Molecular Phylogenetics and Evolution 155 (2021) 107006

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Genetic similarities versus morphological resemblance: Unraveling a polyploid complex in a Mediterranean biodiversity hotspot

Noemí López-González [a,b,*], Javier Bobo-Pinilla [a,b], Nélida Padilla-García [a,b,1], João Loureiro [d], Silvia Castro [d], Blanca M. Rojas-Andrés [c], M. Montserrat Martínez-Ortega [a,b]

[a] Departamento de Botánica y Fisiología Vegetal, University of Salamanca, E-37007 Salamanca, Spain
[b] Biobanco de ADN Vegetal, University of Salamanca, Edificio Multiusos I+D+i, Calle Espejo s/n, 37007 Salamanca, Spain
[c] Department of Molecular Evolution and Plant Systematics & Herbarium (LZ), Institute of Biology, Leipzig University, Johannisallee 21-23, 04103 Leipzig, Germany
[d] Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, 3000-456 Coimbra, Portugal

ARTICLE INFO

Keywords:
Autopolyploidy
Allopolyploidy
Balkan Peninsula
Last Glacial Maximum
Microsatellites
Veronica

ABSTRACT

- The Balkan Peninsula is recognized as one of the hotspots of biodiversity in Europe. This area has shown since the Last Glacial Maximum appropriate conditions for species diversification and hybridization, which has led to the existence of numerous taxonomically unresolved entities. Here, we focus on the Western Balkans and explore the genetic structure and relationships among species belonging to the *V. austriaca* - *V. orbiculata* diploid-polyploid complex, including populations showing intermediate morphologies.
- A combination of nuclear markers (microsatellites), plastid DNA regions (*trnH-psbA*, *ycf6-psbM*) and ploidy level estimations using flow cytometry are employed to assess the genetic structure and evolutionary dynamics of this polyploid complex. To reconstruct the evolutionary history, an approximate Bayesian computation approach is combined with projections of the species distribution models onto the climatic scenarios of the Mid-Holocene (6 ka BP) and Last Glacial Maximum (22 ka BP).
- Four main groups were found: one well-established entity within the diploid level, *V. dalmatica*, a second diploid-tetraploid group which corresponds to *V. orbiculata*, a hexaploid cluster harboring *V. austriaca* subsp. *jacquinii* individuals, and an enigmatic tetraploid group. According to the molecular data obtained, this latter cluster represents an allopolyploid cryptic lineage —with *V. orbiculata* and *V. dalmatica* as putative parents— morphologically similar to *V. orbiculata*, but genetically more related to *V. austriaca* subsp. *jacquinii*. *Veronica dalmatica* and this "uncertain tetraploid" group are involved in the formation of the hexaploid taxon *V. austriaca* subsp. *jacquinii*, with the possibility of recent gene flow among different cytotypes.
- The present study supports a scenario of diversification from a diploid common ancestor leading to two different but interrelated lineages. The first one would correspond with the diploid *V. orbiculata* plus tetraploid individuals of this species arising through allo- and autopolyploidization, and the second one would involve all ploidy levels with allopolyploidization being prevalent.

* Corresponding author at: Departamento de Botánica y Fisiología Vegetal, University of Salamanca, E-37007 Salamanca, Spain.
E-mail address: noe_lg@usal.es (N. López-González).
[1] Present address: Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, 128 01 Prague 2, Czech Republic.

https://doi.org/10.1016/j.ympev.2020.107006
Received 1 March 2020; Received in revised form 20 October 2020; Accepted 29 October 2020
Available online 5 November 2020

1. Introduction

The species-rich Balkan Peninsula is recognized as one of the hotspots of biodiversity in Europe (Griffiths et al., 2004). This high diversity relies on several factors, namely, its location at the intersection of several major floras (Central European, Mediterranean, Anatolian, and Pontic) (Albach, 2006), its topographic complexity, and the variety of geological and palaeoclimatic events that have affected the area (Thompson, 2005). Particularly, the genetic structuring of many plant and animal species within this area was driven by the Pliocene tectonic events and the Quaternary climatic oscillations (e.g. Surina et al., 2011; Psonis et al., 2018).

In particular, mountain ranges have played a determining role in the survival of biota during glacial periods (Médail and Diadema, 2009) due to the presence of nunataks and corridors that granted accessibility to locations showing suitable conditions for survival during the Quaternary glaciations. Thus, the relative lesser impact of the glaciations in the Balkans (as compared with other European regions) linked with the presence of mountain ranges have contributed with refugia for both plants (e.g. Kutnjak et al., 2014) and animals (e.g. Krystufek et al., 2007). These historical and geological factors originated dramatic changes in species distribution ranges, promoting opportunities for species diversification and hybridization (Thompson, 2005). As a result of contacts among diverging lineages, taxonomically unresolved entities that mirror complex genetic relationships are frequently found in the area (e.g. Kučera et al., 2010; Lakušić et al., 2013; Rešetnik et al., 2016a), particularly in the Dinaric Alps of the western Balkan region, where the "refugia-within-refugia hypothesis" has been confirmed (Surina et al., 2011; Kutnjak et al., 2014). These complex evolutionary patterns within this region have passed historically unnoticed, possibly due to restricted access to a conflict zone (Krystufek et al., 2007).

Polyploidy is recognized as a significant force in plant evolution and its consequences are not completely understood due to its complexity (Soltis et al., 2007). The evolutionary importance of polyploidy relies on its potential to generate discontinuous phenotypic and physiological effects (Levin, 2002), to mediate reproductive isolations (Ramsey and Ramsey, 2014) and to change genome structure (Adams and Wendel, 2005). These structural and functional changes are recognized as sources of variation that facilitate evolution (Soltis and Soltis, 1999) and promote speciation (Zhan et al., 2016). However, there are inconsistent patterns of evolution following polyploidy and, thus, a "unified theory" of polyploidy has yet to be established (Soltis et al., 2016).

Based on their origin, two types of polyploids are generally recognized, i.e., autopolyploids, formed by intraspecific crosses and duplication of similar genomes and, allopolyploids, originated through interspecific hybridization and chromosome doubling of diverging genomes. Distinguishing between these mechanisms is crucial to assess the importance and evolutionary pathways of hybridization and polyploidization within particular plant groups (Madlung, 2013; Barker et al., 2016).

Veronica subsection *Pentasepalae* Benth. is a recently diversified polyploid complex (ca. 20 taxa) in which genetic isolation barriers are not definitely established (Martínez-Ortega et al., 2004; Rojas-Andrés et al., 2015; Padilla-García et al., 2018). The group is composed by several clades, among them a "core clade", which comprises most of the species from northern and central Europe, Italy and the Balkans (Rojas-Andrés et al., 2015). Four dominant even ploidy levels are found within it (from di- to octoploid), although the decaploid level has been reported once (Peev, 1972). Also, a few examples of homoploid hybridization, as well as auto- and allopolyploidization can be found, which greatly contribute to blur species limits (Padilla-García et al., 2018). Within the "core clade" of *V.* subsection *Pentasepalae*, the *V. austriaca* - *V. orbiculata* polyploid complex, whose diversification center is located in the Western Balkan Peninsula, includes: (i) a recently described diploid cryptic species endemic to the region, i.e., *Veronica dalmatica* Padilla-García, Rojas-Andrés, López-González & M.M.Mart.Ort., which has been traditionally confused with *V. austriaca* subsp. *jacquinii* (Baumg.) Watzl due to their morphological resemblance; (ii) an additional endemic species −*V. orbiculata* A. Kern− that includes morphologically indistinguishable di- and tetraploid individuals; (iii) the tetra- or hexaploid (exceptionally octo- and decaploid, according to relatively old literature references; Albach et al., 2008) *V. austriaca* subsp. *jacquinii* that is more widely distributed in central and SE Europe and the Caucasus; and (iv) individuals of unknown origin showing several ploidy levels (from di- to hexaploids) and morphologies (either intermediate between the previously presented taxa or related to one of them). The tetraploids that have been recorded as *V. austriaca* subsp. *jacquinii* / *V. orbiculata* due to their transitional morphology have been proposed to be allopolyploids, with diploid *V. dalmatica* and *V. orbiculata* as putative parental species (Rojas-Andrés et al., 2015; Padilla-García et al., 2018). Despite of this knowledge, the phylogenetic relationships among the taxonomic entities that conform the *V. austriaca* - *V. orbiculata* complex are not yet fully understood.

Although with low frequency, mixed-ploidy populations have been found for *V.* subsection *Pentasepalae*. Rojas-Andrés et al. (2020) reported 13 mixed-ploidy populations at a European scale, of which seven are from the Balkan region. From these, the four located in its Western part [composed of $2x + 4x$ (two), $2x + 6x$ (one), and $4x + 6x$ (one) cytotypes] have been included in the present study. The existence of mixed-ploidy populations suggests the presence of mechanisms that enable the co-existence of different cytotypes within a population (Kolář et al., 2017), but they may also suggest recurrent polyploid emergence or dynamic contact zones. In agreement with other authors (e.g. Liepelt et al., 2009; Krystufek et al., 2007) Rojas-Andrés et al. (2020) proposed the existence of primary (mixed-ploidy populations originated by autopolyploid events) and secondary contact zones (allopolyploidization events between pairs of more or less isolated taxa and further contact) in the Balkan Peninsula.

This work focuses on the primary and secondary contact zones of the Western Balkans and aims to unravel the evolutionary dynamics of the taxonomically intricate *V. austriaca* - *V. orbiculata* polyploid complex. For this, a combined approach is applied using data from nuclear markers (Simple Sequence Repeats, SSRs), plastid DNA (cpDNA) and ploidy levels estimated by flow cytometry (FCM). Present and past potential distributions are also evaluated through species distribution models (SDMs) and evolutionary reconstruction is performed through Approximate Bayesian Computation analysis (ABC). The specific goals of this study are: (1) to investigate the genetic relationships among different lineages and cytotypes to understand their evolutionary histories and correspondence with currently accepted taxa; (2) to assess the genetic affinities of the populations composed of individuals which are *a priori* difficult to assign to previously described taxa; (3) to establish hypotheses on the possible origins and ways of formation of the tetra- and hexaploid cytotypes; and (4) to reconstruct possible contact zones among taxa and cytotypes.

