Patterns of Interploidy Admixture in Polyploid Complexes: Insights from *Thymus* sect. *Mastichina* (Lamiaceae)

Francisco José García-Cárdenas[1] | María Ángeles Ortiz[1] | José Carlos del Valle[1] | David Doblas-Pruvost[1] | Manuel de la Estrella[2] | Diego Nieto-Lugilde[2] | Lisa Pokorny[3] | Regina Berjano[1]

[1]Departamento de Biología Vegetal y Ecología, Universidad de Sevilla, Avda. Reina Mercedes s/n, 41012 Seville, Spain | [2]Departamento de Botánica, Ecología y Fisiología Vegetal, Universidad de Córdoba, Campus de Rabanales, Edificio Celestino Mutis, E-14071 Córdoba, Spain. | [3]Real Jardín Botánico (RJB), CSIC, 28014 Madrid, Spain

Lisa Pokorny and Regina Berjano should be considered joint senior authors

Correspondence: Regina Berjano (regina@us.es)

Running title: Interploidy Admixture in Polyploid Complexes

ABSTRACT

Understanding gene flow between ploidy levels in polyploid complexes is essential for species delimitation and conservation. This study explores evolutionary dynamics in the polyploid complex *Thymus* sect. *Mastichina* (Lamiaceae), comprising three taxa: *T. mastichina* subsp. *mastichina*, *T. mastichina* subsp. *donyanae*, and the endangered *T. albicans*. Using Hyb-Seq data, phylogenomics (nuclear orthologs), and population genomics (SNPs), we confirm the section consists of two sister groups with distinct ploidy levels: a diploid and a tetraploid one. The tetraploid group shows low genetic differentiation among its populations, probably indicating rapid expansion across diverse environments. In contrast, the diploid group exhibits more complex genetic structuring, potentially shaped by geomorphology, interploidy introgression, and incipient isolation. Four diploid subgroups (Algarve, Cádiz, Doñana, and Hercynian) are identified, with reticulate evolution. The dense reticulation observed is compatible with incomplete lineage sorting in diploid lineages, due to recent and rapid divergence events. Phylogeographic analyses suggest isolation-by-distance, with two major riverbeds maybe playing a role in shaping genetic differentiation, while interploidy gene

flow detected could have facilitated ancient and/or ongoing admixture between diploid and tetraploid lineages, despite geographic isolation. These findings highlight cryptic genetic diversity and emphasise the need for an integrative taxonomy that includes multiple lines of evidence: morphological, cytological, genomic, and ecological. Conservation efforts should prioritise protecting the four diploid subgroups, aided by flow cytometry, since they may harbour critical adaptive potential to both specific habitat types and/or environmental conditions. This work contributes to advancing our knowledge of evolution in polyploid complexes, by combining genomic approaches and highlighting cryptic diversity in *Thymus* species. Future research should investigate morphometric and chemical data, hybridisation events, divergence times, diversification dynamics, and relationships with other *Thymus* species to further understand polyploid evolution and its impact on biodiversity.

Keywords: Angiosperms353 | Hyb-Seq | Interploidy Admixture | Phylogenomics | Polyploid Bridges | Population Genomics

1 | Introduction

Understanding the mechanisms driving speciation is a central focus of evolutionary biology. Polyploidy is considered one such mechanism of speciation with profound implications for plant evolution and ecology, which is mediated through whole-genome duplication (WGD; Van de Peer *et al.* 2017, 2021). Although this process has often been regarded as an evolutionary 'dead end', due to the limited evidence of ancient polyploidisation events (Arrigo and Barker 2012; Mayrose *et al.* 2015; but see Leebens-Mack *et al.* 2019), the existence of many polyploids of relatively recent origin supports its potential role as a driver of plant diversification (Vanneste *et al.* 2014a, 2014b; Soltis *et al.* 2015; Van de Peer *et al.* 2017, 2021). Even considering that polyploid lineages frequently face extinction, it seems that at least in rare cases they successfully establish and undergo significant diversification, as the "rarely successful hypothesis" postulates (Arrigo and Barker 2012; Hagen and Beaulieu 2024).

Polyploidy can arise through WGD within a single species (autopolyploidy) or after hybridisation (allopolyploidy), a well-established and prevalent process in plants that facilitates gene flow between distinct lineages that contributes to the emergence of novel species (Soltis and Soltis 2009; Sheidai and Koohdar 2023). Although newly established polyploids may initially experience reduced fertility and fitness due to genomic instability or changes in gene expression and epigenetic changes (Comai 2005; Porturas *et al.* 2019), once these issues are overcome, their potential for rapid adaptation can confer a selective advantage. At short-term, the success of polyploids can be explained by increased genetic variation, mediated by changes in gene expression and epigenetic remodelling, which influence the morphology, physiology, and ecology of newly formed polyploids (Schoenfelder and Fox 2015), and also by increased genetic robustness (Van de Peer *et al.* 2017). These changes, in turn, may affect interspecies interactions, confer tolerance to a broader range of ecological and environmental conditions, facilitate the colonisation of new habitats, and increase the potential for invasiveness (te Beest *et al.* 2012; Soltis *et al.* 2014; Van de Peer *et al.* 2021).

Polyploidy can promote sympatric speciation in plants, by acting as a natural reproductive barrier, as offspring from newly formed polyploids and their lower-ploidy progenitor(s) often exhibit unbalanced chromosome numbers, leading to sterility (the so-called *triploid block*; Ramsey and Schemske 1998; Otto and Whitton 2000; Adams and Wendel 2005). Nevertheless, this barrier to interploidy hybridization, it is sometimes leaky (for a review see Salony *et al.* 2025). Thus, the establishment of polyploid species in sympatry likely depends on prezygotic barriers between newly formed polyploids and their diploid progenitor(s), along with shifts from cross- to self-pollination or from sexual to asexual reproduction, which can mitigate cytotype minority exclusion of recently formed polyploids (Van de Peer *et al.* 2017). However, reproductive barriers following WGD may not be complete, particularly in cases where intermediate cytotypes are viable, which could facilitate the backcrossing with parental cytotypes (the so-called *triploid bridge pathway*; Ramsey and Schemske 1998; Husband 2004; Bartolić *et al.* 2025). Such backcrossing events may enable bidirectional gene flow, if the formed fertile interploidy hybrids are capable of

backcrossing with their parentals (Kolář *et al.* 2017; Karbstein *et al.* 2022; Hodač *et al.* 2023; Bartolić *et al.* 2024). Otherwise, unreduced gametes could also mediate unidirectional gene flow from diploids to tetraploids, thus enabling interploidy gene flow (Leal *et al.* 2024).

Interploidy gene flow between lineages has turned out to characterise the evolution of many polyploid complexes (Nieto Feliner *et al.* 2020 and references therein), in which species with different ploidy levels interchange genetic material without apparently altering the chromosome number. Introgression and polyploidisation, although generally addressed in the literature as separate biological phenomena, are not mutually exclusive and often coexist in nature (Leal *et al.* 2024). These introgression events allow receptor species to accumulate new alleles not only through introgression with closely diploid progenitors, but also from more distantly related diploid or polyploid donor lineages, contributing to genetic diversity and facilitating adaptive allele emergence (Suarez-Gonzalez *et al.* 2018; Seear *et al.* 2020; Nieto Feliner *et al.* 2020; Schmickl and Yant 2021; Nibau *et al.* 2022). Despite growing recognition of their interplay, significant gaps remain in our understanding of how interploidy admixture in polyploid complexes shape evolutionary trajectories. Additionally, as polyploid complexes and introgression often blur phylogenetic relationships and obscure phenotypic distinctions (along with other factors such as incomplete lineage sorting, ILS), traditional taxonomic species differentiation can be challenging (Mallet 2007; Soltis *et al.* 2009; Al-Shehbaz 2012; Han *et al.* 2015). This is especially relevant for cryptic and threatened species because accurate identification assures the correct assessment and implementation of conservation strategies (Bickford *et al.* 2007; Fišer *et al.* 2018; Hending 2025).

Polyploidisation, introgression, and hybridisation, are common phenomena in several plant families, including Lamiaceae, and particularly in thyme—genus *Thymus* L. (e.g., Šoštarić *et al.* 2012; Yousefi *et al.* 2015). This genus is notable for its high species diversity (WFO 2025), significant economic value, high level of endemisms, and ecological importance, as thyme formations represent one of the most characteristic communities of Mediterranean shrublands (Morales 1986, 2002, 2010). In the Iberian Peninsula, *Thymus* sect. *Mastichina* presents an excellent model for studying evolutionary patterns driven by polyploidisation and introgression processes, which could be shaping relationships among their components. The section includes two closely related species (three taxa), whose populations are sometimes quite challenging to differentiate morphologically: *Thymus mastichina* (L.) L. and *T. albicans* Hoffmanns. & Link. *Thymus mastichina* is one of the most widely distributed endemic thyme species in the Iberian Peninsula. This species is known to hybridise with a large number of taxa with overlapping ranges (Morales 1986, 2010). Two subspecies have been described within it: *T. mastichina* subsp. *mastichina* (tetraploid), an Iberian widespread taxon, and *T. mastichina* subsp. *donyanae* (diploid), whose populations are mainly located within Doñana National and Natural Parks, in SW Spain (Morales 2010). The other species is *T. albicans* (diploid), found along the southwestern coast of the Iberian Peninsula across habitats highly threatened by coastal urban development (Valdés Castrillón *et al.* 1999). Taxonomists have differentiated these three taxa using morphological traits, mainly based on inflorescence diameter and calyx length (Morales 1986, 2002). However, this group exhibits a high degree of phenotypic variation, at both morphological (Morales 2010;

Méndez-Tovar *et al.* 2015) and chemotypic levels (e.g., Miguel *et al.* 2003; Filipe *et al.* 2017). Therefore, unravelling the intricate relationships of this polyploid complex—essential for clarifying its taxonomic status—should be a priority, as it would help guide future conservation strategies (Hending 2025).

As High-Throughput Sequencing (HTS) technologies have become more accessible (Metzker 2010), population genomics methods have become an essential tool for reconstructing the evolutionary patterns that shape complex species histories (DePristo *et al.* 2011; Ellegren 2014). Among HTS techniques, genome skimming and target-enrichment have revolutionised our understanding of plant systematics (Mamanova *et al.* 2010; Dodsworth 2015; Pezzini *et al.* 2023). Hyb-Seq, which combines these two approaches (Weitemier *et al.* 2014; Dodsworth *et al.* 2019), has gained prominence in plant phylogenomics due to its affordability (reasonable per-sample cost), reproducibility, and replicability, while offering high resolution regardless of tissue-template quality (Villaverde *et al.* 2018; Brewer *et al.* 2019; Hale *et al.* 2020; Shee *et al.* 2020; Zuntini *et al.* 2024). While other genomic techniques, such as restriction site-associated DNA sequencing (RADseq) or genotyping-by-sequencing (GBS), have been used extensively to explore population structure and genetic diversity (Baird *et al.* 2008; Elshire *et al.* 2011; Peterson *et al.* 2012; Campos *et al.* 2024), the potential application of Hyb-Seq in population genetics still remains largely underexplored (Baker *et al.* 2021; McDonnell *et al.* 2021; but see Ousmael and Hansen 2024, Woudstra *et al.* 2025, or Viruel *et al.* 2025 for recent applications). However, current computational tools in phylogenomics and population genomics have been shown to effectively bridge micro- and macroevolutionary levels, using the same Hyb-Seq data, which can be encoded as either nuclear ortholog sequences or single-nucleotide polymorphism (SNPs; Beck *et al.* 2021; Slimp *et al.* 2021; Wenzell *et al.* 2021; Cooper *et al.* 2023; Phang *et al.* 2023; Tiley *et al.* 2024; Rincón Barrado *et al.* 2024a, 2024b; Balant *et al.* 2025; Pardo Otero *et al.* 2024; Freund *et al.* 2025; Liang *et al.* 2025; Stefani *et al.* 2025). By combining these analytical methods, insights into genetic connectivity, introgression, and the microevolutionary patterns of populations can be gained, which can be useful for identifying populations at risk of inbreeding or loss of adaptive potential and help in conservation and biodiversity management.

In this study, we aim to understand the evolutionary dynamics and genetic structure underlying the *Thymus* sect. *Mastichina* polyploid complex, with particular focus on the extent and patterns of interploidy admixture. To achieve this, we employ Hyb-Seq data in combination with phylogenomics and population genomics approaches to (i) resolve phylogenetic relationships among taxa in this section, (ii) assess genetic differentiation and admixture among diploid and tetraploid taxa within the complex, (iii) investigate signals of gene flow and introgression within and across ploidy levels, including potential pathways leading to interploidy admixture, and (iv) evaluate the implications of these patterns for taxonomy and conservation of *Thymus* sect. *Mastichina* taxa. This integrative framework seeks to clarify processes at the micro-macroevolutionary frontier shaping species boundaries in the presence of gene flow across ploidy levels in megadiverse Mediterranean scrubland endemics.

2 | Materials and Methods

2.1 | Study Group and Taxon Sampling

Thymus sect. *Mastichina* comprises two species and three taxa, all endemic to the Iberian Peninsula, differing in their distribution patterns across the region. *Thymus mastichina* subsp. *mastichina*, is a tetraploid taxon (2n = 56, 58, 60) widely distributed throughout the Iberian Peninsula across a variety of habitats such as roadside slopes, abandoned farmlands, pine forests, and mountain scree, growing on loose siliceous substrates but also on calcareous soils. *Thymus mastichina* subsp. *donyanae* is a diploid taxon (2n = 30) mainly located in coastal sandbanks and stabilised dunes within Doñana National and Natural Parks, in SW Spain (Huelva province). *Thymus albicans* is also a diploid taxon (2n = 28, 30) inhabiting stabilised subcostal dunes and coastal pine forests on sandy substrates along the coastline in Cádiz province (Spain) and Algarve (Portugal), SW Iberian Peninsula (Morales 2010). *Thymus albicans* is included in the Spanish national and regional Red Lists (Valdés Castrillón *et al.* 1999; Cabezudo *et al.* 2005), and is as well listed in the International Union for Conservation of Nature (IUCN) Red List (Buira *et al.* 2017).

We sampled a total of 40 populations from May to June 2023 across the Iberian Peninsula (Table 1). Twenty of those populations were the same studied in Girón *et al.* (2012). Individuals of each population were identified following Flora iberica (Morales 2010). Herbarium vouchers of each population were deposited at SEV herbarium (Table 1). Leaves of 8-12 individuals per population were dried in silica gel and stored at room temperature prior to DNA extraction. Additionally, we included three other species of *Thymus* from MA herbarium vouchers as outgroups (*T. carnosus* Boiss., *T. capitellatus* Hoffmanns. & Link and *T. camphoratus* Hoffmanns. & Link).

2.2 | Genome Size and Ploidy Level Assessments

Nuclear DNA amount was assessed from 344 individuals from the 40 studied populations. For these individuals, flow cytometry was performed either on fresh or silica-dried leaf material collected from the same plants used in the genomic workflow, or from seeds originating from those plants. *Lycopersicon esculentum* L. Stupické polní tyčkové rané cultivar (2C = 1.96 pg), and *Raphanus sativus* Saxa cultivar (2C = 1.11 pg) (Doležel *et al.* 1992), provided by the Centre of Plant Structural and Functional Genomics (Institute of Experimental Botany of the Czech Academy of Sciences), were used as internal standard. Leaf tissue of sample as well as internal standard were chopped together with a razor blade in a Petri dish in 1.2 ml of LB01 lysis buffer (Doležel *et al.* 1989), supplemented with Triton X-100 (8%; Girón *et al.* 2012) and 50 μg/ml RNase A, and stained with 50 μg/ml of propidium iodide. The nuclei suspension obtained was filtered through a 30 μm mesh CellTrics disposable filter (Partec GmbH, Munster, Germany), kept on ice for 5–30 min and measured in a flow cytometer CYTOMICS FC500-MP (Beckman Coulter, Fullerton, CA, USA) equipped with a 20 mW Ar-Ion laser at 488 nm. Acquisition was automatically stopped at 20,000 events. The total nuclear DNA content was calculated in each sample by multiplying the DNA content of the

standard by the quotient between the 2C peak positions of the sample and the standard in the histogram of fluorescence intensities. As *Raphanus sativus* slightly but significantly overestimated genome size compared with *Lycopersicon esculentum* ($t = -6.6302$, $df = 49$; $p < 0.01$), we applied a correction factor of 0.95 to all measurements obtained with this standard, based on 50 samples analysed with both standards. Genome size assessments with both fresh material and dried leaves were highly correlated between the two methods (Deming Regression Fit, Pearson's $r = 0.99$; $N = 16$) and that no significant differences were found between paired samples analysed with either method (Paired t-test, $t = -0.19$, $df = 15$, $p = 0.853$). Mean 2C values and standard deviations were calculated for each individual plant and for each population. Ploidy levels were assigned following Girón *et al.* (2012). Additionally, we estimated ploidy level from allelic ratios inferred from SNPs called from our genomic sequencing data (see section 2.4.4 below).

2.3 | Molecular Protocols

2.3.1 | DNA Extraction and Genomic Library Preparation

DNA extraction was performed in eight individuals of each population. Between 15 and 20 mg of silica-dried foliar tissue per specimen was kept at −80ºC for 15 min and then milled with a Mixer Mill MM301 (Retsch Inc., Haan, NRW, Germany). We used the DNeasy Plant Mini Kit (Qiagen, Hilden, NRW, Germany) and the GeneMATRIX DNA Purification Kit (EURx, Warsaw, Poland) for extraction of total cellular DNA, following manufacturer's protocols. Precipitated DNA pellets were resuspended in Milli-Q ultrapure water (Merck KGaA, Darmstadt, HE, Germany). Relative DNA concentration (ng/µL) was measured with a dsDNA Quantification Assay Kit in a Qubit™ 3.0 fluorometer (ThermoFisher Scientific, Waltham, MA, USA). DNA extracts were sonicated with an ultrasonicator Bioruptor® Pico (Diagenode Inc., Liège, Belgium), for 10 cycles with 30" on and 30" off shearing times, to obtain an average fragment size of ~250 bp. Dual-indexed genomic libraries, for Illumina®, were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit and the NEBNext® Multiplex Oligos (Dual Index Primers Sets 1 and 2) from New England Biolabs (Ipswich, MA, USA), at half of the recommended volumes (Hale *et al.* 2020). Size selection was done with AMPure XP magnetic beads (Agencourt, Beverly, MA, USA) and the indexing PCR had 10 cycles in total. Concentration of all genomic libraries was checked using the Qubit® fluorometer, with 10% libraries being profiled for fragment distribution on a 4200 TapeStation System (Agilent Technologies, Santa Clara, CA, USA) with D5000 ScreenTape assays and reagents. Lastly, genomic libraries were normalised (10 nM) and pooled (~24 libs/pool). Each pool contained ~1500 ng DNA for an average fragment size of ~290 bp.

