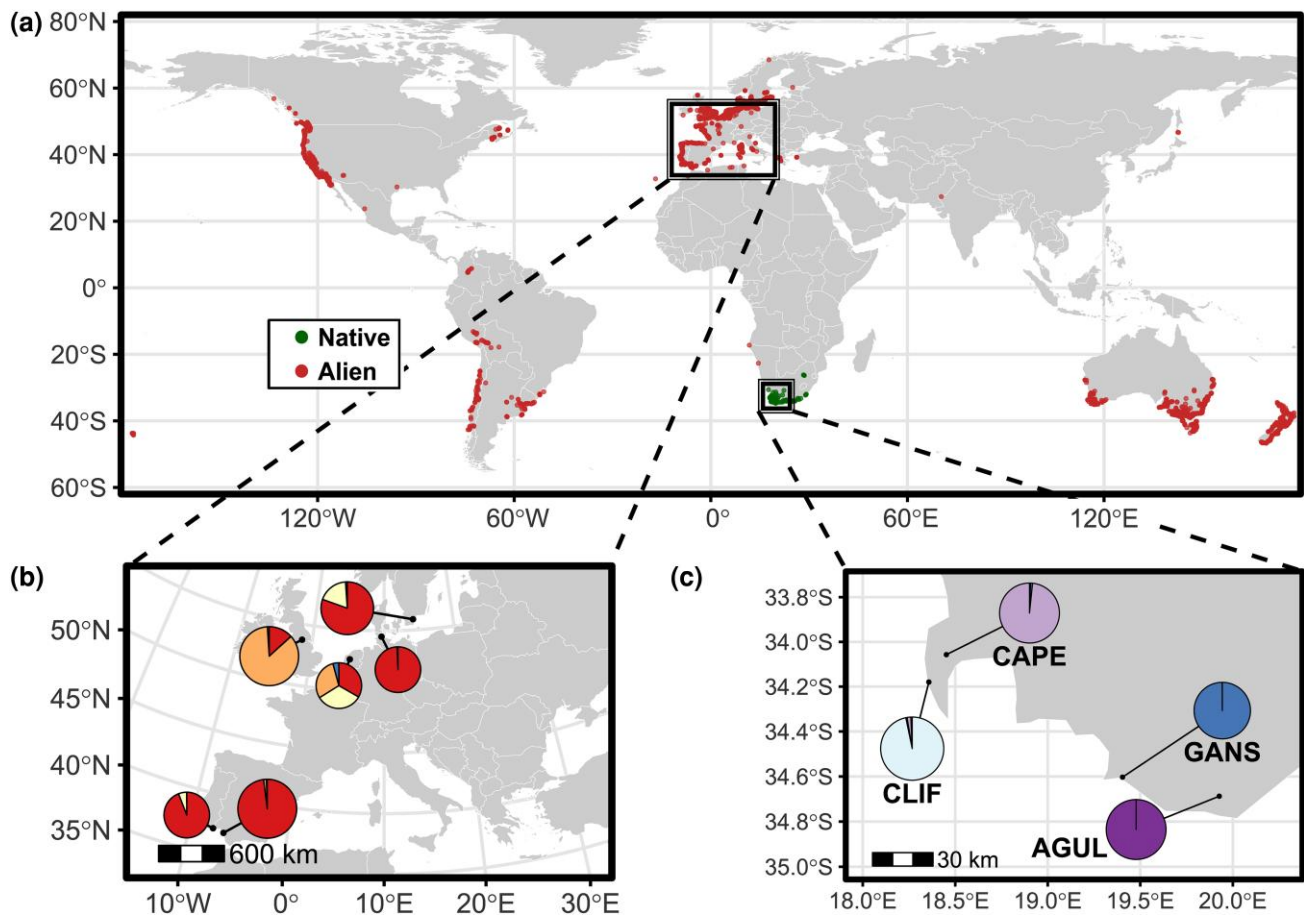


Fig. 1. (a) Occurrence records (18,646 observations) of *C. coronopifolia* in its native and alien ranges (Global Biodiversity Information Facility, GBIF, 2015 to 2025). Records were retrieved via the *rgbif* R package; occurrences within South Africa were classified as native, whereas all others were classified as alien. Only unique, georeferenced presence records passing basic quality filters are shown. Panels (b) and (c) show the results of genetic assignments (pie charts) based on the Bayesian clustering analyses implemented in *STRUCTURE* for (b) European (5,693 SNPs) and (c) South African (13,079 SNPs) populations. For European populations, pie charts represent the mean probability of assignment to each genetic cluster for samples from each country.

Netherlands (~1742), Sweden (~1853), and later Spain (~1886) and the United Kingdom (~1869) (Ridley 1930; van der Toorn 1980; Sánchez-Gullón and García-Murillo 2004; Wilkinson 2023). The species can exclude native plants and continues to expand in Europe, along the Baltic coast and into newly created wetlands (Tomasson 2020). Early introductions may have occurred via shipping, potentially through ballast water transport between South Africa and Europe, or through seed transport in wool (van der Toorn 1980; Verloove and Gonggrijp 2025). Additionally, it may also have spread from gardens (Wilkinson 2023), and is currently sold in Europe as a pond plant. However, its current distribution, spanning both coastal and inland wetlands, is difficult to explain solely by human-mediated dispersal. Seeds can disperse on water, but field observations have long suggested that waterbirds

may contribute to its active dispersal (van der Toorn 1980). The role of migratory birds as competent dispersal vectors has now been empirically demonstrated, with studies confirming that *C. coronopifolia* seeds survive gut passage and that their germination is enhanced thereafter, which has important implications for understanding the species' invasion ecology and spread in its introduced range (Raulings et al. 2011; Lovas-Kiss et al. 2019; Sánchez-García et al. 2024).

Several characteristics of *C. coronopifolia* align directly with mechanisms proposed to resolve the genetic paradox of invasion. Its capacity for self-fertilization fits the predictions of Baker's law, enabling establishment from single propagules in novel habitats, while vegetative propagation ensures the persistence and local spread of successful genotypes, even in the absence of high genetic diversity. These traits are consistent with

Baker's (1965) concept of the "ideal weed" and the "general-purpose genotype," which emphasize broad ecological tolerance, reproductive assurance, and the ability to survive under a wide range of environmental conditions. In *C. coronopifolia*, tolerance to variable hydrological and salinity regimes suggests that phenotypic plasticity contributes to its invasive potential (Costa et al. 2009; Marfella et al. 2023). Finally, its capacity for long-distance dispersal by waterbirds (Lovas-Kiss et al. 2019; Sánchez-García et al. 2024) increases the likelihood of colonization events, reinforcing its status as a species with a "general-purpose genotype" capable of colonizing and persisting in diverse habitats.

Despite its nearly cosmopolitan distribution and early introduction into Europe, surprisingly little is known about the invasion history of *C. coronopifolia*. Most previous research has focused on its ecology and morphology (e.g. Costa et al. 2009; Marfella et al. 2023), while key genetic and demographic aspects underlying its invasive success remain unexplored. To fill these gaps, we sampled and obtained genome-wide data from populations of *C. coronopifolia* across both its native range in South Africa and its invasive range in Europe. Specifically, we address three global questions: (i) What does genomic evidence reveal about the colonization history of *C. coronopifolia* in Europe, particularly regarding the number, origin, and timing of introduction events that shaped its current genetic makeup; (ii) Does *C. coronopifolia* exhibit reproductive traits traditionally associated with Baker's "ideal weed" concept—specifically self-compatibility and high selfing rates—that may have facilitated its establishment and expansion in the invasive range?; and (iii) What genetic footprints have colonization history, spread by secondary dispersal, and reproductive mode left on European populations? By integrating phylogenomic inference, demographic reconstructions, estimates of selfing rates, and quantification of population genetic structure, diversity and admixture, our study provides a comprehensive population-genomic assessment of the colonization and spread of a facultatively selfing species, contributing to the open debate on how reproductive assurance and ecological generalism interact to shape invasion success despite extraordinarily low levels of genetic diversity.

Results

Genomic Datasets

We obtained genotyping by sequencing (GBS) data from 266 plants sampled across 30 populations, including 26 from Europe and four from South Africa (Table 1).

The average number of reads per individual was 9,253,942 (range = 528,852 to 34,055,893; median = 10,530,730). After the different filtering steps in IPYRAD v. 0.9.93 (Eaton and Overcast 2020), the dataset including all genotyped individuals contained 6,661 unlinked SNPs represented in >90% of the individuals and an average proportion of missing data per individual of 6.5% (range = 0.6% to 67.4%, median = 1.59%; Fig. S1). For the dataset with European populations, we obtained a total of 5,692 unlinked SNPs with an average of 6.5% missing data (range = 0.6% to 66.3%; median = 1.32%), while the dataset only including South African populations contained 13,079 unlinked SNPs with 3.98% of missing data (range = 0.9% to 11.6%; median = 2.79%). The number of SNPs differed between the dataset containing only populations from Europe and the one including only plants from South Africa, owing to the higher genetic diversity observed in the latter (Table 1; Wilcoxon rank-sum test, $W = 940$; $n = 266$, $P < 0.001$). Although outgroup individuals of *Cotula* sp. yielded a high number of reads (mean = 24,095,551; range = 21,399,119 to 28,644,221; median = 22,760,471), they exhibited a substantially higher proportion of missing data than ingroup samples (mean = 60.5%; range = 56.4% to 67.1%; median = 58.4%). This pattern is consistent with expectations for GBS datasets that include phylogenetically distant taxa, as sequence divergence at restriction sites reduces locus recovery and increases allelic dropout, typically leading to elevated levels of missing data in the outgroup. All pairs of genotyped individuals had negative or very low values of relatedness (ranging from 0.01 to -3.76), which excludes the possibility that we had sampled clones or close relatives (Manichaikul et al. 2010).