2. Materials and methods

2.1. Plant material

Details about sampling sites and initial taxonomic assignment are provided in Table A.1. The spatial distribution of the selected populations is displayed in Fig. 1. Fresh leaf material was collected and stored in silica gel. The complete dataset comprises 452 individuals corresponding to 28 populations. Populations usually show a low number of individuals, but whenever possible 15–22 individuals were sampled. In those locations where two morphotypes occurred, 15–22 individuals per morphotype were included (Table A.1: populations 17, 24, 26, and 28). Initial plant identification was based on the most recent taxonomic treatment (Rojas-Andrés and Martínez-Ortega, 2016) with the modifications added in Padilla-García et al. (2018). According with this, individuals were assigned to three taxa: *V. austriaca* subsp. *jacquinii*,

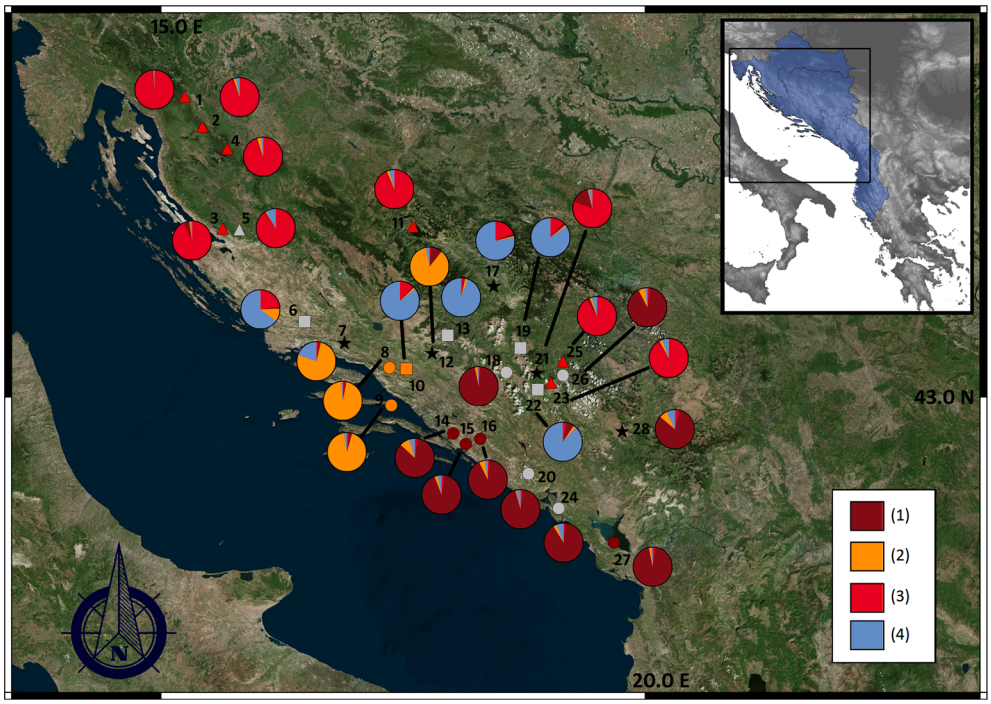

Fig. 1. Spatial distribution of the populations included in the present study. Ploidies are coded by shapes (2*x*: points; 4*x*: squares; 6*x*: triangles). Mixed-ploidy populations are indicated by black stars. Colors of these symbols correspond to initial determinations (orange: *V. orbiculata*; red: *V. austriaca* subsp. *jacquinii*; dark red: *V. dalmatica*; grey: "uncertain"). Pies represent the population genetic structure according to the Bayesian clustering. The segments of the pie denote the probability of belonging to each cluster: (1) *V. dalmatica* cluster; (2) *V. orbiculata* cluster; (3) *V. austriaca* subsp. *jacquinii* cluster; and, (4) "uncertain tetraploid" cluster. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

V. dalmatica and *V. orbiculata.* Individuals that could not be assigned to any of these three taxa were labelled as "uncertain", and catalogued according to their morphological affinity to these three taxa (Table A.1). Vouchers are deposited in the SALA herbarium (abbreviation according to Thiers, continuously updated).

2.2. DNA ploidy level estimations

DNA-ploidy levels of all individuals included in this study were estimated by flow cytometry from silica gel dried leaves. At least three individuals from each population were analyzed separately and the rest by pooled sampling strategy (except for the mixed-ploidy populations where all the individuals were analyzed separately). Nuclear suspensions were prepared following the method described by Galbraith et al. (1983) in which leaf tissue of each individual (up to five individuals in the pooled sampling) was chopped together with leaf tissue from an internal standard using a sharp razor blade in a Petri dish containing a buffer solution (in this case Woody Plant Buffer, Loureiro et al., 2007). Several internal standards were used depending on the C-value and standard availability, despite that in most cases *Raphanus sativus* was the chosen one (Table A.2). The suspension of isolated nuclei was filtered through a 48 μm nylon mesh, incubated with RNase and stained with a saturating solution of propidium iodide following Loureiro et al. (2007) and Rojas-Andrés et al. (2015). For each individual, one run of 5000 counts was made on a CyFlow SL (Partec GmbH, Münster, Germany; equipped with a 488 nm solid-state laser) or a CyFlow Space (Partec GmbH, Münster, Germany; equipped with a 532 nm solid-state laser). Results were acquired using Partec FloMax software v2.4d (Partec GmbH, Münster, Germany). DNA-ploidy level was estimated for each sample based on the C-value and the available chromosome counts for the studied species (Martínez-Ortega et al., 2004; Albach et al., 2008; Rojas-Andrés et al., 2015; Delgado et al., 2018).

2.3. Generation of SSR data and plastid DNA sequences

Total genomic DNA was extracted from silica-gel-dried material following the CTAB protocol (Doyle and Doyle, 1987) with slight modifications. The quality of the extracted DNA was checked on 1% TAE-agarose gels and the amount of DNA was estimated using a Nanodrop 2000C Spectrophotometer (Thermo Scientific). DNA extractions are deposited at the Plant DNA Biobank of the University of Salamanca (Spain).

2.3.1. SSR amplification, fragment analysis and genotyping.

Twelve SSR polymorphic primer pairs were employed (see Table 1 in López-González et al., 2015). Following the procedure developed by Schuelke (2000), the sequence-specific forward primers were marked at the 5′ end with an M13 tail (5′-TGTAAAACGACGGCCAGT-3′) (Eurofins) and then labeled with 5-FAM, VIC, NED, or PET fluorescent dyes (Life Technologies). All PCR reactions were performed on a Mastercycler-Pro thermocycler (Eppendorf). The extended version of the materials and methods, including PCR conditions, primers and reactive volumes are detailed in Appendix A (laboratory procedures).

PCR products were visualized on 2.5% TBE-agarose gels and multiplexed for genotyping. Fragment analysis was conducted at the *Unidad de Genómica - Campus Moncloa* (*Universidad Complutense de Madrid*, Spain) using the internal GeneScan 500 LIZ Size Standard (Applied Biosystems) on a multi-capillary sequencer ABI Prism 3730 (Applied Biosystems). Genotyping was carried out through GeneMarker AFLP/Genotyping Software version 1.8 (SoftGenetics, State College, Pennsylvania, USA). Reproducibility tests were performed over 5–10% of the total number of individuals. Two to three independent runs at different times were carried out. The outcomes indicated an average of 94.5%

Table 1
Summary of the total number, mean number, private number and shared number of SSRs alleles found across populations by cytotype. Shared alleles are given as percentages.

	2*x*	4*x*	6*x*	mixed-ploidy
Range	23–54	30–60	36–75	43–66
Mean	41.6	49	62.5	52.6
Private alleles	2	3	8	0
Shared alleles	**2*x***	**4*x***	**6*x***	**mixed-ploidy**
(vs. 4*x* pops.)	53.17	–		
(vs. 6*x* pops.)	57.94	61.11	–	
(vs. mixed-ploidy pops.)	60.32	50.73	69.05	–

(92–98%) match in the results, suggesting that microsatellite analyses are highly reproducible.

2.3.2. Plastid DNA amplification and sequencing

The cpDNA regions (*trnH-psbA* and *ycf6-psbM)* were selected according to Rojas-Andrés et al. (2015) due to their levels of variability. All amplified fragments were visualized on 1% TBE-agarose gels and purified with ExoSap-IT (USB Corporation) following the manufacturer's instructions. All PCR reactions were performed on a Mastercycler-Pro thermocycler (Eppendorf). PCR conditions, primers and reactive volumes are detailed in Appendix A (laboratory procedures).

PCR products were sequenced by Macrogen Inc. (Seoul, Korea) using an ABI Prism 3730XL DNA analyser (Applied Biosystems). Sequence identities were confirmed using the NCBI public BLAST service as implemented in Geneious Pro version 5.5.8 (Biomatters). Sequences were edited in Geneious and aligned with SATé-II version 2.2.5 (Liu et al., 2011). The total dataset comprises 194 *trnH-psbA* sequences and 194 *ycf6-psbM* sequences with five to ten individuals per population (all individuals in the case of populations 3 and 27). From these, 180 sequences were newly generated, whereas the other 14 were obtained from Rojas-Andrés et al. (2015). Voucher information and GenBank accession numbers are given in Table A.1.

2.4. Data analyses: SSR markers

Genotyping SSR data corresponding to polyploids present an important drawback, because it is not possible to distinguish the exact number of copies for a given allele. Thus, the scoring was done indicating which alleles were present, and the rest of the copies were coded as missing data. For those computer programs that allow ambiguous genotypes, codominant data were used as input (Structure, SPAGeDi). However, for some procedures microsatellite loci were treated as dominant markers by conversion to a binary matrix containing presence and absence of alleles (*K*-means, 'adegenet', 'hierfstat' and 'poppr' R packages).

2.4.1. Population genetic structure

To infer population structure and assign individuals to groups based on the SSR genotypes, two different approaches were considered: non-hierarchical *K*-means clustering (Hartigan and Wong 1979) and Bayesian clustering analysis based on the MCMC algorithm using Structure v.2.3.4 (Pritchard et al., 2000). These two methods were applied since it was not possible to check if the populations under study −most of them with ambiguous genotypes− follow the Hardy-Weinberg model. Non-hierarchical *K*-means clustering, which does not assume Hardy-Weinberg equilibrium, was performed using the R script of Arrigo et al. (2010) setting the number of repetitions to 500,000 and the number of groups (*K*) from 1 to 12. To determine the most probable value of *K*, the criterion described by Evanno et al. (2005) was followed. Structure uses model-based clustering and a Bayesian approach to identify clusters assuming fit to Hardy-Weinberg and linkage equilibrium. Multiple runs of Structure were performed by setting *K* from 1 to 12. The burn-in time and MCMC replication number were both set to 1,000,000 for each run and each run was replicated ten times. To determine the most probable value of *K*, Structure Harvester (Earl and vonHoldt, 2012) was used following the criterion described by Evanno et al. (2005). Results of independent Structure runs were summarized using Clumpp v.1.1.2b (Jakobsson and Rosenberg, 2007). For the visualization of results, bar-plots were generated using the R package 'strataG' (Archer et al., 2017) and a map displaying the probability of belonging to a cluster by population. For the latter, the software QGIS 2.10.1 Pisa (http://qgis.osgeo.org) and the package 'ggplot2' (Wickham, 2010) were employed.