2.3.2 | Hyb-Seq Implementation

Library pools were hybridised with the Angiosperms353 (Johnson *et al.* 2019) expert enrichment panel (Daicel Arbor Biosciences, Ann Arbor, MI, USA). A day-long hybridisation was carried out in an Applied Biosystems™ Veriti™ Thermal Cycler (Waltham, MA, USA), with settings 95ºC 5 min, 65ºC 5 min, and

65ºC ∞. The hybridised, biotin-labelled baits were then captured with streptavidin-coated magnetic beads and further enriched with KAPA HiFi 2X HotStart ReadyMix PCR Kit (Roche, Basel, Switzerland), for 16 cycles, using the i5 and i7 forward and reverse "reamp" primers described in Meyer and Kircher (2010). Amplified PCR products were cleaned with Agencourt AMPure XP magnetic beads, quantified with a Qubit® fluorometer, and profiled on a 4200 TapeStation System. The enriched pools were then normalised and multiplexed, together with the plain genomic libraries, for a full Hyb-Seq implementation (Hale *et al.* 2020; Baker *et al.* 2022). Finally, this multiplexed Hyb-Seq pool (75% enriched library pools + 25% plain genomic libraries) was sequenced on an Illumina® NovaSeq X System producing 2 × 150 bp paired end reads at Macrogen Inc. (Seoul, South Korea).

2.4 | Bioinformatic Analyses

2.4.1 | Data Mining, Target Assembly, and Data Matrix Refinement

Demultiplexed, raw sequencing reads were quality-checked with FastQC 0.12.1 (Andrews 2010) and MultiQC 1.25 (Ewels *et al.* 2016). Low-quality bases, short reads, and remaining adapters were filtered out using fastp 0.23.4 (Chen 2023). Nuclear targets and plastid coding DNA sequences (CDSs) were assembled with HybPiper 2.0.2 (Johnson *et al.* 2016), using the *mega353* target file (McLay *et al.* 2021). Here, we designed an ad hoc plastid CDS target file from the plastome of eight *Thymus* species available from an open access online repository (NCBI 1988). On average, out of ~11M reads/individual, ~1.18M reads/individual mapped to our 353 intended nuclear targets (i.e., ~10.33% reads on target). Target success (measured as percentage of reads mapping to target loci in HybPiper) ranged between 4% and 54%, with ~10.5% target success on average. Potential paralogy, checked with HybPiper, was not detected. Low-coverage (<⅓ median) samples and genes, as measured by the max_overlap.R script (see Shee *et al.* 2020), were removed. Finally, out of 323 individuals initially processed, we retained 304 accessions (301 from *Thymus* sect *Mastichina* plus 3 outgroups) and 332 supercontig sequences per accession, comprising both the exons and their flanking non-coding regions, for subsequent analyses. Detailed statistics from HybPiper and max_overlap.R can be found as a Zenodo repository (see Data Availability Statement below).

2.4.2 | Multiple Sequence Alignment (MSA) and Filtering

Balanced data matrices were aligned with MAFFT 7.520 (Katoh and Standley 2013), using the E-INS-i algorithm, and quality checked with AMAS (Borowiec 2016). Resulting MSAs were used to infer exploratory gene trees with FastTree 2.1.11 (Price *et al.* 2010). TreeShrink 1.2.1 (Mai and Mirarab 2018) was applied to automate removal of outlier branches from FastTree topologies and their corresponding MSAs. AMAS was used to select the threshold that maximised the proportion of parsimony-informative characters (P_{PIC}), while minimising the percent of missing data. Target orthologs with P_{PIC} values <⅓ of the median were excluded. As a result, 301 individuals (298 from *Thymus* sect *Mastichina* plus 3 outgroups) and 329 supercontigs were retained for downstream analysis. Finally, the selected outlier-filtered MSAs were

realigned again with MAFFT and further refined with trimAl 1.4.1 (Capella-Gutierrez *et al.* 2009). AMAS summaries can be found in the Zenodo repository (see Data Availability Statement).

2.4.3 | Gene and Species Tree Inference

Nuclear ML gene trees were estimated with IQ-TREE 2.2.0 (Minh *et al.* 2020). Selection of best-fit model of sequence evolution was done by ModelFinder Plus (Kalyaanamoorthy *et al.* 2017) and support values were calculated (1,000 bootstrap replicates) using both UFBoot2 (Hoang *et al.* 2018) and SH-aLRT (Anisimova *et al.* 2011) algorithms. Unsupported branch bipartitions (BS = 0) across ML gene trees were collapsed (see Mirarab 2019; Simmons and Gatesy 2021) using the Newick utilities 1.6.0 package (Junier and Zdobnov 2010). The species tree was estimated with ASTRAL III 5.7.8 (Zhang *et al.* 2018) under the multispecies coalescent (MSC) framework with branch support estimated as local posterior probabilities (LPP). Finally, the MSC species tree was scaled to present branch lengths in substitutions per site using RAxML-NG 1.2.2 (Kozlov *et al.* 2019). This scaled species tree was visualised using FigTree 1.4.4 (Rambaut 2018). As the plastome represents a canonical coalescent gene (c-gene; Doyle 2022), outlier-filtered, trimmed supercontig MSAs were concatenated using AMAS. Then, a plastid ML tree was inferred using IQ-TREE 2.2.0.

2.4.4 | SNP Calling

All *Thymus* sect. *Mastichina* accessions included in the phylogenomic analyses (298) were initially used for SNP calling, adapting the workflow designed by Slimp *et al.* (2021), with modifications (see Balant *et al.* 2025). Outgroups were excluded as this analysis was intended to examine genetic variation that segregates exclusively within the study populations. The variant detection was carried out with GATK4 4.5.0.0 (McKenna *et al.* 2010), using the longest (most-complete) supercontig sequence (assembled with HybPiper) for each of the 353 targets as a pseudo-genome reference. We opted for a conservative approach at SNP calling, by eliminating multi-allelic variants. A total of 82,137K single nucleotide polymorphisms (SNPs) were initially called. After a robust filtering, the SNP matrix was processed using BCFtools 1.20 (Danecek *et al.* 2021) and VCFtools 0.1.16 (Danecek *et al.* 2011). Linkage disequilibrium was avoided using PLINK 2.0 (Chang *et al.* 2015) with the command *--indep-pairwise* 50 5 0.2. Specifically, the command works within 50 variant sliding windows, by a step of 5 SNPs. Inside these windows, when two SNPs show r^2 correlation values higher than 0.2, one of them is removed. The final filtered SNP data matrix consisted of 254 individuals and 2,223 unlinked SNPs.

2.4.5 | Ploidy Level Estimation Based on Genomic Allelic Ratios

We estimated the ploidy level from our samples following the workflow designed by Viruel *et al.* (2019, 2023). Briefly, allelic frequencies were estimated with nQuire (Weiß *et al.* 2018). These frequencies were used to calculate allelic ratios and delta log-likelihood values per sample using the dplyr 1.1.4 (Wickham *et al.* 2023) and devEMF 4.5 packages (Johnson 2024) in R 4.3.2 (R Core Team 2022). Ploidy was then

inferred using model-based approaches (Corrêa dos Santos *et al.* 2017), selecting the ploidy level corresponding to the model with the lowest log-likelihood value. These estimates were compared to our flow cytometry measurements (see section 2.2 above).

2.4.6 | Population Structure Analyses and Summary Statistics

Using the SNP matrix above, population structure analyses were carried out in ADMIXTURE 1.3.0 (Alexander *et al.* 2009; Alexander and Lange 2011), and the optimal number of clusters was estimated based on cross-validation (CV) error (metric that reflects the model's predictive accuracy for different values of K). In parallel, STRUCTURE 2.3.4 (Pritchard *et al.* 2000), as implemented in the ipyrad toolkit 0.9.102 (Eaton and Overcast 2020), was also used to estimate the optimal number of genetic clusters. A range of K values (2–10) was tested in ten independent runs and the most likely number of clusters was selected by detecting the highest values of the ΔK statistic (Evanno *et al.* 2005). ADMIXTURE results for each K value were plotted using the pophelperShiny 2.1.1 R package (Francis 2017). R packages mapmixture 1.2.0 (Jenkins 2024) and ggplot2 (Wickham 2016) were used to obtain pie charts and project them onto a map obtained from QGIS 3.22 (QGIS.org 2024). Principal Component Analysis (PCA) and Discriminant Analysis of Principal Components (DAPC) were conducted in R package adegenet 1.3–1 (Jombart and Ahmed 2011). R packages poppr 2.9.6 (Kamvar *et al.* 2014) and ape 5.0 (Paradis and Schliep 2019) were also used to visualise the results.

The fixation index (F_{ST}), inbreeding coefficient (F_{IS}), and pairwise identity-by-descent (IBD) were calculated for each pair of the detected phylogenetic groupings and for each population using PLINK 2.0 (Hudson *et al.* 1992; Bhatia *et al.* 2013; Chang *et al.* 2015). Genetic variability within and among phylogenetic groupings was tested using analyses of molecular variance (AMOVA) with the poppr 2.9.6 R package. A Mantel test was performed using the vegan 2.6-8 R package (Oksanen *et al.* 2025), based on Pearson's product-moment correlation. Nucleotide diversity (π) and Tajima's D (Tajima 1989) were calculated using VCFtools. To identify private and shared SNPs among the phylogenetic groupings, we used the VennDiagram 1.7.3 R package (Chen and Boutros 2011). Analyses were performed using a random subsample of 11 individuals per grouping, corresponding to the sample size of the smallest grouping, in order to standardise grouping sizes and thus reduce bias in the number of private SNPs detected due to unequal sample size across groups.

2.4.7 | Introgression and Gene Flow Analyses

We reconstructed a Neighbour-Net phylogenetic network with SplitsTree6 (Huson and Bryant 2024) using both the concatenated gene matrix used for phylogenomic analyses and the filtered SNP matrix. This analysis allowed us to visualize the potential reticulation events among samples. Additionally, an ABBA-BABA test was performed among all genetic groupings using Dsuite 0.5 r58, following Malinsky *et al.* (2021), which relies on Patterson's D statistic to estimate the genome-wide excess of shared derived alleles between two given groupings, in order to test for introgression events.

To further evaluate the presence of interploidy gene flow and to quantify its directionality, we applied fastsimcoal 2.8 (Excoffier *et al.* 2021), using Site Frequency Spectrum (SFS) files and the easySFS software (Gutenkunst *et al.* 2009). We compared four scenarios: (1) no gene flow (NOGF), (2) unidirectional gene flow from tetraploids to diploids (MIG21), or (3) from diploids to tetraploids (MIG12), and (4) bidirectional gene flow (MIGBI). We performed 50 runs, estimating parameters using ECM optimization cycles and coalescent simulations. The best-fitting model was selected based on the lowest median AIC value and verified using likelihood distributions. Finally, we used block-bootstrapping with 100 replicates to calculate confidence intervals for the parameter estimates, ensuring the reliability of our findings (Canty and Ripley 2024).

3 | Results

3.1. | Genome Size Measurements

DNA content estimated by flow cytometry revealed mean 2C values ranging from 1.32 to 3.07 pg (mean ± SD) in each analysed individual (Table S1). Mean 2C values in populations sampled by Girón *et al.* (2012) were on average 6% lower than those obtained in our study (Table 1), although those differences were low in terms of DNA amount (mean = 0.1 pg). Mean 2C values from sampled individuals exhibited a bimodal distribution: samples ranging from 1.32 to 1.84 pg were considered diploid, while those with mean 2C values from 2.31 to 3.07 pg were considered as tetraploid (Table S1), following Girón *et al.* (2012).

3.2 | Phylogenomic Relationships in *Thymus* sect. *Mastichina*

The resulting species tree (Figures 1a and 2) presented *Thymus* sect. *Mastichina* divided into two phylogenetic groups (LPP = 0.78), respectively corresponding to diploid and tetraploid accessions. Following Morales (2010), the diploid group was composed of accessions identified as either *T. albicans* or *T. mastichina* subsp. *donyanae*, as well as accessions identified as *T. mastichina* subsp. *mastichina* in some populations (Table 1). The tetraploid group was composed of accessions identified as *T. mastichina* subsp. *mastichina*, distributed throughout the Iberian Peninsula, but with high variability in calyx length among and within individuals, also fitting the descriptions *T. mastichina* subsp. *donyanae* and *T. albicans* in relation to calyx length. The diploid group revealed two sister subgroups (LPP = 0.83). On one hand, the Hercynian subgroup, composed of diploid individuals phenotypically fitting *T. mastichina* subsp. *mastichina* from two geographically different areas—Serra da Estrela Natural Park (central Portugal; LPP = 0.98) and Córdoba (S Spain; LPP = 1)—and, on the other hand, the rest of the diploid populations, all of them located in SW Iberian Peninsula, that were further subdivided into three weakly-supported (LPP < 0.6), albeit geographically-structured subgroups (Figure 1a). The Algarve subgroup included individuals from the Algarve region (S Portugal), with phenotype fitting with *T. albicans*. The Doñana subgroup was mainly composed of individuals from Doñana National and Natural Parks (Huelva province, Andalusia, S Spain), with phenotype corresponding to *T. mastichina* subsp. *donyanae*, as well as nearby populations, some of

which presented intermediate characters between *T. mastichina* subsp. *donyanae* and *T. albicans* while the westernmost populations of this group fitted with *T. mastichina* subsp *mastichina* (Table 1). Finally, the Cádiz subgroup mostly consisted of individuals from Cádiz province (Andalusia, S Spain) (Table S1). The plastome tree, based on 72 plastid CDSs showed a mixed distribution of individuals, with no clear groupings (Figure S1).

3.3 | Ploidy Level Estimates from Allelic Ratios

Estimated allelic ratios per individual (Figure S2) ranged between 1 and 5.5, with median values from 1.24 to 2.67, in the diploid group, and from 1.77 to 3.04, in the tetraploid group. Model approaches classified the median allelic ratios per individual (Figure S2) into three levels: (a) low (1.24–1.80), (b) medium (1.70–2.25), and (c) high (1.80–3.04). These three allelic ratio levels fit the three ploidy models for which the nQuire pipeline specifically tests: diploid, triploid, and tetraploid, respectively. The tree categories were found in the diploid group and the two latter in the tetraploid group (Figure 1b). 2C DNA content (see section 3.1.) significantly differed among the three putative ploidy level estimates inferred from allelic ratio values (Kruskal-Wallis χ^2 =84.47, df = 2, $p < 0.001$), although these differences were non-significant when considering only the diploid group (Kruskal-Wallis χ^2 = 3.63, df = 2, F = 0.945, p = 0.163). Interestingly, the percentage of admixture of the tetraploid group into the diploid subgroups was significantly correlated with the mean allelic ratio per individual (Spearman rank correlation, □ = 0.742; S = 226265, $p < 0.001$).

3.4 | Phylogeographic Structure and Interploidy Admixture

The genetic structure of all sampled individuals is presented in Figure 1b. Optimal and suboptimal number of population structure groupings ranged between two and five, using the CV-error method, while the ΔK value supported that the optimal number of clusters was five (Figure S3). Population structure analysis for K = 2 distinguished tetraploid and diploid individuals into two distinct genetic groups. For K = 3 and K = 4, the tetraploid samples remained assigned to the same genetic group, and diploid samples were split into two and three genetic groupings, respectively. The likeliest five-cluster scheme (K = 5) was consistent with the phylogenetic groupings described above: the tetraploid *T. mastichina* group; and the diploid group, composed of the same four genetic subgroups (Algarve, Cádiz, Doñana, and Hercynian) observed in the phylogeny (Figure 1a,b), showing a distinct geographic pattern (Figure 1c,d). Interploidy admixture was observed across all K-schemes, with 78% of populations showing evidence of it in at least some individuals for K = 5 (Figure 1b).

The first five principal components of the PCA accounted for 94% of the variance and jointly supported the population structure for K = 5 (Figure 3a). Clustering analysis with DAPC (Figure 3b) was also coherent with the aforementioned five-cluster pattern, and largely consistent with the PCA. The two first linear discriminants (LD) of the DAPC explained >95% of cumulative variance. LD1 clearly separated tetraploid *vs.* diploid samples. LD1 and LD2 separated the four diploid subgroups.

Pairwise fixation indexes (F_{ST}) indicated a moderate differentiation among the five identified genetic groupings ($0.036 \leq F_{ST} \leq 0.101$; Figure 3c), with the Hercynian subgroup showing the highest mean F_{ST}, when compared with other subgroups ($0.071 \leq F_{ST} \leq 0.101$). In contrast, the Doñana subgroup showed the lowest FST ($0.036 \leq F_{ST} \leq 0.071$; Figure 3c). When comparing populations, the lowest F_{ST} index values were found within the tetraploid group and the diploid Doñana subgroup (Figure S4). The AMOVA test (Table S2) showed significant genetic structure in the four tested scenarios (K2-K5). The K2 clustering exhibited the highest levels of genetic variance among groups (28%; $p < 0.001$), followed by K3 and K5 (both 25%; $p < 0.001$).

The Mantel test showed a significant isolation-by-distance pattern among populations within the diploid genetic group ($r = 0.83$; $p = 0.001$), but no geographical isolation was detected among the more geographically spread populations of the tetraploid group ($p = 0.161$). This significant isolation-by-distance pattern was maintained when considering only the three diploid subgroups (Algarve, Doñana, and Cádiz) restricted to the SW Iberian Peninsula ($r = 0.75$; $p = 0.001$).

3.5 | Genetic Diversity

Genetic diversity was analysed using different indicators. The mean inbreeding coefficient (F_{IS}) per genetic grouping (Figure S5 and Table S3) ranged between −0.278 and 0.222. Tetraploid *T. mastichina* individuals showed the lowest mean inbreeding values ($F_{IS} = -0.278$), while diploid individuals showed in general higher inbreeding values ($0.003 \leq F_{IS} \leq 0.222$), especially the Hercynian subgroup (Table S3). The highest mean proportion of identical-by-descent (IBD) alleles was found in the tetraploid *T. mastichina* group (IBD = 0.223), while the lowest values were found in the diploid Cádiz subgroup (IBD = 0.055; Table S3). The nucleotide diversity (π) index showed, in general, low values ($0.0012 \leq \pi \leq 0.0025$; Figure 4a, Table S4); tetraploid populations had the highest mean π values ($0.0015 < \pi < 0.0025$; Figure 4a, Table S4). Tajima's D values (Figure 4b, Table S4), ranged between −0.667 and 0.440, negative mean values were found in all populations within the Doñana subgroup (Figure 4b, Table S4). Private and shared SNPs in a balanced random subsample of analysed individuals showed a skewed pattern of SNP distribution among genetic groupings (Figure 5). On one hand, the tetraploid *T. mastichina* group shared more than 50% of the analysed SNPs with all the diploid subgroups (Figure 5). On the other, 44 private SNPs (2%) were detected in the tetraploid *T. mastichina* group, while in the diploid subgroups the number of private SNPs was much lower (3–11 SNPs/subgroup). The number of shared SNPs among groups and subgroups was also variable (Figure 5). The highest value of shared SNPs corresponded to the tetraploid group and the diploid group minus the Hercynian subgroup (172 SNPs; 8%), the second highest value was found for the tetraploid group and the diploid group excluding Cádiz subgroup (95 SNPs; 4%).

