

Population genomics of *Ulex europaeus* in the Northern Andes: insights into an invasive species in high-mountain ecosystems

Anneth Díaz-Reyes¹, Natalia Aguirre-Acosta^{1,3,4},
Carolina Feuillet-Hurtado^{1,2}, Ghennie Tatiana Rodríguez-Rey^{1,2}

- 1 Grupo de Investigación Biodiversidad y Recursos Naturales (BIONAT), Universidad de Caldas, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas y Naturales, Manizales, Caldas, Colombia
- 2 Universidad de Caldas, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas y Naturales, Manizales, Caldas, Colombia
- 3 Grupo de Investigación Conservación y Biotecnología, Pontificia Universidad Javeriana, Departamento de Ciencias Naturales y Matemáticas, Facultad de Ingeniería y Ciencias, Cali, Valle del Cauca, Colombia
- 4 Pontificia Universidad Javeriana, Departamento de Ciencias Naturales y Matemáticas, Facultad de Ingeniería y Ciencias, Cali, Valle del Cauca, Colombia

Corresponding author: Ghennie Tatiana Rodríguez-Rey (ghennie.rodriguez@ucaldas.edu.co)

Academic editor: Myriam Heuertz ♦ Received 15 July 2025 ♦ Accepted 8 December 2025 ♦ Published 25 February 2026

Abstract

Background and aims – *Ulex europaeus* is an invasive allopolyploid (hexaploid) plant considered a global threat that has successfully colonized various regions, including the Northern Andes. This study aimed to assess its diversity and genetic structure in six sampling sites from the Central and Eastern Andes of the Northern Andes of Colombia, located in conservation-priority highland landscapes, using single nucleotide polymorphisms.

Material and methods – Given the complex inheritance patterns of polyploids, both diploid and hexaploid datasets were analysed to estimate genetic diversity (e.g. H_o , H_e , F_{IS} , and private alleles) and population structure (using e.g. F_{STP} , AMOVA, STRUCTURE, PCA, DAPC analyses). Additionally, the introduction history of the species in the Northern Andes, particularly the introduction from the Eastern to the Central Andes, was investigated using ABC-RF.

Key results – Diploid and hexaploid datasets showed consistent clustering patterns, supporting predominantly disomic inheritance. Populations exhibited high heterozygosity (diploid: $H_o = 0.302$, $H_e = 0.252$; hexaploid: $H_o = 0.483$, $H_e = 0.222$). Genetic structure analyses showed moderate differentiation ($F_{STP} = 0.118$ for diploid; $F_{STP} = 0.074$ for hexaploid) and significant isolation by distance (diploid: $r = 0.479$, p value = 0.022; hexaploid: $r = 0.507$, p value = 0.012), but without a clearly defined spatial pattern, suggesting restricted gene flow influenced by external factors. ABC-RF analyses indicated at least two independent introduction events in the Eastern Andes, followed by multiple dispersal events into the Central Andes.

Conclusion – *Ulex europaeus* populations in the Northern Andes maintain high heterozygosity and restricted gene flow. Polyploidy likely contributes to preserving genetic diversity, while multiple introduction events and human-mediated dispersal shape population structure, underscoring the complex invasion dynamics of this species in high-mountain ecosystems.

Keywords

gene flow barriers, genetic structure, invasion history, SNPs, Andean paramo, polyploidy

INTRODUCTION

In recent decades, the number of introduction events and the successful establishment of invasive species have increased, primarily driven by climate change and human activities (Chiu et al. 2023; Laginhas et al. 2023; Ángel-Vallejo et al. 2024). At the same time, growing interest in understanding the causes and biological impacts of invasions has prompted an increasing number of studies addressing invasion dynamics from ecological, evolutionary, and genetic perspectives (Chiu et al. 2023; Laginhas et al. 2023). In particular, genetic analyses provide critical insights not only into differences between native and introduced populations but also into the spatial and demographic processes operating within invaded regions, especially in complex landscapes or when multiple introduction events are suspected (Ellstrand and Schierenbeck 2000; Hagenblad et al. 2015).

Assessing genetic diversity and structure within invaded regions allows us to understand how introduced species respond to local environmental filters, landscape fragmentation, and stochastic demographic processes (Allendorf and Lundquist 2003; Estoup et al. 2016). This is particularly relevant in mountain ecosystems, where geographic isolation and ecological heterogeneity may limit gene flow, promote divergence, and shape the evolutionary trajectories of invasive populations (Pauchard et al. 2009; Hirsch et al. 2019).

Genetic studies on invasive species have revealed contrasting patterns of diversity depending on introduction history, reproductive strategies, and life history traits (Jiang et al. 2023). While some invasive plant populations exhibit significantly reduced genetic diversity compared to their conspecifics in the native range, mainly due to drastic reductions in population size resulting from founder effects, inbreeding, and reproductive strategies such as self-fertilization or vegetative reproduction (Barrett et al. 2008; Estoup et al. 2016; Wang et al. 2017; Sun et al. 2018), others maintain or even exceed variability of native populations (Lavergne and Molofsky 2007; Ray and Quader 2014; Tang and Ma 2020; Sapkota et al. 2022). Such high genetic diversity in introduced populations is frequently shaped by both biological and environmental factors, including multiple introductions, hybridization, mutation, gene flow and polyploidy (Te Beest et al. 2012; Schrieber and Lachmuth 2017; Madelón et al. 2021; Aguirre-Acosta et al. 2023; Jiang et al. 2023).

Reductions in genetic diversity often occur during the establishment stage, particularly when propagules originate from a single introduction event or a limited source population, potentially constraining adaptive potential in the novel environment (Verhoeven et al. 2011; Schrieber and Lachmuth 2017). Nevertheless, many introduced populations overcome this through phenotypic plasticity, high propagule pressure (i.e. the number and frequency of introduced individuals into the non-native range), and anthropogenic dispersal, which enhance gene flow and promote persistence

(Zimmermann et al. 2010; Hagenblad et al. 2015; Stout et al. 2015; Estoup et al. 2016). In contrast, high genetic diversity in some introduced populations can enhance evolutionary potential, facilitating success across multiple invasion stages (Lavergne and Molofsky 2007).

Among the mechanisms that enhance invasion success, polyploidy stands out as a key factor, particularly in novel environments (Te Beest et al. 2012). In general, polyploid species benefit from their higher genome copy number, which makes them less affected by drastic reductions in population size and genetic drift, and provides greater tolerance to abiotic stress and broader adaptive allelic diversity (Baduel et al. 2018). Polyploidy also enhances clonal reproduction and phenotypic plasticity, traits associated with rapid establishment and spread (Feng et al. 2024) and can remove genetic constraints to promote the development of competitive traits such as increased vegetative growth and seed production (Kirk et al. 2011; Te Beest et al. 2012; Baker et al. 2017; Shang et al. 2019; Mounger et al. 2021; Moura et al. 2021; Bellot et al. 2023). In the case of allopolyploids, the combination of divergent genomes provides fixed heterozygosity that allows more effective masking of deleterious recessive mutations, helping maintain fitness under inbreeding, while disomic pairing can preserve stable heterosis across generations (Otto and Whitton 2000; Comai 2005; Soltis and Soltis 2009; Madlung 2013).

Ulex europaeus L. (gorse) is a widespread invasive shrub and an ideal model to study the influence of polyploidy and genetic variability on invasion success. It is an allopolyploid that originated from hybridization between two genetically distinct ancient lineages within the *Ulex* genus with wide climatic ranges. These lineages separated around 5 million years ago (Mya), and after a prolonged period of independent evolution, they gave rise to the hexaploid allopolyploid *U. europaeus* between ca 0.7 and 2 Mya (Hornoy et al. 2013; Bellot et al. 2023). This species is native to western Europe, reproduces both sexually and asexually, and has successfully colonized over 40 countries due to its broad ecological tolerance, high reproductive capacity, and difficulty of eradication (Broadfield and McHenry 2019; Roberts and Florentine 2021). Much of this spread was initially facilitated by deliberate human introductions, which contributed to its establishment outside its native range, and it is therefore considered a high global invasion risk (Hornoy et al. 2013). In Colombia, *U. europaeus* was introduced in the 1950s to the Eastern Hills of Bogotá D.C., in the Eastern Andes (Ríos 2005; Anderson and Anderson 2010; Baptiste et al. 2010). It has primarily invaded high Andean and subpáramo forests of the Northern Andes, particularly in the Eastern and Central Andes (Ríos 2005; Vargas et al. 2009; Baptiste et al. 2010; Rodríguez et al. 2019; Ángel-Vallejo et al. 2024), which are separated by the inter-Andean valley, an area with edaphoclimatic conditions that are not favourable for its establishment. *Ulex europaeus* thrives in temperate environments with mean annual temperatures between ~4–22°C and moderate to

high rainfall (> 300 mm), conditions that contrast with the warmer and drier climate of the valley (Christina et al. 2020; Ángel-Vallejo et al. 2024).

Previous genetic studies on *U. europaeus* have been conducted both in its native range (i.e. Spain, France, and Scotland) and in invaded regions (i.e. Chile, Réunion, New Zealand, and the United States). Hornoy et al. (2013) evaluated the genetic diversity and structure of various populations using microsatellite markers, finding relatively high genetic diversity worldwide and only moderate structure, with similar or even higher variability in introduced populations due to lineage admixture. More recently, Hozawa and Nawata (2021) analysed invasive populations in Maui, California, Hawaii, and New Zealand, also employing microsatellite markers, and reported high genetic similarity among regions, suggesting low genetic differentiation. Both studies highlight the usefulness of microsatellites in assessing genetic variation in this species, although available information remains limited. To date, no invasive populations have been genetically evaluated in the Northern Andes, a key region due to its unique high-mountain ecological conditions, where *U. europaeus* has shown increasing invasive behaviour (Roberts and Florentine 2021; Ángel-Vallejo et al. 2024).

Furthermore, given the hybrid origin of *U. europaeus* from divergent lineages, the species may exhibit a predominantly disomic inheritance pattern, in which only homologous chromosomes pair and recombine during meiosis, producing gametes with allelic combinations similar to those of diploid species (Soltis et al. 1993; Ramsey and Schemske 2002; Scott et al. 2023). Such disomic behaviour has been documented in other allopolyploids, including *Capsella bursa-pastoris* (L.) Medik., a tetraploid with fully disomic inheritance (Cornille et al. 2016; Duan et al. 2024), and *Triticum aestivum* L. (common wheat), a hexaploid that exhibits diploid-like segregation despite its genomic complexity (Geleta and Ortiz 2016; Bian et al. 2018). However, disomic inheritance may not be uniform across the entire genome of *U. europaeus*. Ribosomal markers show substantial variation among lineages, suggesting that some genomic regions could deviate from strictly disomic segregation. This heterogeneity may reflect a partially segmental inheritance pattern, in which certain loci behave diploid-like while others display more complex segregation depending on the degree of homology between homologous and homeologous chromosomes (Catalán et al. 2006; Scott et al. 2023). Similar intermediate or mixed inheritance systems have been reported in other polyploid taxa, such as *Citrus* species, which display combinations of disomic, tetrasomic, and intermediate segregation patterns (Fan et al. 2022), and in *Ipomoea batatas* (L.) Lam., a segmental allohexaploid showing locus-dependent mixed inheritance patterns (Gao et al. 2024). This potential variability in inheritance modes poses challenges for population genomic analyses, as many available tools assume strictly diploid behaviour. Therefore, evaluating both diploid and polyploid datasets

provides a more accurate and comprehensive view of genetic diversity and population structure in *U. europaeus*.