Population Genetic Structure and Admixture

STRUCTURE v. 2.3.3 (Pritchard et al. 2000) analyses including all genotyped individuals from both Europe and South Africa (6,661 SNPs) identified $K = 2$ as the most likely number of genetic clusters based on the ΔK criterion, while $\text{LnPr}(X|K)$ reached a plateau at $K = 5$ (Fig. S2a). For $K > 5$, no further genetic structure was detected, yielding "ghost clusters" (i.e. clusters with no population or individual assigned to them; Guillot et al. 2005). For $K = 2$, European individuals split into two main groups with low levels of admixture, whereas all South African populations are admixed between these two genetic clusters (Fig. S3). STRUCTURE analyses for $K = 5$ showed the presence of three genetic clusters in Europe with limited correspondence with the geographical location of populations (Fig. 1b). One cluster (hereafter, RED lineage) was represented in multiple populations from Sweden, Germany, the Netherlands,

Table 1 Locality/lineage, code, latitude, longitude, number of genotyped individuals (n), observed heterozygosity (H_O), self-fertilization rate (s), and its standard deviation (SD) for each sampled population and lineage (RED, ORANGE, and YELLOW) within the alien range as inferred by the Bayesian clustering method implemented in *STRUCTURE*.

| Locality/lineage | Code | Latitude | Longitude | n | H_O | Selfing rate (s) \pm SD |
|--------------------------|---------|-----------|-----------|-----|--------|-------------------------------|
| Nyköping | SU-NYKO | 58.71493 | 17.08754 | 8 | 0.0043 | 0.829 \pm 0.099 |
| Halland | SU-HALL | 57.02229 | 12.33151 | 8 | 0.0045 | 0.722 \pm 0.118 |
| Södsvik | SU-SODV | 57.03497 | 16.92443 | 9 | 0.0013 | 0.777 \pm 0.100 |
| Pulken | SU-PULK | 55.88327 | 14.20601 | 18 | 0.0063 | 0.905 \pm 0.046 |
| Reesholm | DE-REES | 54.5206 | 9.6284 | 9 | 0.0036 | 0.940 \pm 0.069 |
| Marshside | UK-MARS | 53.67522 | -2.98218 | 18 | 0.0028 | 0.472 \pm 0.102 |
| Martin Mere | UK-MEER | 53.62133 | -2.86725 | 9 | 0.0114 | 0.938 \pm 0.069 |
| Fairburn Ings | UK-FAIR | 53.74045 | -1.32074 | 9 | 0.0029 | 0.619 \pm 0.129 |
| St Aidan's | UK-AIDA | 53.74416 | -1.41199 | 9 | 0.0026 | 0.536 \pm 0.133 |
| Hoylake | UK-HOYL | 53.39336 | -3.18603 | 9 | 0.0021 | 0.760 \pm 0.103 |
| Frodsham | UK-FROD | 53.30621 | -2.73925 | 1 | - | - |
| Lincolnshire Pond Plants | UK-POND | 53.44444 | -0.19528 | 1 | - | - |
| Exminster, Devon | UK-DEVO | 50.66458 | -3.47183 | 9 | 0.0024 | 0.382 \pm 0.143 |
| Leeuwarden | NL-LEEU | 53.2145 | 5.8723 | 9 | 0.0056 | 0.887 \pm 0.080 |
| Leihoeek | NL-LEIH | 52.75039 | 4.64611 | 9 | 0.0266 | 0.880 \pm 0.081 |
| Marker Wadden | NL-MARK | 52.58471 | 5.37034 | 9 | 0.0023 | 0.364 \pm 0.220 |
| Ribeira | PT-RIBE | 37.16667 | -7.62044 | 10 | 0.0028 | 0.792 \pm 0.091 |
| Castro Marim | PT-CMAR | 37.23419 | -7.44253 | 9 | 0.0028 | 0.755 \pm 0.104 |
| Odiel | ES-ODIE | 37.23922 | -7.01028 | 16 | 0.0023 | 0.352 \pm 0.112 |
| Marismas del Rocío | ES-ROCI | 37.11986 | -6.48378 | 8 | 0.0014 | 0.758 \pm 0.111 |
| Dehesa de Abajo | ES-DEHE | 37.10925 | -6.41384 | 9 | 0.0011 | 0.233 \pm 0.137 |
| El Martinazo | ES-MART | 37.02806 | -6.43806 | 8 | 0.0016 | 0.662 \pm 0.127 |
| Puente del duque | ES-PDUQ | 36.99444 | -6.44056 | 9 | 0.0025 | 0.693 \pm 0.113 |
| Laguna Dulce | ES-DULC | 36.9811 | -6.48487 | 7 | 0.0016 | 0.567 \pm 0.151 |
| Salinas de Cetina | ES-CADI | 36.57528 | -6.14222 | 9 | 0.0019 | 0.725 \pm 0.109 |
| El Palmar | ES-PALM | 36.23333 | -6.06667 | 1 | - | - |
| Cape Town | ZA-CAPE | -34.05714 | 18.4521 | 9 | 0.0089 | 0.652 \pm 0.124 |
| Misty Cliffs | ZA-CLIF | -34.1795 | 18.35728 | 10 | 0.0079 | 0.985 \pm 0.053 |
| Gansbaai | ZA-GANS | -34.60304 | 19.40398 | 8 | 0.0050 | 0.730 \pm 0.116 |
| Cape Agulhas | ZA-AGUL | -34.68755 | 19.92657 | 9 | 0.1141 | 0.169 \pm 0.146 |
| RED | - | - | - | 131 | 0.0016 | 0.577 \pm 0.038 |
| ORANGE | - | - | - | 63 | 0.0026 | 0.465 \pm 0.054 |
| YELLOW | - | - | - | 14 | 0.0024 | 0.674 \pm 0.086 |

The first two characters of each locality indicate the corresponding country, following the International Organization for Standardization (ISO). Summary statistics for each lineage within the alien range were calculated only including non-admixed individuals (*STRUCTURE* q -values $>$ 0.99).

Spain, Portugal and one population from the UK, another cluster (hereafter, ORANGE lineage) was predominantly present in the UK, plus one population from the Netherlands and one individual from Spain, and a third cluster (hereafter, YELLOW lineage) was represented in one population from each of Sweden and the Netherlands and some admixture in population PT-CMAR from Portugal (Fig. 2b). South African populations consisted of two genetic clusters, one of them represented in the population ZA-AGUL (hereafter, BLUE lineage) and another represented in the populations ZA-CLIF, ZA-GANS and ZA-CAPE.

Some individuals scattered across different European populations showed a certain degree of admixed ancestry between the RED and YELLOW lineages or between the RED lineage and the BLUE lineage, mostly present in the South African population ZA-AGUL. Only the

population PT-CMAR from Portugal showed a considerable degree of admixed ancestry between two lineages (RED \times YELLOW) across most individuals. A commercial individual—acquired in the Lincolnshire Pond Plants Ltd. (Brookenby, UK)—labeled UK-POND was the only specimen showing evidence of admixture among three genetic clusters (ORANGE, YELLOW, and BLUE), suggesting a distinct genetic background compared to European field-sampled populations, likely associated with its horticultural origin. Among South African populations, only ZA-CAPE had an admixed ancestry between one lineage only present in South Africa and another (ORANGE) mostly present in European populations.

STRUCTURE analyses exclusively focused on European populations (5,692 SNPs) identified $K=2$ as the most likely number of genetic clusters according to the ΔK criterion, whereas $\text{LnPr}(X|K)$ reached a plateau at $K=4$

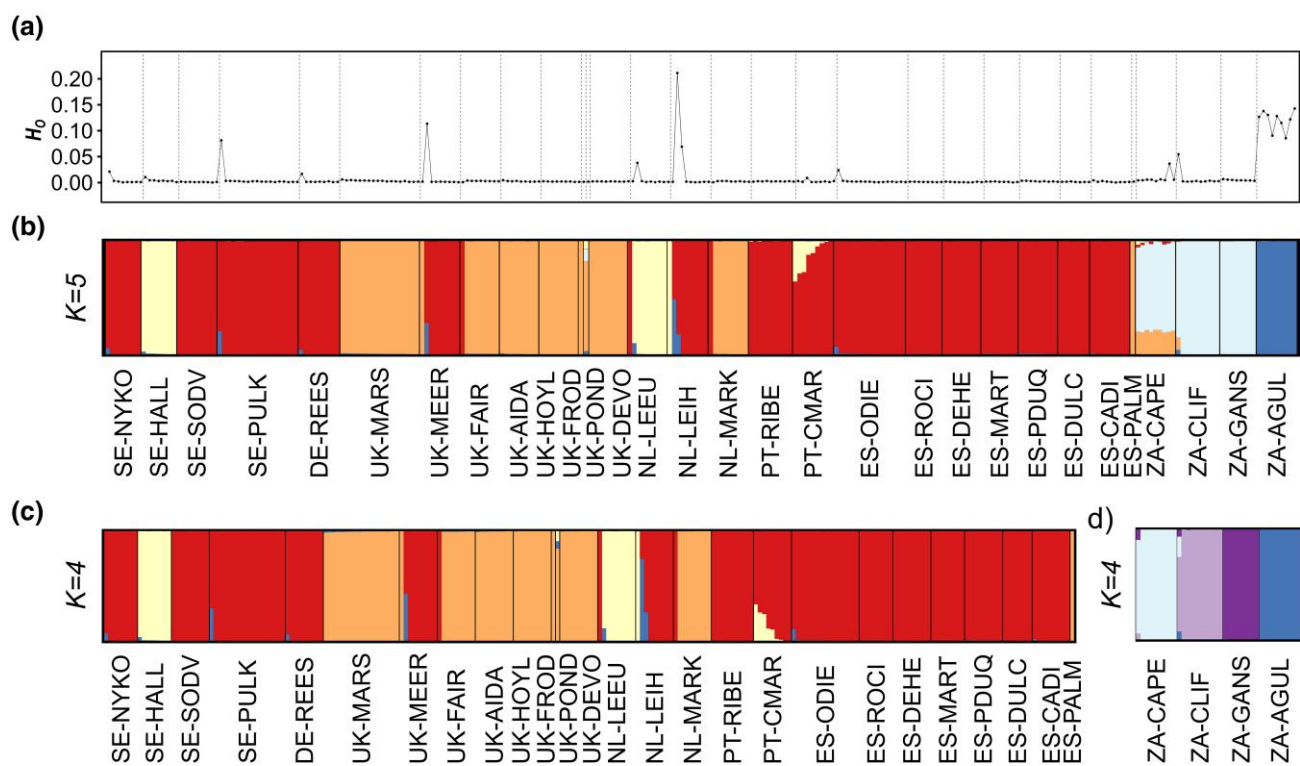


Fig. 2. (a) Observed heterozygosity (H_0) calculated for each individual of *C. coronopifolia*. Panels (b) to (d) show the results of genetic assignments based on the Bayesian clustering analyses implemented in STRUCTURE for datasets containing (b) all 266 individuals (6,661 SNPs), (c) 230 individuals from Europe (5,692 SNPs), and (d) 36 individuals from South Africa (13,079 SNPs). Each individual is represented by a vertical bar, which is partitioned into K coloured segments showing the individual's probability of belonging to the cluster with that colour. Thin vertical black lines separate individuals from different populations. Population codes as described in Table 1.