3. 6 | Reticulation, Introgression and Gene Flow Directionality

The Neighbour-Net phylogenetic network analysis provided a detailed view of the reticulations among the studied samples (Figure S6). Overall, the genetic structure of tetraploid and diploid groups was well defined; however, a dense reticulate core was observed in both networks for the concatenated supercontigs and SNP datasets. A high level of phylogenetic conflict was particularly evident among the diploid samples. Additionally, several reticulate connections were detected between the tetraploid clade and multiple diploid groups, especially the Doñana and Hercynian subgroups, as well as with the central core of the network.

The ABBA-BABA test provided evidence of introgression between the tetraploid *T. mastichina* group and the diploid Hercynian subgroup (Z score > 5.04; $p < 0.0001$; Figure 6 and Table S5), also supported by the high values of f4-ratio (Table S5) and F-branch statistic, which had a value of 1. A lower but significant introgression was also detected between the tetraploid *T. mastichina* group and the Doñana (Z-score = 2.76; $p < 0.005$) and Cádiz (Z-score = 2.349, $p = 0.019$) subgroups. Additionally, the diploid Algarve and Hercynian subgroups also showed significant introgression levels (Figure 6; Table S5).

Coalescent simulations also detected interploidy gene flow in all tested scenarios (Figure 7, Table S6). Bidirectional gene flow was the best model for gene flow between tetraploids and the Algarve, Cádiz, and Doñana diploid subgroups, whereas unidirectional gene flow was the most plausible model from tetraploids to the Hercynian diploid subgroup. In all scenarios, overall migration intensity was consistently higher from tetraploids to diploids.

4 | Discussion

Our integrative analysis uncovers a complex evolutionary history in *Thymus* sect. *Mastichina*, with a clear diploid–tetraploid split and fine-scale phylogeographic structure among diploid populations, with four geographically defined diploid subgroups consistently recovered (Figures 1–3). Yet, plastid–nuclear discordance (Figures 1 and S1), allelic ratio variation (Figure S2), and evidence of gene flow—including interploidy introgression (Figures 6-7)—reveal dynamic and complex evolutionary processes. These patterns highlight ongoing diversification in this Iberian polyploid complex.

4.1 | Ploidy Differentiation and Origin

Flow cytometry revealed a bimodal 2C DNA content in *Thymus* sect. *Mastichina*, matching diploid and tetraploid individuals (Girón *et al*., 2012). In parallel, the phylogenomic reconstruction confirmed that this section is composed of two sister clades—tetraploids vs. diploids—, the latter with four geographically-restricted subgroups (Algarve, Cádiz, Doñana, and Hercynian; Figures 1a and 2). SNP-based clustering methods (ADMIXTURE, PCA, DAPC) also supported this structure (Figures 1b and 3) and so did the Phylogenetic Neighbour-Net network (Figure S6). All tetraploids showed traits of *T. mastichina* subsp. *mastichina* (*sensu* Morales 2010), while diploids encompassed all section morphospecies: *T. mastichina*

subsp. *mastichina* (in Hercynian and Doñana subgroups), *T. mastichina* subsp. *donyanae* (in Doñana), and *T. albicans* (in Algarve, Cádiz, and Doñana). Girón *et al.* (2012) also found diploids matching *T. mastichina* subsp. *mastichina* in Serra da Estrela (Hercynian) and Huelva (Doñana). Thus, the current taxonomy does not reflect the evolutionary history of the section.

Morales (1986) proposed that the tetraploid taxon originated from diploid progenitors in SW Iberian Peninsula and later colonised the entire peninsula, adapting to continental climates during Pleistocene cold phases. However, our results do not support this hypothesis. Additionally, he was unaware of *mastichina*-like diploid individuals in the Hercynian region (Morales, R., pers. com.). Our data reveal that most tetraploids share a percentage of genetic ancestry with the Hercynian diploids. Yet, Hercynian ancestry is not dominant, and *T. mastichina* appears as a genetically distinct lineage, not as a mixture. For that reason, we cannot establish a reliable hypothesis regarding the role of Hercynian diploids in the origin of the tetraploid group. The observed genetic pattern in the tetraploid group could reflect autopolyploidy from an extinct (ghost) diploid, followed by interploidy admixture (see Leal *et al.* 2024). Alternatively, an allopolyploid origin via hybridisation between extant Hercynian diploids and an unknown extant or ghost progenitor (Luo *et al.* 2017), followed by WGD and subgenome dominance could also be plausible (Bird *et al.* 2018; Liu *et al.* 2025). Further evidence, especially on potential diploid progenitors from other sections, is required to assess a putative allopolyploid origin of the tetraploid *T. mastichina*.

Polyploid species may exhibit higher allelic mutation rates across their multiple genomic complements, leading over time to greater genetic variation and, consequently, higher genetic diversity than diploids (Segarra-Moragues and Catalán 2003; Segarra-Moragues *et al.* 2004; Clo 2022; Campos *et al.* 2024). This enhanced variation may confer broader ecological tolerance, enabling polyploids to colonise new habitats more readily than diploids (Stebbins 1985; te Beest *et al.* 2012), and promoting diversification (Meudt *et al.* 2021). In *Thymus* sect. *Mastichina*, the widespread tetraploid *T. mastichina* group shows higher nucleotide diversity (π) than the geographically restricted diploid subgroups. Additionally, no isolation-by-distance pattern is detected in the tetraploid group, likely reflecting higher gene flow and connectivity among its populations. This is also suggested by the low F_{ST} and F_{IS} values in these populations, although we are aware that our study system does not meet all Hardy-Weinberg equilibrium assumptions (e.g., diploid populations and neutral genetic variation in all loci) and this could bias these estimates (Baudry and Depaulis 2003; McTavish and Hillis 2015; Liang and Dutheil 2025). Also, homeologous regions from ancient duplications may cause apparent excess heterozygosity and confound interpretations of variation and structure (Hohenlohe *et al.* 2010), our analyses are based on SNPs coming from nuclear orthologs (exons and their flanking regions), minimising such effects and ensuring more accurate estimates (Balant *et al.* 2025). Alternatively, the reduced differentiation among distant tetraploid populations may reflect rapid range expansion after their emergence, likely due to a greater adaptive potential (Ehrendorfer 1980; Hull□Sanders *et al.* 2009; Treier *et al.* 2009; te Beest *et al.* 2012; Kiedrzyński *et al.* 2021). A plausible scenario involves diploid range contraction and habitat fragmentation (presumably *Quercus* forests; Valdés Castrillón *et al.*

1999), while tetraploids expanded their distribution by invading new habitats, as has been proposed for polyploids (Lumaret 1988; Kiedrzyński *et al.* 2021).

4.2 | Diploid Lineage Structure and Phylogeographic Patterns

The phylogeographic structure observed in diploid populations of *Thymus* sect. *Mastichina* likely reflects a complex history of colonisation and diversification tied to the geological evolution of the Iberian Peninsula. The Hercynian subgroup, comprising two populations in the ancient Iberian Massif (Eguíluz *et al.* 2000), consistently appears as sister to the other diploid subgroups in both the phylogenomic tree and ADMIXTURE analyses (Figure 1). This group may represent a relict lineage that originated and persisted in geologically stable environments. Their higher F_{IS} and F_{ST} values point to low connectivity between populations and a marked population differentiation (Young *et al.* 1996; Charlesworth 2003). The absence of *Thymus* sect. *Mastichina* representatives in North Africa, despite floristic similarity across the Strait of Gibraltar (Rodríguez-Sánchez *et al.* 2008), supports its origin postdating the strait's formation (~5.3 Mya; Krijgsman *et al.* 1999). In fact, divergence time estimates for *Thymus* suggest a relatively recent origin at the Plio-Pleistocene boundary (~2.58 Mya; Drew and Sytsma 2012). If *Thymus* section *Mastichina* originated somewhere in the Iberian Massif, a colonisation of the Algarve (a Mesozoic basin later restructured by the Alpine Orogeny; Ramos *et al.* 2020) by the most recent common ancestor (MRCA) of all diploid lineages, is plausible. Another likely migration could have occurred southward into the Cádiz lowlands, southwest of the Betic Mountains, formed in Alpine Orogeny (Vergés *et al.* 2019). In both cases, this southward dispersal into more meridional latitudes is a well-documented biogeographical pattern in plant lineages, linked to glacial–interglacial dynamics. In fact, Algarve and Cádiz are recognised as putative Western Mediterranean refugia (Gómez and Lunt 2007; Médail and Diadema 2009).

In contrast, the Doñana region, between Algarve and Cádiz, is geomorphologically younger. Its current landscape was shaped by Guadalquivir River sediments and sea-level rise (Rodríguez-Ramírez *et al.* 2019) with stabilization during the Holocene (Rodríguez Vidal 2005; Morales-Molino *et al.* 2011), making it one of the last regions within the current distribution of *Thymus* sect. *Mastichina* to become ecologically available. Negative Tajima's D values in several Doñana populations may indicate a post-bottleneck expansion. The phylogenetic proximity with the Cádiz subgroup could point to a recent colonisation of Doñana's sand dunes by the MRCA of these sister subgroups. However, the low branch support in our phylogeny limits inference on the colonisation source after Doñana dune formation. Overall, genomic and geological evidence suggests a scenario in which the diploid group originated inland (Iberian Massif) and expanded southwards to Algarve and Cádiz refugia and, more recently, to Doñana. Future work estimating divergence times and diversification rates will help test and refine this hypothesis.

Populations of the three diploid subgroups in SW Iberian Peninsula (Algarve, Cádiz, and Doñana) show clear genetic differentiation and significant west-east isolation-by-distance pattern, despite their adjacent distribution. Besides their colonization history, geographic features like the Guadiana and Guadalquivir

riverbeds, between Algarve-Doñana and Doñana-Cádiz, respectively, could be acting as barriers to gene flow (see Ortiz *et al.* 2008). The absence of significant introgression among these adjacent groups (Table S5) aligns with their low gene flow. Also, the high inbreeding levels (Table S3), especially in the Cádiz subgroup, could indicate low functional connectivity among populations (Young *et al.* 1996; Charlesworth 2003). Interestingly, a population of the Doñana subgroup (Pinar Algaida) is currently located on the eastern margin of the Guadalquivir estuary (Cádiz province). However, its location likely reflects local coastal dynamics, not cross□riverbed dispersal. Around 6.5 Kya, a marine transgression (peaking at current sea level) filled the estuary with sediments, reshaping the shoreline into sandy barriers and marshes. Then, ~5 Kya, an accelerated shoreline progradation built a promontory connecting the western margin of the ancient estuary. Later erosional pulses (e.g., storms, tsunamis) isolated this promontory as a small island and sculpted it into present-day Algaida, nowadays on the eastern estuary bank (Rodríguez-Ramírez 2008).

4.3 | Reticulate Evolution, Admixture and Gene Flow

We found patterns among diploid and tetraploid lineages that point to a complex evolutionary history on *Thymus* sect. *Mastichina*. The observed shared heterozygosity in ADMIXTURE analysis (Figure 1b) and the evidence of reticulate evolution observed in Neighbour-Net networks (Figure S6). Such a structure is consistent with ILS following recent or rapid divergence events, especially among diploid lineages. Additionally, signals of interploidy gene flow between diploid and tetraploid lineages of *Thymus* sect. *Mastichina* provide evidence for secondary contact between ploidy levels throughout the evolutionary history of the group. ABBA-BABA tests identified significant introgression from tetraploid *T. mastichina* into several diploid subgroups—especially the Hercynian, but also Doñana and Cádiz (Figure 6). Coalescent-based analyses (Figure 7) also detected consistent patterns of interploidy gene flow across ploidy levels in all diploid subgroups. Although bidirectional gene flow emerged as the best-supported model between tetraploids and most diploid groups, the overall intensity of gene flow was consistently higher from tetraploids to diploids, pointing to asymmetric interploidy dynamics. This pattern is further supported by allele frequency distributions (Figure S2), allele-ratio-based ploidy inference, SNP-based ancestry proportions (Figures 1 and 3), and the Neighbour-Net network (Figure S6). Altogether, these results suggest recurrent and geographically widespread gene flow between ploidy levels, particularly in contact zones.

The clear bimodal distribution of 2C DNA content in *Thymus* sect. *Mastichina* supports a diploid–tetraploid distinction, yet our allele-ratio-based ploidy model also identifies triploid and tetraploid profiles within the diploid cytometric group. This mismatch suggests that this approach (Viruel *et al*. 2019, 2023), relying on nQuire (Weiß *et al*. 2018) and ploidyNGS (Corrêa dos Santos *et al*. 2017), may not be fully reliable in polyploid complexes with known or suspected interploidy admixture, as no 2C DNA content differences were found among inferred ploidy categories. Other allele-ratio-based models, such as gbs2ploidy (Gompert and Mock 2017), nQuack (Gaynor *et al*. 2024), or vcf2ploidy (DeWaters 2020), could face similar limitations. However, it is important to note that these models are not fully equivalent to nQuire, as they rely on different mathematical assumptions, and their performance would need to be tested empirically to assess

how they behave in cases of interploidy admixture. Rather than reflecting current ploidy levels, the triploid and tetraploid categories inferred by the Viruel *et al.* (2019, 2023) workflow may reflect past or ongoing introgression (Nieto Feliner *et al*. 2020; Wang *et al.* 2023; Brown *et al.* 2024), with the fixation of homeolog loci from tetraploid genomes into diploid lineages. This hypothesis is further supported by significant correlations between the frequency assignment of diploidy to the tetraploid genetic group and allele ratios, and by introgression signals detected through the ABBA-BABA test, which revealed gene flow between tetraploid *T. mastichina* and the diploid Cádiz, Doñana, and Hercynian subgroups. The observed pattern could suggest not only recent but also ancient episodes of introgression, potentially dating back to periods soon after the tetraploid emergence.

Historical gene flow is likely in contact zones where diploids and tetraploids have long coexisted, enabling hybridisation and backcrossing. Over time, introgressed alleles may become fixed in recipient genomes, leaving persistent genomic signatures that may not necessarily reflect active gene exchange (Soltis and Soltis 2009; Nieto Feliner *et al.* 2020; Stull *et al.* 2023). This could explain interploidy admixture in isolated diploid populations such as Serra da Estrela (Hercynian subgroup). In contact zones of diploid and tetraploid populations, ongoing low-level gene exchange, possibly mediated by triploid bridges, could be plausible. Our coalescent-based demographic analyses, indicating gene flow mainly from tetraploids into diploids, also support this hypothesis. While the extent of recent interploid admixture remains uncertain, our results suggest that historical hybridisation and subsequent backcrossing have left a lasting imprint on the genomic architecture of several diploid populations (e.g., North Cartaya, Villamanrique, Serra da Estrela). Diploid individuals with traits from different morphospecies further suggest repeated interploid admixture and complex introgressed genotypes. These patterns align with reticulate evolution, where diploid–tetraploid barriers are porous under certain spatial-ecological conditions (Kauai *et al.* 2024; Bartolić *et al.* 2024, 2025). Introgressed alleles from tetraploids may be fixed or remain at intermediate frequencies in diploids, increasing heterogeneity and blurring taxon boundaries. The overall genomic cohesion of tetraploid *T. mastichina* may suggest limited recent diploid introgression, also detected by the coalescent-based demographic modelling, supporting asymmetrical, context-dependent interploid admixture mediated by population density, ecological compatibility, and hybrid fertility (Vallejo Marín *et al.* 2016; Kauai *et al.* 2024; Leal *et al.* 2024; Bartolić *et al.* 2024, 2025).

4. 4 | Evolutionary and Conservation Implications

Our results offer valuable insights and serve as a guiding example for the conservation of taxonomically complex or cryptic species (Bickford *et al.* 2007; Fišer *et al*. 2018; Hending 2025), where morphology alone proves inadequate (e.g., phenotypic plasticity; Matesanz *et al.* 2010). In *Thymus albicans* s.l. (referring to all diploids in *Thymus* sect. *Mastichina*), morphospecies poorly reflect underlying biological lineages. We recommend flow cytometry in conservation assessments to identify and monitor cytotypes and genetic structure (Heslop-Harrison *et al.* 2023). This rapid, cost-effective method can be easily integrated into routine prospections of new populations for conservation purposes. Given growing threats from tourism,

coastal development, and land-use change (Valdés Castrillón *et al.* 1999; Cabezudo *et al.* 2005; Morales 2010), conservation should prioritise the four genetic subgroups identified within *T. albicans* as distinct units, as each lineage preserves the species' genetic and ecological legacy.

Low genetic diversity, high inbreeding, and isolation—especially in Hercynian and Cádiz populations—suggest reduced gene flow and potential reproductive barriers, which may compromise long-term population viability (Young *et al.* 1996; Charlesworth 2003; Campos *et al.* 2024). This is especially relevant in gynodioecious species like the endangered *T. albicans*, where the effects of such cross-fertilisation barriers may be exacerbated in populations with female-biased sex ratios, as observed in many *Thymus* species (Karbstein *et al.* 2019) and also many of the studied populations (Berjano, pers. obs.). Future conservation-oriented analyses should incorporate additional variables, such as ecological features, population threat levels, and protection status, to provide more specific management recommendations.

5 | Conclusions

Our integrative genomic approach confirms that *Thymus* sect. *Mastichina* comprises two sister groups with distinct ploidy levels. The tetraploid *T. mastichina* likely arose from related diploids, and shows no genetic structuring, consistent with recent expansion. In contrast, diploid populations exhibit complex structuration shaped by historical geomorphology, isolation, ILS and occasional interploidy introgression. Despite genetic divergence, phenotypic overlap and signs of reticulate evolution challenge current taxonomy, revealing cryptic diversity and underscoring the need for an integrative taxonomic framework combining morphological, cytogenetic, genomic, and geographic data. We advocate a taxonomic re-evaluation using multi-source evidence and emphasise the urgent conservation of the distinct diploid lineages of the endangered *T. albicans* s.l.—including not only the '*albicans*' morphotype (Algarve, Cádiz) but also the '*mastichina*' and '*donyanae*' diploid morphotypes of the Hercynian and Doñana subgroups—, which represent unique evolutionary trajectories and may harbour critical adaptive potential. Future work should expand flow cytometry sampling across the Hercynian range and reassess morphological distinctions between diploid and tetraploid '*mastichina*' morphotypes and select the most relevant populations for the effective conservation of the lineages. Additionally, estimates of divergence times, diversification rates, and phylogenetic networks are needed to fully resolve the evolutionary history of these ecologically and economically valuable Iberian taxa.