Here, we evaluated the population genomics of *U. europaeus* in high mountain ecosystems of the Central and Eastern Andes using single nucleotide polymorphisms (SNPs). Specifically, we addressed the following questions: (1) How does the genetic diversity of *U. europaeus* vary across the Northern Andes? (2) Is gene flow restricted among sampling sites in this region? (3) What was the likely introduction history of *U. europaeus* from the Eastern to the Central Andes? We hypothesize that polyploidy may contribute to high heterozygosity in these populations, which could influence the species' genetic diversity. Additionally, we expect that biogeographic barriers limit genetic connectivity and promote population differentiation due to unsuitable conditions for *U. europaeus* in adjacent ecosystems.

MATERIAL AND METHODS

Study area and sampling design

For the population genomic analysis, samples were collected in six sampling sites located in the Central and Eastern Andes, within the Northern Andes of Colombia, where the species has established naturalized, non-planted populations and with individuals in reproductive stage (Fig. 1; Suppl. material 1, Table S1).

In the Central Andes, three sampling sites were selected. One site is located in the Lagunilla sector of Murillo, Tolima Department (Central-TOL1), within the buffer zone of Los Nevados National Natural Park and characterized by a subpáramo ecosystem. The other two sites are situated in high Andean forests: the Alaska sector of Murillo, Tolima (Central-TOL2), and San Felix, Caldas Department (Central-CAL). Both sites are subject to disturbance from livestock activity.

In the Eastern Andes, the selected sampling sites include Bogotá D.C. (East-DC), the capital of Colombia and the area closest to the Eastern Hills, where *U. europaeus* was first recorded in the Northern Andes; Chía, Cundinamarca Department (East-CUN); and Ventaquemada, Boyacá Department (East-BOY). All three are located within high Andean Forest ecosystems and are embedded in a landscape matrix that has been modified and fragmented by human activities.

At each sampling site, mature leaves were collected from 30 adult individuals (N = 180), each separated by at least five meters to reduce the likelihood of sampling clonal individuals. Samples were labelled and preserved in silica gel. Subsequently, leaf tissue was macerated with liquid nitrogen and stored at -20°C until DNA extraction.

DNA extraction and SNP genotyping

Genomic DNA (gDNA) was extracted using the CTAB II protocol (Doyle and Doyle 1987), following a standardized procedure for the collected plant material.

DNA integrity was assessed by electrophoresis on 0.8% agarose gels. DNA concentration was quantified using a spectrophotometer and purity was evaluated by the A260/A280 and A260/A230 ratios. High-quality gDNA from 180 individuals was sent to LGC Genomics (Berlin, Germany) for genotyping by sequencing (GBS). The ddRAD-seq libraries were prepared by digesting gDNA with PstI and ApeKI restriction enzymes, and sequencing was performed using 150 bp paired-end reads on Illumina NextSeq 500/550 v.2 and NovaSeq 6000 platforms.

Sequencing data were demultiplexed using Illumina bcl2fastq v2.20, allowing one or two mismatches (or N) in barcode recognition, depending on barcode uniqueness. Adapter sequences were removed and reads not matching with restriction sites at the 5' end were discarded. Reads were trimmed with Trimmomatic v.0.39 (Bolger et al. 2014) using a sliding window of 10 bases, retaining only those with an average Phred quality score ≥ 20 . Reads containing Ns or shorter than 20 bases were also removed. Quality control of the final reads was performed using FastQC v.0.11.9 (Andrews 2010).

Due to the absence of a reference genome for *U. europaeus*, a de novo clustering approach was used. Combined reads from all individuals were clustered using CD-HIT-EST v.4.6.1 (Fu et al. 2012), allowing a maximum sequence divergence of 5%. Singletons and clusters supported by fewer than 20 reads were discarded. These clusters, representing consensus ddRAD loci assembled

from multiple individuals, served as an artificial reference that enabled consistent alignment and comparison across samples in the absence of a complete genome. Trimmed reads were then aligned to the reference clusters using Bowtie2 v.2.2.3 (Langmead and Salzberg 2012). Although Bowtie2 is commonly used for mapping against long contigs, it is also efficient and accurate for aligning short Illumina reads to fragmented or reduced representations of the genome and is therefore appropriate for ddRAD-seq consensus clusters (see Bowtie2 manual; Chambers et al. 2023). Finally, all alignments were merged with SAMtools v.1.17 (Danecek et al. 2021) and represent ddRAD loci shared among samples.

Given that *U. europaeus* may exhibit predominantly disomic inheritance with potential segmental components, parallel analyses were performed under both inheritance assumptions to capture its genomic complexity. Because many population genomic tools have limited compatibility with polyploid data, analyses were conducted using diploid and hexaploid datasets when possible (ploidy = 2 and ploidy = 6, respectively). In cases where polyploid data could not be processed, only the diploid dataset was used. Comparing both approaches allowed us to assess the robustness of genetic patterns inferred under different inheritance models.

Variant calling and SNP genotyping for both datasets were performed using FreeBayes v.1.3.6 (Garrison and Marth 2012) on the high-performance computing cluster

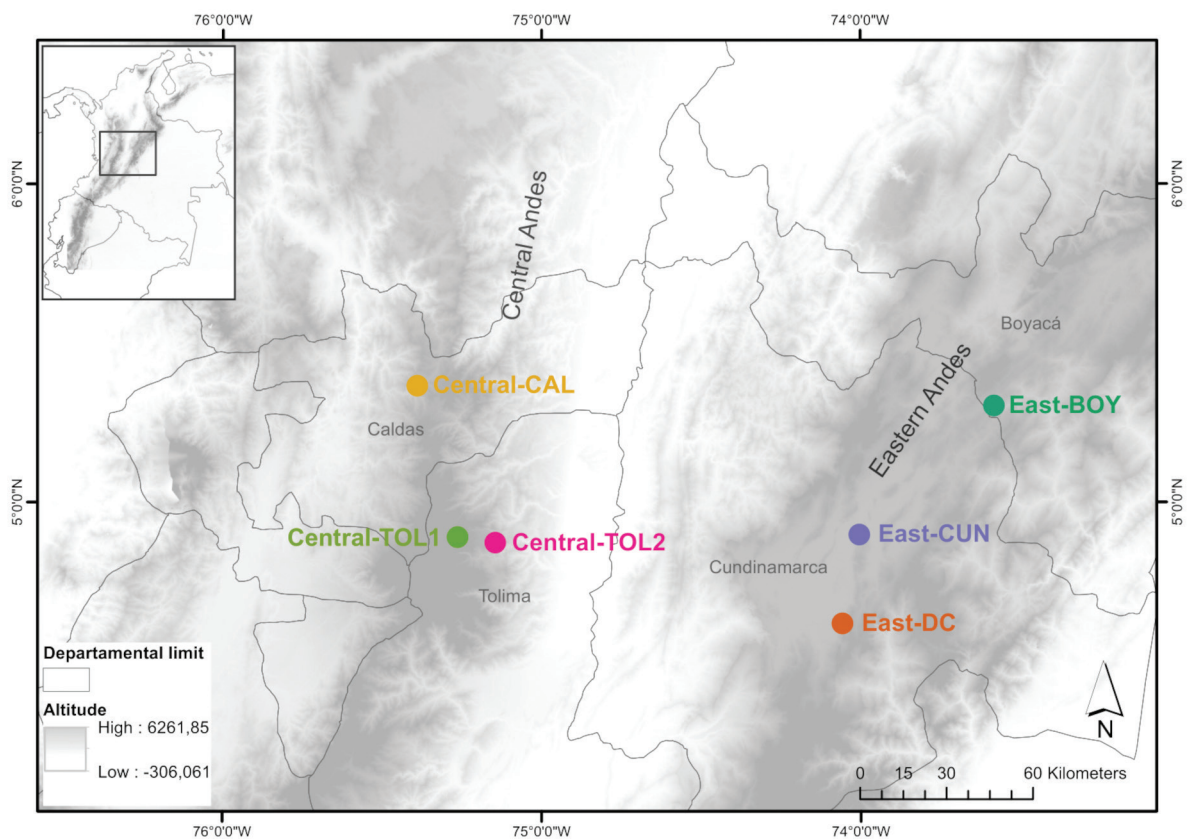


Figure 1. Map of the study area for the collection of *Ulex europaeus* in the Central and Eastern Andes of the Northern Andes. Coloured circles indicate the sampling sites.

at the Bioinformatics and Computational Biology Center (BIOS) in Manizales, Colombia. Resulting VCF files were initially filtered to retain only bi-allelic SNPs using BCFtools v.1 (Danecek et al. 2021). Subsequently, a two-step quality filtering process was applied using the filter function. First, SNPs with quality score (QUAL) < 20 were removed. For the diploid dataset, genotypes with sequencing depth (DP) < 5 or genotype quality (GQ) < 15 were marked as missing, whereas for the hexaploid dataset the DP threshold was set to < 20 according to Onoue et al. (2022). Then, SNPs with more than 10% missing data and those with a minor allele frequency (MAF) ≤ 0.05 were excluded. Finally, for the diploid dataset, SNPs showing suspiciously high heterozygosity (i.e. > 90% of individuals heterozygous) were identified and excluded using custom R scripts. A conservative threshold was applied to avoid removing potentially informative SNPs for population structure, as can occur when all loci deviating from Hardy–Weinberg equilibrium across the entire dataset are excluded (Pearman et al. 2022). The filtering process was applied separately to diploid and hexaploid datasets, resulting in two final VCF files (Suppl. material 2 and 3). All filtering steps, except the heterozygosity filter, were implemented using the following BCFtools commands: `bcftools view -m2 -M2 -v snps input.vcf | bcftools filter -S -e 'QUAL<20 || FMT/DP<5 || FMT/GQ<15'` (for diploid dataset) or `'QUAL<20 || FMT/DP<20 || FMT/GQ<15'` (for the hexaploid dataset) | `bcftools filter -e 'F_MISSING > 0.1 || MAF <= 0.05' -O v -o filtered_snps.vcf`.

Genetic diversity, population structure, and introduction history

Genetic analyses were conducted using a multi-locus, genome-wide approach that enabled a comprehensive assessment of genetic diversity within sampling sites and of the population genetic structure. This approach provides robust insights into the evolutionary history of populations by including loci with different genealogies, thereby increasing statistical power, reducing false discovery rates, and improving the precision of population genetic parameter estimates (Carling and Brumfield 2007).