Discriminant analysis of principal components (DAPC) (Jombart et al., 2010) was used as an additional approach for evaluating genetic groupings. DAPC is a multivariate method that combines principal component analysis with discriminant analysis. It provides a clear distinction of clusters by optimizing between-cluster variation, while minimizing within-cluster variation. *A priori* assignments, which are required, were based on the initial determinations (considering also ploidy levels). Analyses were performed using the R package 'adegenet' (Jombart, 2008).

The percentages of variance explained by the different clustering options were also compared by AMOVA analyses using the R package 'poppr' (Kamvar et al., 2014). Monte-Carlo tests with 999 repetitions were performed for calculating *p*-values. For these analyses the following groupings were considered: cytotype, taxon (initial identifications), and taxon-cytotype. These groupings were compared with the clustering resulting by non-hierarchical *K*-means clustering, Bayesian clustering, and supported clades resulting for the neighbor-joining (see below in this section). All the scripts used to analyze the data and generate the graphics are available on GitHub (https://github.com/NoeLG4/SSR).

Differentiation between pairs of populations as a function of their geographic distance was evaluated using the pairwise genetic distance statistic *Rho* (Ronfort et al., 1998). *Rho* is an inter-class relatedness coefficient allowing comparison among ploidy levels. Furthermore, *Rho* is unbiased with respect to inheritance pattern (Servick et al., 2015). Distances between all population pairs regarding cytotype (2*x*, 4*x*, 6*x*, and mixed-ploidy population pairs) were calculated. The same calculations were performed for the genetic groups obtained by Bayesian clustering. Isolation by distance (IBD) was then inspected through the regression of *Rho* values between population pairs on geographic distances (logarithmic scale). Statistical significance was tested using Mantel tests with 9999 permutations. *Rho* was calculated using SPAGeDi (Hardy and Vekemans, 2002), and regressions and Mantel tests were performed in R platform with the package 'vegan' (Oksanen et al., 2007).

In addition, pairwise among-population fixation indices (*F*st) were computed and 10,000 matrices were generated by bootstrapping over loci. These matrices were used to build a neighbor-joining tree (NJ). *F*st was calculated using 'hierfstat' package (Goudet, 2005) (the code employed for the generation of the matrices is available on GitHub). The neighbor-joining tree and bootstrap values were obtained through the programs Neighbor and Consense from Phylip software (Felsenstein, 2005). The resulting tree was visualized using SplitsTree software (http://www.splitstree.org/).

2.5. Data analyses: Plastid DNA (DIYABC)

For a comparison of the different scenarios which would explain the present relationships among the entities involved, an ABC (Approximate Bayesian Computation) statistical approach was applied through the software DIYABC v2.1 (Cornuet et al., 2014). DIYABC allows testing posterior probabilities of alternative scenarios simulating a large number of data sets for each case. The logistic regression procedure (Fagundes et al., 2007) assesses each scenario among simulated datasets allowing selecting the one which better fits the observed data.

Five groups were considered as a starting point for DIYABC. These groups were based on the four genetic clusters derived from Bayesian clustering, and DAPC and AMOVA results (see Results section 3.3). Additionally, *V. orbiculata* was subdivided into two groups, diploids and autopolyploids. Within *V. orbiculata* cluster, one individual suspected to be of allopolyploid origin was removed from the initial dataset to avoid excessive complexity in the alternative scenarios tested. Two additional individuals were deprecated due to probable initial misidentification (ind. 142, ind. 452; see Results section 3.3). Thus, the final groups considered were: diploid *V. orbiculata*, autotetraploid *V. orbiculata*, "uncertain tetraploid", diploid *V. dalmatica* and hexaploid *V. austriaca* subsp. *jaquinii*. A set of 36 alternative scenarios (Fig. B.1) was constructed for testing all plausible evolutionary hypotheses about: (1) the ancestral group and, (2) the formation of polyploid entities through

hybridization processes.

As an initial approach, parameters and *prior* distributions were chosen with large intervals due to the lack of ancestral information. Group sizes were set equally in all cases. Divergence times were unrestricted allowing the software to set the most likely value. Kimura two-parameter model (Kimura, 1980) was chosen and the mutation rate was sampled from a uniform distribution with limits set to [10^{-9}- 10^{-7}]. Because DIYABC selects the optimized rate from a given mutation range, we assigned mutation rates using a conservative confidence interval set between an extremely 'slow' rate (1.0×10^{-9} per site per generation) and an extremely 'fast' rate (1.0×10^{-7} per site per generation).

One million datasets were simulated for each scenario (Cornuet et al., 2008, 2010). The selection of the best scenario was made calculating the posterior probabilities of each one by performing a polychotomous weighted logistic regression on the 1% of simulated data sets closest to the observed data set (Cornuet et al., 2008, 2010). Subsequent distributions of parameters were evaluated under the best scenario using a local linear regression on the 1% closest simulated data sets with a logit transformation. Confidence in the selected scenario was checked by evaluating Type I and Type II error rates (Cornuet et al., 2010). The model adequacy for the best scenario was evaluated comparing the similarity between real data and simulated data through the posterior distribution of the parameter values.

2.6. Species distribution models (SDMs)

SDMs were performed considering the localities of *V. orbiculata* and *V. dalmatica*, once their identity was confirmed by the Bayesian clustering (see Results section 3.3). Additionally, three more localities of *V. dalmatica* were appended based on confirmed data by Padilla-García et al. (2018). Regarding *V. orbiculata*, ten more localities from herbarium specimens that could be unequivocally identified were added (Table A.3). Determinations were carried out by B. M. Rojas-Andrés and M. M. Martínez-Ortega. Vouchers are deposited in the herbaria K, LD, W, and WU. The potential current models were projected to reconstruct past distributions and search for possible contact zones between these two taxa during the Mid-Holocene (MidH, 6 ka BP) and the Last Glacial Maximum (LGM, 22 ka BP). The corresponding present and past Bioclim climatic layers available at www.worldclim.com (Hijmans et al., 2005) were downloaded to model climatic-suitability at past times. Climatic variables were downloaded at 2.5 arc-minute resolution to fit the spatial resolution of the LGM data set. Correlation analysis among bioclimatic variables was performed applying a threshold of 0.6 to avoid redundancy. The variance inflation factor value (VIF; Marquardt, 1970) was applied to test multi-collinearity using the 'car' package (Fox and Weisberg, 2011). VIF threshold was set to three. Due to the low number of occurrences no more than two variables should be selected, so it was decided to choose one related to temperature and another to precipitation. This information was combined with theoretical considerations to select the appropriate pair of climatic variables for the modelling. The couple of climatic features that are suspected to have a relevant influence on the niche and distribution limits of these species are: (i) mean temperature of the coldest quarter (bio11) and precipitation of the driest month (bio14) for *V. orbiculata*; (ii) mean diurnal range (bio2) and precipitation seasonality (bio15) for *V. dalmatica*.

The same set of variables was used under two different global circulation models: MIROC (Model for Interdisciplinary Research on Climate; Hasumi and Emori, 2004) and CCSM4 (The Community Climate System Model Version 4; Gent et al., 2011). Systematic sampling was implemented to avoid sampling bias, as described by Fourcade et al. (2014). Afterwards, multiple scenarios were evaluated using the R package 'ENMeval' (Muscarella et al., 2014), which implements the Maximum Entropy algorithm (Phillips et al., 2006). The models were run with the feature classes L, Q, H, LQ and LQH (where L = linear, Q = quadratic, H = hinge), and a regularization multiplier (rm) from 0.5 to 3 by 0.5. The selected method was the leave-one-out strategy (jackknife) to compensate for the low number of presence records (Pearson et al., 2007). The area under the curve (AUC) and the Akaike information criterion corrected for small sample sizes (AICc) were used to evaluate the models. Following these criteria, the model showing the lowest AICc and the best AUC was selected. AUC above 0.75 were considered potentially useful, 0.80–0.90 good, and 0.90–1.0 excellent (Elith, 2002). The palaeodistributions (MidH and LGM) were generated by projecting the best model onto the two past global circulation models using the package 'raster' (Hijmans and van Etten, 2014). The final consensus palaeodistribution was calculated as the average cell value between the CCSM4 and MIROC projections.

3. Results

3.1. DNA ploidy level variation

In this work, ploidy level estimations were newly generated for 368 individuals; the remaining estimations were obtained from Rojas-Andrés et al. (2015, 2020) and Padilla-García et al. (2018). Ploidies ranged from 2*x* to 6*x* (Table A.4). Twenty-three populations were composed of a single cytotype (ten comprising diploids, five tetraploids, and eight hexaploids), whereas five were mixed-ploidy populations. All ploidy estimations made using flow cytometry were in accord with previous chromosome counts for each species or subspecies. The "uncertain" individuals harbored all ploidy levels present in the dataset (2*x*, 4*x* and 6*x*). Heterogeneity in DNA-ploidy level within a taxon was found in *V. orbiculata* (2*x*, 4*x*), with the diploid cytotype being predominant. *Veronica austriaca* subsp. *jacquinii* was hexaploid in all the populations considered here with the exception of population 21 that showed one diploid individual (ind. 142; see 3.3). Populations initially identified as *V. dalmatica* were all diploid, and those assigned to the "uncertain" category and identified as *V. dalmatica / V. orbiculata* were also diploid except for one hexaploid individual (ind. 452; see Results section 3.3).

Significant IBD (at $p < 0.001$) between pairs of ploidy levels was found among diploid populations (i.e. "uncertain" diploid, *V. dalmatica* and diploid *V. orbiculata* populations). Significant IBD values (at $p < 0.05$) were also found in two other cases: diploid vs. tetraploid populations and diploid vs. mixed-ploidy populations (Table A.6).

3.2. Genetic variation: SSR profiles

From the initial set of 12 microsatellites, two were discarded due to genotyping incongruence and indels presence. A total of 126 alleles were amplified from the remaining 10 microsatellite loci that provided reliable genotypes. All loci were found to be highly polymorphic. Information regarding number of alleles, as well as number of private and shared alleles per population considering ploidy level is given in Table 1. The total number of alleles across hexaploid populations did not strongly exceed the number of alleles found in lower ploidy populations. For the 126 alleles, 63 were shared between the three cytotypes and the mixed-ploidy populations. The number of private alleles per population is shown in Table A.5. The highest allele-sharing values were between hexaploids and mixed-ploidy populations (69.05%), followed by hexaploids and tetraploid populations (61.11%). Considering cytotypes, the 13 private alleles were distributed as follows: eight in hexaploid populations, three in tetraploid populations and two in diploid populations. No private allele was found in mixed-ploidy populations.