Author Contributions

R.B., D.N., and M.d.l.E. conceived the project, designed the methodological approach, and secured funding. DD, R.B., J.C.d.V., M.A.O., D.N.L., and F.J.G.-C. carried out the field work. L.P. designed the molecular and computational workflows, assisted by R.B. M.A.O., and F.J.G.-C. R. B., D.D, J.C.d.V., F.J.G-C., and L.P., conducted molecular work. R. B., D.D, J.C.d.V., F.J.G-C and M.A.O conducted flow cytometry analyses. F.J.G.-C., R.B., M.A.O. and L.P. analysed the data. R.B. and L.P. supervised the study. F.J.G.-C.,

R.B, L.P., and M.A.O. drafted the manuscript. All authors reviewed and edited subsequent versions and approved of the final one.

Acknowledgements

This study is part of the Conserva3 project (TED2021-130133B-I00) funded by MCIN/AEI/10.13039/501100011033 with funds from the European Union "NextGenerationEU"/PRTR. LP benefited from a Ramón y Cajal grant (RYC2021-034942-I) funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/PRTR.

Collection permits were obtained from Consejería de Sostenibilidad, Medio Ambiente y Economía Azul, Junta de Andalucía (Spain), Doñana National Park, and from Departamento de Conservação da Natureza e Biodiversidade (Portugal, permits no. 576–580/REC/2023). In Doñana National Park, logistic and technical support was provided by ICTS-RBD-CSIC, Ministry of Science and Innovation, and co-financed by FEDER Funds.

Authors would like to thank Javier Jiménez-Lopez, Javier López Tirado, Billy James Williams, Esther Martín Carretié, Andrés Melero, as well as the staff of the Environmental Agency and the Botanical Network of the Andalusian Government (Junta de Andalucía) and of the Doñana National Park for their support in population location and sample collection. Ricarda Riina provided DNA templates for *Thymus* outgroups.

Genomic libraries were prepared at the DNA lab of the Herbarium Research Services (CITIUS, University of Seville). Flow cytometry measurements were performed at the facilities of the Biology Research Services (CITIUS, University of Seville). Part of the costs of those Research Services were funded by 2023 and 2024 calls of VII PPIT of the University of Sevilla. Computational workflows were implemented in the Hercules supercomputer at the Scientific Computer Centre of Andalucia (CICA). Additional thanks go to Manica Balant, Matt Johnson, Rafa Albadalejo, Xavier Picó, Isaac Overcast, Paolo Bartolić, Marcial Escudero, and Ana Valdés, for their analytical advice, and to four anonymous reviewers that helped improve the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The raw Hyb-Seq data are available on NCBI Sequence Read Archive (SRA) under BioProject PRJNA1326772. The code pertaining to the implemented analyses, the ad hoc plastid CDS target file, output

datasets, and spreadsheets with summary statistics are available through Zenodo repository 10.5281/zenodo.15331851.

References

Adams, K. L., and J. F. Wendel. 2005. Polyploidy and Genome Evolution in Plants. *Current Opinion in Plant Biology* **8**, no. 2: 135–141. https://doi.org/10.1016/j.pbi.2005.01.001

Alexander, D. H., and K. Lange. 2011. Enhancements to the ADMIXTURE Algorithm for Individual Ancestry Estimation. *BMC Bioinformatics* **12**: 246. https://doi.org/10.1186/1471-2105-12-246

Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Research* **19**, no. 9: 1655–1664. https://doi.org/10.1101/gr.094052.109

Al-Shehbaz, I. A. 2012. A Generic and Tribal Synopsis of the Brassicaceae (Cruciferae). *Taxon* **61**, no. 5: 931–954. https://doi.org/10.1002/tax.615002

Andrews, S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Anisimova, M., M. Gil, J.-F. Dufayard, C. Dessimoz, and O. Gascuel. 2011. Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. *Systematic Biology* **60**: 685–699. https://doi.org/10.1093/sysbio/syr041

Arrigo, N., and M. S. Barker. 2012. Rarely Successful Polyploids and Their Legacy in Plant Genomes. *Current Opinion in Plant Biology* **15**: 140–146. https://doi.org/10.1016/j.pbi.2012.03.010

Baird, N. A., P. D. Etter, T. S. Atwood *et al.* 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* **3**, no. 10: e3376. https://doi.org/10.1371/journal.pone.0003376

Baker, W. J., S. Dodsworth, F. Forest *et al.* 2021. Exploring Angiosperms353: An Open, Community Toolkit for Collaborative Phylogenomic Research on Flowering Plants. *American Journal of Botany* **108**, no. 7: 1059–1065. https://doi.org/10.1002/ajb2.1703

Baker, W. J., P. Bailey, V. Barber *et al.* 2022. A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology* **71**, no. 2: 301–319. https://doi.org/10.1093/sysbio/syab035

Balant, M., D. Vitales, Z. Wang *et al.* 2025. Integrating Target Capture with Whole Genome Sequencing of Recent and Natural History Collections to Explain the Phylogeography of Wild-Growing and

Cultivated *Cannabis*. *Plants, People, Planet* **7**, no. 6: 1771-1788. https://doi.org/10.1002/ppp3.70043

Bartolić, P., E. J. Morgan, N. Padilla-García, and F. Kolář. 2024. Ploidy as a Leaky Reproductive Barrier: Mechanisms, Rates and Evolutionary Significance of Interploidy Gene Flow. *Annals of Botany* **134**, no. 4: 537–550. https://doi.org/10.1093/aob/mcae096

Bartolić, P., A. Voltrová, L. Macková *et al.* 2025. Overcoming Ploidy Barriers: The Role of Triploid Bridges in the Genetic Introgression of *Cardamine amara*. *Molecular Ecology* **34**, no. 7: e17702. https://doi.org/10.1111/mec.17702

Baudry, E. and F. Depaulis. 2003. Effect of Misoriented Sites on Neutrality Tests With Outgroup. *Genetics* **165**, no 3: 1619–1622, https://doi.org/10.1093/genetics/165.3.1619

Beck, J. B., M. L. Markley, M. G. Zielke *et al.* 2021. Are Palmer's Elm-Leaf Goldenrod and the Smooth Elm-Leaf Goldenrod Real? The Angiosperms353 Kit Provides Within-Species Signal in *Solidago ulmifolia* s. l. *Systematic Botany* **46**, no. 4: 1107–1113. https://doi.org/10.1600/036364421X16370109698740

te Beest, M., J. J. Le Roux, D. M. Richardson *et al.* 2012. The More the Better? The Role of Polyploidy in Facilitating Plant Invasions. *Annals of Botany* **109**, no. 1: 19–45. https://doi.org/10.1093/aob/mcr277

Bhatia, G., N. Patterson, S. Sankararaman, and A. L. Price. 2013. Estimating and Interpreting F_{ST}□: The Impact of Rare Variants. *Genome Research* **23**, no. 9: 1514–1521. https://doi.org/10.1101/gr.154831.113

Bickford, D., D. J. Lohman, N. S. Sodhi *et al.* 2007. Cryptic Species as a Window on Diversity and Conservation. *Trends in Ecology and Evolution* **22**, no. 3: 148–155. https://doi.org/10.1016/j.tree.2006.11.004

Bird, K. A., R. VanBuren, J. R. Puzey, and P. P. Edger. 2018. The Causes and Consequences of Subgenome Dominance in Hybrids and Recent Polyploids. *New Phytologist* **220**, no.1: 87–93. https://doi.org/10.1111/nph.15256

Borowiec, M. L. 2016. AMAS: A Fast Tool for Alignment Manipulation and Computing of Summary Statistics. *PeerJ* **4**: e1660. https://doi.org/10.7717/peerj.1660

Brewer, G. E., J. J. Clarkson, O. Maurin *et al.* 2019. Factors Affecting Targeted Sequencing of 353 Nuclear Genes from Herbarium Specimens Spanning the Diversity of Angiosperms. *Frontiers in Plant Science* **10**: 1102. https://doi.org/10.3389/fpls.2019.01102

Brown, M. R., R. J. Abbott, and A.D. Twyford. 2024. The Emerging Importance of Cross□Ploidy Hybridisation and Introgression. *Molecular Ecology* **33**, no. 8: e17315. https://doi.org/10.1111/mec.17315

Buira, A., A. Carapeto, P. G. García Murillo, and T. Monteiro-Henriques. 2017. *Thymus albicans*. The IUCN Red List of Threatened Species: e.T103496011A103496020. https://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T103496011A103496020.en

Cabezudo, B., S. Talavera, G. Blanca *et al.* 2005. Lista Roja de la flora vascular de Andalucía. Sevilla: Junta de Andalucía. Consejería de Medio Ambiente. ISBN 84-96329-62-3

Campos, M., E. Pérez□Collazos, A. Díaz□Pérez *et al*. 2024. Repeated Migration, Interbreeding and Bottlenecking Shaped the Phylogeography of the Selfing Grass *Brachypodium stacei*. *Molecular Ecology* **33**, no. 19: e17513. https://doi.org/10.1111/mec.17513

Canty, A., and Ripley, B. 2024. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-31. https://cran.r-project.org/web/packages/boot/boot.pdf

Capella-Gutierrez, S., J. M. Silla-Martinez, and T. Gabaldón. 2009. trimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* **25**, no. 15: 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *Gigascience* **4**, no. 1: s13742-015-0047–8. https://doi.org/10.1186/s13742-015-0047-8

Charlesworth, D. 2003. Effects of Inbreeding on the Genetic Diversity of Populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **358**, no. 1434: 1051–1070. https://doi.org/10.1098/rstb.2003.1296

Chen, S. 2023. Ultrafast One□Pass FASTQ Data Preprocessing, Quality Control, and Deduplication Using fastp. *iMeta* **2**, no. 2: e107. https://doi.org/10.1002/imt2.107

Chen, H., and P. C. Boutros. 2011. VennDiagram: A Package for the Generation of Highly-Customizable Venn and Euler Diagrams in R. *BMC Bioinformatics* **12**, no. 1: 35. https://doi.org/10.1186/1471-2105-12-35

Clo, J. 2022. The evolution of the additive variance of a trait under stabilizing selection after autopolyploidization, *Journal of Evolutionary Biology* **35**, no. 6: 891–897. https://doi.org/10.1111/jeb.14010

Comai, L. 2005. The Advantages and Disadvantages of Being Polyploid. *Nature Reviews Genetics* **6**, no. 11: 836–846. https://doi.org/10.1038/nrg1711

Cooper, B. J., M. J. Moore, N. A. Douglas, W. L. Wagner, and M. G. Johnson. 2023. Target Enrichment and Extensive Population Sampling Help Untangle the Recent, Rapid Radiation of *Oenothera* Sect. *Calylophus*. *Systematic Biology* **72**, no. 2: 249–263. https://doi.org/10.1093/sysbio/syac032

Corrêa dos Santos, R. A., G. H. Goldman, and D. M. Riaño-Pachón. 2017. ploidyNGS: Visually Exploring Ploidy with Next Generation Sequencing Data. *Bioinformatics* **33**, no. 16: 2575–2576. https://doi.org/10.1093/bioinformatics/btx204

Danecek, P., A. Auton, G. Abecasis *et al.* 2011. The Variant Call Format and VCFtools. *Bioinformatics* **27**, no. 15: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., J. K. Bonfield, J. Liddle *et al.* 2021. Twelve Years of SAMtools and BCFtools. *GigaScience* **10**, no. 2: giab008. https://doi.org/10.1093/gigascience/giab008

DePristo, M. A., E. Banks, R. Poplin *et al.* 2011. A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data. *Nature Genetics* **43**, no. 5: 491–498. https://doi.org/10.1038/ng.806

DeWaters, D. 2020. vcf2ploidy. Waldorf (MD): GitHub. https://github.com/dandewaters/vcf2ploidy.

Dodsworth, S. 2015. Genome Skimming for Next-Generation Biodiversity Analysis. *Trends in Plant Science* **20**, no. 9: 525–527. https://doi.org/10.1016/j.tplants.2015.06.012

Dodsworth, S., L. Pokorny, M. G. Johnson *et al.* 2019. Hyb-Seq for Flowering Plant Systematics. *Trends in Plant Science* **24**, no. 10: 887–891. https://doi.org/10.1016/j.tplants.2019.07.011

Doyle, J. J. 2022. Defining Coalescent Genes: Theory Meets Practice in Organelle Phylogenomics. *Systematic Biology* **71**, no. 2: 476–489. https://doi.org/10.1093/sysbio/syab053

Doležel, J., P. Binarová, and S. Lucretti. 1989. Analysis of Nuclear DNA Content in Plant Cells by Flow Cytometry. *Biologia Plantarum* **31**, no. 2: 113–120. https://doi.org/10.1007/BF02907241

Doležel, J., S. Sgorbati, and S. Lucretti. 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* **85**, no. 4: 625–631. https://doi.org/10.1111/j.1399-3054.1992.tb04764.x

Drew, B. T., and K. J. Sytsma. 2012. Phylogenetics, Biogeography, and Staminal Evolution in the Tribe Mentheae (Lamiaceae). *American Journal of Botany* **99**, no. 5: 933–953.

https://doi.org/10.3732/ajb.1100549

Eaton, D. A. R., and I. Overcast. 2020. ipyrad: Interactive Assembly and Analysis of RADseq Datasets. Bioinformatics, 36, no. 8: 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Eguíluz, L., J. I. G. Ibarguchi, B. Ábalos, and A. Apraiz. 2000. Superposed Hercynian and Cadomian Orogenic Cycles in the Ossa-Morena Zone and Related Areas of the Iberian Massif. *Geological Society of America Bulletin* **112**, no. 9: 1398–1413. https://doi.org/10.1130/0016-7606(2000)112<1398:SHACOC>2.0.CO;2

Ehrendorfer, F. 1980. Polyploidy and Distribution. In, *Polyploidy*, edited by W. H. Lewis, 45–60. Boston, MA: Springer US

Ellegren, H. 2014. Genome Sequencing and Population Genomics in Non-Model Organisms. *Trends in Ecology and Evolution* **29**, no. 1: 51–63. https://doi.org/10.1016/j.tree.2013.09.008

Elshire, R. J., J. C. Glaubitz, Q. Sun *et al.* 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **6**, no. 5: e19379. https://doi.org/10.1371/journal.pone.0019379

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the Number of Clusters of Individuals Using the Software s□□□□□□□□: a Simulation Study. *Molecular Ecology* **14**, no. 8: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Ewels, P., M. Magnusson, S. Lundin, and M. Käller. 2016. MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report. *Bioinformatics* **32**, no. 19: 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Excoffier, L., N. Marchi, D. A. Marques, R. Matthey-Doret, A. Gouy and V. C. Sousa. 2021. Fastsimcoal2: Demographic Inference under Complex Evolutionary Scenarios. *Bioinformatics* 37(24): 4882–4885. https://doi.org/10.1093/bioinformatics/btab468

Filipe, A., J. C. R. Cardoso, G. Miguel *et al.* 2017. Molecular Cloning and Functional Characterization of a Monoterpene Synthase Isolated from the Aromatic Wild Shrub *Thymus albicans*. *Journal of Plant Physiology* **218**, 35–44. https://doi.org/10.1016/j.jplph.2017.07.013

Fišer, C., C. T. Robinson, and F. Malard. 2018. Cryptic Species as a Window into the Paradigm Shift of the Species Concept. *Molecular Ecology* **27**, no. 3: 613–635. https://doi.org/10.1111/mec.14486

Francis, R. M. 2017. □□□□□□□□□□: An R Package and Web App to Analyse and Visualize Population Structure. *Molecular Ecology Resources* **17**, no. 1: 27–32. https://doi.org/10.1111/1755-0998.12509

Freund, F. D., D. Gates, M. G. Johnson, and C. J. Rothfels. 2025. Phylogenetics and Population Structure of the Western North American Endemic Pacific Laurasian Clade of *Isoëtes*. *American Journal of Botany* **112**, no. 4: e70030. https://doi.org/10.1002/ajb2.70030

Gaynor, M. L., J. B. Landis, T. K. O'Connor *et al.* 2024. nQuack: An R Package for Predicting Ploidal Level from Sequence Data Using Site based Heterozygosity. *Applications in Plant Sciences* **12**, no. 4: e11606. https://doi.org/10.1002/aps3.11606

Girón, V., T. Garnatje, J. Vallès, E. Pérez-Collazos, P. Catalán, and B. Valdés. 2012. Geographical Distribution of Diploid and Tetraploid Cytotypes of *Thymus* sect. *Mastichina* (Lamiaceae) in the Iberian Peninsula, Genome Size and Evolutionary Implications. *Folia Geobotanica* **47**: 441–460. https://doi.org/10.1007/s12224-012-9126-0

Gómez, A., and D. H. Lunt. 2007. Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula. In, *Phylogeography of Southern European Refugia*, edited by S. Weiss, and N. Ferrand, 155–188. Dordrecht: Springer Netherlands. https://doi.org/10.1007/1-4020-4904-8_5

Gompert, Z., and K. E. Mock. 2017. Detection of Individual Ploidy Levels with Genotyping by Sequencing (GBS) Analysis. *Molecular Ecology Resources* **17**, no. 6: 1156–1167. https://doi.org/10.1111/1755-0998.12657

Gutenkunst, R. N., R. D. Hernández, S. H. Williamson and C. D. Bustamante. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP data. *PLoS Genetics* 5: e1000695. https://github.com/isaacovercast/easySFS

Hagen, E. R., and J. M. Beaulieu. 2024. New Beginnings for Dead Ends: Polyploidy, -SSE Models and the Dead-End Hypothesis. *Annals of Botany* **134**, no. 6: 923–932. https://doi.org/10.1093/aob/mcae143

Hale, H., E. M. Gardner, J. Viruel, L. Pokorny and M. G. Johnson. 2020. Strategies for Reducing per-Sample Costs in Target Capture Sequencing for Phylogenomics and Population Genomics in Plants. *Applications in Plant Sciences* **8**, no. 4: e11337. https://doi.org/10.1002/aps3.11337