Observed (H_o) and expected (H_e) heterozygosity were estimated at the overall level for both diploid and hexaploid datasets using the `poppr` v.2.9.4 and `vcfR` v.1.14.0 packages in R v.4.2.1 (Kamvar et al. 2014; Knaus and Grünwald 2017; R Core Team 2021). For the diploid dataset, H_o , H_e , and the number of private (i.e. unique) alleles (A_p) were estimated per sampling site using Arlequin v.3.5.1 (Excoffier and Lischer 2010). In addition, the fixation index F_{IS} , Hardy-Weinberg equilibrium test (P_{HWE}), and heterozygote excess test (P_{E-Het}) were calculated per sampling site using `genepop` v.1.2.3 in R (Rousset 2008). Overall, F_{IS} values for each dataset were estimated using `hierfstat` v.0.5-11 in R (Goudet 2005).

To evaluate genetic differentiation, Wright's fixation indices (F_{ST}) were estimated pairwise among sampling

sites for both diploid and hexaploid datasets using 10,000 permutations in the `StAMPP` v.1.6.3 package in R (Pembleton et al. 2013). Overall genetic differentiation for both datasets was assessed using the unweighted mean of pairwise F_{ST} values (F_{STP}), since the Weir and Cockerham (1984) estimators are not directly applicable to polyploid data. A false discovery rate correction (Benjamini and Hochberg 1995) was applied to adjust for multiple testing. To determine the percentage of variance explained within and among sampling sites, an analysis of molecular variance (AMOVA) was performed on the diploid dataset using Arlequin, with 10,000 permutations. A total of 201 grouping scenarios (hypotheses) were evaluated, representing all possible combinations of sampling sites into 2 to 5 groups, to explore alternative patterns of population structure. The models with significant among-groups variation components (F_{CT}) were retained for interpretation.

The number of genetic clusters (K), representing the number of inferred populations to which individuals could be assigned, was estimated using the `Structure` v.2.3.4 tool (Pritchard et al. 2010) implemented in `ipyrad` v.0.9.92 (Eaton and Overcast 2020), which uses Bayesian clustering methods. Analyses were conducted separately for diploid and hexaploid datasets, including all individuals. VCF files were converted to `hdf5` format using the `vcf_to_hdf5` tool. Each run included a burn-in of 10,000 iterations followed by 100,000 MCMC, testing K values from 2 to 6. The most likely number of clusters was determined using both the average log-likelihood (`estLnProbMean`) and the ΔK method (Evanno et al. 2005). Genetic assignment barplots for $K = 2-6$ were generated using the `toyplot` v.0.18.0 tool.

To explore patterns of genetic variation among individuals and populations, visualize population structure, and assess the degree of differentiation between sampling sites in the Northern Andes, principal component analysis (PCA) and discriminant analysis of principal components (DAPC) were also performed for both datasets using the `adegenet` v.2.0.1 package in R (Jombart 2008; Jombart and Ahmed 2011). For DAPC, 300 principal components (PCs) were initially retained and validated through cross-validation, dividing the data into 90% training and 10% testing sets using the `xvalDapc` function with. Thirty repetitions were performed to ensure robustness. One hundred PCs were retained and a value of $K = 6$ was set based on the Bayesian information criterion (BIC) (Jombart and Collins 2015).

The relationship between geographic distance and genetic differentiation among sampling sites was evaluated for both datasets (diploid and hexaploid) using a Mantel test (Mantel 1967), based on Spearman's correlation and 10,000 permutations with a free permutation approach (Jackson and Somers 1989). Geographic distances among sampling sites were calculated using Google Earth Pro. Genetic differentiation matrices (pairwise F_{ST}) were transformed to $F_{ST}/(1-F_{ST})$ to linearize the relationship with geographic distance (Rousset 1997; Jiao et al. 2024).

The Mantel test was conducted using the *vegan* v.2.6-4 package in R (Oksanen et al. 2022). In addition, clustering trees based on Kosman/Leonard genetic distance were constructed for each dataset using the *stats* v.4.2.1, *PopGenReport* v.3.0.7 (Adamack and Gruber 2014), and *dendextend* v.1.19.1 (Galili 2015) packages in R.

To infer the introduction history of *U. europaeus* from the Eastern to the Central Andes, an Approximate Bayesian Computation analysis based on the diploid dataset was performed using the Random Forest algorithm (ABC-RF), implemented in *DIYABC-RF* v.1.2.1 (Collin et al. 2021). The main objective was to reconstruct the direction and pattern of introduction into the Central Andes, originating from previously established populations in the Eastern Andes, which historical records identify as the first introduction sites in Colombia (Ríos 2005; Baptiste et al. 2010). Although the initial introduction from the native range was not explicitly modelled, an unsampled “ghost ancestral population” was included to represent the unknown genetic source introduced into the country (Lee and Son 2022). To reduce model complexity and enhance statistical power of scenario selection, the sequential approach proposed by Fraimout et al. (2017) was adopted. This approach involves the evaluation of two consecutive sets of scenarios, which were defined based on historical records and previous genetic structure analyses that revealed clear clustering among the sampling sites.

In a first step, three introduction scenarios were tested for the Eastern Andes sites East-CUN and East-DC, located near the Eastern Hills of Bogotá, where the earliest records of the species in Colombia are concentrated. The scenarios evaluated were: (a) independent introductions of East-CUN and East-DC from a shared ancestral population; (b) introduction to East-DC from the ancestral population, followed by introduction to East-CUN from East-DC; and (c) introduction to East-CUN from the ancestral population, and a subsequent introduction to East-DC from East-CUN (Suppl. material 1, Fig. S1). This step aimed to reconstruct the initial spread of the species within the Eastern Andes before reaching the Central Andes. The East-BOY site was excluded from this step because it did not provide relevant information regarding the likely invasion route into the Central Andes. Genetic structure analyses indicated that East-BOY represents a highly structured and genetically distinct population, exhibiting even greater divergence from Central Andes sites.

Based on the most likely scenario identified in the first step and on genetic clustering patterns revealed in previous analyses, four alternative introduction scenarios were subsequently modelled for the Central Andes sampling sites (Central-TOL2, Central-TOL1, and Central-CAL). These scenarios evaluated whether the sites were established through three independent introduction events (scenarios 1 and 2), or through two introduction events: one involving the Tolima sites (Central-TOL2 and Central-TOL1), and another involving the Caldas site

(Central-CAL; scenarios 3 and 4; Suppl. material 1, Fig. S2).

In each modelling step, 30,000 and 80,000 simulated datasets were generated under a uniform prior, respectively. Effective population size priors ranged from 10 to 10,000 for sampling sites and from 10 to 100,000 for the unsampled ghost ancestral population. Divergence time priors ranged between 10 and 1,000 generations, with the constraint that the most recent divergence time was smaller than the older ones ($t_1 < t_2$ for the first step, and $t_1 < t_2 < t_3 < t_4$ for the second step). Population size changes following the introductions were also included. The data were summarized using summary statistics for SNPs and linear discriminant analysis (LDA). The best scenario was selected using the Random Forest algorithm with 10 replicates of 1,000 trees. Posterior probabilities and prior error rates were estimated to evaluate model performance (Fraimout et al. 2017).

In summary, both diploid and hexaploid datasets were used to estimate overall heterozygosity indices (H_o and H_e), F_{ST} , Bayesian clustering, PCA, DAPC, Mantel test, and clustering trees. In contrast, only the diploid dataset was used for AMOVA, per-site estimates of H_o , H_e , A_p , F_{IS} , P_{HWE} , and P_{E-Het} , and ABC-RF analyses (Suppl. material 1, Fig. S3), due to software limitations with polyploid data.

RESULTS

SNP sequencing, filtering, and dataset summary

A total of 472.5 million raw paired-end reads were obtained from the 180 individuals, averaging approximately 2.5 million reads per sample. The reference cluster assembly consisted of 359,271 contigs. Variant calling generated 394,021 SNPs for the diploid dataset and 82,388 SNPs for the hexaploid dataset. After quality control and filtering, 3,891 SNPs were retained in the diploid dataset (with 5.6% missing data) and 2,994 SNPs in the hexaploid dataset (with 6.1% missing data). These variants were distributed across 1,867 contigs in the diploid dataset and 1,582 contigs in the hexaploid dataset. Both datasets included all 180 individuals, and quality control confirmed high sequencing quality across all samples (Suppl. material 1, Table S2).

Genetic diversity

At the overall level, observed heterozygosity (H_o) exceeded expected heterozygosity (H_e) in both datasets: $H_o = 0.302$ and $H_e = 0.252$ for the diploid dataset, and $H_o = 0.483$ and $H_e = 0.222$ for the hexaploid dataset (Table 1). In the diploid dataset, this pattern was consistent across all sampling sites (Table 1). As mentioned in the methodology, all other diversity estimates (except for overall heterozygosity) were based solely on the diploid dataset due to software limitations. The Hardy-Weinberg equilibrium test showed highly significant deviations ($P_{EHW} < 0.0001$) both at the overall level and per sampling

Table 1. Genetic diversity statistics by sampling site and overall, estimated from diploid and hexaploid datasets for *Ulex europaeus* in the Central and Eastern Andes of the Northern Andes. Private alleles (A_p), observed heterozygosity (H_o), expected heterozygosity (H_e), p value of the Hardy-Weinberg equilibrium test (P_{EHW}), p value of the test for excess heterozygosity (P_{E-Het}), and inbreeding coefficient (F_{IS}). *Highly significant p value (< 0.0001).

Andes	Sampling site	A_p	H_o	H_e	P_{EHW}	F_{IS}
					P_{E-Het}	
Central	Central-TOL1	5	0.383	0.288	<0.0001*	-0.369
	Central-TOL2	1	0.343	0.264	<0.0001*	-0.364
	Central-CAL	2	0.388	0.281	<0.0001*	-0.443
Eastern	East-CUN	1	0.346	0.268	<0.0001*	-0.344
	East-DC	0	0.345	0.262	<0.0001*	-0.356
	East-BOY	2	0.460	0.319	<0.0001*	-0.493
Overall (Diploid)	-	11	0.302	0.252	<0.0001*	-0.325
Overall (Hexaploid)	-	-	0.483	0.222	-	-

site, likely due to an excess of heterozygotes ($P_{E-Het} < 0.0001$; Table 1). These results were further supported by negative inbreeding coefficients (F_{IS}), which persisted even after filtering loci with fixed heterozygotes, suggesting that the observed heterozygosity excess underlies these patterns.

Among sampling sites, East-BOY exhibited the highest observed heterozygosity ($H_o = 0.460$) and the lowest F_{IS} (-0.493), suggesting high genetic diversity and excess of heterozygotes. In contrast, East-DC showed no private alleles, while Central-TOL1 had the highest number ($A_p = 5$), highlighting differences in allelic richness across sampling sites (Table 1).