3.3. Genetic structure and population differentiation

According to the method proposed by Evanno et al. (2005) and Arrigo et al. (2010) (for Structure and *K*-means, respectively), the best value for K is 2 (Fig. A.1a) and the second best K value is $K = 4$ (Fig. A.1a). However, the AMOVA analyses indicate that the clustering proposed by Structure and *K*-means at $K = 4$ explains the highest percentage of variance among groups (ca. 20.1%; Table 2) and the next best

Table 2
Analysis of molecular variance (AMOVA) of microsatellite data under different clustering options.

Criterion (*K*)	Source of variation	Df	Variation (%)[1]
No grouping	Among populations (considering subpopulations)	27 (31)	**29.609** (30.380)
STRUCTURE algorithm ($K = 4$)	Among clusters	3	**20.082**
	Among samples within clusters	38	15.096
	Within samples	410	64.822
STRUCTURE algorithm ($K = 2$)	Among clusters	1	**15.677**
	Among samples within clusters	33	19.731
	Within samples	417	64.592
K-means algorithm ($K = 4$)	Among clusters	3	**20.090**
	Among samples within clusters	40	15.168
	Within samples	408	64.742
K-means algorithm ($K = 2$)	Among clusters	1	**17.338**
	Among samples within clusters	30	18.779
	Within samples	420	63.883
NJ ($K = 3$)	Among clusters	2	**19.177**
	Among samples within clusters	33	16.534
	Within samples	416	64.289
Cytotype ($K = 3$)	Among clusters	2	**17.433**
	Among samples within clusters	30	17.581
	Within samples	419	64.986
Initial identifications ($K = 7$)	Among clusters	7	**16.047**
	Among samples within clusters	28	15.992
	Within samples	416	68.023
Initial identifications considering cytotype ($K = 11$)	Among clusters	10	**17.920**
	Among samples within clusters	27	14.690
	Within samples	414	67.390
30 km distance ($K = 14$)	Among clusters	13	4.128
	Among samples within clusters	14	25.693
	Within samples	424	70.178
50 km distance ($K = 10$)	Among clusters	9	8.162
	Among samples within clusters	18	22.113
	Within samples	424	69.725

K: number of groups; df: degrees of freedom
[1] $p < 0.001$ in cases indicated with bold numbers (i.e. except 30 km distance and 50 km distance)

value is at $K = 3$ (19.18%; Table 2). The Evanno method identifies the highest level of population structure; therefore, *K* may be underestimated if there is hierarchical structure (Janes et al., 2017). According to the Structure manual (Pritchard et al., 2010), higher values of *K* are justified when individuals are strongly assigned to all clusters (Fig. A.1c), and especially, if the population assignments make biological sense (Gilbert et al., 2012). As we are not searching for the *true K* but exploring genetic structure, it is preferable to examine more than one value of *K* (see e.g. Rosenberg et al., 2002).

***K* = 2**. When considering 2 as the possible number of clusters, *K*-means mostly differentiates between diploids and polyploids (percentage of variance explained 17.34%). Structure generates a group that contains all populations identified as *V. dalmatica* plus the diploid individuals initially assigned to the "uncertain" category and a second group harboring the remaining populations (i.e., *V. orbiculata*, *V. austriaca* subsp. *jacquinii* and "uncertain" tetra- and hexaploids) (Fig. A.1b). According to the AMOVA analyses this grouping showed lower percentage of variance explained (15.67%) than the *K*-means one.

***K* = 3**. Grouping considering $K = 3$ is coincident for Structure and *K*-means clustering methods, as well as the relationships between populations shown by the neighbor-joining phylogram (Fig. 2). One cluster corresponds to *V. orbiculata* (2*x*, 4*x*), a second cluster corresponds to *V. dalmatica* plus virtually all the diploids within the "uncertain" category, and a third cluster is composed by the remaining polyploids. Regarding the neighbor-joining phylogram internal relationships were overall poorly resolved (with few exceptions, bootstrap values (BS) < 50). *Veronica dalmatica* and *V. orbiculata* form two clearly supported groups (BS > 85), with *V. orbiculata* 2*x* and 4*x* receiving high statistical support (BS = 98.5). Polyploids considered together were also retrieved as a single supported group (BS = 90.8) but, considered separately ("uncertain tetraploids" and hexaploids) none of them form a strongly supported clade.

As regards the AMOVA analyses that considers cytotypes (i.e., three groups), it explains 17.43% of the total genetic variation. This value is similar to that obtained by the grouping that overall differentiates between diploids and polyploids (17.34%).

***K* = 4**. In this case the clusters revealed by Structure and *K*-means are overall coincident with 95.35% of the individuals identically assigned. This clustering identifies the following genetic groups: (1) *V. dalmatica* plus virtually all the diploids within the "uncertain" category; (2) the diploid individuals initially assigned to *V. orbiculata* and four tetraploid individuals, which belong to the mixed-ploidy populations (7 and 12) of *V. orbiculata*; (3) the individuals initially assigned to *V. austriaca* subsp. *jacquinii* plus the hexaploids within the category "uncertain" (population 5); and, (4) another cluster with tetraploid individuals of the "uncertain" category, a tetraploid population initially identified as *V. orbiculata* (population 10), and three tetraploid individuals which belong to a mixed-ploidy population of *V. orbiculata* (population 7) (Fig. 1). The assessment of differentiation among populations through DAPC also led to the described four main groups (Fig. 3). The "uncertain tetraploids" are placed between *V. austriaca* subsp. *jacquinii* and *V. orbiculata* clusters, partially overlapping with the latter, making difficult to establish a clear boundary among these entities.

Considering pairs of clusters, significant IBD (at $p < 0.05$) was found among cluster 1 populations (*V. dalmatica* plus "uncertain" diploids), between cluster 1 and cluster 2 (*V. orbiculata*), and between cluster 1 and cluster 3 ("uncertain tetraploid" populations) (Table 3).

Deviations from *K* = 4 clustering. Some individuals do not fit exactly in the previously described scheme (4.87% and 5.97% of the samples for Structure and *K*-means, respectively). These individuals mostly belong to three of the "uncertain" populations: one tetraploid population (population 6; *V.* aff. *austriaca* subsp. *jacquinii*) and two diploid populations (populations 24a and 28a, initially catalogued as *V. dalmatica* / *V. orbiculata*). Population 6 –"uncertain tetraploid"– shows four individuals (out of 16) within the *V. austriaca* subsp. *jacquinii* group following the Bayesian clustering but just one according to *K*-means. Five individuals (out of 21) in population 24a –"*V. dalmatica* cluster"– belong to *V. orbiculata* according to *K*-means, while considering Structure two of them cluster with the tetraploid individuals of the "uncertain" category and the other three belong to *V. dalmatica*. Population 28a –"*V. dalmatica* cluster"– shows three individuals (out of 15) within the *V. orbiculata* cluster according to both Structure and *K*-means. Following Structure there is an additional individual that clusters with tetraploid individuals of the "uncertain" category. However, these four individuals correspond to the *V. dalmatica* cluster according to DAPC analysis.

Another individual (ind. 452) from this population, which is hexaploid and therefore expected to cluster with *V. austriaca* subsp. *jacquinii*, belongs to the *V. dalmatica* cluster according to the Bayesian clustering. However, in the DAPC analysis it clusters with the *V. austriaca* subsp. *jacquinii* group (Fig. 3). A diploid individual (ind. 142) from population 21 initially determined as *V. austriaca* subsp. *jacquinii,* belongs to the *V. dalmatica* cluster according to both Structure and *K*-means. This individual also clusters with *V. dalmatica* according to the DAPC analysis (Fig. 3). These two last *exceptions/deviations* are most likely cases of field misidentifications.

***K* > 4**. Additional groupings considering geographical distance, initial identifications, and initial identifications considering also cytotypes showed lower values than the grouping based on $K = 4$.

3.4. Inference on population history using DIYABC models

Sequence information regarding total number of sites, polymorphic sites and levels of variation is given in Table 4. The list of all parameters,

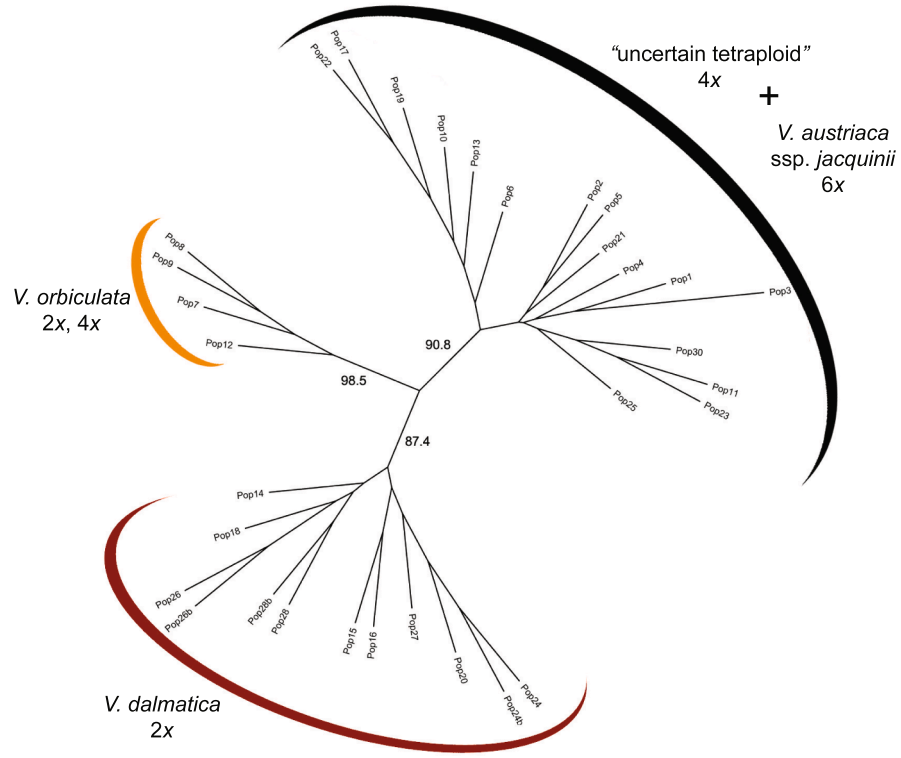

Fig. 2. Neighbor-joining analysis based on pairwise among-population fixation index of microsatellite data. Numerals by nodes are bootstrap values over 50%.