Han T.-S., Q. Wu Q, X.-H. Hou, Z.-W. Li, Y.-P. Zou, S. Ge, and Y.-L. Guo. 2015. Frequent Introgressions from Diploid Species Contribute to the Adaptation of the Tetraploid Shepherd's Purse (*Capsella bursa-pastoris*). *Molecular Plant* **8**, 427–438. https://doi.org/10.1016/j.molp.2014.11.016

Hending, D. 2025. Cryptic Species Conservation: A Review. *Biological Reviews* **100**, no. 1: 258–274. https://doi.org/10.1111/brv.13139

Heslop-Harrison, J. S. (Pat), T. Schwarzacher, and Q. Liu. 2023. Polyploidy: Its Consequences and Enabling Role in Plant Diversification and Evolution. *Annals of Botany* **131**, no. 1: 1–10.

https://doi.org/10.1093/aob/mcac132

Hoang, D. T., O. Chernomor, A. Von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**, no. 2: 518–522. https://doi.org/10.1093/molbev/msx281

Hodač, L., K. Karbstein, S. Tomasello, J. Wäldchen, J. P. Bradican and E. Hörandl. 2023. Geometric Morphometric Versus Genomic Patterns in a Large Polyploid Plant Species Complex. *Biology* 12, no. 3: 418. https://doi.org/10.3390/biology12030418

Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. 2010. Population Genomics of Parallel Adaptation in Threespine Stickleback Using Sequenced RAD Tags. *PLoS Genetics* **6**, no. 2: e1000862. https://doi.org/10.1371/journal.pgen.1000862

Huson, D.H., and D. Bryant. 2024. The SplitsTree App: interactive analysis and visualization using phylogenetic trees and networks. *Nature Methods* **21**: 1773–1774. https://doi.org/10.1038/s41592-024-02406-3

Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. Estimation of Levels of Gene Flow from DNA Sequence data. *Genetics* **132**: 583–589. https://doi.org/10.1093/genetics/132.2.583

Hull□Sanders, H. M., R. H. Johnson, H. A. Owen and G. A. Meyer. 2009. Effects of Polyploidy on Secondary Chemistry, Physiology, and Performance of Native and Invasive Genotypes of *Solidago gigantea* (Asteraceae). *American Journal of Botany* **96**, no. 4: 762–770. https://doi.org/10.3732/ajb.0800200

Husband, B. C. 2004. The Role of Triploid Hybrids in the Evolutionary Dynamics of Mixed-Ploidy Populations. *Biological Journal of the Linnean Society* **82**, no. 4: 537–546. https://doi.org/10.1111/j.1095-8312.2004.00339.x

Jenkins, T. L. 2024. □□□□□x□□□□: An R Package and Web App for Spatial Visualisation of Admixture and Population Structure. *Molecular Ecology Resources* **24**, no. 4: e13943. https://doi.org/10.1111/1755-0998.13943

Johnson, P. 2024. devEMF: EMF Graphics Output Device. https://github.com/plfjohnson/devEMF

Johnson, M. G., E. M. Gardner, Y. Liu *et al.* 2016. HybPiper: Extracting Coding Sequence and Introns for Phylogenetics from High□Throughput Sequencing Reads Using Target Enrichment. *Applications in Plant Sciences* **4**, no. 7: 1600016. https://doi.org/10.3732/apps.1600016

Johnson, M. G., L. Pokorny, S. Dodsworth *et al.* 2019. A Universal Probe Set for Targeted Sequencing of

353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering. *Systematic Biology* **68**, no. 4: 594–606. https://doi.org/10.1093/sysbio/syy086

Jombart, T., and I. Ahmed. 2011. *adegenet 1.3-1*: New Tools for the Analysis of Genome-Wide SNP Data. *Bioinformatics* **27**, no. 11: 3070–3071. https://doi.org/10.1093/bioinformatics/btn129

Junier, T., and E. M. Zdobnov. 2010. The Newick Utilities: High-Throughput Phylogenetic Tree Processing in the Unix Shell. *Bioinformatics* **26**, no. 13: 1669–1670. https://doi.org/10.1093/bioinformatics/btq243

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nature Methods* **14**, no. 6: 587–589. https://doi.org/10.1038/nmeth.4285

Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. *Poppr*: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction. *PeerJ* **2**: e281. https://doi.org/10.7717/peerj.281

Karbstein, K., S. Tomasello, and K. Prinz. 2019. Desert□like Badlands and Surrounding (Semi□)Dry Grasslands of Central Germany Promote Small□Scale Phenotypic and Genetic Differentiation in *Thymus praecox*. *Ecology and Evolution* **9**, no. 24: 14066–14084. https://doi.org/10.1002/ece3.5844

Karbstein, K., S. Tomasello, L. Hodač *et al.* 2022. Untying Gordian Knots: Unraveling Reticulate Polyploid Plant Evolution by Genomic Data using the Large *Ranunculus auricomus* Species Complex. *New Phytologist* **235**, no. 5: 2081–2098. https://doi.org/10.1111/nph.18284

Katoh, K., and D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, no. 4: 772–780. https://doi.org/10.1093/molbev/mst010

Kauai, F., Q. Bafort, F. Mortier, M. Van Montagu, D. Bonte, and Y. Van de Peer. 2024. Interspecific Transfer of Genetic Information Through Polyploid Bridges. *Proceedings of the National Academy of Sciences of the United States of America* **121**, no. 21: e2400018121. https://doi.org/10.1073/pnas.2400018121

Kiedrzyński, M., K. M. Zielińska, I. Jędrzejczyk *et al.* 2021. Tetraploids expanded beyond the mountain niche of their diploid ancestors in the mixed-ploidy grass *Festuca amethystina* L. *Scientific Reports* **11**: 18735. https://doi.org/10.1038/s41598-021-97767-6

Kolář, F., M. Čertner, J. Suda, P. Schönswetter, and B. C. Husband. 2017. Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. *Trends in Plant Science* **22**, no. 12: 1041–1055.

https://doi.org/10.1016/j.tplants.2017.09.011

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: A Fast, Scalable and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference. *Bioinformatics* **35**, no. 21: 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Krijgsman, W., F. J. Hilgen, I. Raffi, F. J. Sierro, and D. S. Wilson. 1999. Chronology, Causes and Progression of the Messinian Salinity Crisis. *Nature* **400**, no. 6745: 652–655. https://doi.org/10.1038/23231

Leal, J. L., P. Milesi, E. Hodková *et al.* 2024. Complex Polyploids: Origins, Genomic Composition, and Role of Introgressed Alleles. *Systematic Biology* **73**, no. 2: 392–418. https://doi.org/10.1093/sysbio/syae012

Leebens-Mack, J. H., M. S. Barker, E. J. Carpenter *et al.* 2019. One Thousand Plant Transcriptomes and the Phylogenomics of Green Plants. *Nature* **574**, no. 7780: 679–685. https://doi.org/10.1038/s41586-019-1693-2

Liang, J. and J. Y. Dutheil. 2025. Polarizing SNPs without Outgroup. *bioRxiv*: https://doi.org/10.1101/2025.01.31.635919

Liang, K. J., A. Shepherd-Clowes, A. Agut, O. Hidalgo, P. Tejero Ibarra, and J. Viruel. 2025. Conservation Implications for the Iberian Narrow Endemic *Androsace cantabrica* (Primulaceae) Using Population Genomics with Target Capture Sequence Data. *Ecology and Evolution* **15**, no. 8: e71901. https://doi.org/10.1002/ece3.71901

Liu, J., D. Wang, M. Li *et al.* 2025. Ancient Allopolyploidy and Specific Subgenomic Evolution Drive Adaptive Radiation in Poplars and Willows. *Research Square*. https://doi.org/10.21203/rs.3.rs-5852798/v1

Lumaret, R. 1988. Adaptive Strategies and Ploidy Levels. *Acta Œcologica, Œcologia Plantarum* **9**, no. 1: 83–93

Luo, X., Q. Hu, P. Zhou, D. Zhang, Q. Wang, R. J. Abbott, and J. Liu. 2017. Chasing Ghosts: Allopolyploid Origin of *Oxyria Sinensis* (Polygonaceae) from Its Only Diploid Congener and an Unknown Ancestor. *Molecular Ecology* **26**, no, 11: 3037–3049. https://doi.org/10.1111/mec.14097

Mai, U., and S. Mirarab. 2018. TreeShrink: Fast and Accurate Detection of Outlier Long Branches in Collections of Phylogenetic Trees. *BMC Genomics* **19**, no. S5: 272. https://doi.org/10.1186/s12864-018-4620-2

Malinsky, M., M. Matschiner, and H. Svardal. 2021. Ds□□□□□ Fast *D*□Statistics and Related Admixture Evidence from VCF Files. *Molecular Ecology Resources* **21**, no. 2: 584–595. https://doi.org/10.1111/1755-0998.13265

Mallet, J. 2007. Hybrid Speciation. *Nature* **446**, no. 7133: 279–283. https://doi.org/10.1038/nature05706

Mamanova, L., A. J. Coffey, C. E. Scott *et al.* 2010. Target-Enrichment Strategies for Next-Generation Sequencing. *Nature Methods* **7**, no. 2: 111–118. https://doi.org/10.1038/nmeth.1419

Matesanz, S., E. Gianoli, and F. Valladares. 2010. Global Change and the Evolution of Phenotypic Plasticity in Plants. *Annals of the New York Academy of Sciences* **1206**, no. 1: 35–55. https://doi.org/10.1111/j.1749-6632.2010.05704.x

Mayrose, I., S. H. Zhan, C. J. Rothfels, N. Arrigo, M. S. Barker, L. H. Riesenberg, and S. P. Otto. 2015. Methods for Studying Polyploid Diversification and the Dead-End Hypothesis: A Reply to Soltis *et al*. (2014). *New Phytologist* **206**, no. 1: 27–35. https://doi.org/10.1111/nph.13192

McDonnell, A. J., W. J. Baker, S. Dodsworth *et al.* 2021. Exploring Angiosperms353: Developing and Applying a Universal Toolkit for Flowering Plant Phylogenomics. *Applications in Plant Sciences* **9**, no. 7: aps3.11443. https://doi.org/10.1002/aps3.11443

McKenna, A., M. Hanna, E. Banks *et al.* 2010. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data. *Genome Research* **20**, no. 9: 1297–1303. https://doi.org/10.1101/gr.107524.110

McLay, T. G. B., J. L. Birch, B. F. Gunn *et al.* 2021. New Targets Acquired: Improving Locus Recovery from the Angiosperms353 Probe Set. *Applications in Plant Sciences* **9**, no. 7: aps3.11420. https://doi.org/10.1002/aps3.11420

McTavish, E.J. and D. M. Hillis. 2015. How do SNP Ascertainment Schemes and Population Demographics Affect Inferences about Population History? *BMC Genomics* **16**: 266. https://doi.org/10.1186/s12864-015-1469-5

Médail, F., and K. Diadema. 2009. Glacial Refugia Influence Plant Diversity Patterns in the Mediterranean Basin. *Journal of Biogeography* **36**, no. 7: 1333–1345. https://doi.org/10.1111/j.1365-2699.2008.02051.x

Méndez-Tovar, I., H. Martín, Y. Santiago, A. Ibeas, B. Herrero, and M. C. Asensio-S.-Manzanera. 2015. Variation in Morphological Traits Among *Thymus Mastichina* (L.) L. Populations. *Genetic Resources and Crop Evolution* **62**: 1257–1267. https://doi.org/10.1007/s10722-015-0229-3

Metzker, M. L. 2010. Sequencing Technologies — the Next Generation. *Nature Reviews Genetics* **11**, no. 1: 31–46. https://doi.org/10.1038/nrg2626

Meudt, H. M., D. C. Albach, A. J. Tanentzap *et al.* 2021. Polyploidy on Islands: Its Emergence and Importance for Diversification. *Frontiers in Plant Science* **12**: 637214. https://doi.org/10.3389/fpls.2021.637214

Meyer, M., and M. Kircher. 2010. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harbor Protocols* **2010**, no. 6: pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Miguel, M. G., A. C. Figueiredo, M. M. Costa, D. Martins, J. Duarte, J. G. Barroso, and L. G. Pedro. 2003. Effect of the Volatile Constituents Isolated from *Thymus albicans*, *Th. mastichina*, *Th. carnosus* and *Thymbra capitata* in Sunflower Oil. *Nahrung* **47**, no. 6: 397–402. https://doi.org/10.1002/food.200390089

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, no. 5: 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mirarab, S. 2019. Species Tree Estimation Using ASTRAL: Practical Considerations. *arXiv*. https://doi.org/10.48550/arXiv.1904.03826

Morales, R. 1986. Taxonomía de los Géneros *Thymus* (Excluida la Sección *Serpyllun*) y *Thymbra* en la Península Ibérica. *Ruizia. Monografías del Jardín Botánico* **3**: 1–324

Morales, R. 2002. The History, Botany and Taxonomy of the Genus *Thymus*. In, *Thyme: The Genus* Thymus, edited by Stahl-Biskup, E., and F. Sáez, 1–43. London: CRC Press

Morales, R. 2010. *Thymus* L. In, *Flora Iberica. Plantas vasculares de la Península Ibérica e Islas Baleares. Verbenaceae-Labiatae-Callitrichaceae*, edited by Morales, R., A. Quintanar, F. Cabezas, A. J. Pujadas, and S. Cirujano, 349–409. Madrid: Consejo Superior de Investigaciones Científicas

Morales-Molino, C., J. M. Postigo-Mijarra, M. García-Antón, and C. Zazo. 2011. Vegetation and Environmental Conditions in the Doñana Natural Park Coastal Area (SW Iberia) at the Beginning of the Last Glacial Cycle. *Quaternary Research* **75**, no. 1: 205–212. https://doi.org/10.1016/j.yqres.2010.08.009

NCBI. 1988. National Center for Biotechnology Information. Bethesda (MD): National Library of Medicine (US). https://www.ncbi.nlm.nih.gov/

Nibau, C., A. Gonzalo, A. Evans, W. Sweet-Jones, D. Phillips, and A. Lloyd. 2022. Meiosis in Allopolyploid *Arabidopsis suecica*. *The Plant Journal* **111**, no. 4: 1110–1122. https://doi.org/10.1111/tpj.15879

Nieto Feliner, G., J. Casacuberta, and J. F. Wendel. 2020. Genomics of Evolutionary Novelty in Hybrids and Polyploids. *Frontiers in Genetics* **11**: 792. https://doi.org/10.3389/fgene.2020.00792

Oksanen, J., G. L. Simpson, F. G. Blanchet *et al.* 2025. vegan: Community Ecology Package. R Package Version 2.4–3. https://github.com/vegandevs/vegan

Ortiz, M. Á., K. Tremetsberger, A. Terrab *et al.* 2008. Phylogeography of the Invasive Weed *Hypochaeris radicata* (Asteraceae): From Moroccan Origin to Worldwide Introduced Populations. *Molecular Ecology* **17**, no. 16: 3654–3667. https://doi.org/10.1111/j.1365-294X.2008.03835.x

Otto, S. P., and J. Whitton. 2000. Polyploid Incidence and Evolution. *Annual Review of Genetics* **34**, no. 1: 401–437. https://doi.org/10.1146/annurev.genet.34.1.401

Ousmael, K. M., and O. K. Hansen. 2024. From phylogenomics to breeding: Can universal target capture probes be used in the development of SNP markers for kinship analysis? *Applications in Plant Sciences* 13, no. 1: e11624. https://doi.org/10.1002/aps3.11624

Paradis, E., and K. Schliep. 2019. *ape* 5.0: an Environment for Modern Phylogenetics and Evolutionary Analyses in R. *Bioinformatics* **35**, no. 3: 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pardo Otero, E., M. Pimentel, E. Sahuquillo Balbuena, and R. Piñeiro. 2024. Phylogenomic support for the allopolyploid origin of the northwest Iberian endemic orchid *Dactylorhiza cantabrica* with Hyb-Seq. *Journal of Systematics and Evolution* **63**, no. 2: 319–330. https://doi.org/10.1111/jse.13131

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE* **7**, no. 5: e37135. https://doi.org/10.1371/journal.pone.0037135

Pezzini, F. F., F. Ferrari, L. L. Forrest, M. L. Hart, K. Nishii, and C. A. Kidner. 2023. Target Capture and Genome Skimming for Plant Diversity Studies. *Applications in Plant Sciences* **11**, no. 4: e11537. https://doi.org/10.1002/aps3.11537

Phang, A., F. F. Pezzini, D. F. R. P. Burslem, G. S. Khew, D. J. Middleton, M. Ruhsam, and P. Wilkie. 2023. Target capture sequencing for phylogenomic and population studies in the Southeast Asian genus *Palaquium* (Sapotaceae). *Botanical Journal of the Linnean Society* **203**, no. 2: 134–147. https://doi.org/10.1093/botlinnean/boad022

Porturas, L. D., T. J. Anneberg, A. E. Curé, S. Wang, D. M. Althoff, and K. A. Segraves. 2019. A meta analysis of whole genome duplication and the effects on flowering traits in plants. *American Journal of Botany* **106**, no. 3: 469–476. https://doi.org/10.1002/ajb2.1258

Price, M. N., P. S. Dehal, and A. P. Arkin. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **5**, no 3: e9490. https://doi.org/10.1371/journal.pone.0009490

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, no. 2: 945–959. https://doi.org/10.1093/genetics/155.2.945

QGIS.org. 2024. QGIS Geographic Information System. QGIS Association. http://www.qgis.org

R Core Team. 2022. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rambaut, A. 2018. FigTree v1.4.4. https://github.com/rambaut/figtree

Ramos, A., O. Fernández, P. Terrinha, J. A. Muñoz, and Á. Arnaiz. 2020. Paleogeographic Evolution of a Segmented Oblique Passive Margin: The Case of the SW Iberian Margin. *International Journal of Earth Sciences* **109**, no. 6: 1871–1895. https://doi.org/10.1007/s00531-020-01878-w

Ramsey, J., and D. W. Schemske. 1998. Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants. *Annual Review of Ecology and Systematics* **29**, no. 1: 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467

Rincón Barrado, M., M. F. Perez, T. Villaverde, C. García Verdugo, J. Caujapé Castells, R. Riina, and I. Sanmartín. 2024a. Phylogenomics and Phylogeographic Model Testing Using Convolutional Neural Networks Reveal a History of Recent Admixture in the Canarian *Kleinia neriifolia*. *Molecular Ecology* **33**, no. 21: e17537. https://doi.org/10.1111/mec.17537

Rincón Barrado, M., T. Villaverde, M. F. Perez, I. Sanmartín, and R. Riina. 2024b. The Sweet Tabaiba or There and Back Again: Phylogeographical History of the Macaronesian *Euphorbia balsamifera*. *Annals of Botany* **133**, nos. 5–6: 883–904. https://doi.org/10.1093/aob/mcae001