Population genetic structure

The evaluation of genetic differentiation revealed moderate population structuring in the diploid dataset and weaker structuring in the hexaploid dataset. In the diploid dataset, the overall fixation index estimated using Weir and Cockerham's method ($F_{ST} = 0.100$) was nearly identical to the unweighted mean of pairwise values ($F_{ST,p} = 0.118$), indicating consistent differentiation estimates regardless of the approach used. In contrast, the hexaploid dataset showed a lower overall differentiation ($F_{ST,p} = 0.074$), reflecting weaker population structure. Significant pairwise F_{ST} values were observed among all sampling sites in both datasets (Fig. 2), with the strongest differentiation between East-BOY/Central-CAL and East-BOY/Central-TOL1, and the lowest between East-DC/East-CUN.

In the Central Andes, Central-TOL1 and Central-TOL2, the two geographically closest sites, exhibited relatively low genetic differentiation (diploid: $F_{ST} = 0.084$; hexaploid: $F_{ST} = 0.054$) compared to their higher

differentiation with Central-CAL (diploid: $F_{ST} = 0.152$ and 0.136; hexaploid: $F_{ST} = 0.081$ and 0.079). Similarly, in the Eastern Andes, East-BOY showed high genetic differentiation from all other sites, while East-DC and East-CUN, being geographically closer, presented lower levels of differentiation (diploid: $F_{ST} = 0.066$; hexaploid: $F_{ST} = 0.051$; Fig. 2).

The Mantel test revealed a moderate but significant positive correlation between genetic and geographic distance, suggesting a pattern of isolation-by-distance (diploid: $r = 0.479$, p value = 0.022; hexaploid: $r = 0.507$, p value = 0.012; Suppl. material 1, Fig. S4). AMOVA results showed that most of the genetic variation was found within sampling sites (86–87%), with lower proportions of variation among sites within groups (7–9%) and among groups (3–5%). A total of six structuring scenarios showing significant F_{CT} values were retained. The sampling site East-BOY clustered separately from the other populations in five scenarios. In contrast, Central-TOL1 and Central-TOL2 were consistently grouped together in four scenarios (Suppl. material 1, Table S3).

In the structure analysis, although the ΔK method identified $K = 3$ as the most likely number of genetic clusters for both datasets, the average log-likelihood (estLnProbMean) supported $K = 6$, which reflects the finer-scale genetic structure observed among sampling sites (Suppl. material 1, Figs S5, S6). In this context, Q represents the proportion of ancestry (membership coefficient) of each individual assigned to each genetic cluster inferred by Structure (Pritchard et al. 2000). In the diploid dataset, results showed a clearly defined assignment pattern in most populations (Fig. 3A).

Central-TOL1 and Central-TOL2 exhibited complete assignment of individuals to a single cluster (100% with $Q > 0.8$), followed by East-BOY with 93.3% of individuals predominantly assigned. East-DC (76.7%) and Central-CAL (70%) showed intermediate levels of assignment,

while East-CUN presented the most diffuse pattern, with only 60% of individuals strongly assigned. These results suggest a partially defined genetic structure, with well-differentiated populations in the Central Andes and higher levels of admixture among populations from the Eastern

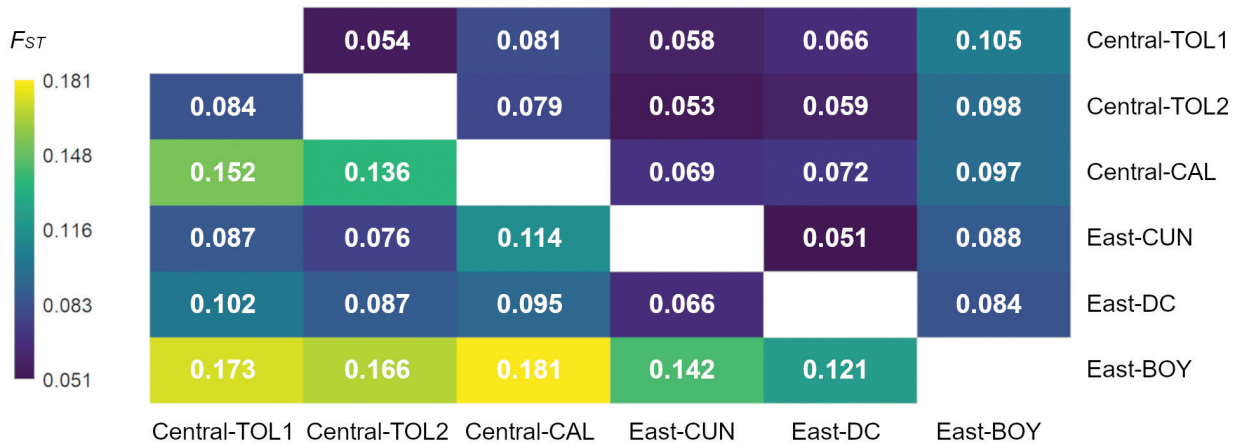


Figure 2. Pairwise estimates of F_{ST} among sampling sites from the diploid (below the diagonal) and hexaploid (above the diagonal) datasets for *Ulex europaeus* in the Central and Eastern Andes of the Northern Andes. All pairwise estimates F_{ST} were significant after False Discovery Rate correction (p value < 0.0001).

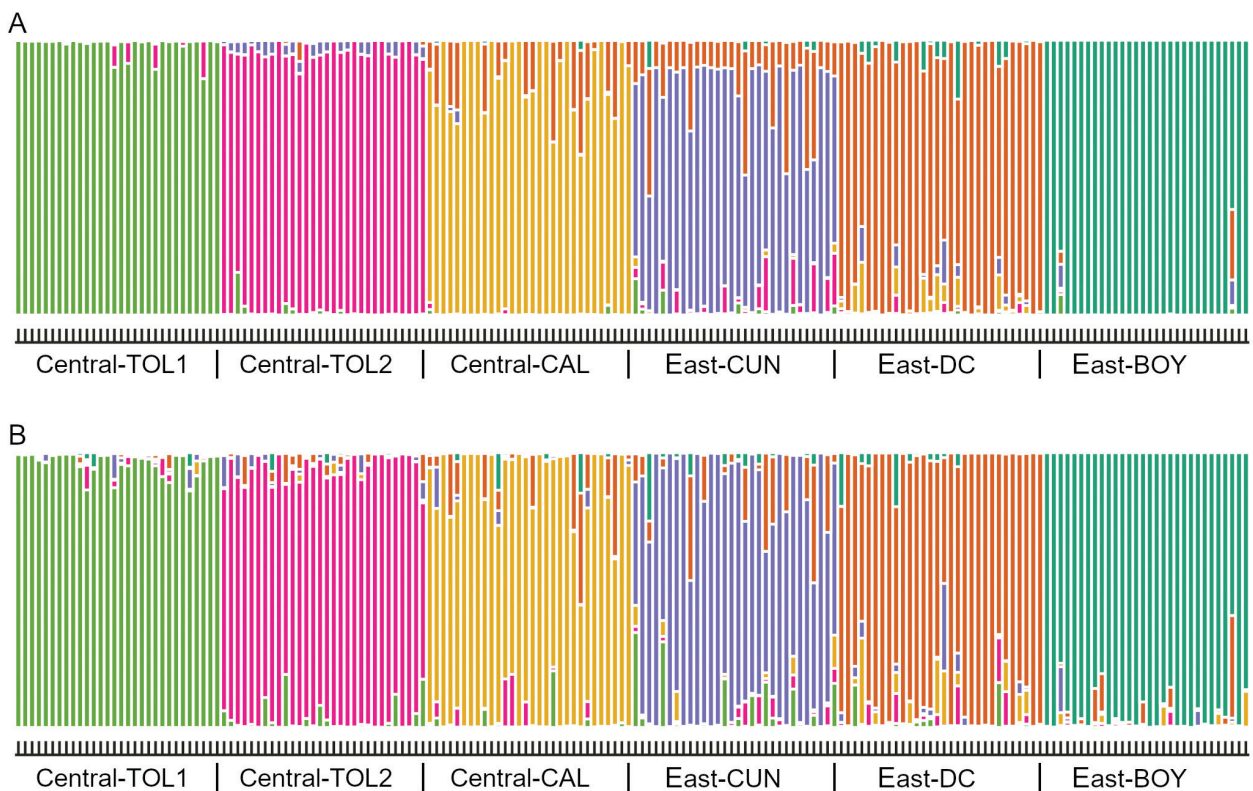


Figure 3. Population structure inferred by structure ($K = 6$) from the diploid (A) and hexaploid (B) datasets for *Ulex europaeus* in the Central and Eastern Andes of the Northern Andes. The colours represent the genetic clusters; each vertical line corresponds to an individual and indicates the proportion of membership to each group. The labels below the bars indicate the sampling site of each individual, grouped in sets of 30.

Andes. In the hexaploid dataset, overall patterns were similar (Fig. 3B). Central-TOL1 again showed complete assignment (100%), while East-BOY and Central-TOL2 had high assignment levels (93.3% and 86.7%, respectively). In contrast, Central-CAL, East-CUN, and East-DC showed lower proportions of strongly assigned individuals (66.7% in all three cases), indicating a higher degree of admixture compared to the diploid dataset. Overall, both datasets revealed a general correspondence between genetic structure and sampling sites. Central-TOL1, Central-TOL2, and East-BOY displayed greater differentiation, whereas Central-CAL, East-CUN, and East-DC showed higher genetic admixture. The detected

admixture reveals genetic connections among some Central and Eastern populations, indicating recently shared ancestry or restricted gene flow, consistent with the significant F_{ST} values observed.