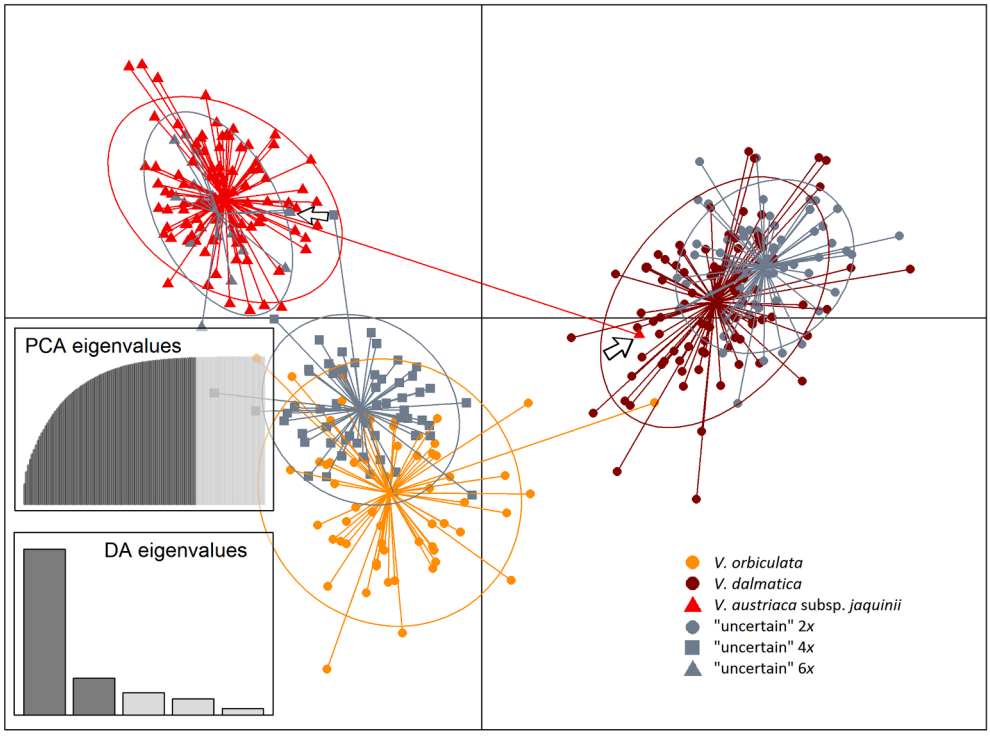

Fig. 3. Discriminant Analysis of Principal Components (DAPC). Initial assignments correspond to initial identifications according to Table A1. Colors correspond to initial determinations, "uncertain" category is represented in grey with ploidies coded by shapes (2*x*: points; 4*x*: squares; 6*x*: triangles). The two individuals indicated with an arrow are indiv. 142 from pop. 21 (red triangle), and indiv. 452 from pop. 28 (grey triangle). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

number of generations, and *prior/posterior* distributions used to model scenarios is summarized in Table B.1. The number of generations is remarkably low, although this number cannot be transformed into years automatically even less in the case of a perennial herb. The scenario with the highest posterior probability is Scenario 1 (P = 0.25 [0.23–0.27]) followed by Scenario 3 (P = 0.17 [0.15–0.19]), Scenario 5 (P = 0.16 [0.15–0.17]) and Scenario 7 (P = 0.14 [0.13–0.16]) (Fig. 4). The best scenario showed an initial divergence from a common ancestor between the diploids *V. orbiculata* and *V. dalmatica*. The "uncertain tetraploid" group seems to have originated from hybridization between the two diploid species with *V. dalmatica* as the greatest contributor (*r1* = 0.92). The following evolutionary event –a hybridization process too– is

Table 3
Isolation by distance tests. Correlation between pair-wise distances among populations (by Bayesian clustering) using *Rho* coefficients and the logarithm of the geographical distance.

C1				C2			C3		C4	
C1	C2	C3	C4	C2	C3	C4	C3	C4	C4	
0.222	0.246	0.011	0.197	0.553	0.047	0.012	~ 0	0.059	~ 0	*R-squared*
0.193	0.221	~0	0.182	0.441	0.019	~ 0	~ 0	0.043	~ 0	*Radj(*)*
*	*	**NS**	*	·	**NS**	**NS**	**NS**	**NS**	**NS**	***p***
0.011	**0.004**	**0.392**	**0.003**	**0.09**	**0.203**	**0.582**	**0.846**	**0.560**	**0.72**	

Abbreviations: C1 = *V. dalmatica* plus "uncertain" diploid individuals cluster; C2 = *V. orbiculata* cluster; C3 = *V. austriaca* subsp. *jacquinii* cluster; C4 = "uncertain tetraploid" individuals cluster; NS = not significant
· $p < 0.10$
* $p < 0.05$

Table 4
cpDNA sequence information.

Region	No. of sequences	Total no. of sites*	PolymorphicSites	Singleton variable sites	Parsimony informative sites
ycf6-psbM	194	623	24	1	23
trnH-psbA	194	291	10	3	7

* excluding sites with gaps / missing data

detected in the origin of the *V. austriaca* subsp. *jacquinii* group, with *V. dalmatica* and the "uncertain tetraploid" cluster as putative parents. In this case, *V. dalmatica* is also the main contributor to the allopolyploidization event ($r2 = 0.81$). Scenario 3 differs from scenario 1 in the progenitors implicated in the second allopolyploidization event. For this scenario, the diploid *V. orbiculata* –instead of *V. dalmatica*– together with the "uncertain tetraploid" group would be involved in the formation of the hexaploid *V. austriaca* subsp. *jacquini*. Scenario 5 is similar to scenario 1 but considers that the "uncertain tetraploid" cluster arose by autopolyplodization of *V. dalmatica*. Likewise, scenario 7 is parallel to scenario 3, but considers that the "uncertain tetraploid" cluster is the result of an autopolyplodization event of the diploid *V. orbiculata*. All scenarios indentify an autotetraploid origin for the *V. orbiculata* tetraploid individuals included and none identifies its participation in the origin of *V. austriaca* subsp. *jacquinii*.

Type I error rate for the best scenario was 43%, indicating that the probability that data sets simulated under this scenario are assigned to other scenarios. Type II error rate, which represents the probability that data sets simulated under other scenarios are assigned to the best one, was 12%. These relatively elevated values may be consequence of high similarities among scenarios. The similarity between real and simulated data sets for the best scenario is shown in Fig. B.2. From a total of 13 summary statistics only two cases of statistics diverged from the simulated ones (p-value < 0.05).

3.5. Modelling species distributions at LGM, MidH and present

The models corresponding to potential current distributions of the species showed high predictive accuracy, with AUC values of 0.95 and 0.94 for *V. orbiculata* and *V. dalmatica* respectively. Both models showed the lowest AICc values. For *V. orbiculata* the model selected was LQ (rm = 0.5). In the case of *V. dalmatica* the model was H (rm = 1). Considering variable importance for *V. orbiculata*, bio14 (precipitation of the driest month) shows slightly more explanatory power (53.47%) than bio11 (mean temperature of the coldest quarter; 45.53%). The highest explanative power for *V. dalmatica* relies on bio2 [mean of monthly (max temp - min temp); 83.35%].

The predicted potential current distributions of the species are displayed in Fig. A.2. *Veronica dalmatica* occurs in both coastal and interior environments. Suitable locations are also found in southern Greece and Balkan Peninsula inland mountain ranges. *Veronica orbiculata* appears restricted to coastal Mediterranean influenced environments. For both species there are ecological suitable localities in the Apennine Peninsula. Fig. 5 and Fig. A.3 show the areas that are suitable for the species during the LGM and MidH, respectively. Distribution areas predicted for the MidH are quite similar to the potential current distributions; thus, the possibilities of contact among populations are comparable. During the LGM suitable areas for *V. dalmatica* could have been displaced to the west in the Italian Peninsula and Sicily, to the southwestern Balkan coast and to the eastern inland mountains in the Balkans. For *V. orbiculata* suitable localities were mainly related to coastal environments in the northwestern areas of the Balkan Peninsula.

Although these models have to be taken with caution since the number of presences is remarkably low, these results do not support *a priori* hybridization between *V. dalmatica* and *V. orbiculata* during the LGM, given that the best suitable areas for those species do not overlap. Nevertheless, contact zones between populations could have appeared at the post-glacial period and during the Mid Holocene Climatic Optimum, when climatic conditions were quite similar to the current ones.

4. Discussion

4.1. On the origin of the diploids and the role of the Neretva river basin as a barrier to gene flow within the Balkan Peninsula

The *V. austriaca* - *V. orbiculata* polyploid complex is a system formed by diploids, autopolyploids, and allopolyploids that interact among them. The best clustering option (Fig. A.1, Table 2) revealed four distinct lineages: the diploid *V. dalmatica*, the diploid-tetraploid *V. orbiculata*, the "uncertain tetraploid" lineage and the hexaploid *V. austriaca* subsp. *jacquinii*. The most probable early evolutionary scenario according to the ABC analysis (Fig. 4) involves an initial divergence of two diploid lineages from a common ancestor: *V. orbiculata* and *V. dalmatica*. An additional possibility to explain the origin of *V. orbiculata* considering its position in the NJ (Fig. 2) and the introgression pattern observed in some individuals when considering $K = 2$ (e.g. population 12; Fig. A.1a) would be via homoploid hybridization, with *V. dalmatica* as one of the putative parental species. The genetic structure analyses based on AFLP-data for the whole subsection *Pentasepalae* (Padilla-García et al., 2018; see Fig. S2C) would also support this hypothesis. However, an accurate evaluation of this hypothesis with the available data is impossible due to the absence of extant candidates for the second parental species.