Rodríguez-Ramírez, A. 2008. Geomorfología del Parque Nacional de Doñana. In, *Geología de Huelva: Lugares de interés geológico*, edited by Facultad de Ciencias, 150–155. Universidad de Huelva

Rodríguez-Ramírez, A., J. J. R. Villarías-Robles, J. N. Pérez-Asensio, and S. Celestino-Pérez, S. 2019. The Guadalquivir Estuary: Spits and Marshes. In, *The Spanish Coastal Systems*, edited by Morales, J. A., 517–541. Cham, Switzerland: Springer International Publishing AG. https://doi.org/10.1007/978-3-319-93169-2_22

Rodríguez Vidal, J. 2005. La Evolución Geomorfológica de Doñana. In, *Doñana, agua y biosfera*, edited by García Novo, F., and C. Marín Cabrera, 129–132. Madrid: Ministerio de Agricultura, Alimentación y Medio Ambiente, Confederación Hidrográfica del Guadalquivir. ISBN: 84-609-6326-8

Rodríguez-Sánchez, F., R. Pérez-Barrales, F. Ojeda, P. Vargas, and J. Arroyo. 2008. The Strait of Gibraltar as a Melting Pot for Plant Biodiversity. *Quaternary Science Reviews* **27**, no. 23–24: 2100–2117. https://doi.org/10.1016/j.quascirev.2008.08.006

Salony, S., J. Clo, F. Kolář, and C. Lafon Placette. 2025. A survey of triploid block inconsistency: potential mechanisms and evolutionary implications on polyploid speciation. *Journal of Experimental Botany* **76**, no. 8: 2096–2111. https://doi.org/10.1093/jxb/eraf050

Schmickl, R., and L. Yant. 2021. Adaptive Introgression: How Polyploidy Reshapes Gene Flow Landscapes. *New Phytologist* **230**, no. 2: 457–461. https://doi.org/10.1111/nph.17204

Schoenfelder, K. P., and D. T. Fox. 2015. The Expanding Implications of Polyploidy. *Journal of Cell Biology* **209**, no. 4: 485–491. https://doi.org/10.1083/jcb.201502016

Seear, P. J., M. G. France, C. L. Gregory, D. Heavens, R. Schmickl, L. Yant, and J. D. Higgins. 2020. A Novel Allele of Asy3 Is Associated with Greater Meiotic Stability in Autotetraploid *Arabidopsis lyrata*. *PLoS Genetics* **16**, no. 7: e1008900. https://doi.org/10.1371/journal.pgen.1008900

Segarra-Moragues, J. G., and P. Catalán. 2003. Life History Variation Between Species of The Relictual Genus *Borderea* (Dioscoreaceae): Phylogeography, Genetic Diversity, and Population Genetic Structure Assessed by RAPD Markers. *Biological Journal of the Linnean Society* **80**, no. 3: 483–498. https://doi.org/10.1046/j.1095-8312.2003.00247.x

Segarra-Moragues, J. G., M. Palop-Esteban, F. González-Candelas, and P. Catalán 2004. Characterization of Seven $(CTT)_n$ Microsatellite Loci in the Pyrenean Endemic *Borderea pyrenaica* (Dioscoreaceae): Remarks on Ploidy Level and Hybrid Origin Assessed through Allozymes and Microsatellite Analyses. *Journal of Heredity* **95**, no. 2: 177–183. https://doi.org/10.1093/jhered/esh028

Shee, Z. Q., D. G. Frodin, R. Cámara-Leret, and L. Pokorny. 2020. Reconstructing the Complex Evolutionary History of the Papuasian *Schefflera* Radiation Through Herbariomics. *Frontiers in Plant Science* **11**: 258. https://doi.org/10.3389/fpls.2020.00258

Sheidai, M., and F. Koohdar. 2023. Evidence for Ancient Introgression and Gene Flow in the Genus *Tamarix* L. (Tamaricaceae): a Computational Approach. *Genetic Resources and Crop Evolution* **70**, no. 6: 1653–1661. https://doi.org/10.1007/s10722-022-01525-0

Simmons, M. P., and J. Gatesy. 2021. Collapsing Dubiously Resolved Gene-Tree Branches in Phylogenomic

Coalescent Analyses. *Molecular Phylogenetics and Evolution* **158**, 107092. https://doi.org/10.1016/j.ympev.2021.107092

Slimp, M., L. D. Williams, H. Hale, and M. G. Johnson. 2021. On the Potential of Angiosperms353 for Population Genomic Studies. *Applications in Plant Sciences* **9**, no. 7: aps3.11419. https://doi.org/10.1002/aps3.11419

Soltis, D. E., V. A. Albert, J. Leebens Mack *et al.* 2009. Polyploidy and Angiosperm Diversification. *American Journal of Botany* **96**, no. 1: 336–348. https://doi.org/10.3732/ajb.0800079

Soltis, P. S., X. Liu, D. B. Marchant, C.J. Visger, and D. E. Soltis. 2014. Polyploidy and Novelty: Gottlieb's Legacy. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**, no. 1648: 20130351. https://doi.org/10.1098/rstb.2013.0351

Soltis, P. S., D. B. Marchant, Y. Van de Peer, and D. E. Soltis. 2015. Polyploidy and Genome Evolution in Plants. *Current Opinion in Genetics and Development*, **35**: 119–125. https://doi.org/10.1016/j.gde.2015.11.003

Soltis, P. S. and D. E. Soltis. 2009. The Role of Hybridization in Plant Speciation. *Annual Review of Plant Biology* **60**, no. 1: 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Šoštarić, I., Z. Liber, M. Grdiša, P. D. Marin, Z. Dajić-Stevanović, and Z. Šatović. 2012. Genetic Diversity and Relationships Among Species of the Genus *Thymus* L. (section *Serpyllum*). *Flora - Morphology, Distribution, Functional Ecology of Plants* **207**, no. 9: 654–661. https://doi.org/10.1016/j.flora.2012.06.018

Stebbins, G. L. 1985. Polyploidy, Hybridization, and the Invasion of New Habitats. *Annals of the Missouri Botanical Garden* 72, no. 4. 824. https://doi.org/10.2307/2399224

Stefani, S., L. Champoud, L. Fazan, M. Ronikier, M. Perret, G. Kozlowski, and C. Christe. 2025. Genetic divergence of *Clematis alpina* in the Swiss Prealps: a tale of the margins. *Alpine Botany* 135: 167–185. https://doi.org/10.1007/s00035-025-00333-3

Stull, G. W., K. K. Pham, P. S. Soltis, and D. E. Soltis. 2023. Deep Reticulation: the Long Legacy of Hybridization in Vascular Plant Evolution. *The Plant Journal* **114**, no. 4: 743–766. https://doi.org/10.1111/tpj.16142

Suarez-Gonzalez, A., C. Lexer, and Q. C. B. Cronk. 2018. Adaptive Introgression: a Plant Perspective. *Biology Letters* **14**, no. 3: 20170688. https://doi.org/10.1098/rsbl.2017.0688

Tajima, F. 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism.

Genetics **123**, no. 3: 585–595. https://doi.org/10.1093/genetics/123.3.585

Tiley, G. P., A. A. Crowl, T. O. M. Almary *et al.* 2024. Genetic Variation in *Loudetia simplex* Supports the Presence of Ancient Grasslands in Madagascar. *Plants, People, Planet* **6**, no. 2: 315–329. https://doi.org/10.1002/ppp3.10437

Treier, U. A., O. Broennimann, S. Normand, A. Guisan, U. Schaffner, T. Steinger, and H. Müller-Schärer. 2009. Shift in Cytotype Frequency and Niche Space in the Invasive Plant *Centaurea maculosa*. *Ecology* **90**, no. 5: 1366–1377. https://doi.org/10.1890/08-0420.1

Valdés Castrillón, B., Z. M. Díaz Lifante, and R. Parra Martín. 1999. *Thymus Albicans*. In, *Libro Rojo de la Flora Silvestre Amenazada de Andalucía. Tomo I: Especies en Peligro de Extinción*, edited by Valdés Castrillón, B., 286–289. Sevilla: Consejería de Medio Ambiente. Junta de Andalucía.

Vallejo Marín, M., A. M. Cooley, M. Y. Lee, M. Folmer, M. R. McKain, and J. R. Puzey. 2016. Strongly Asymmetric Hybridization Barriers Shape the Origin of a New Polyploid Species and its Hybrid Ancestor. *American Journal of Botany* **103**, no. 7: 1272–1288. https://doi.org/10.3732/ajb.1500471

Van de Peer, Y., T.-L. Ashman, P. S. Soltis, and D. E. Soltis. 2021. Polyploidy: an Evolutionary and Ecological Force in Stressful Times. *The Plant Cell* **33**, no. 1: 11–26. https://doi.org/10.1093/plcell/koaa015

Van de Peer, Y., E. Mizrachi, and K. Marchal. 2017. The Evolutionary Significance of Polyploidy. *Nature Reviews Genetics* **18**, no. 7: 411–424. https://doi.org/10.1038/nrg.2017.26

Vanneste, K., G. Baele, S. Maere, and Y. Van de Peer. 2014a. Analysis of 41 Plant Genomes Supports a Wave of Successful Genome Duplications in Association with the Cretaceous–Paleogene Boundary. *Genome Research* **24**, no. 8: 1334–1347. https://doi.org/10.1101/gr.168997.113

Vanneste, K., S. Maere, and Y. Van de Peer. 2014b. Tangled up in Two: a Burst of Genome Duplications at the End of The Cretaceous and the Consequences for Plant Evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**, no. 1648; 20130353. https://doi.org/10.1098/rstb.2013.0353

Vergés, J., J. C. Kullberg, A. Casas-Sainz *et al.* 2019. An Introduction to the Alpine Cycle in Iberia. In, *The Geology of Iberia: A Geodynamic Approach* Regional Geology Reviews, edited by C. Quesada, and J. T. Oliveira, 1–14. Cham, Switzerland: Springer International Publishing AG.

Villaverde, T., L. Pokorny, S. Olsson *et al.* 2018. Bridging the Micro- and Macroevolutionary Levels in Phylogenomics: Hyb-Seq Solves Relationships from Populations to Species and Above. *New Phytologist* **220**, no. 2: 636–650. https://doi.org/10.1111/nph.15312

Viruel, J., M. Conejero, O. Hidalgo *et al.* 2019. A Target Capture-Based Method to Estimate Ploidy from Herbarium Specimens. *Frontiers in Plant Science* **10**: 937. https://doi.org/10.3389/fpls.2019.00937

Viruel, J., O. Hidalgo, L. Pokorny, F. Forest, B. Gravendeel, P. Wilkin, and I. J. Leitch. 2023. A Bioinformatic Pipeline to Estimate Ploidy Level from Target Capture Sequence Data Obtained from Herbarium Specimens. In, *Plant Cytogenetics and Cytogenomics. Methods in Molecular Biology*, edited by T. Heitkam, and S. Garcia, 115–126. New York, NY: Springer US. https://doi.org/10.1007/978-1-0716-3226-0_5

Viruel, J., C. J. Sweeney, R. Day *et al.* 2025. Phylogenomic Barcoding of Soil Seed Bank–Persistent and Wind-Dispersed Non-Native Plant Species in South Georgia. *Molecular Ecology Resources* **26**, no. 1: e70068. https://doi.org/10.1111/1755-0998.70068

Wang, T., A. D. J. van Dijk, J. Bucher *et al.* 2023. Interploidy Introgression Shaped Adaptation during the Origin and Domestication History of *Brassica napus*. *Molecular Biology and Evolution* **40**, no. 9: msad199. https://doi.org/10.1093/molbev/msad199

Weiß, C. L., M. Pais, L. M. Cano, S. Kamoun, and H. A. Burbano. 2018. nQuire: A Statistical Framework for Ploidy Estimation Using Next Generation Sequencing. *BMC Bioinformatics* **19**, no. 1: 122. https://doi.org/10.1186/s12859-018-2128-z

Weitemier, K., S. C. K. Straub, R. C. Cronn, M. Fishbein, R. Schmickl, A. McDonnell, and A. Liston. 2014. Hyb-Seq: Combining Target Enrichment and Genome Skimming for Plant Phylogenomics. *Applications in Plant Sciences* **2**, 1400042. https://doi.org/10.3732/apps.1400042

Wenzell, K. E., A. J. McDonnell, N. J. Wickett, J. B. Fant, and K. A. Skogen. 2021. Incomplete reproductive isolation and low genetic differentiation despite floral divergence across varying geographic scales in *Castilleja*. *American Journal of Botany* **108**, no. 7: 1–-19. https://doi.org/10.1002/ajb2.1700

WFO. 2025. World Flora Online. http://www.worldfloraonline.org, accessed on 10 March 2025.

Wickham, H. 2016. *g*gplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023. dplyr: A Grammar of Data Manipulation. R package version 1.1.4. https://github.com/tidyverse/dplyr

Woudstra, Y., A. S. Quatela, N. C. A. M. Wagemaker, S. Ivanovic, T. Slotte, P. Meirmans, B. Gravendeel, & K. J. F. Verhoeven (2025). Universal Hyb-Seq kits capture considerable intraspecific variation: Less is more in herbarium-inclusive molecular ecology. bioRxiv. https://doi.org/10.1101/2025.10.14.682325

Young, A., T. Boyle, and T. Brown. 1996. The Population Genetic Consequences of Habitat Fragmentation for Plants. *Trends in Ecology and Evolution* **11**, no. 10: 413–418. https://doi.org/10.1016/0169-5347(96)10045-8

Yousefi, V., A. Najaphy, A. Zebarjadi, and H. Safari. 2015. Genetic Diversity and Geographic Dispersion in *Thymus* spp, as detected by RAPD markers. *Philippine Journal of Crop Science* **40**, no. 1: 82–88. https://www.ukdr.uplb.edu.ph/journal-articles/4960

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: Polynomial Time Species Tree Reconstruction from Partially Resolved Gene Trees. *BMC Bioinformatics* **19**, no. S6: 153. https://doi.org/10.1186/s12859-018-2129-y

Zuntini, A. R., T. Carruthers, O. Maurin *et al.* 2024. Phylogenomics and the Rise of the Angiosperms. *Nature* **629**, no. 8013: 843–850. https://doi.org/10.1038/s41586-024-07324-0

Tables and Figures

TABLE 1. Voucher information for studied populations, including identification according to Flora Iberica (currently recognised morphospecies), location, geographic coordinates, DNA content, ploidy-level estimation, and herbarium voucher numbers. Asterisks (*) indicate phenotypic variability in calyx length, both among and within individuals, which hinders reliable identification based on this character alone (*sensu* Morales *et al.* 2010).

Species	Population ID	Population abbreviation	Locality	Lat	Long	C DNA (pg)	Ploidy	Herbarium voucher #
Thymus albicans	A01	PUERTO REAL	Spain. Cádiz, Puerto Real. Dehesa de las Yeguas.	36.556	-6.135	1.59 ± 0.01 (8)	2x	SEV290208
Thymus albicans	A02	LA BARROSA	Spain. Cádiz, la Barrosa Suburban Park.	36.363	-6.170	1.60 ± 0.04 (9)	2x	SEV290209
Thymus albicans	A03	ROCHE	Spain. Cádiz, Conil de la Frontera. ZEC Pinar de Roche.	36.316	-6.133	1.63 ± 0.02 (10)	2x	SEV290210
*Thymus albicans**	A04, D03	LA JUNCOSILLA	Spain. Sevilla, Villamanrique. La Juncosilla. Doñana Natural Park.	37.192	-6.329	1.54 ± 0.02 (15)	2x	SEV290225
Thymus albicans	A05	FARO	Portugal. Algarve, Faro. Ria Formosa Natural Park.	37.028	-7.977	1.60 ± 0.02 (9)	2x	SEV290212
Thymus albicans	A06	VILLARRASA	Spain. Huelva, Villarrasa. Dehesa del Duque.	37.356	-6.604	1.50 ± 0.01 (5)	2x	SEV290213
Thymus albicans	A07	BONARES	Spain. Huelva, Bonares.	37.259	-6.725	1.52 ± 0.01 (10)	2x	SEV290214
*Thymus albicans**	A08	PINAR ALGAIDA	Spain. Doñana Natural Park. Pinar de la Algaida.	36.855	-6.308	1.53 ± 0.01 (11)	2x	SEV290215
Thymus albicans	A09	QUARTEIRA	Portugal. Algarve, Quarteira. Vale do Lobo.	37.062	-8.080	1.54 ± 0.03 (8)	2x	SEV290216
Thymus albicans	A10	VALE DO GARRAO	Portugal. Algarve, Almancil. Vale do Garrao. Ria Formosa Natural Park.	37.044	-8.038	1.55 ± 0.04 (7)	2x	SEV290217
Thymus albicans	A11	NIEBLA	Spain. Huelva, Niebla. Rabo Conejo.	37.489	-6.719	1.55 ± 0.02 (11)	2x	SEV290218

Species	Population ID	Population abbreviation	Locality	Lat	Long	C DNA (pg)	Ploidy	Herbarium voucher #
Thymus albicans	A12	LAS LOMAS	Spain. Cádiz, Vejer de la Frontera. Las Lomas.	36.299	-5.909	1.54 ± 0.05 (8)	2x	SEV290219
Thymus albicans	A13	EL SOTO	Spain. Cádiz, Vejer de la Frontera. El Soto.	36.242	-5.929	1.6 ± 0.02 (9)	2x	SEV290220
Thymus albicans	A14	CHICLANA	Spain. Cádiz, Chiclana. Pinar del Hierro.	36.388	-6.122	1.57 ± 0.02 (8)	2x	SEV290221
*Thymus albicans**	A15	VILLAMANRIQUE	Spain. Sevilla. Villamanrique de la Condesa. Doñana Natural Park. Dehesa de Gato.	37.192	-6.329	1.61 ± 0.03 (6)	2x	SEV290222
*Thymus mastichina subsp. donyanae**	D01	ALMONTE	Spain. Huelva, Almonte. El Porretal. ZEC Doñana North and West.	37.167	-6.497	1.50 ± 0.02 (9)	2x	SEV290223
*Thymus mastichina subsp. donyanae**	D02	HINOJOS	Spain. Huelva, Hinojos. ZEC Doñana North and West.	37.287	-6.446	1.51 ± 0.01 (10)	2x	SEV290224
*Thymus mastichina subsp. donyanae**	D04	PINHEROS DO MARIM	Portugal. Algarve, Quelfes. Pinheiros do Marim. Ria Formosa Natural Park.	37.034	-7.818	1.44 ± 0.02 (9)	2x	SEV290226
*Thymus mastichina subsp. donyanae**	D05	ABALARIO	Spain. Huelva, Abalario. Doñana Natural Park.	37.072	-6.651	1.51 ± 0.02 (8)	2x	SEV290227
*Thymus mastichina subsp. donyanae**	D06	RIBETEHILOS	Spain. Huelva, Ribetehilos. Doñana Natural Park.	37.150	-6.643	1.51 ± 0.01 (10)	2x	SEV290228
*Thymus mastichina subsp. donyanae**	D07	CORRAL LIEBRE	Spain. Huelva, Corral de la Liebre. Doñana National Park.	36.908	-6.414	1.47 ± 0.02 (9)	2x	SEV290229
*Thymus mastichina subsp. donyanae**	D08	CORRALILLO OSCURO	Spain. Huelva. Corralillo Oscuro. Doñana National Park.	36.987	-6.414	1.51 ± 0.01 (8)	2x	SEV290230
*Thymus mastichina subsp. mastichina**	M01	CAZALLA	Spain. Sevilla, Cazalla de la Sierra. Sierra Morena de Sevilla Natural Park.	37.908	-5.802	2.64 ± 0.06 (8)	4x	SEV290231
Thymus mastichina subsp. mastichina	M02	CORDOBA	Spain. Córdoba, El Vacar. ZEC Guadalmellato.	38.100	-4.833	1.58 ± 0.01 (9)	2x	SEV290232