Principal component analyses (PCA) confirmed the genetic differentiation among sampling sites, revealing six distinguishable groups in both datasets. The first two principal components explained 15.92% of the total variation in the diploid dataset and 12.08% in the hexaploid dataset (Fig. 4A, B). In the Central Andes, Central-TOL1 and Central-TOL2, both geographically close, showed greater genetic similarity, while Central-CAL appeared more differentiated. In the Eastern Andes,

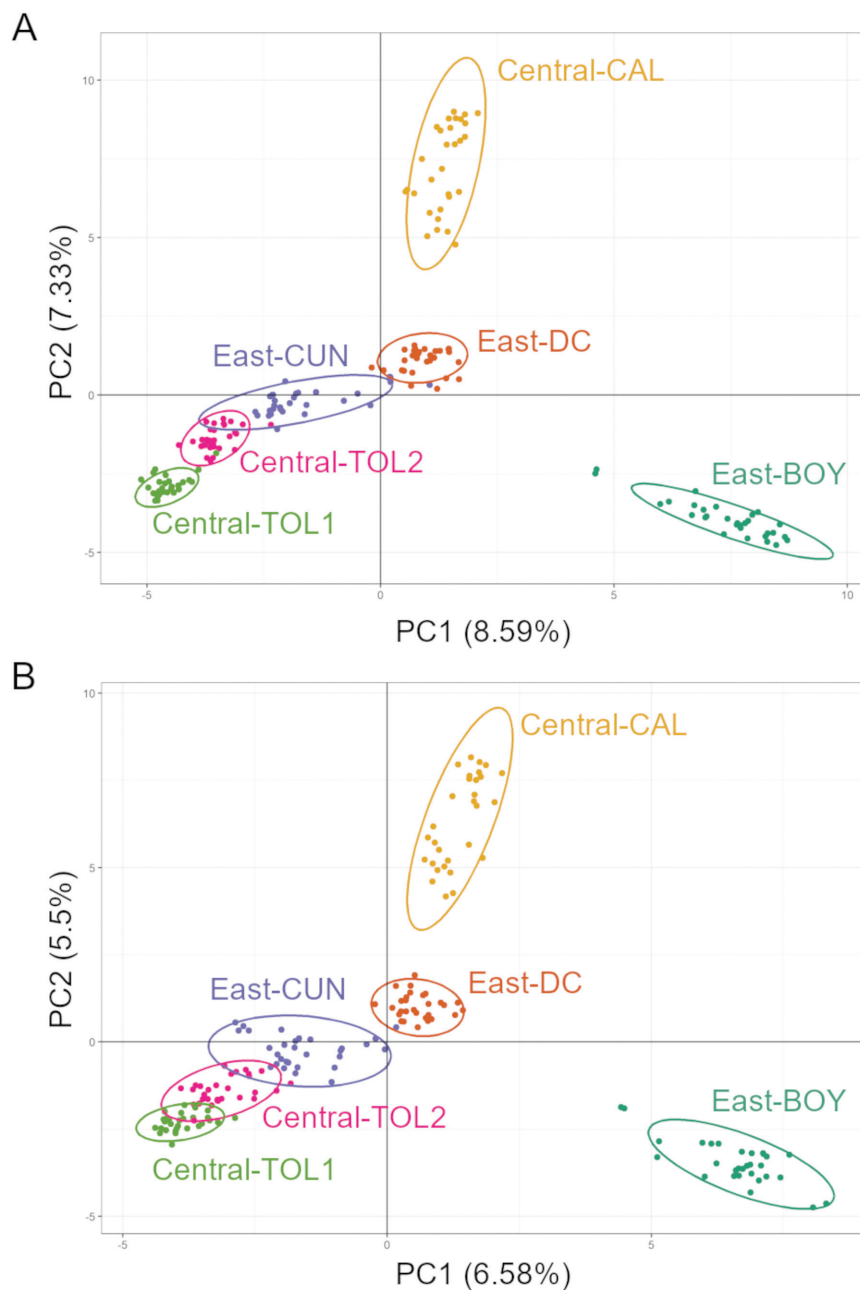


Figure 4. Principal component analysis (PCA) from the diploid (A) and hexaploid (B) datasets for *Ulex europaeus* in the Central and Eastern Andes of the Northern Andes. Each point represents an individual, grouped by colour according to the sampling site.

East-CUN and East-DC, also geographically close, were genetically similar, while East-BOY was more distinct. However, the PCA did not reveal a clear differentiation between the two regions (Eastern vs Central Andes). Discriminant analysis of principal components (DAPC), supported by the BIC, also indicated $K = 6$ as the most likely number of genetic groups (Suppl. material 1, Figs S7, S8).

The Kosman/Leonard dendrograms showed six well-defined genetic groups corresponding to the sampling sites, which were clustered into two main clades. In the diploid dataset, Central-TOL1 and Central-TOL2 (Central Andes) showed a close genetic similarity with East-CUN (Eastern Andes), while East-BOY and East-DC (Eastern Andes), grouped together and exhibited genetic similarity with Central-CAL (Central Andes) (Fig. 5A). A similar clustering pattern was observed in the hexaploid dataset, although Central-TOL2 appeared more genetically similar to East-CUN than Central-TOL1 (Fig. 5B). These patterns indicate closer genetic affinities, i.e. more recently shared ancestry, between some Eastern and Central Andean sampling sites than among sites within each region.

Introduction history

In the ABC-RF analysis, the first step indicated that the most likely scenario for the introduction of *U. europaeus* to the Eastern Hills involved at least two independent introduction events (scenario 1): one into East-CUN and another into East-DC. In both cases, propagules originated from an ancestral, unsampled source population introduced into the country (Fig. 6; Table 2; Suppl. material 1, Fig. S1A).

The second step of the analysis suggests that the most likely origin of the sampling sites in the Central Andes involved two independent introduction events: one into the Tolima sites (Central-TOL1 and Central-TOL2), originating from East-CUN; and another into Central-CAL, originating from East-DC (Fig. 6; Table 2). The most supported scenario (scenario 4) further proposes that the introduction in Tolima occurred first at Central-TOL1 from East-CUN, followed by dispersal to Central-TOL2 from Central-TOL1 (Fig. 6). However, the distinction between this and the alternative scenario (scenario 3), in which Central-TOL2 was colonized first, was only moderate, as both received a similar proportion

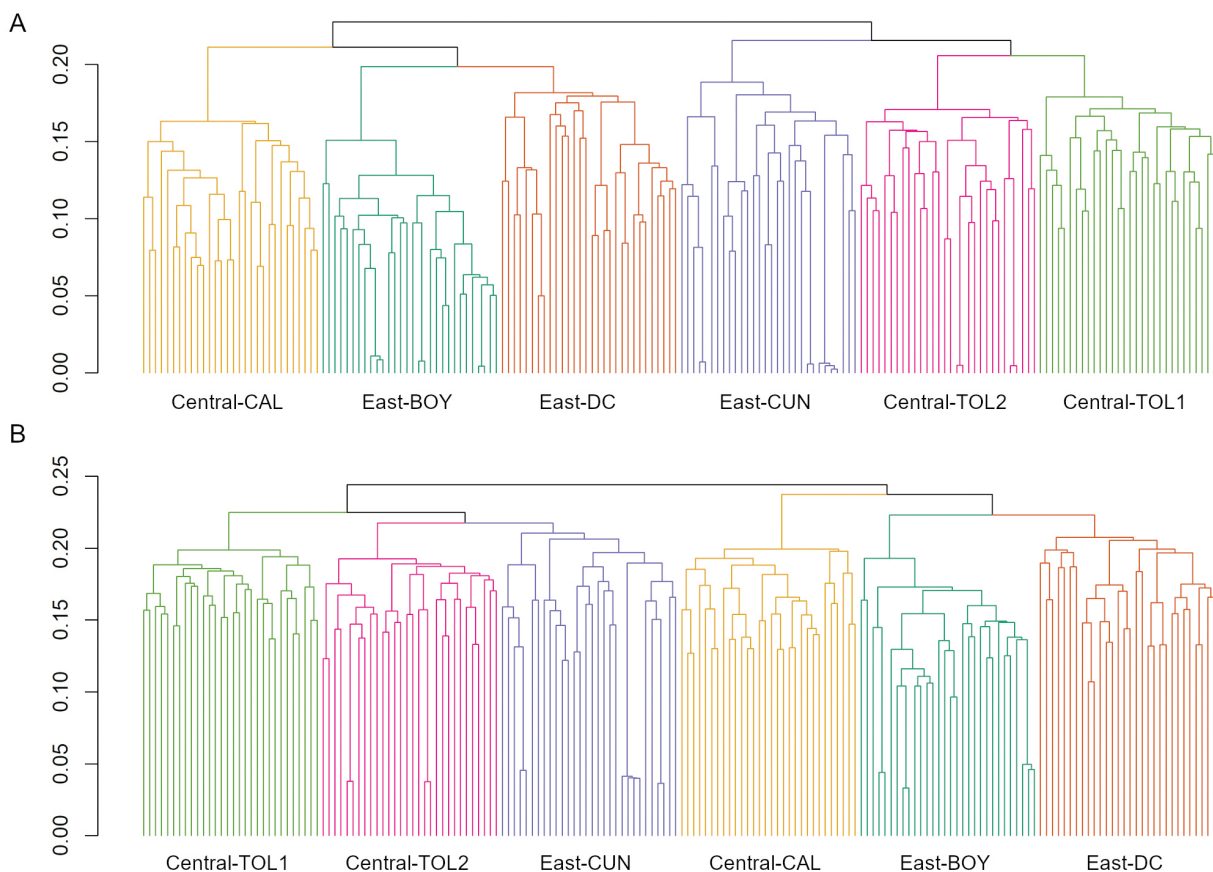


Figure 5. Clustering dendrograms based on Kosman and Leonard genetic distance from the diploid (A) and hexaploid (B) datasets for *Ulex europaeus* in the Central and Eastern Andes of the Northern Andes. Each terminal node represents an individual, grouped by colour according to the sampling site.

of RF votes (Table 2; Suppl. material 1, Fig. S2C–D). These results suggest that the introduction of *U. europaeus* to the Central Andes occurred through at least two pathways, each tracing back to distinct Eastern Andes sources, consistent with a scenario of multiple introductions followed by localized dispersal events.

DISCUSSION

The invasive success of *U. europaeus* in the Northern Andes does not appear to be limited by genetic changes associated with its introduction process. The consistently high levels of heterozygosity across sampling sites may support the species' persistence and adaptive potential in these novel environments. A moderate level of genetic structure was detected, indicating restricted gene flow among sites and revealing populations that are beginning to differentiate. Overall, the observed patterns are consistent with a complex introduction history involving multiple independent introduction events followed by localized dispersal processes.

Comparison between diploid and hexaploid datasets

Genetic diversity and population structure analyses performed with both diploid and hexaploid datasets revealed similar patterns. Diversity estimates showed consistent trends across assumed ploidy models, and the genetic structure analyses indicated similar clustering

among sampling sites. These results suggest that *U. europaeus* populations in the Northern Andes exhibit a predominantly disomic behaviour, such that the diploid model adequately describes the patterns observed in the hexaploid dataset. Any potential segmental behaviour would likely be restricted to specific multicopy genomic regions, such as the ETS and ITS (Bellot et al. 2023), rather than representing a genome-wide pattern.

Genetic diversity and invasion success

Hornoy et al. (2013) proposed a two-stage invasion process for *U. europaeus*: initial dispersal from Spain to Northern Europe, followed by anthropogenic introductions elsewhere. During this process, despite the loss of rare alleles, overall genetic diversity was not significantly reduced. Although native populations were not included in the present study, the results may be consistent with Hornoy et al. (2013) in that the loss of rare alleles does not necessarily lead to a substantial reduction in total genetic diversity. This comparison should, however, be interpreted with caution, as SNP and microsatellite markers differ in mutation rates and evolutionary dynamics, even though both marker types can reveal similar patterns of genetic structure (Coates et al. 2009).