(Fig. S2b). Clustering solutions for $K = 2$ to 4 are virtually identical to inferences obtained for STRUCTURE analyses, including all genotyped individuals (see Fig. 2b, c, Fig. S4). STRUCTURE analyses focused on South African populations (13,079 SNPs) identified $K = 4$ as the most likely number of genetic clusters according to the ΔK criterion, and $\text{LnPr}(X|K)$ reached a plateau at the same K value (Fig. S2c). For $K = 2$, we can see that the main separation is between ZA-AGUL and the other three populations (Fig. S5). For $K = 4$, each population was assigned to a different genetic cluster, with very limited admixture in one individual from ZA-CAPE and another from ZA-CLIF (Fig. 2d).

ADMIXTURE v. 1.3.0 (Alexander et al. 2009) yielded results nearly identical to those obtained with STRUCTURE (Fig. S6). The only notable difference occurred at $K = 5$, where the ZA-GANS lineage split before the YELLOW lineage (which separated at $K = 6$). Evaluation of model fit via EVALADMIX (García-Erill and Albrechtsen 2020) supported $K = 7$ as the most informative clustering solution, showing the strongest reduction in non-zero residual correlations, whereas higher K values did not result in substantial further improvement (Fig. S6).

The degree of admixed ancestry of individuals—estimated as the standard deviation of population membership probabilities across genetic clusters inferred by STRUCTURE and ADMIXTURE—was not correlated with the proportion of missing data (Pearson's correlation tests, all $r < 0.1$, all $P > 0.7$).

Principal component analyses (PCA) of genetic variation were congruent with clustering results, revealing patterns of genetic differentiation among populations and individuals consistent with those inferred by STRUCTURE and ADMIXTURE (Fig. S7).

Phylogenomic Analyses and Estimation of Divergence Time

Analyses A01 in BPP v. 4.1 (Flouri et al. 2018) recovered the five individuals of *Cotula* sp. as the basal lineage, confirming their suitability as an outgroup for SVDQUARTETS analyses (Chifman and Kubatko 2014) (Fig. S8). Phylogenomic reconstructions obtained with SVDQUARTETS and BPP showed slightly different, yet qualitatively congruent, relationships among lineages. Both approaches supported that the South African population

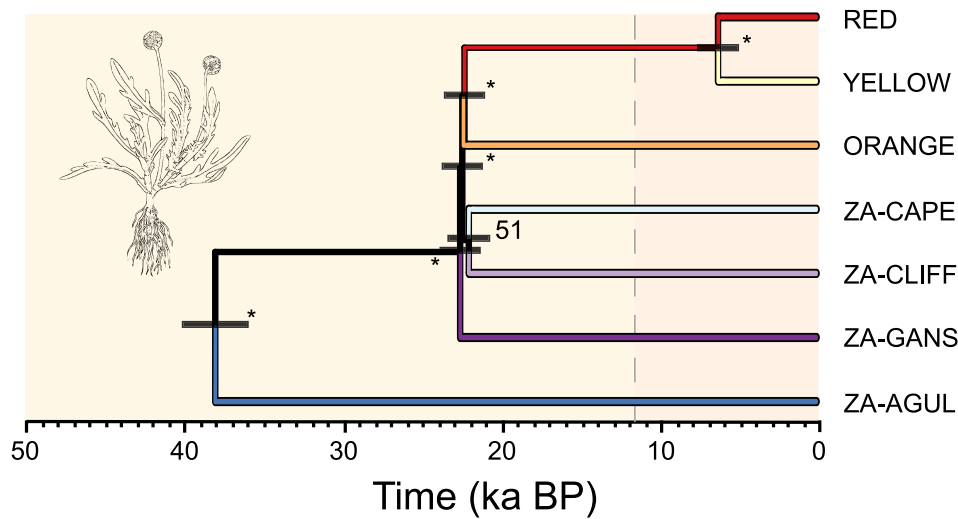


Fig. 3. Phylogenetic tree inferred with *SVDQUARTETS* and divergence times estimated using *BPP* (analysis A00) for lineages of *C. coronopifolia* identified by the Bayesian clustering analyses implemented in *STRUCTURE* in the European invasive range (lineages RED, ORANGE and YELLOW) and the South African native range (populations ZA-CAPE, ZA-CLIFF, ZA-GANS, and ZA-AGUL). Bars on nodes indicate 95% highest posterior densities (HPD) intervals of divergence times estimated considering a genomic mutation rate of 7.0×10^{-9} substitutions per site per generation and a one-year generation time; numbers at nodes indicate bootstrap support values (* = 100). Branch colors correspond to the genetic clusters inferred by *STRUCTURE* (see Fig. 2). Background shading indicates geological divisions of the Quaternary, and the vertical dashed line marks the boundary between the Late Pleistocene (left) and the Holocene (right). Illustration of *C. coronopifolia*: Marina Trillo.

ZA-AGUL was the earliest diverging lineage, splitting from a clade that includes all remaining lineages during the last glacial period (~40 ka BP) (Fig. 3, Figs. S8 and S9). Subsequently, this clade appears to have undergone a rapid diversification around the Last Glacial Maximum (~23 ka), giving rise to five lineages: three South African lineages (ZA-GANS, ZA-CLIFF, and ZA-CAPE) and two European lineages (ORANGE and RED + YELLOW) (Fig. 3, Figs. S8 and S9). Finally, the RED and YELLOW lineages split later, during the Holocene (~6 ka BP) (Fig. 3, Figs. S8 and S9). With the exception of the sister relationship between the two most recently diverged lineages (RED and YELLOW), which received strong support in both *BPP* and *SVDQUARTETS*, the relationships among the remaining lineages within this second major clade were weakly and inconsistently supported across methods (Fig. S8). This pattern likely reflects a hard polytomy resulting from near-simultaneous divergence events, leaving limited phylogenetic signal to confidently resolve branching order. Estimates of divergence times (analyses A00) were highly consistent, regardless of whether the *SVDQUARTETS* or the *BPP* topology was used as the fixed species tree in the dating analyses (Fig. 3, Fig. S9).

Demographic History

STAIRWAY PLOT v. 2.1 analyses suggested that the three European lineages experienced broadly similar demographic trajectories, characterized by an inferred

increase in effective population size (N_e) predating the documented arrival of the species to Europe, followed by demographic stagnation until the present (Fig. 4a to c). In contrast, *STAIRWAY PLOT* suggests that most South African populations have experienced independent demographic trajectories (Fig. 4d to g). The population ZA-CAPE went through a pronounced bottleneck approximately 2 ka BP, reducing its N_e by >90% relative to pre-bottleneck levels (Fig. 4d). The populations ZA-CLIFF and ZA-GANS experienced a sudden increase of N_e during the Holocene (~8 ka and 2 ka BP, respectively), followed by demographic stability (Fig. 4e, f). Finally, ZA-AGUL experienced a gradual decline in N_e since the onset of the Holocene until the present day (Fig. 4g).

Estimates of Population Genetic Diversity and Selfing Rates

Genetic diversity was very low across most genotyped individuals and populations (Table 1 and Fig. 2). The only exceptions were some admixed individuals and the native ZA-AGUL population, which exhibited markedly higher levels of genetic diversity (Table 1 and Fig. 2a). When excluding these, genetic diversity—estimated as observed heterozygosity (H_o)—was significantly higher in populations from the native range than in those from the invasive range in Europe (Wilcoxon rank-sum test, $W=651$; $n=208$, 26; $P<0.001$).

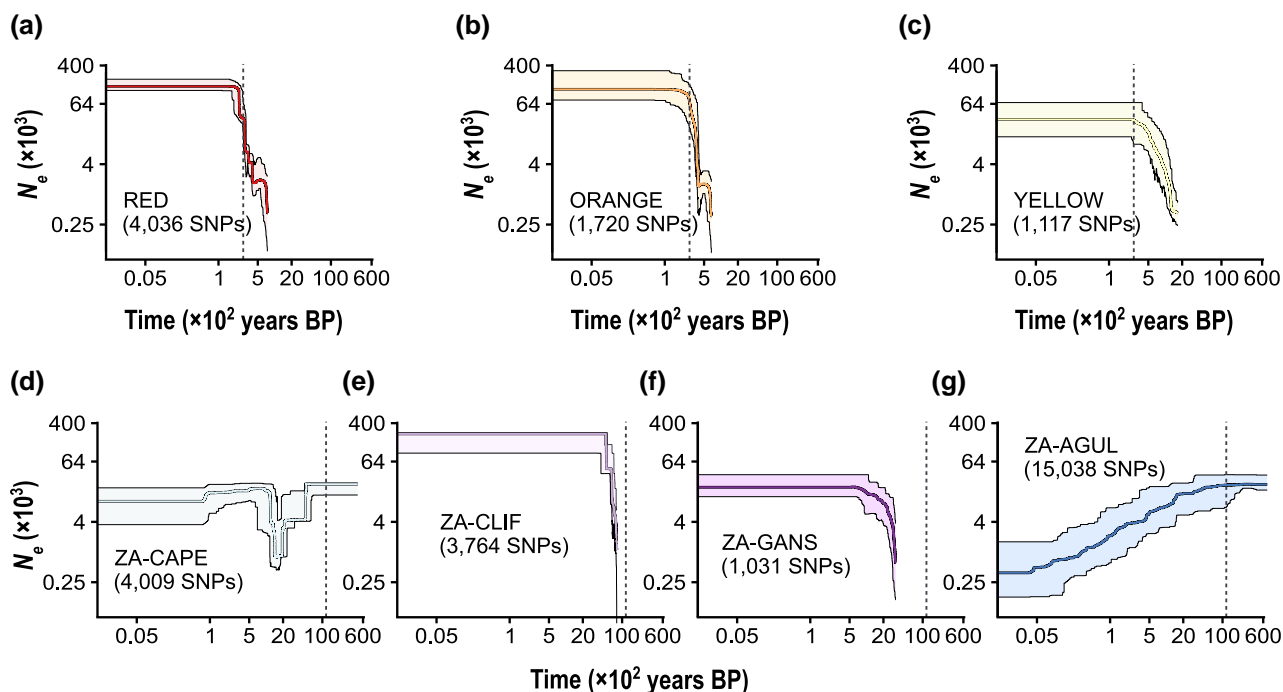


Fig. 4. Demographic reconstructions inferred using STAIRWAY PLOT for each lineage of *C. coronopifolia* identified by Bayesian clustering analyses implemented in STRUCTURE for the European invasive range (lineages RED, ORANGE and YELLOW) and the South African native range (populations ZA-CAPE, ZA-CLIFF, ZA-GANS, and ZA-AGUL) (see Fig. 2). Plots show median (solid lines) and 2.5 and 97.5 percentiles (shaded areas) of effective population size (N_e) through time, estimated assuming a genomic mutation rate of 7.0×10^{-9} per site per generation and one generation per year (both axes in logarithmic scale). Vertical dashed lines indicate the first record of the presence of the species in Europe (1739, 286 years ago) for the European lineages and the onset of the Holocene ($\sim 11,650$ BP) for the South African populations.