Species	Population ID	Population abbreviation	Locality	Lat	Long	C DNA (pg)	Ploidy	Herbarium voucher #
Thymus mastichina subsp. mastichina	M03	SOUTH CARTAYA	Spain. Huelva, Cartaya. Campo Común de Abajo.	37.238	-7.080	1.56 ± 0.04 (8)	2x	SEV290233
*Thymus mastichina subsp. mastichina**	M04	GILENA	Spain. Sevilla, Gilena.	37.281	-4.901	2.68 ± 0.02 (9)	4x	SEV290234
*Thymus mastichina subsp. mastichina**	M05	GERENA	Spain. Sevilla, Gerena.	37.561	-6.159	2.71 ± 0.03 (11)	4x	SEV290235
*Thymus mastichina subsp. mastichina**	M06	MONTE GORDO	Portugal. Algarve, Monte Gordo.	37.186	-7.488	2.78 ± 0.05 (9)	4x	SEV290236
*Thymus mastichina subsp. mastichina**	M07	MADRID	Spain. Madrid. Hoyos de Manzanares. Cuenca Alta del Manzanares Regional Park.	40.602	-3.917	2.74 ± 0.04 (8)	4x	SEV290237
*Thymus mastichina subsp. mastichina**	M08	VILLALUENGA	Spain. Cádiz, Villaluenga del Rosario. Sierra de Grazalema Natural Park.	36.811	-5.391	2.71 ± 0.07 (9)	4x	SEV290238
*Thymus mastichina subsp. mastichina**	M09	RONDA	Spain. Málaga, Ronda, Pinsapar de la Nava. Sierra de las Nieves Natural Park.	36.665	-5.051	2.86 ± 0.06 (8)	4x	SEV290239
*Thymus mastichina subsp. mastichina**	M10	NORTH CARTAYA	Spain. Huelva, Cartaya. Campo Común de Arriba.	37.330	-7.100	1.55 ± 0.02 (11)	2x	SEV290240
*Thymus mastichina subsp. mastichina**	M11	EL PORTIL	Spain. Huelva, El Portil.	37.203	-7.014	1.57 ± 0.02 (12)	2x	SEV290241
*Thymus mastichina subsp. mastichina**	M12	PAJARONCILLO	Spain. Cuenca, Pajaroncillo.	39.930	-1.726	2.73 ± 0.03 (6)	4x	SEV290242
*Thymus mastichina subsp. mastichina**	M13	SAN CLEMENTE	Spain. Cuenca, San Clemente.	39.338	-2.472	2.84 ± 0.06 (8)	4x	SEV290243
*Thymus mastichina subsp. mastichina**	M14	SERRA DA ESTRELA	Portugal. Guarda, Serra da Estrela Natural Park.	40.379	-7.507	1.49 ± 0.04 (7)	2x	SEV290244

Species	Population ID	Population abbreviation	Locality	Lat	Long	C DNA (pg)	Ploidy	Herbarium voucher #
*Thymus mastichina subsp. mastichina**	M15	PONFERRADA	Spain. León, Ponferrada.	42.564	-6.771	2.64 ± 0.04 (6)	4x	SEV290245
*Thymus mastichina subsp. mastichina**	M16	TORO	Spain. Zamora, Toro.	41.516	-5.414	2.7 ± 0.06 (7)	4x	SEV290246
*Thymus mastichina subsp. mastichina**	M17	CERRO DA CABEÇA	Portugal. Loulé, Cerro da Cabeça Alta.	37.146	-8.108	2.64 ± 0.04 (3)	4x	SEV290492
*Thymus mastichina subsp. mastichina**	M18	PARIDERAS	Spain. Albacete, Parideras.	38.582	-2.526	2.52 ± 0.04 (8)	4x	SEV290247
Thymus carnosus	IB0432	OUTGROUP 1	Portugal. Faro. Ria Formosa Natural Park. Praia da ANção.	37.017	-8.017	N/A	N/A	MA887950
Thymus capitellatus	IB0437	OUTGROUP 2	Portugal. Sesimbra. Aldeia do Meco.	38.449	-9.175	N/A	N/A	MA643678
Thymus camphoratus	IB0459	OUTGROUP 3	Portugal. Faro. Vilanova de Milfontes.	37.740	-8.788	N/A	N/A	MA905261

Figure Captions

FIGURE 1 | Phylogenomic relationships and population structure of Iberian endemic *Thymus* sect. *Mastichina* individuals. (a) Multispecies coalescent (MSC) tree obtained with ASTRAL-III (branch lengths in substitutions per site were scaled with RAxML-NG) for 329 nuclear orthologs. Support values, measured as local posterior probabilities (LPPs), under 75% not shown. (b) Population structure inferred with ADMIXTURE from 2,223 filtered and unlinked SNPs for the most likely clustering scenarios (K = 5 is the optimum; see Figure S2). Columns named "a", "b" and "c" correspond to individual ploidy assignments (diploid, triploid and tetraploid, respectively) based on allelic ratio distributions as estimated with nQuire (see Figure S5). (c) Spatial distribution of genetic clusters across the Iberian Peninsula for K = 5. (d) An inset zooming into the SW Iberian Peninsula. Countries are denoted by their 2-digit ISO code.

FIGURE 2 | Unrooted MSC nuclear tree obtained with ASTRAL-III (branch lengths in substitutions per site were scaled with RAxML-NG) for 329 nuclear orthologs, showing the distribution of genetic groupings of *Thymus* sect. *Mastichina*. Colour scheme as in previous figure.

FIGURE 3 | Graphical representation of a dimensionality reduction technique and a multivariate method for 254 samples based on 2,223 filtered and unlinked SNPs. (a) Five-dimensional Principal Component Analysis (PCA), and (b) Discriminant Analysis of Principal Components (DAPC). (c) Pentagonal diagram showing mean pairwise fixation indexes (line thickness proportional to F_{ST} value) between the five genetic groupings identified in the phylogenomic and population structure analyses (see Figures 1 and 2). Colour scheme as in previous figures.

FIGURE 4 | Genetic diversity estimated as (a) nucleotide diversity (π) and (b) Tajima's D for *Thymus* sect. *Mastichina* populations (see Table S3). Colours denote the five genetic groupings identified in the phylogenomic and population structure analyses (see Figure 1). Colour scheme as in previous figures.

FIGURE 5 | Venn diagram showing private SNPs (numbers in the ellipses without any overlap) and shared SNPs (numbers in the overlapping areas among ellipses) of a random subsample of 11 individuals for each genetic grouping within *Thymus* sect. *Mastichina*.

FIGURE 6 | Diagrams showing the level of introgression as revealed by Dsuite (ABBA-BABA test, see Table S4) among the five genetic groupings in *Thymus* sect. *Mastichina*. (a) Sankey diagram illustrating gene flow between genetic groupings. (b) Schematic representation of introgression flow, with p-values indicating statistical significance. In both cases, the width of the bands or connection lines is proportional to the significance of the introgression. Colour scheme as in previous figures.

FIGURE 7 | Strength and direction of interploidy introgression, as inferred with fastsimcoal2, among tetraploid and diploid groups of *Thymus* sect. *Mastichina*. Upper tree diagrams show ABBA-BABA tests of genome-wide interploidy introgression based on Patterson's D statistic for the tetraploid group (4x) with

each of the four diploid subgroups: (a) Algarve (Alg); (b) Cádiz (Cád); (c) Doñana (Doñ); and (d) Hercynian (Herc). 2x represents the remaining diploid subgroups (see Table S4). Significant tests are shown with asterisks (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). Lower tree diagrams show preferred scenarios of isolation-with-migration as inferred from coalescent-based demographic analyses with fastsimcoal2. Bidirectional interploidy gene flow of asymmetrical strength between the tetraploid group and most diploid subgroups (a through c) and unidirectional gene flow from tetraploids to the Hercynian diploid subgroup (d). Migration is indicated forward in time. Columns and arrow thickness, and the number inside them, represent the effective population sizes and the number of migrants, respectively. Values above and below horizontal arrows represent migration rates. The maximum estimated log-likelihood (MaxEstLhood), the maximum observed log-likelihood (MaxObsLhood) and AIC index for each scenario are also shown.

Figure 1

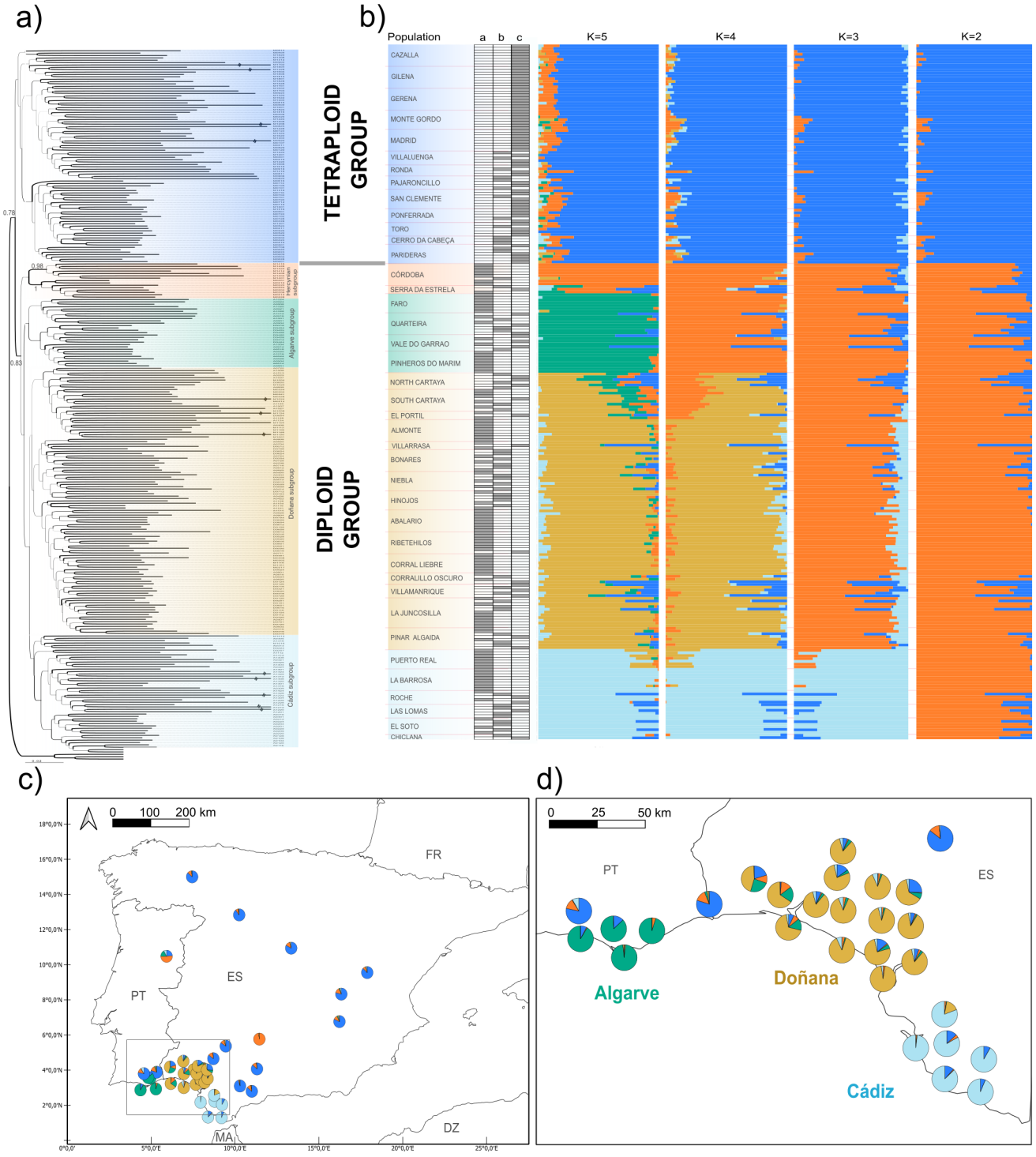

a)
b)
c)
d)
TETRAPLOID GROUP
DIPLOID GROUP
Population
a
b
c
K=5
K=4
K=3
K=2
CAZALLA
GILENA
GERENA
MONTE GORDO
MADRID
VILLALUENGA
RONDA
PAJARONCILLO
SAN CLEMENTE
PONFERRADA
TORO
CERRO DA CABEÇA
PARIDERAS
CÓRDOBA
SERRA DA ESTRELA
FARO
QUARTEIRA
VALE DO GARRAO
PINHEROS DO MARIM
NORTH CARTAYA
SOUTH CARTAYA
EL PORTIL
ALMONTE
VILLARRASA
BONARES
NIEBLA
HINOJOS
ABALARIO
RIBETEHILOS
CORRAL LIEBRE
CORRALILLO OSCURO
VILLAMANRIQUE
LA JUNCOSILLA
PINAR ALGAIDA
PUERTO REAL
LA BARROSA
ROCHE
LAS LOMAS
EL SOTO
CHICLANA
Hercynian subgroup
Algarve subgroup
Doñana subgroup
Cádiz subgroup
0.78
0.98
0.83
0 100 200 km
FR
ES
PT
MA
DZ
0 25 50 km
PT
ES
Algarve
Doñana
Cádiz

Figure 2

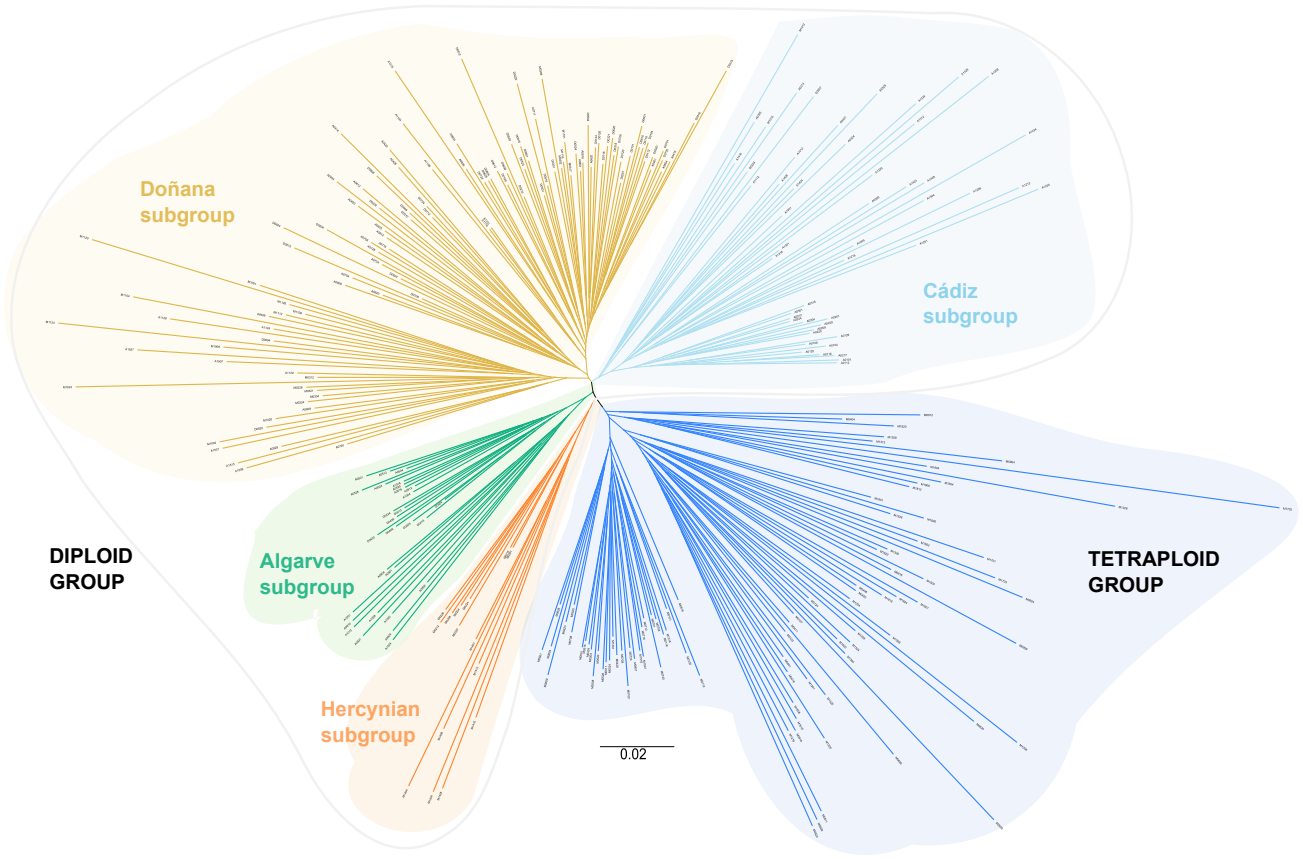

Doñana
subgroup
Cádiz
subgroup
DIPLOID
GROUP
Algarve
subgroup
TETRAPLOID
GROUP
Hercynian
subgroup
0.02