The populations analysed showed no evidence of heterozygosity loss. On the contrary, a consistent excess of heterozygotes was detected, which may be explained by several factors. On the one hand, the founding individuals may have been predominantly heterozygous, providing

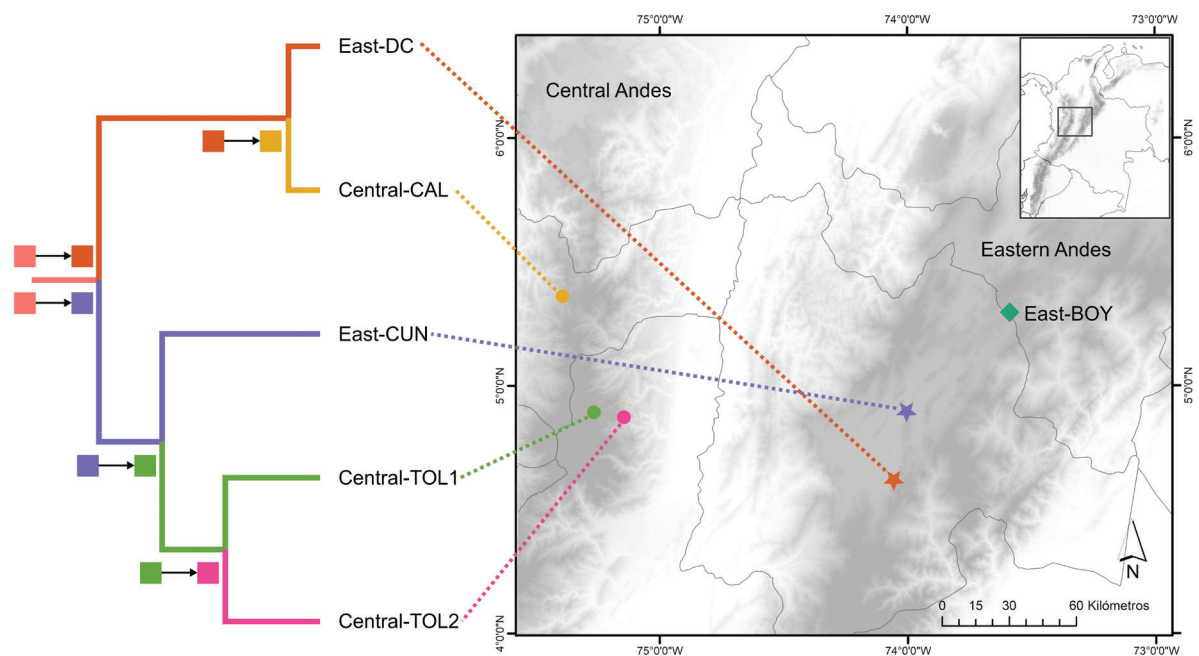


Figure 6. Graphical illustration of the most likely introduction scenario of *Ulex europaeus* from the Eastern Andes to the Central Andes in the Northern Andes. Sampling sites are shown using consistent colours across the tree and the map, with the salmon-coloured symbol representing the unsampled ancestral source population. In the tree, inferred dispersal pathways are indicated at the corresponding nodes. On the map, source populations are represented by stars, whereas recipient populations are represented by dots

Table 2. Scenarios evaluated and results from the two successive steps of the ABC-RF analysis used to infer the introduction history of *Ulex europaeus* from the Eastern Andes to the Central Andes, based on the diploid dataset. The best (most likely) scenario identified at each step is indicated in bold and marked with an asterisk (*). Prior error, RF votes, and posterior probability values correspond to the average of 10 RF replicates.

Scenarios	Prior error (%)	RF votes (%)	Posterior probability
Step 1. Eastern Hills			
Three scenarios, 16 summary statistics, 30,000 simulations, and 1,000 decision trees			
1. Independent introduction of East-CUN and East-DC from an ancestral population*		46.5	0.576
2. Introduction to East-DC from an ancestral population, and posterior dispersal to East-CUN from East-DC	0.371	24.4	-
3. Introduction to East-CUN from an ancestral population, and posterior dispersal to East-DC from East-CUN		29.1	-
Step 2. Introduction to Central Andes			
Four scenarios, 292 summary statistics, 80,000 simulations, and 1,000 decision trees			
1. Three independent introduction events, one for each population: first Central-TOL2 from East-CUN, then Central-TOL1 from East-CUN, and finally Central-CAL from East-DC		4.1	-
2. Three independent introduction events, one for each population: first Central-TOL1 from East-CUN, then Central-TOL2 from East-CUN, and finally Central-CAL from East-DC		0.7	-
3. Two independent introduction events, one for Tolima sampling sites and another for Caldas sampling site: first Central-TOL2 from East-CUN, then Central-TOL1 from Central-TOL2, and finally Central-CAL from East-DC	0.205	44.2	-
4. Two independent introduction events, one for Tolima sampling sites and another for Caldas sampling site: first Central-TOL1 from East-CUN, then Central-TOL2 from Central-TOL1, and finally Central-CAL from East-DC*		51.0	0.530

an initial evolutionary advantage (Bock et al. 2015). Likewise, hexaploidy can facilitate the accumulation and maintenance of genetic variation by allowing multiple chromosomal sets (Mounger et al. 2021). This excess may also be associated with hybrid vigour, as heterozygous individuals often exhibit higher fitness and adaptive capacity. Another possibility is the presence of ongoing positive selection favouring heterozygous genotypes, thereby maintaining high levels of genetic diversity even after introduction events (Mounger et al. 2021).

Beyond these intrinsic mechanisms, demographic processes commonly associated with recent invasions could contribute to the high levels of heterozygosity observed. In particular, introduced populations can exhibit elevated heterozygosity due to the recent admixture of multiple introduced lineages, which rapidly increases genetic variability by combining alleles from different source regions (Keller et al. 2014). However, it remains unknown whether Northern Andes populations descend from multiple European genetic lineages, as native-range populations were not included. Nevertheless, given that the bioclimatic conditions of the Northern Andes differ both from the native range and from other invaded regions, this excess heterozygosity may have facilitated climatic niche expansion (Ángel-Vallejo et al. 2025) and,

consequently, enhanced the invasive success (Keller et al. 2014) of the species in this region.

Population structure and inference of the introduction history

Pairwise F_{ST} values from both datasets were significant among all sampling sites, indicating moderate population structuring and restricted gene flow (non-random mating) even among geographically close sites. Although the Central and Eastern Andes are separated by the Magdalena River Valley and lie above 2600 m a.s.l., a potential biogeographic barrier, no clear genetic clustering by mountain range (Eastern vs Central Andes) was detected. This suggests that some level of gene flow does occur among sites, but that it is not sufficient to fully homogenize the populations.

Natural dispersal in *U. europaeus* occurs mainly through explosive dehiscence, zoochory, and hydrochory (Roberts and Florentine 2021), mechanisms that primarily facilitate short-distance movement. Consequently, long-distance gene flow would require additional processes capable of overcoming the species' inherent dispersal limitations. In the Northern Andes, the limited availability of effective local agents for its spread suggests that the contribution

of natural dispersal to gene flow at larger spatial scales is likely modest.

Limited gene flow within and between mountain ranges may be facilitated by human activities. Intentional dispersal, such as the use of *U. europaeus* in live fences (Ríos 2005), could connect sites across the Central and Eastern Andes, while accidental dispersal through the transport of seeds on machinery, vehicles, or agricultural materials (Gouldthorpe et al. 2006) likely promotes gene flow within each mountain range. Additionally, road edges provide favourable conditions for *U. europaeus*, enhancing colonization, establishment, and expansion, thereby connecting populations that would otherwise remain isolated (León-Cordero et al. 2016). In the Northern Andes, populations are commonly observed along roadsides, which further facilitates local gene flow.

Clustering analyses revealed that the East-DC and East-CUN populations, located near the Eastern Hills of Bogotá, exhibited higher genetic admixture and lower differentiation than the other populations. This pattern suggests that propagules introduced into other sampling sites likely originated from areas near Bogotá (Ríos 2005; Baptiste et al. 2010). This inference is supported by ABC-RF results, which identified at least two independent introduction events of *U. europaeus* into the Eastern Andes, followed by introductions into the Central Andes. In addition, multiple independent introduction events within the Central Andes were detected, including the spread of propagules from East-CUN into Tolima, and a separate introduction from East-DC into Caldas.

Regarding the colonization of *U. europaeus* in the Central Andes, the ABC-RF analysis showed some uncertainty about the sequence of establishment of the Tolima populations. However, the most likely scenario indicates that Central-TOL1 was first established from East-CUN, followed by the colonization of Central-TOL2 from Central-TOL1, whereas Central-CAL originated independently from East-DC more recently. This sequence aligns with the introduction times reported by Ángel-Vallejo et al. (2025), who estimated the time since introduction through interviews with local residents regarding how long the plant had been present on their land. Based on these accounts, Central-TOL1 is the oldest population (~40 years), followed by Central-TOL2 (~33 years), and Central-CAL (~18 years). Together, this evidence supports the inference of independent introduction events in the Central Andes. Although the East-BOY population was not included in the ABC-RF analysis, clustering and genetic structure results indicate that it is genetically distinct, suggesting a separate introduction event, possibly from a different propagule source or colonization pathway.

In conclusion, the results of this study indicate that populations of *U. europaeus* in the Northern Andes maintain consistently high levels of heterozygosity. This pattern supports the hypothesis that polyploidy may contribute to preserving genetic variation and enhancing

the species' adaptability to a variety of environmental conditions, thereby facilitating its persistence and expansion in invaded ecosystems. Genetic structure analyses revealed recent differentiation among populations and restricted gene flow. While biogeographic features such as the inter-Andean valley may limit dispersal in the region, our results did not detect an effect of this barrier on the clustering pattern and instead suggests that other factors, including human-mediated dispersal, play an important role in shaping genetic connectivity. Finally, the inferred introduction history points to at least two independent introduction events in the Eastern Andes, followed by multiple introductions into the Central Andes, highlighting the complexity of *U. europaeus* invasion in high mountain ecosystems. Future studies incorporating environmental data and genome-wide selection analyses could provide insights into the role of local adaptation in the success of this species.

DATA AVAILABILITY STATEMENT

The sequencing data generated and analysed during this study have been deposited in the NCBI BioProject database under the BioProject ID PRJNA1249511 (<https://www.ncbi.nlm.nih.gov/bioproject/1249511>). The data that supports the findings of this study are available in the supplementary material of this article.

ACKNOWLEDGEMENTS

We express our gratitude to the Ministerio de Ciencia, Tecnología e Innovación (Minciencias) and the Universidad de Caldas for supporting the project entitled "Population status, reproductive biology and genetic diversity of the invasive plant *Ulex europaeus* L., in populations of the Colombian Central Cordillera" (projects 1127-852-71470 and 0643120), as well as for funding the ADR master's degree. We also thank the Centro de Bioinformática y Biología Computacional (BIOS) for providing access to their high-performance computing cluster. We thank the Laboratorio de Colecciones Biológicas and Laboratorio de Biología Vegetal y Molecular of the Universidad de Caldas for providing the space and equipment necessary for the laboratory phase of this research. We thank María Camila Ángel-Vallejo and Eliana Jimena García-Marín for their assistance and contributions during fieldwork. We also extend our thanks to the owners of the property in Alaska for allowing us access to their land, where invasive *U. europaeus* populations are present. Finally, we sincerely thank the editor and reviewer for their constructive comments and valuable suggestions, which helped to substantially improve the quality of this manuscript.