Formerly considered as the diploid level of *V. austriaca* subsp. *jacquinii* (Rojas-Andrés et al., 2015), *V. dalmatica* was recognized at the specific rank after exhaustive molecular and morphological analyses that demonstrated its monophyly and independence from *V. austriaca*

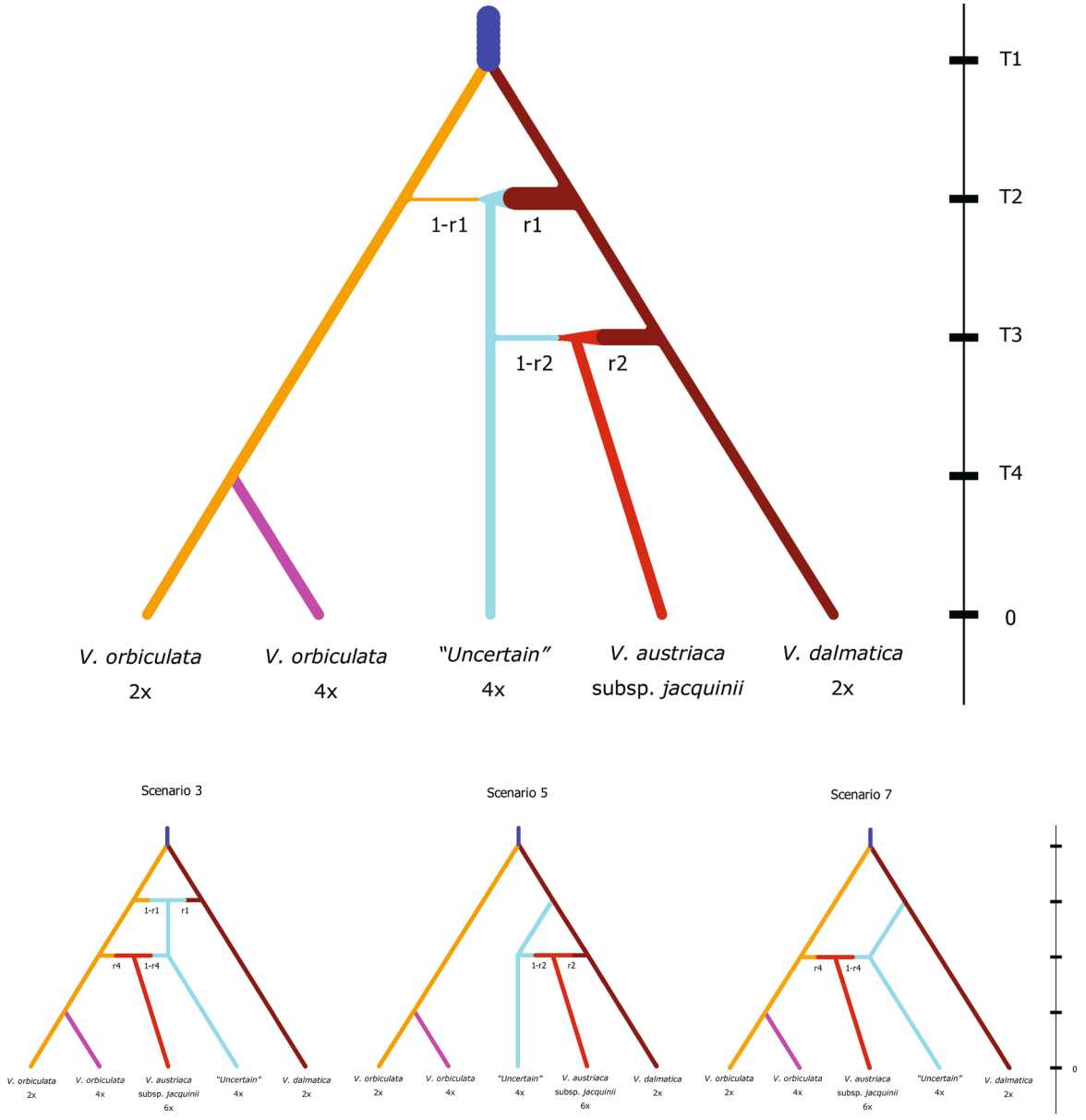

Fig. 4. Approximate Bayesian Computation analysis. Main: best scenario resulting from DIYABC analysis (i.e. scenario 1). Width of the horizontal lines shows the relative contribution to the hybridization events. The next best scenarios are displayed below (i.e. scenarios 3, 5 and 7).

subsp. *jacquinii* (Padilla-García et al., 2018). Here, the status of *V. dalmatica* as an independent species –based on SSR markers and cpDNA sequences– is confirmed (Figs. A.1, 2–4). Although *V. dalmatica* and *V. orbiculata* sometimes share morphological similarities, they appear as two discrete genetic entities. Genetic differentiation between both species shows a pattern of IBD (Table 3, Fig. 4), which is often associated with post-glacial colonization (Pope et al., 2006, Rešetnik et al., 2016b). Apart from the genetic differentiation, the combination of morphology, ploidy level and distribution can be employed for taxonomic delimitation: *Veronica dalmatica* and the "uncertain" diploids appear invariably restricted to the south-east of the Neretva river Basin, which confirms the distribution proposed by Padilla-García et al. (2018). This distribution is similar to other Balkan narrow endemics, such as *Dianthus ciliatus* subsp. *dalmaticus* (Fedorov and Kovanda, 1976), several species belonging to *Knautia* sect. *Trichra* (*Knautia dinarica* subsp. *dinarica*, *Knautia visianii*; Frajman et al., 2016) or the recently described *Iris orjenii* (Bräuchler and Cikovac, 2007) and *Campanula austroadriatica* (Lakušić et al., 2013). The Neretva river valley seems to represent a significant barrier to gene flow for animal (Podnar et al., 2004; Krystufek et al., 2007) and to some extent also for plant species (e.g., *Edraianthus tenuifolius*, Surina et al., 2011; *Alyssum austrodalmaticum*, Španiel et al., 2017). Indeed, this region to the southeast of the Neretva canyon –southern Dinaric Alps– has been highlighted for its endemism richness (Griffiths et al., 2004).

4.2. Formation of the polyploid complex

4.2.1. On the different origins of the tetraploid level

The lineage corresponding to *V. orbiculata* includes diploid, allo- and autotetraploid individuals. The tetraploid individuals of *V. orbiculata* from population 12 are genetically identical to each other –and to some diploids from the same population–, do not show more than two alleles per locus, and are morphologically indistinguishable from diploids of *V. orbiculata*. Thus, they are assumed to be of autopolyploid origin. However, an additional set of tetraploids were detected within the other *V. orbiculata* mixed-ploidy population (population 7). This population

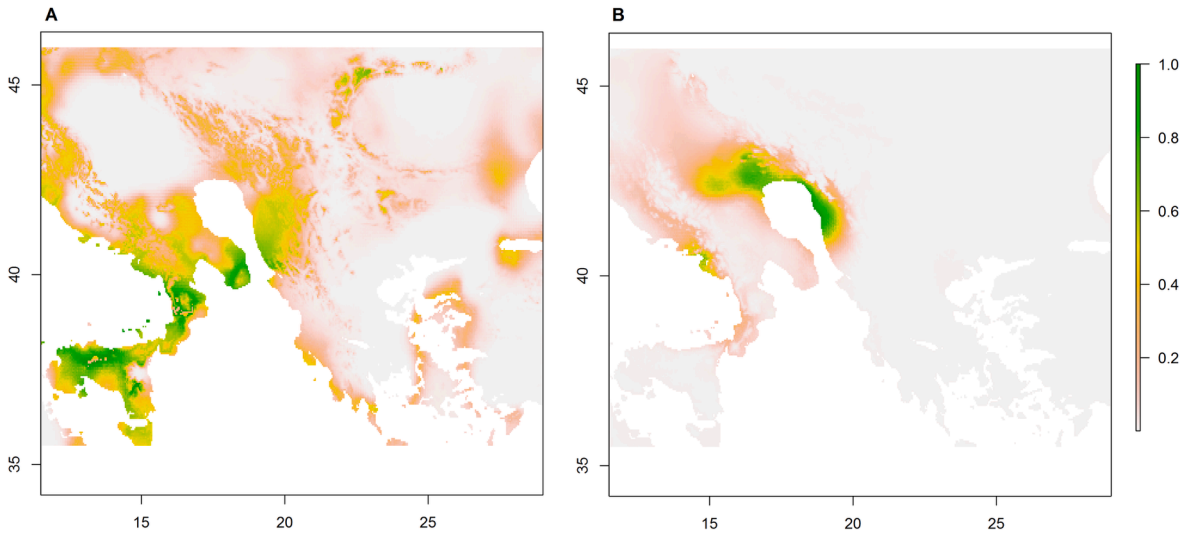

Fig. 5. Predicted environmental suitability for (A) *V. dalmatica* and (B) *V. orbiculata* under LGM conditions.

included five tetraploid individuals, three of them assigned to the "uncertain tetraploid" cluster, plus two that are genetically associated with *V. orbiculata* (Fig. A.1c). These individuals show different alleles than the diploid members of this population and they are assumed to be of allopolyploid origin. These individuals of allopolyploid origin genetically associated with the *V. orbiculata* cluster could have been the result of hybridization between *V. dalmatica* and *V. orbiculata* with a higher contribution of *V. orbiculata* and/or posterior backcrossing with *V. orbiculata*. The present results confirm the constitution of *V. orbiculata* as an entity with two ploidy levels and the origin of the tetraploids through two mechanisms, auto- and allopolyploidy. Interestingly, *V. orbiculata* and these allotetraploid individuals constitute another example of different outcomes that could arise from distinct allopolyploidization events in time and/or space, even when the same parental entities are involved (Steen et al., 2000; Ainouche et al., 2004; Brysting et al., 2007).

In the second lineage detected, allopolyploidy seems to have played a major role (Fig. 4). The diploid level is represented by *V. dalmatica* and horizontal connections are detected. The presence of a cryptic allotetraploid lineage within the complex with *V. dalmatica* and *V. orbiculata* as parental species is here confirmed for the first time. Polyploid complexes are a source of cryptic lineages that have remained unrecognized due to their unclear (e.g., sometimes they are phenotypically similar to the parents and are consequently overlooked) and/or highly variable morphologies. Thus, cryptic lineages are frequently detected when different cytogenetic units are examined in detail or when the population structure of polyploid systems is accurately studied (e.g. Majure et al., 2012; Crowl et al., 2017; Hu et al., 2019; Afonso et al., 2020). We here demonstrate that this is also the case of the *V. austriaca* - *V. orbiculata* complex, within which the existence of cryptic diversity has been neglected to date. This has traditionally obscured the detection of the "uncertain tetraploid" lineage that has acted as a progenitor in the formation of other higher level polyploid taxa within the complex (see section 4.2.2.).