Figure 3

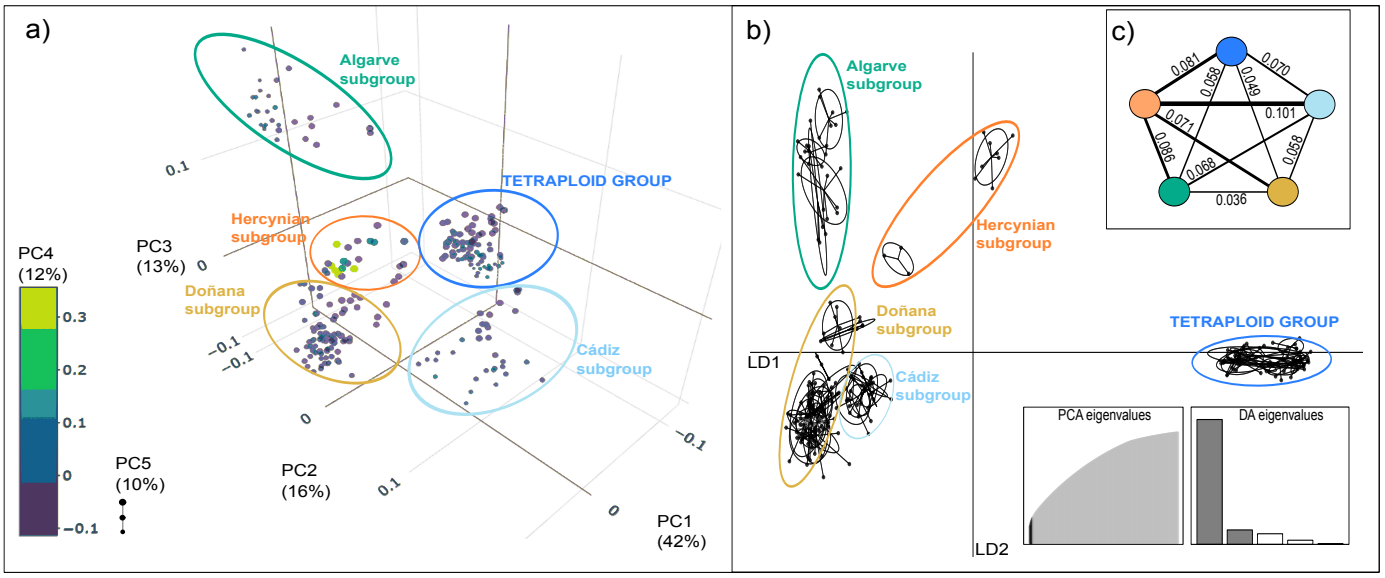

a)
Algarve subgroup
TETRAPLOID GROUP
Hercynian subgroup
Doñana subgroup
Cádiz subgroup
PC4 (12%)
PC3 (13%)
PC5 (10%)
PC2 (16%)
PC1 (42%)
0.3
0.2
0.1
0
−0.1
b)
Algarve subgroup
Hercynian subgroup
Doñana subgroup
Cádiz subgroup
TETRAPLOID GROUP
LD1
LD2
PCA eigenvalues
DA eigenvalues
c)
0.081
0.070
0.058
0.049
0.101
0.071
0.086
0.068
0.058
0.036

Figure 4

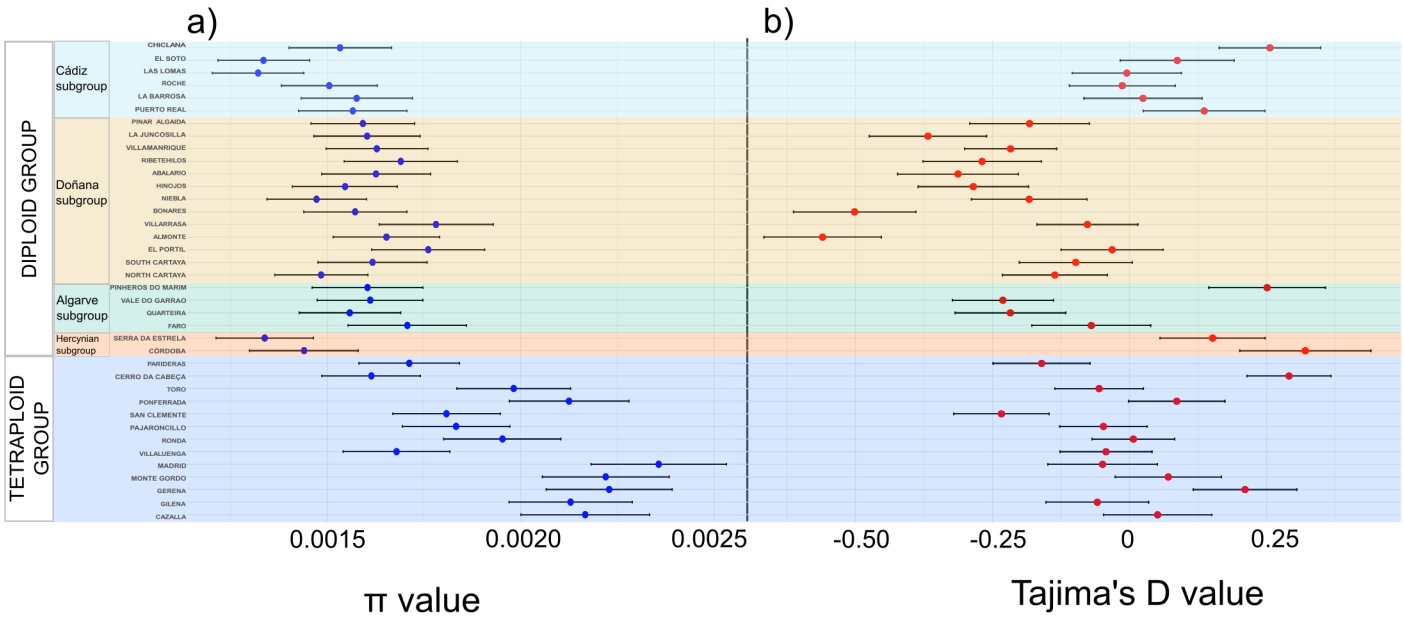

a)
b)
DIPLOID GROUP
TETRAPLOID GROUP
Cádiz subgroup
Doñana subgroup
Algarve subgroup
Hercynian subgroup
CHICLANA
EL SOTO
LAS LOMAS
ROCHE
LA BARROSA
PUERTO REAL
PINAR ALGAIDA
LA JUNCOSILLA
VILLAMANRIQUE
RIBETEHILOS
ABALARIO
HINOJOS
NIEBLA
BONARES
VILLARRASA
ALMONTE
EL PORTIL
SOUTH CARTAYA
NORTH CARTAYA
PINHEROS DO MARIM
VALE DO GARRAO
QUARTEIRA
FARO
SERRA DA ESTRELA
CÓRDOBA
PARIDERAS
CERRO DA CABEÇA
TORO
PONFERRADA
SAN CLEMENTE
PAJARONCILLO
RONDA
VILLALUENGA
MADRID
MONTE GORDO
GERENA
GILENA
CAZALLA
0.0015
0.0020
0.0025
π value
-0.50
-0.25
0
0.25
Tajima's D value

Figure 5

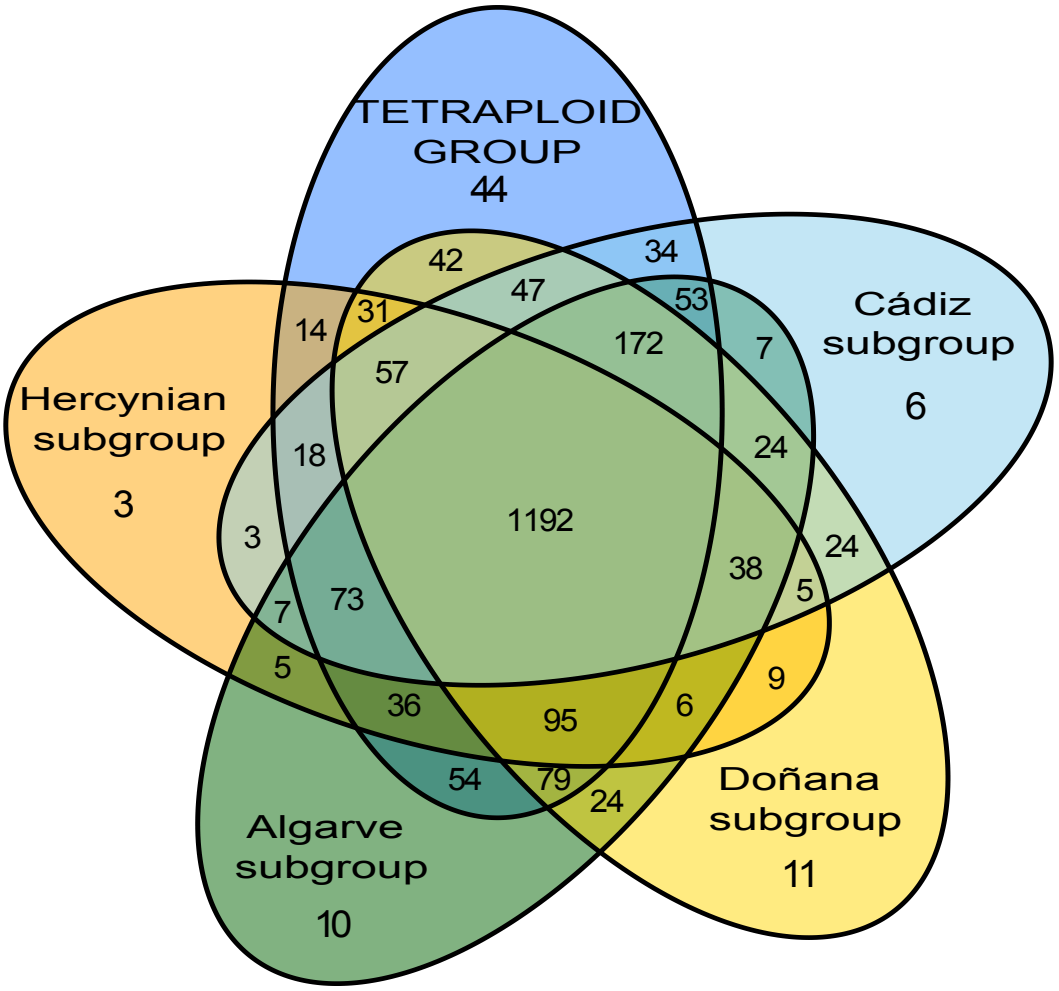

TETRAPLOID GROUP
44
Cádiz subgroup
6
Hercynian subgroup
3
Algarve subgroup
10
Doñana subgroup
11
42
34
47
53
31
14
172
7
57
18
24
1192
3
24
38
5
7
73
5
9
36
95
6
54
79
24

Figure 6

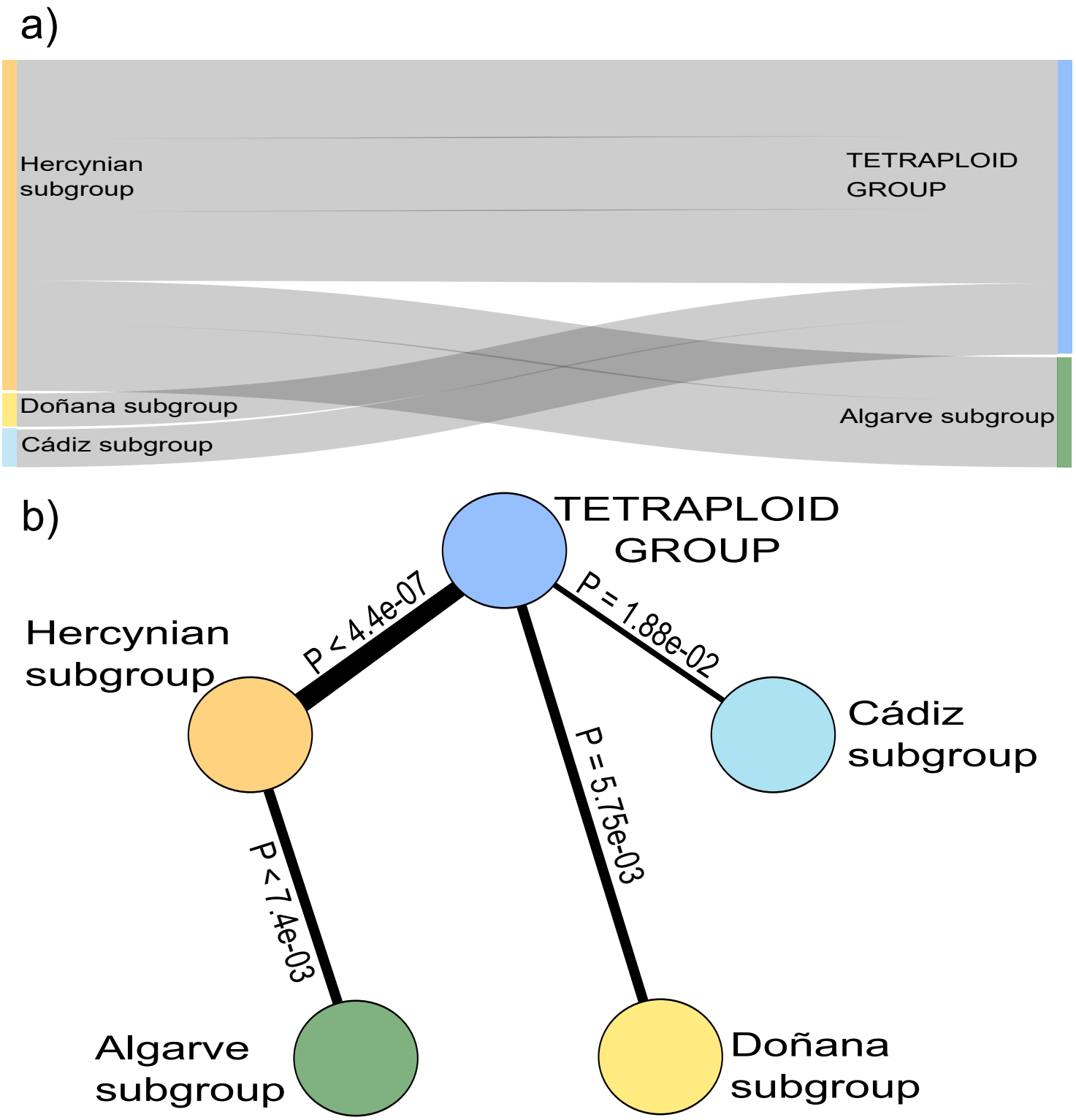
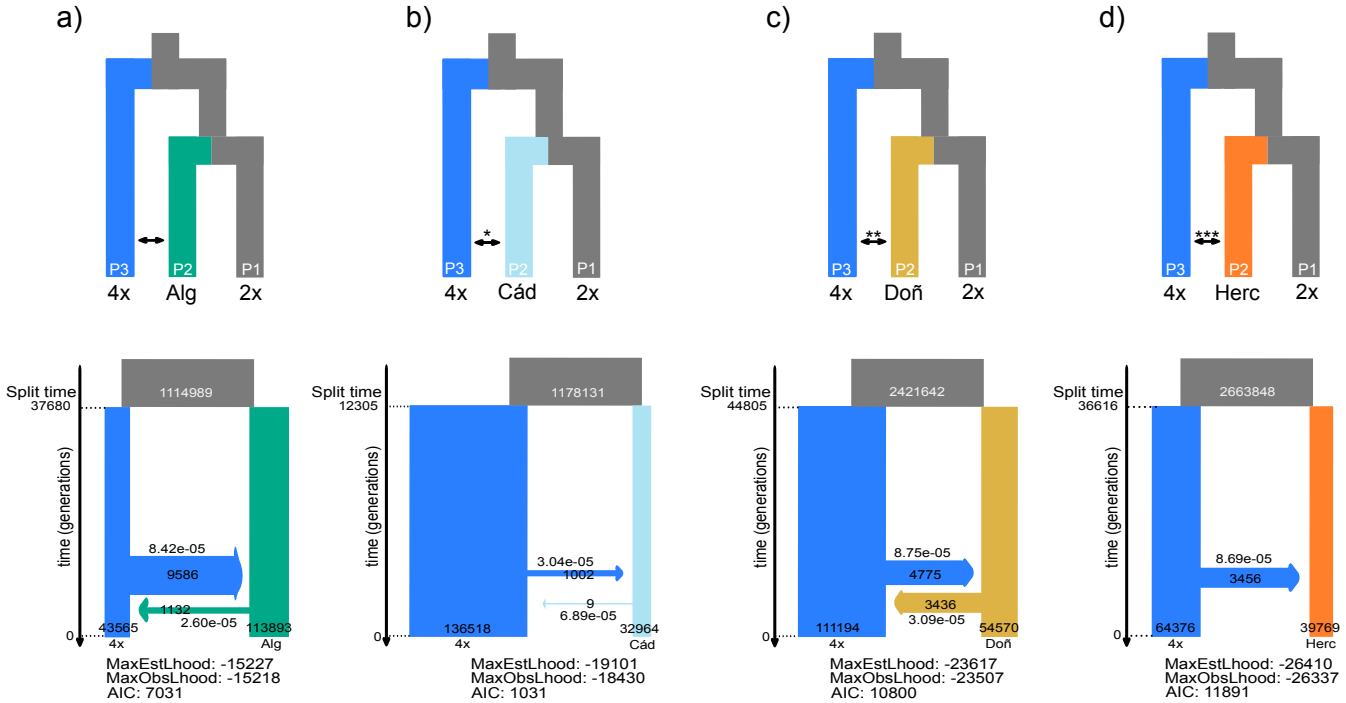

a)
Hercynian subgroup
Doñana subgroup
Cádiz subgroup
TETRAPLOID GROUP
Algarve subgroup
b)
TETRAPLOID GROUP
Hercynian subgroup
Cádiz subgroup
Algarve subgroup
Doñana subgroup
P < 4.4e-07
P = 1.88e-02
P = 5.75e-03
P < 7.4e-03

Figure 7


a)
P3
P2
P1
4x
Alg
2x
Split time
37680
1114989
time (generations)
8.42e-05
9586
1132
2.60e-05
43565
113893
0
4x
Alg
MaxEstLhood: -15227
MaxObsLhood: -15218
AIC: 7031
b)
P3
P2
P1
4x
Cád
2x
Split time
12305
1178131
time (generations)
3.04e-05
1002
9
6.89e-05
136518
32964
0
4x
Cád
MaxEstLhood: -19101
MaxObsLhood: -18430
AIC: 1031
c)
P3
P2
P1
4x
Doñ
2x
Split time
44805
2421642
time (generations)
8.75e-05
4775
3436
3.09e-05
111194
54570
0
4x
Doñ
MaxEstLhood: -23617
MaxObsLhood: -23507
AIC: 10800
d)
P3
P2
P1
4x
Herc
2x
Split time
36616
2663848
time (generations)
8.69e-05
3456
64376
39769
0
4x
Herc
MaxEstLhood: -26410
MaxObsLhood: -26337
AIC: 11891