REFERENCES

Adamack AT, Gruber B (2014) PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology*

- and Evolution 5(4): 384–387. <https://doi.org/10.1111/2041-210X.12158>
- Aguirre-Acosta N, Urdampilleta JD, Otero JT, Aguilar R (2023) Genetic diversity of an invasive tree across time and contrasting landscape conditions. *Forest Ecology and Management* 548: 121429. <https://doi.org/10.1016/j.foreco.2023.121429>
- Allendorf FW, Lundquist LL (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology* 17: 24–30. <https://doi.org/10.1046/j.1523-1739.2003.02365.x>
- Anderson SAJ, Anderson WR (2010) Ignition and fire spread thresholds in gorse (*Ulex europaeus*). *International Journal of Wildland Fire* 19(5): 589–598. <https://doi.org/10.1071/wf09008>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [accessed 02.10.2023]
- Ángel-Vallejo MC, Aguirre-Acosta N, Rodríguez-Rey GT, García-Marín EJ, Álvarez-Mejía LM, Feuillet-Hurtado C (2024) Distribution models in invasive plants with climatic niche expansion: a case study of *Ulex europaeus* L. in Colombian Andes. *Biological Invasions* 26(6): 1919–1930. <https://doi.org/10.1007/s10530-024-03285-7>
- Ángel-Vallejo MC, García-Marín EJ, Feuillet-Hurtado C, Rodríguez-Rey GT, Álvarez-Mejía LM, Aguirre-Acosta N (2025) Spatial variability in early reproductive characteristics of *Ulex europaeus* L. (Fabaceae) invading in the high-mountain of central Colombia. *Acta Oecologica* 128: 104098. <https://doi.org/10.1016/j.actao.2025.104098>
- Baduel P, Bray S, Vallejo-Marín M, Kolář F, Yant L (2018) The “polyploid hop”: shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Frontiers in Ecology and Evolution* 6: 117. <https://doi.org/10.3389/fevo.2018.00117>
- Baker RL, Yarkhunova Y, Vidal K, Ewers BE, Weinig C (2017) Polyploidy and the relationship between leaf structure and function: implications for correlated evolution of anatomy, morphology, and physiology in Brassica. *BMC Plant Biology* 17: 3. <https://doi.org/10.1186/s12870-016-0957-3>
- Baptiste MP, Castaño N, Cárdenas López D, Gutiérrez FDP, Gil D, Lasso CA (2010) Análisis de riesgo y propuesta de categorización de especies introducidas para Colombia. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá.
- Barrett SCH, Colautti RI, Eckert CG (2008) Plant reproductive systems and evolution during biological invasion. *Molecular Ecology* 17(1): 373–383. <https://doi.org/10.1111/j.1365-294X.2007.03503.x>
- Bellot S, Dias PMB, Affagard M, Ainouche M-L, Misset M-T, Ainouche A (2023) Molecular phylogenetics shed light on polyploid speciation in gorses (*Ulex*, Fabaceae: Genisteae) and on the origin of the invasive *Ulex europaeus*. *Botanical Journal of the Linnean Society* 202(1): 52–75. <https://doi.org/10.1093/botlinnean/boac061>
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57(1): 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bian Y, Yang C, Ou X, Zhang Z, Wang B, Ma W, Gong L, Zhang H, Liu B (2018) Meiotic chromosome stability of a newly formed allohexaploid wheat is facilitated by selection under abiotic stress as a spandrel. *New Phytologist* 220(1): 262–277. <https://doi.org/10.1111/nph.15267>
- Bock DG, Caseys C, Cousens RD, Hahn MA, Heredia SM, Hübner S, Turner KG, Whitney KD, Rieseberg LH (2015) What we still don't know about invasion genetics. *Molecular Ecology* 24: 2277–2297. <https://doi.org/10.1111/mec.13032>
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15): 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Broadfield N, McHenry MT (2019) A world of gorse: persistence of *Ulex europaeus* in managed landscapes. *Plants* 8(11): 523–544. <https://doi.org/10.3390/plants8110523>
- Carling MD, Brumfield RT (2007) Gene sampling strategies for multi-locus population estimates of genetic diversity (θ). *PLoS ONE* 2(1): e160. <https://doi.org/10.1371/journal.pone.0000160>
- Catalán P, Segarra-Moragues JG, Palop-Esteban M, Moreno C, González-Candelas F (2006) A Bayesian approach for discriminating among alternative inheritance hypotheses in plant polyploids: the allotetraploid origin of genus *Borderea* (Dioscoreaceae). *Genetics* 172(3): 1939–1953. <https://doi.org/10.1534/genetics.105.042788>
- Chambers EA, Tarvin RD, Santos JC, Ron SR, Betancourth-Cundar M, Hillis DM, Matz MV, Cannatella DC (2023) 2b or not 2b? 2bRAD is an effective alternative to ddRAD for phylogenomics. *Ecology and Evolution* 13(3): e9842. <https://doi.org/10.1002/ece3.9842>
- Chiu JH, Chong KY, Lum SKY, Wardle DA (2023) Trends in the direction of global plant invasion biology research over the past two decades. *Ecology and Evolution* 13(1): e9690. <https://doi.org/10.1002/ece3.9690>
- Christina M, Limbada F, Atlan A (2020) Climatic niche shift of an invasive shrub (gorse, *Ulex europaeus*): a world-scale comparison in native and introduced regions. *Journal of Plant Ecology* 13(1): 42–50. <https://doi.org/10.1093/jpe/rtz041>
- Coates BS, Sumerford DV, Miller NJ, Kim KS, Sappington TW, Siegfried BD, Lewis LC (2009) Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity* 100(5): 556–564. <https://doi.org/10.1093/jhered/esp028>
- Collin FD, Durif G, Raynal L, Lombaert E, Gautier M, Vitalis R, Marin J-M, Estoup A (2021) Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. *Molecular Ecology Resources* 21: 2598–2613. <https://doi.org/10.1111/1755-0998.13413>
- Comai L (2005) The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* 6: 836–846. <https://doi.org/10.1038/nrg1711>
- Cornille A, Salcedo A, Kryvokhyzha D, Glémin S, Holm K, Wright SI, Lascoux M (2016) Genomic signature of successful colonization of Eurasia by the allopolyploid shepherd's purse (*Capsella bursa-pastoris*). *Molecular Ecology* 25(2): 616–629. <https://doi.org/10.1111/mec.13491>
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. *Gigascience* 10(2): giab008. <https://doi.org/10.1093/gigascience/giab008>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19(1): 11–15.
- Duan T, Sicard A, Glémin S, Lascoux M (2024) Separating phases of allopolyploid evolution with resynthesized and natural *Capsella bursa-pastoris*. *eLife* 12: RP88398. <https://doi.org/10.7554/eLife.88398.3>
- Eaton DAR, Overcast I (2020) ipyrad: interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36(8): 2592–2594. <https://doi.org/10.1093/bioinformatics/btz966>
- Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants. *Proceedings of the National Academy of Sciences* 97(1): 7043–7050. <https://doi.org/10.1007/s10681-006-5939-3>

- Estoup A, Ravigné V, Hufbauer R, Vitalis R, Gautier M, Facon B (2016) Is there a genetic paradox of biological invasion? Annual Review of Ecology, Evolution, and Systematics 47: 51–72. <https://doi.org/10.1146/annurev-ecolsys-121415-032116>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14(8): 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Fan M, Gao Y, Wu Z, Haider S, Zhang Q (2022) Evidence for hexasomic inheritance in *Chrysanthemum morifolium* Ramat. based on analysis of EST-SSR markers. Genome 65(2): 75–81. <https://doi.org/10.1139/gen-2020-0155>
- Feng D, Cheng J, Yang X, Tian Z, Liu Y, Zhang Y, Qiang S (2024) Polyploidization-enhanced effective clonal reproduction endows the successful invasion of *Solidago canadensis*. Ecological Applications 34: e2738. <https://doi.org/10.1002/eap.2738>
- Fraimout A, Debat V, Fellous S, Hufbauer RA, Foucaud J, Pudlo P, Marin J-M, Price DK, Cattell J, Chen X, Deprá M, Duyck PF, Guedot C, Kenis M, Kimura MT, Loeb G, Loiseau A, Martinez-Sañudo I, Pascual M, Polihronakis Richmond M, Shearer P, Singh N, Tamura K, Xuéreb A, Zhang J, Estoup A (2017) Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. Molecular Biology and Evolution 34(4): 980–996. <https://doi.org/10.1093/molbev/msx050>
- Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28(23): 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Galili T (2015) dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. Bioinformatics 31(22): 3718–3720. <https://doi.org/10.1093/bioinformatics/btv428>
- Gao M, Hua T, Niu G, Masabni J, Dewalt W (2024) A locus-dependent mixed inheritance in the segmental allohexaploid sweetpotato (*Ipomoea batatas* [L.] Lam). Frontiers in Plant Science 15: 1398081. <https://doi.org/10.3389/fpls.2024.1398081>
- Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. Preprint. arXiv:1207.3907. <https://doi.org/10.48550/arXiv.1207.3907>
- Geleta M, Ortiz R (2016) Molecular and genomic tools provide insights on crop domestication and evolution. Advances in Agronomy 135: 181–223. <https://doi.org/10.1016/bs.agron.2015.09.005>
- Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5(1): 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Gouldthorpe J, Austen L, Moore J, Poulish G, Sandiford L, Ireson J, et al. (2006) Gorse national best practice manual—managing gorse (*Ulex europaeus* L.) in Australia. Department of Agriculture, Hobart.
- Hagenblad J, Hülskötter J, Acharya KP, Brunet J, Chabrierie O, Cousins SAO, Dar PA, Diekmann M, De Frenne P, Hermy M, Jamoneau A, Kolb A, Lemke I, Plue J, Reshi ZA, Graae BJ (2015) Low genetic diversity despite multiple introductions of the invasive plant species *Impatiens glandulifera* in Europe. BMC Genetics 16(1): 103. <https://doi.org/10.1186/s12863-015-0242-8>
- Hirsch H, Castillo ML, Impson FAC, Kleinjan C, Richardson DM, Le Roux JJ (2019) Ghosts from the past: even comprehensive sampling of the native range may not be enough to unravel the introduction history of invasive species—the case of *Acacia dealbata* invasions in South Africa. American Journal of Botany 106(3): 352–362. <https://doi.org/10.1002/ajb2.1244>
- Hornoy B, Atlan A, Roussel V, Buckley YM, Tarayre M (2013) Two colonisation stages generate two different patterns of genetic diversity within native and invasive ranges of *Ulex europaeus*. Heredity 111(5): 355–363. <https://doi.org/10.1038/hdy.2013.53>
- Hozawa M, Nawata E (2021) Assessment of the genetic diversity of *Ulex europaeus* in Maui, California, Hawaii and New Zealand by a method of microsatellite markers. Biology and Life Sciences Forum 4(1): 5. <https://doi.org/10.3390/IECPS2020-08564>
- Jackson DA, Somers KM (1989) Are probability estimates from the permutation model of Mantel's test stable? Canadian Journal of Zoology 67(3): 766–769. <https://doi.org/10.1139/z89-108>
- Jiang H, Zhang Y, Tu W, Sun G, Wu N, Zhang Y (2023) The general trends of genetic diversity change in alien plants. Invasion Plants 12(14): 2690. <https://doi.org/10.3390/plants12142690>
- Jiao X, Wu L, Zhang D, Wang H, Dong F, Yang L, Wang S, Amano HE, Zhang W, Jia C, Rheindt FE, Lei F, Song G (2024) Landscape heterogeneity explains the genetic differentiation of a forest bird across the Sino-Himalayan Mountains. Molecular Biology and Evolution 41(3): msae027. <https://doi.org/10.1093/molbev/msae027>
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24(11): 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21): 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling. <https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf> [accessed 27.01.2026]
- Kamvar ZN, Tabima JE, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2: e281. <https://doi.org/10.7717/peerj.281>
- Keller SR, Fields PD, Berardi, AE, Taylor DR (2014) Recent admixture generates heterozygosity–fitness correlations during the range expansion of an invading species. Journal of Evolutionary Biology 27: 616–627. <https://doi.org/10.1111/jeb.12330>
- Kirk H, Paul J, Straka J, Freeland JR (2011) Long-distance dispersal and high genetic diversity are implicated in the invasive spread of the common reed, *Phragmites australis* (Poaceae), in northeastern North America. American Journal of Botany 98(7): 1180–1190. <https://doi.org/10.3732/ajb.1000278>
- Knaus BJ, Grünwald NJ (2017) vcfr: a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources 17(1): 44–53. <https://doi.org/10.1111/1755-0998.12549>
- Laginhas BB, Fertakos ME, Bradley BA (2023) We don't know what we're missing: evidence of a vastly undersampled invasive plant pool. Ecological Applications 33(2): e2776. <https://doi.org/10.1002/eap.2776>
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proceedings of the National Academy of Sciences 104(10): 3883–3888. <https://doi.org/10.1073/pnas.0607324104>