When admixed individuals were excluded, levels of genetic diversity differed significantly among the three main lineages identified in the invasive range (Kruskal–Wallis rank sum test, $\chi^2 = 55.43$; $df = 2$; $n = 131, 63, 14$; $P < 0.001$). Post hoc pairwise Wilcoxon rank-sum tests indicated that individuals assigned to lineage ORANGE had significantly higher genetic diversity than those in lineage RED ($P < 0.001$) but did not differ significantly from those in lineage YELLOW ($P = 0.402$). Genetic diversity did not differ significantly between the RED and YELLOW lineages ($P = 0.083$).

Admixed individuals, especially those involving the RED and BLUE lineages, had higher levels of genetic diversity compared to non-admixed individuals assigned to either of these two genetic clusters (Wilcoxon rank-sum test, $W = 949$; $n_1 = 134$, $n_2 = 9$; $P < 0.001$). However, individuals from population PT-CMAR, which exhibits considerable genetic admixture between the RED and YELLOW lineages, did not show higher levels of genetic diversity compared to non-admixed European individuals assigned to the same genetic clusters (Wilcoxon rank-sum test, $W = 772$; $n = 131, 14$; $P = 0.426$).

Selfing rates (s) estimated using RMES (David et al. 2007; Miller and Coltman 2014) were significantly

different from zero for most sampled populations ($H_O = s = 0$; $P < 0.001$), with the exception of NL-MARK ($P = 0.221$) and ZA-AGUL ($P = 0.149$) (Table 1). Selfing rates ranged from 23.3% in ES-DEHE to 98.5% in ZA-CLIFF, indicating a high prevalence of selfing across most populations (Table 1). RMES analyses conducted by pooling all non-admixed (STRUCTURE q -values > 0.999) individuals within each lineage inferred by STRUCTURE in Europe also revealed high selfing rates (RED: 57.7%; ORANGE: 46.5%; YELLOW: 67.4%; Table 1).

Hybrid Simulations for the Determination of Recency of Admixture Between Lineages

To distinguish between recent and historical hybridization events among the main lineage pairs for which STRUCTURE analyses detected admixed individuals in the European invasive range (RED–YELLOW and RED–BLUE; Fig. 2), we compared the distribution of ancestry coefficients (q -values) and observed heterozygosity (H_O) with expectations derived from simulated hybrid classes (Fig. 5a). Simulations of crosses between lineages RED and YELLOW indicate that only one individual from population PT-CMAR exhibited levels of genetic diversity and admixture consistent with a recent backcross

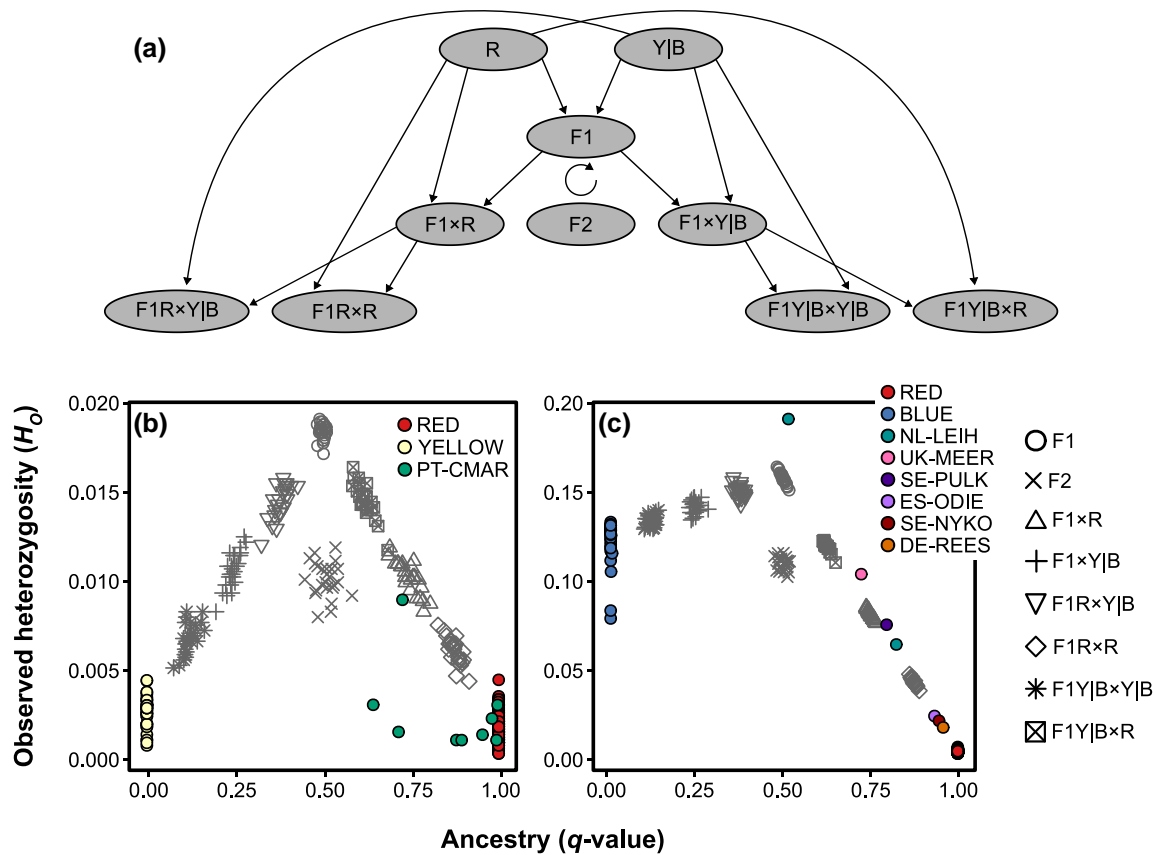


Fig. 5. Simulation of different hybrid classes for crosses between lineages and their expected levels of genetic diversity, estimated as observed heterozygosity (H_o). Panel (a) illustrates a schematic representation of simulated crosses between lineages RED (R) and YELLOW (Y), and between lineages RED (R) and BLUE (B). These correspond to the two lineage pairs for which *STRUCTURE* analyses identified individuals with admixed ancestry among empirical samples in the European invasive range. The vertical bar (|) indicates whether the lineage YELLOW or BLUE was involved in a given outcrossing event (denoted with “x”) with lineage RED. Simulations include first-generation (F1) and second-generation (F2) hybrids, as well as first-generation (F1xR and F1xY|B) and second-generation (F1R x R, F1R x Y|B, F1Y|B x Y|B, and F1Y|B x R) backcrosses. Panels (b) and (c) show observed heterozygosity (H_o) plotted against ancestry proportions inferred by *STRUCTURE* for empirical (colored dots) and simulated (gray symbols) genotypes in crosses between (b) RED and YELLOW and (c) RED and BLUE lineages. Pure individuals of each lineage (RED, YELLOW, and BLUE) from the empirical dataset are shown, and admixed individuals are assigned to their respective populations using distinct color codes. The symmetric distribution of simulated hybrid classes in panel (b) reflects the similar levels of H_o observed in the pure RED and pure YELLOW lineages. In contrast, the asymmetric distribution of simulated classes in panel (c) is driven by significant differences in genetic diversity of parental lineages, with the BLUE lineage exhibiting substantially higher H_o than the RED lineage. Notably, some individuals from population PT-CMAR show levels of H_o considerably lower than expected based on their inferred ancestry proportions, suggesting that they originated from historical hybridization events followed by several generations of selfing.

between an F1 hybrid and the parental RED lineage (Fig. 5b). In contrast, the remaining individuals from PT-CMAR showed substantially lower H_o than expected for any simulated hybrid class with comparable ancestry proportions (Fig. 5b). This pattern suggests that these individuals likely originated from older hybridization events followed by several generations of selfing, which progressively reduced heterozygosity (Fig. 5b). In the case of RED-BLUE crosses, most genotyped individuals displayed levels of H_o consistent with those predicted for simulated hybrid classes with similar ancestry

coefficients (Fig. 5c). This concordance between empirical and simulated expectations supports the interpretation that these individuals resulted from relatively recent hybridization events.

Discussion

By using *C. coronopifolia* as a model system to test Baker’s “ideal weed” hypothesis, our study provides one of the first population-genomic analyses of the colonization history and successful spread of a highly

invasive selfing species. Our results suggest that the species' invasion success in Europe is likely linked to a predominantly selfing mating system. This reproductive strategy, previously suggested but never formally quantified, appears to have contributed to the extremely low within-population diversity and the pronounced genetic differentiation across both native and introduced ranges. These findings are consistent with the predictions of the "ideal weed" framework, which postulates that reproductive assurance through selfing can enhance colonization ability by enabling single individuals to establish viable populations. At the same time, the genetic structure observed among European populations reflects introductions of multiple lineages followed by rapid expansion under selfing, illustrating how reproductive assurance can compensate for low propagule diversity and facilitate large-scale range expansion. Altogether, these results position selfing as a major driver of the genetic landscape of *C. coronopifolia*, which explains observed patterns of genetic diversity and structure and provides important clues to understand the colonization history and successful spread.

Genetic Structure and the Origin of Invasive Lineages

Bayesian clustering analyses revealed a marked genetic structure in both the native and invasive ranges, likely resulting from high selfing rates and extremely limited gene flow among populations and across lineages (Fig. 2). Each of the genotyped populations from the native range of *C. coronopifolia* in South Africa formed a distinct genetic cluster, with very limited signatures of admixture even among populations separated by short geographical distances (e.g. <20 km between ZA-CAPE and ZA-CLIFF; Fig. 1). In the invasive European range, we identified three genetic clusters distributed across different regions of the continent. With a few exceptions, most populations were dominated by individuals belonging to the same genetic cluster. This pattern may reflect a priority effect, whereby first-established lineages reduce the chances of later arrivals to successfully establish, reinforcing the persistence of genetically homogeneous populations and limiting admixture (Fraser et al. 2015).

Phylogenomic reconstructions and divergence time estimates provide important insights into the origin of the European lineages. Dating analyses indicate that ZA-AGUL diverged from other lineages approximately 40 ka BP, while the remaining South African populations and European lineages split ~22 ka BP, with the most recent split, between the RED and YELLOW European lineages, taking place ~6 ka BP. These divergence times (>6 ka) largely predate the first European records of the species by the 18th century, indicating

that the three lineages in the invasive range did not originate through in situ diversification and, instead, most likely colonized Europe after one or more introduction events from unsampled source populations in the native range. The pronounced genetic structure of the species highlights the challenge of identifying the specific ancestral sources of European lineages, a task that would likely require extensive fine-scale sampling across the native range.