4.2.2. Origin and spread of the hexaploid level

The hexaploid level, i.e., *V. austriaca* subsp. *jacquinii* cluster, represents another case of allopolyploidization with *V. dalmatica* and the "uncertain tetraploid" cluster involved in its origin. As already stated, *V. austriaca* subsp. *jacquinii* has been traditionally confused with *V. dalmatica* due to their morphological resemblance. Several examples illustrate that it is not rare that an allopolyploid event originates individuals in which one parental morphology is retained (Alexander-Webber et al., 2016) or at least −as it is in our case− dominates, even though increased fitness and/or sometimes difficult to detect slight phenotypic changes (e.g., increased size) can occur (Buggs et al., 2009; Welles and Ellstrand, 2019). Although increased fitness of *V. austriaca* subsp. *jacquinii* needs to be experimentally tested (e.g. enhanced dispersal abilities, increased ecological tolerances), it is interesting to note its wide distribution range. Thus, while the diploid-tetraploid lineages within *V. austriaca* - *V. obiculata* complex are mostly restricted to the west coast of the Balkans and the closest mountain ranges (Dinaric Alps) (Fig. 1), the hexaploid lineage extends its range across the Balkan Peninsula reaching the Carpathians (Rojas-Andrés and Martínez-Ortega, 2016).

This pattern of restricted diploid and widespread polyploid distribution has been found in many other polyploid complexes (e.g. Stebbins 1971; Lowry and Lester, 2006; Majure et al., 2012). Polyploidization has been suggested as a mechanism for generating new evolutionary lineages with broader ecological tolerances than their diploid progenitors, consequence of their increased genetic diversity (Otto and Whitton, 2000). However, polyploidy (and genetic diversity) *per se* does not explain an enlargement of geographical ranges or niche breadth in other polyploid complexes (e.g. Theodoridis et al., 2013; Castro et al., 2019). Within *V. austriaca* - *V. orbiculata*, the "uncertain tetraploid" lineage shows higher genetic diversity than their diploid relatives (Table 4), but its niche breadth does not seem to be neither distinct nor broader than that of the diploids. Polyploidization generates novelty, but only certain unique combinations of newly generated variation may be suitable to occur in ecological niches different from those occupied by their progenitors. Not broader niches but distinct niches (ecological niche differentiation) have been proposed to be crucial for polyploid success (Baniaga et al., 2020). The ability to occupy niches separate from those of the parental entities may be related with slight morphological differences potentially involving functional significance (Hao et al., 2013; López-Jurado et al., 2019) and this seems to be also the case for the hexaploid *V. austriaca* subsp. *jacquinii*.

4.2.3. The importance of the direction in allopolyploidization events

Conflicting signal between data sources (usually plastid and nuclear information) has been frequently interpreted as a signature of hybridization (Friar et al., 2008). As already explained, the evolutionary reconstruction suggests that *V. dalmatica* is the main donor to both allopolyploidization events (Fig. 4). This conclusion is based on plastid DNA data, which point to *V. dalmatica* as female parental on many more occasions than the "uncertain tetraploids" and *V. orbiculata*. Contrastingly, the analyses based on nuclear data place the allopolyploid entities closer to *V. orbiculata* (Fig. 3) and show that the most divergent cluster

within the polyploid complex is the *V. dalmatica* group (Fig. A.1b). This implies that *V. orbiculata* and the "uncertain tetraploids" could have acted as the predominant pollen donors in both allopolyploidization events. Although not much information on the reproductive biology of these entities is available, our preliminary unpublished data indicate that the style of the tetraploids is slightly longer than that of the diploids. Also consistently larger sized pollen grains corresponding to higher ploidy levels have been described within *V.* subsect *Pentasepalae* (Martínez-Ortega et al., 2000). According to Husband et al. (2002), the consistently high siring success of pollen from tetraploids in *Chamerion angustifolium* appears to be facilitated by differential success of pollen tubes. Surveys of pollen grain size (or volume) in many taxonomic groups including *Veronica* have revealed positive correlations with pistil length and corolla size (Sánchez-Agudo et al., 2009) and even with pollen-tube growth rates (Williams and Rouse, 1990). All these facts together would contribute to reinforce the idea that polyploids would tend to mainly act as pollen donors.

Unidirectional mating from diploids to polyploids is known from a great number of polyploid complexes (e.g. Slotte et al., 2008; Jørgensen et al., 2011; Diallo et al., 2016). Furthermore, it is assumed that crosses will produce viable seeds when the ratio of maternal to paternal genomes is at the required 2:1 proportion in the endosperm (Endosperm Balance Number [EBN] hypothesis; Johnston and Hanneman, 1982). However, this does not imply that the maternal donor has to be necessarily the high ploidy parental in an inter-ploid cross: for correct endosperm function, the EBN but not the ploidy levels should be in a 2:1 maternal to paternal ratio (Lafon-Placette et al., 2016). Studies in *Chamerion angustifolium* (Burton and Husband, 2000), *Impatiens* (Arisumi, 1982), *Solanum* (Johnston and Hanneman, 1982), and *Trifolium* (Parrott and Smith, 1986) support this hypothesis. Several works have found the diploid species acting as the maternal parent in the formation of polyploids (e.g. Popp and Oxelman, 2007; Peng et al., 2010; Schinkel et al., 2017). In fact, in some polyploid complexes seed germination is normal after inter-cytotype crosses and post-zygotic reproductive isolation is only strong when diploids act as pollen donors (e.g. *Leucanthemum pluriflorum* group; Greiner and Oberprieler, 2012).

4.2.4. The tetraploid bridge

Polyploids show increased genetic diversity compared to diploids (Table A.5) and a high degree of shared alleles (61.11%; Table 1). This notable genetic relatedness (Fig. 2) could be the result of recent or even ongoing inter-ploidy gene flow among natural populations. Similarly, the occurrence of distinct tetraploid lineages points to a possible existence of gene flow between diploid and tetraploid cytotypes probably mediated through unreduced gametes. Thus, the tetraploid level may act as a *bridge* not just in the formation of hexaploids but in the potential gene flow among cytotypes. The genetic divergence observed between diploids and hexaploids (Table 2, Fig. 2) suggests that direct gene flow is relatively lower between these cytotypes.

It is well-known nowadays that polyploid events do not necessarily promote instantaneous speciation (Ramsey and Schemske 1998). The absence of reproductive isolation among cytotypes could allow sexual reproduction and, as a consequence, prevent formation of more differentiated genetic groups than the observed ones (Fig. 2). Studies focusing on other polyploid complexes have found weaker postzygotic isolation among high order polyploids than in diploid-polyploid crosses (Sonnleitner et al., 2013; Sutherland and Galloway, 2017). The role of the "uncertain tetraploid" as a *bridge* within the polyploid complex goes in accordance with the proposal that natural polyploidization is not only involved in establishing barriers to hybridization, but also in overthrowing them (Lafon-Placette et al., 2017).

Despite the genetic relatedness, *V. austriaca* subsp. *jacquinii* and the "uncertain tetraploid" allopolyploid clusters can be morphologically differentiated as occurs in population 17, where it was possible to identify individuals belonging to *V. austriaca* subsp. *jacquinii* (pop. 17b, Table A.1) and those showing an intermediate morphology (pop. 17a, Table A.1). Structure assigned the individuals of pop. 17 to two clusters, which is in agreement with the initial determinations.

4.3. The species concept in the framework of polyploid complexes: To assign taxonomic rank or not to assign it, that is the question.

Species recognition and species delimitation are critical aspects for taxonomy with many implications for management and biodiversity conservation. The analyses presented here confirm that *Veronica austriaca* subsp. *jacquinii* and *V. orbiculata* are morphologically recognizable entities. All the diploids of "uncertain" identity initially identified as *V. dalmatica / V. orbiculata* definitely belong to *V. dalmatica*. Although *V. dalmatica* has been traditionally confused with *V. austriaca* subsp. *jacquinii,* Padilla-García et al. (2018) have established their respective geographic ranges and have indicated morphological characters that aid their identification. In addition, the mostly diploid species included in this study (i.e. *V. dalmatica* and *V. orbiculata*) are genetically differentiated (Fig. 2) and occur in separated geographic areas (Fig. 1). The results of the present study indicated clear genetic divergence between the lower ploidy level (diploid) and the higher ploidy levels (allotetra- and hexaploid), but not among the allopolyploids (Fig. 2). Considering the allopolyploids, some individuals from the "uncertain tetraploid" cluster are morphologically undistinguishable from *V. orbiculata*, while others display an intermediate morphology (Table A.1).

Hybrids are mosaics of characters with a wide range of outcomes (Soltis et al., 2014) and represent a challenge to the major forms of species concepts (Ramsey and Ramsey, 2014). Hence, previous studies on *V.* subsection *Pentasepalae* (Padilla-García et al., 2018) followed several lines of evidence for species delimitation, i.e., general lineage species concept (De Queiroz, 2007). The last taxonomic revision considered phylogenetic analysis based on DNA sequence data, ploidy level, morphology, and geographic range (Rojas-Andrés and Martínez-Ortega, 2016) following an integrative taxonomic approach (Dayrat, 2005). Considering the reticulate evolution that has led to the genetic similarities between the allopolyploids and the morphological resemblance of the tetraploids with *V. austriaca* subsp. *jacquinii* and especially with *V. orbiculata*, it is difficult to define clear taxonomic boundaries. Additionally, according to our own field observations, the "uncertain tetraploid" does not seem to have ecological preferences nor a specific geographic range. Being aware of the drawbacks of including morphologically highly similar cytotypes in one species (Soltis et al., 2007), assigning a taxonomic status to the "uncertain tetraploids" may raise more practical problems than solutions in this particular case. Thus, providing a taxonomic assignment is here discarded; instead maintaining the "uncertain tetraploids" −probably derived from hybridization between *V. orbiculata* and *V. dalmatica*− within the variation of *V. austriaca* subsp. *jacquinii*, seems to be more reasonable, unless further evidences and characters that allow their identification are obtained. Nevertheless, to avoid obscuring insights into evolution and speciation, it is crucial to consider the existence of an additional biological entity within the studied species complex: an allotetraploid morphologically cryptic lineage.

4.4. Evolutionary history in a Mediterranean biodiversity hotspot

The Balkan Peninsula is known by its floristic richness and represents one of the most important hotspots of biodiversity in Europe (Griffiths et al., 2004). Despite gaps of information in comparison to the western Mediterranean Basin, a high frequency of polyploids and hybrids has been reported in this area (Marques et al., 2018), linked with its complex geological and historical climatic context with successive changes in land connections and oscillation in climate regimes. In comparison with other European regions, the highest concentration of mixed-ploidy populations for *V.* subsect. *Pentasepalae* is found in the Balkan Peninsula (Rojas-Andrés et al., 2020) and especially on its western part (Fig. 1; Table A.5).

According to paleodistribution models several distinct areas could have acted as refugia for the survival of *V. orbiculata* and *V. dalmatica* during LGM (Fig. 5). The most suitable areas for both species are coincident with previously described refugia (see Nieto-Feliner, 2014).