- Lee S-R, Son DC (2022) Genetic diversity pattern reveals the primary determinant of burcucumber (*Sicyos angulatus* L.) invasion in Korea. *Frontiers in Plant Science* 13: 997521. <https://doi.org/10.3389/fpls.2022.997521>
- León-Cordero R, Torchelsen FP, Overbeck GE, Anand M (2016) Analyzing the landscape characteristics promoting the establishment and spread of gorse (*Ulex europaeus*) along roadsides. *Ecosphere* 7(3): e01201. <https://doi.org/10.1002/ecs2.1201>
- Madelón M, Aguirre-Acosta N, Acosta MC, Montti L, Qi W, Aguilar R (2021) Genetic reconstruction of potential invasion pathways of *Ligustrum lucidum* into Argentina. *Acta Oecologica* 111: 103733. <https://doi.org/10.1016/j.actao.2021.103733>
- Madlung A (2013) Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity* 110: 99–104. <https://doi.org/10.1038/hdy.2012.79>
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27(2_Part_1): 209–220.
- Mounger J, Ainouche ML, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards CL (2021) Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society B* 376(1826): 20200117. <https://doi.org/10.1098/rstb.2020.0117>
- Moura RF, Queiroga D, Vilela E, Moraes AP (2021) Polyploidy and high environmental tolerance increase the invasive success of plants. *Journal of Plant Research* 134: 105–114. <https://doi.org/10.1007/s10265-020-01236-6>
- Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al (2022) *vegan*: community ecology package (Version 2.6-4). <https://doi.org/10.32614/CRAN.package.vegan>
- Onoue N, Kono A, Azuma A, Matsuzaki R, Nagano AJ, Sato A (2022) Inbreeding depression in yield-related traits revealed by high-throughput sequencing in hexaploid persimmon breeding populations. *Euphytica* 218: 125. <https://doi.org/10.1007/s10681-022-03073-1>
- Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401–437. <https://doi.org/10.1146/annurev.genet.34.1.401>
- Pauchard A, Kueffer C, Dietz H, Daehler CC, Alexander J, Edwards PJ, Arévalo JR, Cavieres LA, Guisan A, Haider S, Jakobs G, McDougall K, Millar CI, Naylor BJ, Parks CG, Rew LJ, Seipel T (2009) Ain't no mountain high enough: plant invasions reaching new elevations. *Frontiers in Ecology and the Environment* 7(9): 479–486. <https://doi.org/10.1890/080072>
- Pearman WS, Urban L, Alexander A (2022) Commonly used Hardy–Weinberg equilibrium filtering schemes impact population structure inferences using RADseq data. *Molecular Ecology Resources* 22: 2599–2613. <https://doi.org/10.1111/1755-0998.13646>
- Pembleton LW, Cogan NOI, Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* 13: 946–952. <https://doi.org/10.1111/1755-0998.12129>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Pritchard JK, Wen X, Falush D (2010) Documentation for Structure Software: Version 2.3. University of Chicago, Chicago, IL.
- R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/> [accessed 02.10.2023]
- Ramsey J, Schemske DW (2002) Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* 33: 589–639. <https://doi.org/10.1146/annurev.ecolsys.33.010802.150437>
- Ray A, Quader S (2014) Genetic diversity and population structure of *Lantana camara* in India indicates multiple introductions and gene flow. *Plant Biology* 16(3): 651–658. <https://doi.org/10.1111/plb.12087>
- Ríos H (2005) Guía Técnica para la Restauración Ecológica de Áreas Afectadas por Especies Vegetales Invasoras en el Distrito Capital. Subdirección Científica Grupo de Ecología de la Restauración. Jardín Botánico de Bogotá José Celestino Mutis.
- Roberts J, Florentine S (2021) Biology, distribution and control of the invasive species *Ulex europaeus* (Gorse): a global synthesis of current and future management challenges and research gaps. *Weed Research* 61(4): 272–281. <https://doi.org/10.1111/wre.12491>
- Rodríguez D, Ocampo R, Malambo N, Barrera JI, Rojas JE, Contreras SM, Moreno AC (2019) Plan de Prevención, Manejo y Control de Retamo Espinoso (*Ulex europaeus*) y Retamo Liso (*Genista monspessulana*) en la Jurisdicción CAR. Pontificia Universidad Javeriana y Corporación Autónoma Regional de Cundinamarca.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145(4): 1219–1228. <https://doi.org/10.1093/genetics/145.4.1219>
- Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8(1): 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Sapkota S, Boggess SL, Trigiano RN, Klingeman WE, Hadziabdic D, Coyle DR, Nowicki M (2022) Microsatellite loci reveal high genetic diversity, mutation, and migration rates as invasion drivers of callery pear (*Pyrus calleryana*) in the southeastern United States. *Frontiers in Genetics* 13: 861398. <https://doi.org/10.3389/fgene.2022.861398>
- Schrieber K, Lachmuth S (2017) The genetic paradox of invasions revisited: the potential role of inbreeding × environment interactions in invasion success. *Biological Reviews* 92(2): 939–952. <https://doi.org/10.1111/brv.12263>
- Scott AD, Van de Velde JD, Novikova PY (2023) Inference of polyploid origin and inheritance mode from population genomic data. In: Van de Peer Y (Ed.) *Polyploidy: Methods and Protocols*. Springer US, New York, 279–295. https://doi.org/10.1007/978-1-0716-2561-3_15
- Shang L, Li L-F, Song Z-P, Wang Y, Yang J, Wang C-C, Qiu S-Y, Huang J-X, Nie M, Wolfe LM, Li B (2019) High genetic diversity with weak phylogeographic structure of the invasive *Spartina alterniflora* (Poaceae) in China. *Frontiers in Plant Science* 10: 1467. <https://doi.org/10.3389/fpls.2019.01467>
- Soltis DE, Soltis PS, Rieseberg LH (1993) Molecular data and the dynamic nature of polyploidy. *CRC Critical Reviews in Plant Sciences* 12: 243–273. <https://doi.org/10.1080/07352689309701903>
- Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60: 561–588. <https://doi.org/10.1146/annurev.arplant.043008.092039>
- Stout JC, Duffy KJ, Egan PA, Harbour M, Hodkinson TR (2015) Genetic diversity and floral width variation in introduced and native populations of a long-lived woody perennial. *AoB Plants* 7: plu087. <https://doi.org/10.1093/aobpla/plu087>
- Sun S, Lu B, Lu X, Huang S (2018) On reproductive strategies of invasive plants and their impacts on native plants. *Biodiversity Science* 26(5): 457–467. <https://doi.org/10.17520/biods.2017294>

- Tang J-S, Ma M (2020) Genetic diversity and genetic differentiation of invasive weed *Xanthium italicum* in China. *Comptes Rendus. Biologies* 343(1): 63–72. <https://doi.org/10.5802/crbio.7>
- Te Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubešová M, Pyšek P (2012) The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany* 109(1): 19–45. <https://doi.org/10.1093/aob/mcr277>
- Vargas O, León O, Díaz-Espinosa A (2009) Restauración Ecológica en Zonas Invasidas por Retamo Espinoso y Plantaciones Forestales de Especies Exóticas. Universidad Nacional de Colombia, Facultad de Ciencias Departamento de Biología, Bogotá, D.C., 1–305.
- Verhoeven KJF, Macel M, Wolfe LM, Biere A (2011) Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B: Biological Sciences* 278(1702): 2–8. <https://doi.org/10.1098/rspb.2010.1272>
- Wang Y-J, Müller-Schärer H, van Kleunen M, Cai A-M, Zhang P, Yan R, Dong B-C, Yu F-H (2017) Invasive alien plants benefit more from clonal integration in heterogeneous environments than natives. *New Phytologist* 216(4): 1072–1078. <https://doi.org/10.1111/nph.14820>
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38(6): 1358–1370. <https://doi.org/10.2307/2408641>
- Zimmermann H, Ritz CM, Hirsch H, Renison D, Wesche K, Hensen I (2010) Highly reduced genetic diversity of *Rosa rubiginosa* L. populations in the invasive range. *International Journal of Plant Sciences* 171(4): 435–446. <https://doi.org/10.1086/651244>

SUPPLEMENTARY MATERIALS

Supplementary material 1

Additional tables and figures supporting the analyses.

<https://doi.org/10.5091/plecevo.165188.suppl1>

Supplementary material 2

VCF file for the diploid dataset.

<https://doi.org/10.5091/plecevo.165188.suppl2>

Supplementary material 3

VCF file for the hexaploid dataset.

<https://doi.org/10.5091/plecevo.165188.suppl3>