Demographic History

Demographic reconstructions in STAIRWAY PLOT show that the four South African populations experienced independent demographic trajectories, which is not surprising given their marked genetic differentiation and distinct reproductive systems (i.e. demographic independence). Although the European lineages were not detected in our limited sampling from South Africa, and direct comparisons of genetic diversity between source and invasive populations are therefore not possible, non-admixed European individuals exhibited significantly lower heterozygosity than non-admixed individuals from predominantly selfing native populations (i.e. ZA-CAPE, ZA-CLIF, and ZA-GANS). This suggests that colonization may have contributed to an additional reduction in genetic diversity in the invasive range. However, because source populations were likely already genetically depauperate, the magnitude of this reduction appears moderate (38.4% relative to native selfing populations), corresponding to a decrease in mean observed heterozygosity (H_O) from 0.005 to 0.002. STAIRWAY PLOT further indicates that the three European lineages experienced relative demographic stability beginning around 300 years ago, roughly coinciding with their introduction to Europe. In highly selfing species, effective population size (N_e) is strongly influenced by reduced recombination and elevated homozygosity, which can generate patterns of apparent demographic stability even when census population sizes (N_c) expand rapidly. Because very low genetic diversity in predominantly selfing species can distort coalescent patterns and may violate the assumptions of demographic models, these reconstructions should be interpreted with extreme caution (Hartfield et al. 2017). Nevertheless, our results are consistent with the hypothesis that selfing may buffer founder effects during colonization, maintaining relatively stable N_e despite rapid spatial invasion. This pattern contrasts with expectations for highly invasive outcrossing organisms, which typically experience strong bottlenecks followed by explosive demographic expansions during the invasion phase (e.g. Ortego et al. 2021).

Introduction and Invasion History

Our data point to multiple introductions and the arrival of at least three different lineages. The three lineages recovered in the invasive range of *C. coronopifolia* in Europe are distributed across different regions in the continent, with limited congruence between genetic structure and geography. This pattern likely reflects the interplay between introduction history and subsequent expansion facilitated by long-distance dispersal (LDD). The earliest record of *C. coronopifolia* in the wild in Europe (1739 in the Netherlands, in a locality that is now in Germany) suggests a possible introduction via shipping (either through ballast water or through wool transport) after the Dutch colonized South Africa in 1652. However, historical records from the UK (D.M., Wilkinson, C.D. Preston & A.J. Green, in prep.) indicate cultivation in several gardens prior to 1739, providing an alternative pathway for introduction. These records also suggest that the species first became firmly established in the wild in the UK in north-west England, originally by 1886 in marshes near Leasowe Castle where it was planted in the garden (Wilkinson 2023), in a locality ca. 4 km from UK-HOYL. By 1933 it had spread 25 km north to a locality near UK-MARS, and by 1959 to UK-FAIR 125 km to the east. The dominance of the ORANGE lineage in these and other UK sites suggests that expansion across the UK has been driven mainly by spread from north-west England. In contrast, the southern Iberian coast is dominated by the RED lineage, likely originating from a population initially reported in Cádiz province in 1886 (Sánchez-Gullón and García-Murillo 2004). This lineage is also dominant in Denmark and in eastern Sweden, where it is currently expanding rapidly (Tomasson 2020). It is dominant at one UK site (UK-MEER), which notably was only created in the 1970s, specifically to attract migratory waterbirds. The third, YELLOW lineage is only dominant in populations from the Netherlands and Western Sweden, consistent with a separate introduction into the Netherlands. Indeed, our three Dutch populations were each dominated by a different lineage, making this region uniquely diverse and a possible hub for introductions. Recently, Verloove and Gonggrijp (2025) speculated that this plant may have colonized Europe naturally from Africa, with a south-to-north colonization pathway (e.g. from the Iberian Peninsula to Scandinavia), in which case we would expect to find all lineages in southern Europe and only a subset in more recently colonized northern regions. However, neither our data nor historical records support this scenario.

The expansion of the species in recent decades may also have been reinforced by its ongoing sale as an

ornamental pond plant. However, we found that the commercial plant that we genotyped (UK-POND) showed an admixed ancestry involving three lineages, in contrast with field-sampled individuals, which were either fully assigned to a single genetic cluster or exhibited admixture between two of them. This suggests that horticulture has generated admixture in commercially available plants that is not represented in any of the field populations we sampled. Therefore, the deliberate spread of ornamental plants appears unlikely to have played a major role in the recent European expansion of this alien species. In contrast, there is good evidence for the effective dispersal of *C. coronopifolia* by gut passage inside migratory waterbirds (i.e. endozoochory) into wetlands with a broad salinity range (Raulings et al. 2011; Lovas-Kiss et al. 2019; Sánchez-García et al. 2024). Previous studies have found evidence that migratory waterbirds shape genetic connectivity for native plants in Europe (Green et al. 2023), including *Juncus bufonius* populations in sites where we sampled *C. coronopifolia* (Williams-Marland et al. 2026), and it is likely that avian vectors have also played a key role in the invasion success of *C. coronopifolia*. Common flyways make it possible for birds to directly disperse seeds from e.g. Spain to Sweden (Lovas-Kiss et al. 2019). The Netherlands, where the three European lineages are present, acts as a hub for many migratory bird species providing strong connections with all the other European areas (Delany et al. 2009). Accordingly, most of the field sites we sampled hold major concentrations of migratory birds and are protected for that reason. Seed dispersal by floating along currents could also have an important role in colonization along coastlines (Costa et al. 2009), whereas livestock may also carry seeds both externally and through gut passage (Tomasson 2020).

A Predominantly Selfing Mating System

Our results reveal a predominantly selfing mating system in *C. coronopifolia*, with a mean self-fertilization rate of ~70% (ranging from 23.3% to 98.5%; Table 1). As frequently found in other flowering plants, our results indicate predominant but not obligate selfing, with the presence of variable rates of outcrossing across populations (Siol et al. 2008; Whitehead et al. 2018; Jullien et al. 2021). One remarkable exception is the native population ZA-AGUL from South Africa, where selfing was undetectable and which presented much higher levels of genetic diversity compared to the rest of the populations (Table 1; Fig. 1a). Phylogenomic reconstructions and dating analyses showed that the lineage present in this population was the earliest diverged (~40 ka BP; Fig. 4), suggesting an evolutionary

transition to selfing from an ancestral outcrossing or predominantly outcrossing mating system (Ortiz et al. 2007). Accordingly, the evolution of selfing from obligate cross-fertilization is the most frequent reproductive transition in angiosperms and has occurred repeatedly in different lineages over a range of contrasting evolutionary timescales (Stebbins 1957; Barrett 2008; Cutter 2019; e.g. genus *Primula*: Wang et al. 2021; genus *Capsella*: Slotte et al. 2012; Brandvain et al. 2013). In the same vein, the lack of selfing in the European population NL-MARK, belonging to a lineage in which the rest of the populations present high selfing rates, indicates that the reproductive mode is probably highly labile and context dependent (Whitehead et al. 2018). Thus, our results suggest the presence of different reproductive strategies across and within lineages and populations of *C. coronopifolia*, in line with plasticity in mating-system documented in other flowering plants (Whitehead et al. 2018; Suijkerbuijk et al. 2025; e.g. *Arabidopsis lyrata*: Mable and Adam 2007; Willi and Määtänen 2010). Multiple factors could explain contrasting selfing rates across populations and local selection for a particular reproductive strategy, including relative abundances of efficient pollinators or different sources of abiotic, biotic or human-induced stressors, among others (Mable and Adam 2007; Willi and Määtänen 2010; Whitehead et al. 2018; Cutter 2019; Suijkerbuijk et al. 2025). In addition, populations with high-selfing rates are more likely to undergo effective dispersal by birds and colonize new sites via a single seed, which may represent an important evolutionary advantage during range expansions following initial colonization of non-native areas (Baker 1955; Razanajatovo et al. 2019).

The consequences of selfing are evident in observed patterns of genetic diversity and structure across populations of *C. coronopifolia*. Selfing is well known to reduce within population genetic diversity while amplifying genetic differentiation among populations, effects that are patent in our data and mirror previous studies where selfers often show much lower heterozygosity and higher structure relative to outcrossers (Charlesworth and Wright 2001; Glémin et al. 2006; Siol et al. 2008; Glémin and Ronfort 2013; Hartfield et al. 2017). Accordingly, we found consistently very low levels of genetic diversity across all selfing populations, which contrast with the higher diversity in population ZA-AGUL, where outcrossing dominates and selfing is probably absent or very limited (Table 1; Fig. 1a). Genetic diversity in population NL-MARK, where we found no significant selfing rates, was comparable to that of predominantly selfing populations. This suggests that outbreeding had no positive impact on the genetic diversity of this population, likely because it originated

from an already genetically depauperate selfing population and has experienced no gene flow from other populations with different genetic backgrounds.

Selfing also appears to be the main driver of observed genetic structure in *C. coronopifolia*, with three and four distinct lineages in Europe and South Africa, respectively, each largely isolated and with limited admixture. This suggests that selfing has facilitated lineage diversification by limiting gene flow and maintaining distinct genetic clusters over short geographic scales, even allowing their co-existence within certain populations (Loveless and Hamrick 1984; Wright et al. 2013; Hu 2015). Accordingly, our data revealed the co-occurrence of different lineages in several invasive populations in Europe (UK-MEER, UK-FAIR, NL-LEUU, NL-LEIH, and NL-MARK), where individuals assigned to different genetic clusters co-existed without evidence of genetic admixture (Fig. 2).

Although selfing seems to dominate in most populations, our Bayesian clustering analyses also revealed the presence of some admixed individuals resulted from sporadic outcrossing events involving different lineages (Fig. 2b). As expected, many of these individuals presented substantially higher levels of heterozygosity compared to non-admixed individuals (Fig. 2a), illustrating how outcrossing events can rapidly increase individual genetic diversity. However, our data also showed that several individuals with admixed genetic ancestry exhibit heterozygosity levels as low as those of non-admixed individuals from the parental selfing lineages. Simulations further indicated that, although initial admixture can elevate genetic diversity, successive generations of selfing rapidly reduce heterozygosity, eventually restoring values comparable to non-admixed selfing populations. This is the case of PT-CMAR population, which includes individuals with varying levels of admixture between RED and YELLOW lineages (Fig. 2c), yet most of them show levels of heterozygosity as low as those found in non-admixed populations from either parental lineage (Fig. 5b). Although the increased genetic diversity in first generation admixed individuals is often diminished by subsequent selfing, our data also demonstrate effective allele exchange across selfing lineages, potentially enabling adaptive gene flow if introgressed variants provide local fitness advantages (Suarez-Gonzalez et al. 2018; Burgarella et al. 2019; Lewis et al. 2023).

Conclusions

In summary, our population-genomic analyses reveal that the invasion success of *C. coronopifolia* in Europe is probably linked to its colonization history and the predominance of selfing as a reproductive strategy.

The species' genetic makeup—characterized by a strong population structure and limited gene flow—suggests multiple introductions, or at least the introduction of multiple lineages followed by rapid expansion under a selfing regime. This combination of reproductive assurance and dispersal versatility has enabled *C. coronopifolia* to establish and persist across diverse European habitats, despite its reduced genetic diversity. By placing these findings within Baker's framework of the "ideal weed" and the "general-purpose genotype," our study highlights how traits such as self-compatibility, vegetative reproduction, and ecological generalism can overcome the genetic constraints typically associated with colonization bottlenecks. Our study exemplifies how mating system evolution can mediate the balance between colonization ability and long-term adaptability during invasion. Future research should expand this framework by incorporating whole-genome sequencing across native and invaded *C. coronopifolia* ranges to determine the genomic basis of alternative mating systems, assess adaptive differentiation among lineages, and explore the evolutionary consequences of selfing, including the purging of deleterious alleles and the development of selfing-syndrome traits (Barrett 2008; Brandvain et al. 2013; Hartfield 2016). Comparing European populations with those from other invaded continents will help disentangle primary versus secondary colonization pathways, whereas investigating the role of migratory waterbirds and livestock as potential dispersal vectors will further clarify how the interplay between natural and anthropogenic processes have shaped the species' global spread.

Materials and Methods

Plant Material and Sampling Sites

From July 2021 to April 2023, we sampled 30 populations of *C. coronopifolia* across its native and alien range in South Africa ($n = 4$) and Europe ($n = 26$), respectively (Table 1). We collected mature plants along transects with a spacing of 1 to 3 m between individuals and preserved them in silica gel until needed for DNA extraction. Additionally—to see if there are any differences between plants collected in the wild and those from horticulture—we acquired a cultivated individual from the Lincolnshire Pond Plants Ltd. (Brookenby, UK). Five individuals of *Cotula* spp. collected in Knysna, Leisure Island (−34.06517, 23.06142), South Africa, were used as an outgroup in phylogenomic analyses. We used a global positioning system (GPS) or Google Earth to record spatial coordinates for each site. Geographical coordinates and other details of sampling sites are indicated in Table 1.

Sequencing and Processing of Genomic Data

We used the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) to extract and purify DNA from leaf tissue, according to the manufacturer's instructions. We determined genomic integrity on 1% agarose gels and quantified DNA concentration using a Qubit Fluorometer (Thermo Fisher Scientific®). Genotyping by sequencing (GBS) libraries were prepared by Ecogenics GmbH (Switzerland) according to Elshire et al. (2011) and using the restriction enzymes *EcoRI* and *MseI*. The resulting PCR products were sequenced on an Illumina NovaSeq platform using an S2 flow cell (2×100 bp).

Demultiplexed sequence data were assembled de novo using default parameters in IPYRAD v. 0.9.93 (Eaton and Overcast 2020). The filtering step was performed by allowing a maximum of five low-quality base calls per read and a minimum read length of 35 bp after adapter trimming. For alignment, we used a clustering threshold of 90%, allowing a maximum of two alleles per site in the consensus sequences. Unless otherwise indicated, all downstream analyses were performed using unlinked (one SNP per locus) single nucleotide polymorphism (SNP) data sets. We used the option *relatedness2* in vcftools v. 0.1.17 to calculate the relatedness among all pairs of genotyped individuals and to exclude the possibility that we had sampled clones ($\phi \sim 0.5$) or close relatives within each study population (Manichaikul et al. 2010; Danecek et al. 2011). To avoid potential biases arising from differences in allele frequencies among genetically differentiated populations, relatedness analyses were conducted separately for each population.

Population Genetic Structure and Admixture

We assessed population genetic structure and admixture using the Bayesian Markov chain Monte Carlo clustering method implemented in the program STRUCTURE v. 2.3.3 (Pritchard et al. 2000). To fully explore population genetic structure and admixture, we initially analyzed the data from all populations jointly and, subsequently, we ran independent analyses exclusively focused on populations from South Africa (native range) and Europe (invasive range). We ran STRUCTURE analyses assuming correlated allele frequencies and admixture and without using prior population information (Hubisz et al. 2009). To estimate the most likely number of genetic clusters, we ran several independent runs for each value of K (from $K = 1$ to $K = 8$) with 200,000 MCMC cycles, following a burn-in step of 100,000 iterations. We retained the 10 runs having the highest likelihood for each value of K . As recommended by Janes et al. (2017), we reported results across multiple K values,

and used two statistics to guide interpretation of the most likely number of genetic clusters: the log probability of the data [$\text{LnPr}(X|K)$; Pritchard et al. 2000] and the ΔK statistic (Evanno et al. 2005), calculated in *STRUCTURE HARVESTER* (Earl and vonHoldt 2012). We used *PONG* v. 1.5 (Behr et al. 2016) to visualize the cluster membership of individuals as bar plots.

To assess the robustness of ancestry estimates inferred with *STRUCTURE*, we performed complementary analyses using the maximum-likelihood approach implemented in *ADMIXTURE* v. 1.3.0 (Alexander et al. 2009). We evaluated K values from 2 to 8 with 20 independent replicates per K and selected the replicate with the highest log-likelihood for downstream analyses. Model fit was further evaluated using *EVALADMIX* (García-Erill and Albrechtsen 2020) by estimating pairwise residual correlations to determine whether inferred ancestry proportions adequately captured the underlying genetic structure. As for *STRUCTURE*, results were visualized using *PONG*.

In addition, we performed principal component analyses (PCA) as implemented in the *R* v. 4.5.0 (*R* Core Team 2025) package *adegenet* (Jombart 2008). Prior to performing PCAs, missing genotypes were imputed using the mean allele frequency across all samples (Jombart 2008).

Phylogenomic Analyses

First, we reconstructed the phylogenetic relationships among lineages using two independent approaches: the multispecies coalescent model (analysis A01) implemented in *BPP* v. 4.1 (Flouri et al. 2018), and the quartet-based method implemented in *SVDQUARTETS* (Chifman and Kubatko 2014). In both analyses, we included five representative individuals for each lineage identified by *STRUCTURE* analyses (*STRUCTURE* q -values > 0.999; see Results section) and used five individuals of *Cotula* spp. as an outgroup. For tree inference under analysis A01 in *BPP*, the *loci* file generated by *IPYRAD* was edited and converted into a *BPP* input file using custom scripts (J. Ortego, <https://github.com/OrtegoLab/ipyrad2bpp>). Due to high computational demands, we retained a subset of 1,000 loci for *BPP* analyses. We specified a uniform prior on rooted trees for the species tree, enabled automatic fine-tuning of parameters, selected the diploid option to indicate that the input sequences were unphased, and adjusted the inverse-gamma priors for θ ($\alpha = 3$, $\beta = 0.04$) and τ ($\alpha = 3$, $\beta = 0.07$) according to empirical estimates based on the number of segregating sites per site (Huang et al. 2020). *BPP* does not require the specification of an outgroup, as the program jointly samples the root position along with the other nodes of the tree. However, we included the five individuals of

Cotula spp. to confirm their suitability as an outgroup for *SVDQUARTETS* analyses, which require an outgroup to root the tree. We performed two independent replicate runs in *BPP* for 200,000 generations, sampling every 2 generations, after a burn-in of 10,000 generations. To reconstruct phylogenomic relationships in *SVDQUARTETS* (Chifman and Kubatko 2014), we used a matrix of unlinked SNPs, exhaustively evaluated all possible quartets, and performed nonparametric bootstrapping with 100 replicates to quantify uncertainty in relationships.

Estimation of Divergence Time

We used analysis A00 in *BPP* to estimate the posterior distribution of divergence times (τ ; Flouri et al. 2018). The phylogenetic tree inferred from (i) either *SVDQUARTETS* or (ii) *BPP* (analyses A01) was specified as the tree prior. For divergence time estimation, the outgroup was excluded from the fixed topology because *BPP* does not require explicit outgroup specification and jointly estimates the root position along with the remaining internal nodes. Moreover, excluding phylogenetically distant taxa minimizes locus dropout and the preferential retention of conserved loci typical of reduced-representation datasets, which may bias parameter estimates, including divergence times (Leaché et al. 2015; Huang and Knowles 2016). For each topology, we conducted two independent replicate runs of 1,000,000 generations, sampling every 2 generations after a burn-in of 100,000 generations using the same dataset and parameters settings described above for tree inference analyses. The estimation of divergence times was conducted using the equation $\tau = 2\mu t$, where τ indicates the divergence in substitutions per site as estimated by *BPP*, μ is the per-site mutation rate per generation, and t is the absolute divergence time in years (Huang et al. 2020). We assumed a mutation rate of 7.0×10^{-9} per site per generation, as estimated for *Arabidopsis thaliana* (Ossowski et al. 2010).

Demographic History

We reconstructed the past demographic history of each lineage using *STAIRWAY PLOT* v. 2.1, an approach which implements a flexible multi-epoch demographic model based on the site-frequency spectrum (SFS) and that does not require whole-genome sequence data or reference genome information (Liu and Fu 2020). We calculated the SFS for each lineage inferred by *STRUCTURE* (see Results and Fig. 2) and ran *STAIRWAY PLOT* considering one generation per year, the mutation rate of 7.0×10^{-9} per site per generation estimated for *Arabidopsis thaliana* (Ossowski et al. 2010), and 200 bootstrap replicates to estimate 95% confidence intervals.

Estimates of Population Genetic Diversity and Selfing Rates

To quantify the genetic diversity, we calculated observed heterozygosity (H_O) for each individual as implemented in the R package *hierfstat* (Goudet 2005). Then, we ran Wilcoxon–Mann–Whitney tests in R to compare H_O values between pure and admixed individuals identified by STRUCTURE analyses (see Results). Selfing rates were calculated for each population (see Table 1) and European lineages inferred by STRUCTURE analyses (STRUCTURE q -values > 0.999) using the robust multilocus estimate of selfing (r_{MES}), with 10,000 iterations to generate the P -values (David et al. 2007; Miller and Coltman 2014). Since r_{MES} software uses the heterozygosity variance across the target population, we generated a separate dataset for each population or lineage. Due to computational limitations, we selected 1,000 random SNPs for each dataset (e.g. Adhikari et al. 2021).

Hybrid Simulations to Determine the Recency of Admixture Between Lineages

STRUCTURE analyses revealed that some individuals present different degrees of admixed ancestry, suggesting either contemporary or historical hybridization between lineages (see Results). More specifically, admixed individuals primarily showed different ancestry proportions involving either two lineages distributed in Europe or one lineage distributed in Europe and another distributed in South Africa (see Results). To evaluate whether admixed individuals match or deviate from genetic diversity and admixture expectations corresponding to different hybrid classes resulting from recent hybridization between lineages (i.e. F1, F2 and different backcrosses), as opposed to historical events, we performed hybrid simulations using HYBRIDLAB (Nielsen et al. 2006; e.g. Salces-Castellanos et al. 2021). We performed independent simulations considering the two pairs of lineages primarily involved in the formation of admixed individuals (RED and YELLOW; RED and BLUE). We selected as reference parental lineages pure individuals with $\geq 99.9\%$ probability of cluster membership, as determined by the STRUCTURE analyses on our empirical dataset (see Results and Fig. 2) and used them to simulate 25 individuals corresponding to each of six different hybrid classes, as illustrated in Fig. 5a. Finally, we ran STRUCTURE analyses for datasets including both empirical and simulated individuals, calculated their admixture coefficients and plotted them against their respective estimates of observed heterozygosity (H_O ; calculated as detailed previously). If admixture is the consequence of recent hybridization between lineages, we would expect that patterns of genetic diversity and admixture of empirical individuals align to those

expected for a particular simulated hybrid class. On the contrary, if observed admixture is the outcome of historical hybridization, we would expect lower levels of genetic diversity than expected from admixture proportions for a given simulated hybrid class, especially if several generations of selfing have taken place after the hybridization event.

Supplementary Material

Supplementary material is available at [Genome Biology and Evolution](#) online.

Acknowledgments

We thank for the sampling permits to the Junta de Andalucía (project 2021/13), English Nature, The Wildfowl & Wetlands Trust, The Royal Society for the Protection of Birds, Instituto da Conservação da Natureza e Florestas, and the competent environmental authorities in and all the Netherlands. We are grateful to Salinas de Cetina for granting permission to access ES-CADI. Logistic support was provided by Laboratorio de Ecología Molecular (LEM-EBD) at Estación Biológica de Doñana (CSIC). We also thank Centro de Supercomputación de Galicia (CESGA) and Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) for access to computational resources. We thank Prof. Glynis Goodman-Cron for the taxonomic clarification of the specimens collected in South Africa, and two anonymous reviewers for their constructive and valuable comments on an earlier version of the manuscript.

Author Contributions

RSG, JO, MAO and AJG conceived and designed the study. RSG, AJG, MAO, FH, CR, JR, EAS, LT, KT and CHAL collected the samples. RSG analysed the data, with input from JO. RSG wrote the first draft of the manuscript. JO and AJG revised the first and later drafts. MAO, CG, EAS, KT and CHAL contributed to later drafts. All authors read and approved the final manuscript.

Funding

This research was funded by the Spanish Ministerio de Ciencia e Innovación (project number PID2020–112774GB-I00/AEI/10.13039/501100011033). RSG was supported by an FPI fellowship from the Spanish Ministerio de Ciencia, Innovación y Universidades (PRE2021-099466). Sampling in Portugal was funded by FCT—Foundation for Science and Technology (UID/04326/2025 and LA/P/0101/2020).

Data Availability

Raw Illumina reads have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1431730. Input files for all analyses are available on Zenodo (<https://doi.org/10.5281/zenodo.18861187>).

Literature Cited

- Adhikari S, Revolinski SR, Eigenbrode SD, Burke IC. Genetic diversity and population structure of a global invader Mayweed chamomile (*Anthemis cotula*): management implications. *AoB Plants*. 2021;13:plab049. <https://doi.org/10.1093/aobpla/plab049>.
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–1664. <https://doi.org/10.1101/gr.094052.109>.
- Allendorf FW, Lundquist LL. Introduction: population biology, evolution, and control of invasive species. *Conserv Biol*. 2003;17:24–30. <https://doi.org/10.1046/j.1523-1739.2003.02365.x>.
- Baker H. The genetics of colonizing species. Academic Press; 1965.
- Baker HG. Self-compatibility and establishment after “long-distance” dispersal. *Evolution*. 1955;9:347–349. <https://doi.org/10.2307/2405656>.
- Barrett SCH. Major evolutionary transitions in flowering plant reproduction: an overview. *Int J Plant Sci*. 2008;169:1–5. <https://doi.org/10.1086/522511>.
- Behr AA, Liu KZ, Liu-Fang G, Nakka P, Ramachandran S. Pong: fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics*. 2016;32:2817–2823. <https://doi.org/10.1093/bioinformatics/btw327>.
- Bock DG, et al. What we still don't know about invasion genetics. *Mol Ecol*. 2015;24:2277–2297. <https://doi.org/10.1111/mec.13032>.
- Brandvain Y, Slotte T, Hazzouri KM, Wright SI, Coop G. Genomic identification of founding haplotypes reveals the history of the selfing species *Capsella rubella*. *PLoS Genet*. 2013;9:e1003754. <https://doi.org/10.1371/journal.pgen.1003754>.
- Burgarella C, et al. Adaptive introgression: an untapped evolutionary mechanism for crop adaptation. *Front Plant Sci*. 2019;10:4. <https://doi.org/10.3389/fpls.2019.00004>.
- Charlesworth D, Wright SI. Breeding systems and genome evolution. *Curr Opin Genet Dev*. 2001;11:685–690. [https://doi.org/10.1016/S0959-437X\(00\)00254-9](https://doi.org/10.1016/S0959-437X(00)00254-9).
- Chifman J, Kubatko L. Quartet inference from SNP data under the coalescent model. *Bioinformatics*. 2014;30:3317–3324. <https://doi.org/10.1093/bioinformatics/btu530>.
- Costa JC, Neto C, Arsénio P, Capelo J. Geographic variation among Iberian communities of the exotic halophyte *Cotula coronopifolia*. *Alpine Bot*. 2009;119:53–61. <https://doi.org/10.1007/s00035-009-0056-2>.
- Cutter AD. Reproductive transitions in plants and animals: selfing syndrome, sexual selection and speciation. *New Phytol*. 2019;224:1080–1094. <https://doi.org/10.1111/nph.16075>.
- Danecek P, et al. The variant call format and vcf tools. *Bioinformatics*. 2011;27:2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- David P, Pujol B, Viard F, Castella V, Goudet J. Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol*. 2007;16:2474–2487. <https://doi.org/10.1111/j.1365-294X.2007.03330.x>.
- Delany S, Scott D, Dodman T, Stroud D, editors. An atlas of wader populations in Africa and western eurasia. Wetlands International; 2009.
- Dlugosch KM, Anderson SR, Braasch J, Cang FA, Gillette HD. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Mol Ecol*. 2015;24:2095–2111. <https://doi.org/10.1111/mec.13183>.
- Earl DA, vonHoldt BM. Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv Genet Resour*. 2012;4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>.
- Eaton DAR, Overcast I. Ipyrad: interactive assembly and analysis of RADseq datasets. *Bioinformatics*. 2020;36:2592–2594. <https://doi.org/10.1093/bioinformatics/btz966>.
- Elshire RJ, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*. 2011;6:e19379. <https://doi.org/10.1371/journal.pone.0019379>.
- Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol*. 2005;14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Flouri T, Jiao X, Rannala B, Yang Z. Species tree inference with bpp using genomic sequences and the multispecies coalescent. *Mol Biol Evol*. 2018;35:2585–2593. <https://doi.org/10.1093/molbev/msy147>.
- Fraser CI, Banks SC, Waters JM. Priority effects can lead to underestimation of dispersal and invasion potential. *Biol Invasions*. 2015;17:1–8. <https://doi.org/10.1007/s10530-014-0714-1>.
- García-Erill G, Albrechtsen A. Evaluation of model fit of inferred admixture proportions. *Mol Ecol Resour*. 2020;20:936–949. <https://doi.org/10.1111/1755-0998.13171>.
- Glémin S, Bazin E, Charlesworth D. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proc Biol Sci*. 2006;273:3011–3019. <https://doi.org/10.1098/rspb.2006.3657>.
- Glémin S, Ronfort J. Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evolution*. 2013;67:225–240. <https://doi.org/10.1111/j.1558-5646.2012.01778.x>.
- Goudet J. *Hierfstat*, a package for r to compute and test hierarchical F-statistics. *Mol Ecol Notes*. 2005;5:184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>.
- Green AJ, et al. Dispersal of aquatic and terrestrial organisms by waterbirds: a review of current knowledge and future priorities. *Freshw Biol*. 2023;68:173–190. <https://doi.org/10.1111/fwb.14038>.
- Guillot G, Estoup A, Mortier F, Cosson JF. A spatial statistical model for landscape genetics. *Genetics*. 2005;170:1261–1280. <https://doi.org/10.1534/genetics.104.033803>.
- Hao JH, Qiang S, Chrobok T, van Kleunen M, Liu QQ. A test of Baker's law: breeding systems of invasive species of Asteraceae in China. *Biol Invasions*. 2011;13:571–580. <https://doi.org/10.1007/s10530-010-9850-4>.
- Hartfield M. Evolutionary genetic consequences of facultative sex and outcrossing. *J Evol Biol*. 2016;29:5–22. <https://doi.org/10.1111/jeb.12770>.
- Hartfield M, Bataillon T, Glémin S. The evolutionary interplay between adaptation and self-fertilization. *Trends Genet*. 2017;33:420–431. <https://doi.org/10.1016/j.tig.2017.04.002>.
- Hu X. Mating system as a barrier to gene flow. *Evolution*. 2015;69:1158–1177. <https://doi.org/10.1111/evo.12660>.
- Huang H, Knowles LL. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Syst Biol*. 2016;65:357–365. <https://doi.org/10.1093/sysbio/syu046>.
- Huang JP, Hill JVG, Ortego J, Knowles LL. Paraphyletic species no more: genomic data resolve a Pleistocene radiation and

- validate morphological species of the *Melanoplus scudderi* complex (Insecta: Orthoptera). *Syst Entomol.* 2020;45:594–605. <https://doi.org/10.1111/syen.12415>.
- Hubisz MJ, Falush D, Stephens M, Pritchard J. Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour.* 2009;9:1322–1332. <https://doi.org/10.1111/men.2009.9.issue-5>.
- Janes JK, et al. The $K=2$ conundrum. *Mol Ecol.* 2017;26:3594–3602. <https://doi.org/10.1111/mec.14187>.
- Jombart T. *Adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics.* 2008;24:1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>.
- Jullien M, Ronfort J, Gay L. How and when does outcrossing occur in the predominantly selfing species *Medicago truncatula*? *Front Plant Sci.* 2021;12:619154. <https://doi.org/10.3389/fpls.2021.619154>.
- Leaché AD, Banbury BL, Felsenstein J, Nieto-Montes de Oca A, Stamatakis A. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Syst Biol.* 2015;64:1032–1047. <https://doi.org/10.1093/sysbio/syv053>.
- Lewis JA, Kandala P, Penley MJ, Morran LT. Gene flow accelerates adaptation to a parasite. *Evolution.* 2023;77:1468–1478. <https://doi.org/10.1093/evolut/qpad048>.
- Liu J, et al. Invasive alien plants in China: role of clonality and geographical origin. *Biol Invasions.* 2006;8:1461–1470. <https://doi.org/10.1007/s10530-005-5838-x>.
- Liu X, Fu YX. Stairway plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biol.* 2020;21:280. <https://doi.org/10.1186/s13059-020-02196-9>.
- Lovas-Kiss Á, et al. Shorebirds as important vectors for plant dispersal in Europe. *Ecography.* 2019;42:956–967. <https://doi.org/10.1111/ecog.04065>.
- Loveless MD, Hamrick JL. Ecological determinants of genetic structure in plant populations. *Annu Rev Ecol Syst.* 1984;15:65–95. <https://doi.org/10.1146/annurev.es.15.110184.000433>.
- Mable BK, Adam A. Patterns of genetic diversity in outcrossing and selfing populations of *Arabidopsis lyrata*. *Mol Ecol.* 2007;16:3565–3580. <https://doi.org/10.1111/j.1365-294X.2007.03416.x>.
- Manichaikul A, et al. Robust relationship inference in genome-wide association studies. *Bioinformatics.* 2010;26:2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>.
- Marfella L, et al. Distribution of the invasive alien species *Cotula coronopifolia* L. (Asteraceae) relating to water salinity and sodicity in the Variconi wetland (Campania, southern Italy). *Hydrobiologia.* 2023;850:1653–1668. <https://doi.org/10.1007/s10750-023-05175-w>.
- Miller JM, Coltman DW. Assessment of identity disequilibrium and its relation to empirical heterozygosity–fitness correlations: a meta-analysis. *Mol Ecol.* 2014;23:1899–1909. <https://doi.org/10.1111/mec.12707>.
- Nielsen EE, Bach LA, Kotlicki P. Hybridlab (version 1.0): a program for generating simulated hybrids from population samples. *Mol Ecol Notes.* 2006;6:971–973. <https://doi.org/10.1111/j.1471-8286.2006.01433.x>.
- Ortego J, Céspedes V, Millán A, Green AJ. Genomic data support multiple introductions and explosive demographic expansions in a highly invasive aquatic insect. *Mol Ecol.* 2021;30:4189–4203. <https://doi.org/10.1111/mec.16050>.
- Ortiz MÁ, Tremetsberger K, Talavera S, Stuessy T, García-Castaño JL. Population structure of *Hypochaeris salzmanniana* DC. (Asteraceae), an endemic species to the Atlantic coast on both sides of the Strait of Gibraltar, in relation to Quaternary sea level changes. *Mol Ecol.* 2007;16:541–552. <https://doi.org/10.1111/j.1365-294X.2006.03157.x>.
- Ossowski S, et al. The Rate and Molecular Spectrum of Spontaneous Mutations in *Arabidopsis thaliana*. *Science.* 2010;327:88–92. <https://doi.org/10.1126/science.1180512>.
- Pannell JR, et al. The scope of Baker's law. *New Phytol.* 2015;208:656–667. <https://doi.org/10.1111/nph.13539>.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155:945–959. <https://doi.org/10.1093/genetics/155.2.945>.
- Pyšek P, et al. Scientists' warning on invasive alien species. *Biol Rev Camb Philos Soc.* 2020;95:1511–1534. <https://doi.org/10.1111/brv.12627>.
- R Core Team. 2025. R: A Language and Environment for Statistical Computing. [Computer software]. Version 4.5.0. R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Raulings E, Morris K, Thompson R, Mac Nally R. Do birds of a feather disperse plants together? *Freshw Biol.* 2011;56:1390–1402. <https://doi.org/10.1111/j.1365-2427.2011.02576.x>.
- Razanajatovo M, et al. Autofertility and self-compatibility moderately benefit island colonization of plants. *Global Ecol Biogeogr.* 2019;28:341–352. <https://doi.org/10.1111/gcb.12854>.
- Ridley HN. The dispersal of plants throughout the world. L. Reeve & Co.; 1930.
- Rius M, Bourne S, Hornsby HG, Chapman MA. Applications of next-generation sequencing to the study of biological invasions. *Curr Zool.* 2015;61:488–504. <https://doi.org/10.1093/czoolo/61.3.488>.
- Rius M, Darling JA. How important is intraspecific genetic admixture to the success of colonising populations? *Trends Ecol Evol.* 2014;29:233–242. <https://doi.org/10.1016/j.tree.2014.02.003>.
- Rollins LA, et al. High genetic diversity is not essential for successful introduction. *Ecol Evol.* 2013;3:4501–4517. <https://doi.org/10.1002/ece3.824>.
- Roman J, Darling JA. Paradox lost: genetic diversity and the success of aquatic invasions. *Trends Ecol Evol.* 2007;22:454–464. <https://doi.org/10.1016/j.tree.2007.07.002>.
- Salces-Castellano A, et al. Long-term cloud forest response to climate warming revealed by insect speciation history. *Evolution.* 2021;75:231–244. <https://doi.org/10.1111/evo.14111>.
- Sánchez-García R, Green AJ, Tomasson L, Hortas F, Ortiz MA. Invasive Buttonweed *Cotula coronopifolia* (Asteraceae) is halotolerant and has high potential for dispersal by endozoochory. *Plants.* 2024;13:2915. <https://doi.org/10.3390/plants13162219>.
- Sánchez-Gullón E, García-Murillo P. *Cotula coronopifolia* L. In: Sanz Elorza M, Dana Sánchez ED, Sobrino Vesperinas E, editors. Atlas de las plantas alóctonas invasoras de España. Ministerio de Medio Ambiente; 2004. p. 134–135. https://www.miteco.gob.es/content/dam/mitesco/es/biodiversidad/temas/inventarios-nacionales/cotula_coronopifolia_tcm30-70081.pdf
- Schrieber K, Lachmuth S. The genetic paradox of invasions revisited: the potential role of inbreeding × environment interactions in invasion success. *Biol Rev Camb Philos Soc.* 2017;92:939–952. <https://doi.org/10.1111/brv.12263>.
- Simberloff D, et al. Impacts of biological invasions: what's what and the way forward. *Trends Ecol Evol.* 2013;28:58–66. <https://doi.org/10.1016/j.tree.2012.07.013>.
- Siol M, Prospero JM, Bonnini I, Ronfort J. How multilocus genotypic pattern helps to understand the history of selfing populations:

- a case study in *Medicago truncatula*. *Heredity* (Edinb). 2008;100:517–525. <https://doi.org/10.1038/hdy.2008.5>.
- Slotte T, Hazzouri KM, Stern D, Andolfatto P, Wright SI. Genetic architecture and adaptive significance of the selfing syndrome in *Capsella*. *Evolution*. 2012;66:1360–1374. <https://doi.org/10.1111/j.1558-5646.2011.01540.x>.
- Stebbins GL. Self-fertilization and population variability in the higher plants. *Am Nat*. 1957;91:337–354. <https://doi.org/10.1086/281999>.
- Stroh PA, Walker KJ, Humphrey TA, Pescott O, Burkmar RJ. *Plant atlas 2020: mapping changes in the distribution of the British and Irish Flora*. Princeton University Press; 2023.
- Suarez-Gonzalez A, Lexer C, Cronk QCB. Adaptive introgression: a plant perspective. *Biol Lett*. 2018;14:20170688. <https://doi.org/10.1098/rsbl.2017.0688>.
- Suijkerbuijk HAC, Ramos SE, Poelman EH. Plasticity in plant mating systems. *Trends Plant Sci*. 2025;30:424–436. <https://doi.org/10.1016/j.tplants.2024.10.013>.
- Tomasson L. *Cotula coronopifolia*—invasive or just another alien species? Master's Thesis. Department of Ecology, Swedish University of Agricultural Sciences; 2020.
- van der Toorn J. On the ecology of *Cotula coronopifolia* L. and *Ranunculus sceleratus* L. 1. Geographic distribution, habitat, and field observations. *Acta Bot. Neerl*. 1980;29:385–396. <https://doi.org/10.1111/j.1438-8677.1980.tb01244.x>.
- Verloove F, Gonggrijp S. *Damasonium polyspermum* (Alismataceae), new to The Netherlands, with thoughts on the nativeness of plant species allegedly introduced by migratory birds. *Gorteria—Dutch Botanical Archives*. 2025;47:10–15. [https://www.friscris.be/en/publications/damasonium-polyspermum-alismataceae-new-to-the-netherlands-with-thoughts-on-the-nativeness-of-plant-species-allegedly-introduced-by-migratory-birds\(16c91406-e808-4988-b9d2-a7a8af246c88\).html](https://www.friscris.be/en/publications/damasonium-polyspermum-alismataceae-new-to-the-netherlands-with-thoughts-on-the-nativeness-of-plant-species-allegedly-introduced-by-migratory-birds(16c91406-e808-4988-b9d2-a7a8af246c88).html).
- Wang XJ, et al. The genomic selfing syndrome accompanies the evolutionary breakdown of heterostyly. *Mol Biol Evol*. 2021;38:168–180. <https://doi.org/10.1093/molbev/msaa199>.
- Whitehead MR, Lanfear R, Mitchell RJ, Karron JD. Plant mating systems often vary widely among populations. *Front Ecol Evol*. 2018;6:38. <https://doi.org/10.3389/fevo.2018.00038>.
- Wilkinson DM. Plants on the wing: waterbirds and plant dispersal in Britain. *Br. Wildl*. 2023;35:167–171.
- Willi Y, Määtänen K. Evolutionary dynamics of mating system shifts in *Arabidopsis lyrata*. *J Evol Biol*. 2010;23:2123–2131. <https://doi.org/10.1111/j.1420-9101.2010.02073.x>.
- Williams WI, Friedman JM, Gaskin JF, Norton AP. Hybridization of an invasive shrub affects tolerance and resistance to defoliation by a biological control agent. *Evol Appl*. 2014;7:381–393. <https://doi.org/10.1111/eva.12134>.
- Williams-Marland B, et al. 2026 Feb 25. All for one or one for all? Disentangling the *Juncus bufonius* complex through morphometrics, cytometry and genomics [preprint]. *bioRxiv*. <https://doi.org/10.64898/2026.02.24.707752>
- Wright SI, Kalisz S, Slotte T. Evolutionary consequences of self-fertilization in plants. *Proc R Soc B Biol Sci*. 2013;280:20130133. <https://doi.org/10.1098/rspb.2013.0133>

Associate editor: Sangeet Lamichhane