In the case of *V. orbiculata* the northern coastal region of the Western Balkans could have acted at that time as refugium. Several studies have found evidence of a microrefugium in the Kvarner bay and the Istrian Peninsula (Magri et al., 2006; Resetnik et al., 2016a).

For *V. dalmatica* inland mountain ranges of the central Balkan Peninsula (Southern Carpathians) appear as suitable areas during the LGM (Fig. 5). However, the presence of *V. dalmatica* in this place is unlikely due to the distance to the Dalmatian region and the lack of adequate intermediate places to act as corridors. The most plausible refugial areas for *V. dalmatica* are: southern parts of the West Balkans (in comparison to *V. orbiculata*); and/or southern regions of Italy (Fig. 5). The access to these latter regions could have been facilitated by the available corridors between the Balkan and Apennine Peninsulas. During the LGM the Adriatic Sea reached its lowest level, which resulted in a narrow (15–20 km), shallow basin called Meso-Adriatic Depression (Correggiari et al., 1996). These sea-level conditions extend southwards to the Otranto Strait in the southern Adriatic, which was also considerably narrower than nowadays, connecting the eastern and western Adriatic coasts along a large area, thus allowing migration of *V. dalmatica* to more suitable locations. There is one record of the presence of a plant identified as "*V. austriaca* subsp. *jacquinii*" in the region near Matera in the Basilicata region of Italy (Rojas-Andrés and Martínez-Ortega, 2016). Unfortunately, we have failed at finding these plants in the field despite of all the efforts; as we suspected that these plants are actually *V. dalmatica*, this record would support our hypothesis of migration of this species during the glaciations. If this is confirmed, it will represent a further example of amphi-Adriatic distribution in the subsection, as is the case of *V. kindlii* (Padilla-García et al., 2018).

Veronica dalmatica could have also taken shelter in the southernmost part of the Dinaric Alps, an area of special topographic complexity within the Western Balkans that acted as a Quaternary microrefugium for plant (e.g. Kutnjak et al., 2014) and animal species (e.g. Krystufek et al., 2007). If this was the case, the survival of *V. orbiculata* and *V. dalmatica* in two separate refugia during the LGM would represent another example supporting the "refugia-within-refugia hypothesis" in the Balkan region (Podnar et al., 2004; Surina et al., 2011).

The divergence of *V. orbiculata* and *V. dalmatica* from a common ancestor could be the result of different ecological conditions in northern and southern refugia at the time of the LGM (a split that geographically coincides with the lower Neretva river valley in the Western Balkans) (Lakušić et al., 2013). Posterior differentiation is supported by the IBD pattern observed (Table 3). Range expansion in postglacial periods could explain the presence of secondary contacts in the southern Dinaric Alps between *V. orbiculata* and *V. dalmatica,* which would explain the first event of allopolyploidization within the complex (Fig. 4). The second hybridization event could have occurred in recent times according to the low number of generations (Table 5) −as suggested by previous studies (Rojas-Andrés et al. 2015; Padilla-García et al., 2018)− in places where the three entities appear intermixed.

5. Conclusions

Polyploidy and hybridization are key processes in plant evolution and in the last decades they are lively topics of research. This study provides conclusive evidence on the formation of the *V. austriaca* - *V. orbiculata* complex as a result from both auto- and allopolyploid processes. The obtained results indicate that there are two well-established entities at the diploid level with a clear genetic differentiation between diploid and hexaploid cytotypes, and an allotetraploid cryptic lineage involved in the formation of the hexaploid cytotype. However, it is complicated to establish boundaries among the tetraploid individuals initially determined as "uncertain" and the rest of the entities involved. Morphological resemblance with its relatives and genetic similarities with its hexaploid descendant blurs species delimitation. However, the recognition of the allopolyploid cluster as an intermediate biological entity, allows its consideration on future evolutionary and biodiversity studies.

The complex geological and historical climatic context of the Balkan Peninsula promotes opportunities for species diversification and hybridization. Different refugia are hypothesized for the survival of *V. orbiculata* and *V. dalmatica* during the LGM. Different ecological conditions in northern and southern parts of the Western Balkans and posterior differentiation by IBD might be responsible for the divergence of these two diploid species. Allopolyploidization events occurred afterwards as a result of *V. orbiculata* and *V. dalmatica* range expansion during postglacial periods.

CRediT authorship contribution statement

Noemí López-González: Methodology, Software, Visualization, Formal analysis, Investigation, Writing - original draft. **Javier Bobo-Pinilla:** Formal analysis, Visualization, Writing - review & editing. **Nélida Padilla-García:** Writing - review & editing. **João Loureiro:** Supervision, Resources, Writing - review & editing. **Silvia Castro:** Supervision, Resources, Writing - review & editing. **Blanca M. Rojas-Andrés:** Methodology, Supervision, Conceptualization, Writing - review & editing. **M. Montserrat Martínez-Ortega:** Methodology, Resources, Funding acquisition, Project administration, Supervision, Conceptualization, Writing - review & editing.

Acknowledgements

We want to thank the colleagues who helped us in the fieldwork campaigns: S. Andrés-Sánchez, M. Santos-Vicente and X. Giráldez. We thank B. Frajman for providing material. Thanks to Zuzana Münzbergová, Pamela Soltis, and Douglas Soltis for their help during the first steps of this work. We are especially grateful to Teresa Malvar for her priceless help in the lab. This work was supported by the Spanish *Ministerio de Economía y Competitividad* (projects CGL2009-07555 and CGL2012-32574); and the Spanish *Ministerio de Educación* (Ph.D. grant to NLG). The Portuguese Foundation for Science and Technology with POPH/FSE funds financed the work of SC (starting grant IF/01267/2013). This research has made use of the computational resources available from *Centro de Supercomputación de Castilla y León* (SCAYLE http://www.scayle.es) financed by the European Regional Development Fund (ERDF).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2020.107006.

References

Adams, K.L., Wendel, J.F., 2005. Polyploidy and genome evolution in plants. Curr. Opin. Plant Biol. 8, 135–141. https://doi.org/10.1016/j.pbi.2005.01.001.

Afonso, A., Loureiro, J., Arroyo, J., Olmedo-Vicente, E., Castro, S., 2020. Cytogenetic diversity in the polyploid complex *Linum suffruticosum* sl (Linaceae). Bot. J. Linn. Soc. boaa060 https://doi.org/10.1093/botlinnean/boaa060.

Ainouche, M.L., Baumel, A., Salmon, A., Yannic, G., 2004. Hybridization, polyploidy and speciation in *Spartina* (Poaceae). New Phytol. 161, 165–172. https://doi.org/10.1046/j.1469-8137.2003.00926.x.

Albach, D.C., 2006. Evolution of *Veronica* (Plantaginaceae) on the Balkan Peninsula. Phytologia Balcanica 12, 231–244.

Albach, D.C., Martínez-Ortega, M.M., Delgado, L., Weiss-Schneeweiss, H., Özgökce, F., Fischer, M.A., 2008. Chromosome numbers in *Veroniceae* (Plantaginaceae): review and several new counts. Ann. Mo. Bot. Gard. 95, 543–567. https://doi.org/10.3417/2006094.

Alexander-Webber, D., Abbott, R.J., Chapman, M.A., 2016. Morphological convergence between an allopolyploid and one of its parental species correlates with biased gene

expression and DNA loss. J. Hered. 107, 445–454. https://doi.org/10.1093/jhered/esw035.
Archer, F.I., Adams, P.E., Schneiders, B.B., 2017. Stratag: An R package for manipulating, summarizing and analysing population genetic data. Mol. Ecol. Resour. 17, 5–11. https://doi.org/10.1111/1755-0998.12559.
Arisumi, T., 1982. Endosperm balance numbers among New Guinea-Indonesian *Impatiens* species. J. Hered. 73, 240–242. https://doi.org/10.1093/oxfordjournals.jhered.a109629.
Arrigo, N., Felber, F., Parisod, C., Buerki, S., Alvarez, N., David, J., Guadagnuolo, R., 2010. Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. New Phytol. 187, 1170–1180. https://doi.org/10.1111/j.1469-8137.2010.03328.x.
Baniaga, A.E., Marx, H.E., Arrigo, N., Barker, M.S., 2020. Polyploid plants have faster rates of multivariate niche differentiation than their diploid relatives. Ecol. Lett. 23, 68–78. https://doi.org/10.1111/ele.13402.
Barker, M.S., Arrigo, N., Baniaga, A.E., Li, Z., Levin, D.A., 2016. On the relative abundance of autopolyploids and allopolyploids. New Phytol. 210, 391–398. https://doi.org/10.1111/nph.13698.
Bräuchler, C., Cikovac, P., 2007. *Iris orjenii* (Iridaceae)—a new species from the littoral Dinaric Alps. Willdenowia 37, 221–228. https://doi.org/10.3372/wi.37.37112.
Brysting, A.K., Oxelman, B., Huber, K.T., Moulton, V., Brochmann, C., 2007. Untangling complex histories of genome mergings in high polyploids. Syst. Biol. 56, 467–476. https://doi.org/10.1080/10635150701424553.
Buggs, R.J., Soltis, P.S., Soltis, D.E., 2009. Does hybridization between divergent progenitors drive whole-genome duplication? Mol. Ecol. 18, 3334–3339. https://doi.org/10.1111/j.1365-294X.2009.04285.x.
Burton, T.L., Husband, B.C., 2000. Fitness differences among diploids, tetraploids, and their triploid progeny in *Chamerion angustifolium:* mechanisms of inviability and implications for polyploidy evolution. Evolution 54, 1182–1191. https://doi.org/10.1111/j.0014-3820.2000.tb00553.x.
Castro, M., Loureiro, J., Serrano, M., Tavares, D., Husband, B.C., Siopa, C., Castro, S., 2019. Mosaic distribution of cytotypes in a mixed-ploidy plant species, *Jasione montana*: nested environmental niches but low geographical overlap. Bot. J. Linn. Soc. 190, 51–66. https://doi.org/10.1093/botlinnean/boz007.
Cornuet, J.M., Santos, F., Beaumont, M.A., Robert, C.P., Marin, J.M., Balding, D.J., Guillemaud, T., Estoup, A., 2008. Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. Bioinformatics 24, 2713–2719. https://doi.org/10.1093/bioinformatics/btn514.
Cornuet, J.M., Ravigné, V., Estoup, A., 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC *Bioinformatics* 11, 401. https://doi.org/10.1186/1471-2105-11-401.
Cornuet, J.M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J. M., Estoup, A., 2014. DIYABC v2.0: a software to make Approximate Bayesian Computation inferences about population history using Single Nucleotide Polymorphism. DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189. https://doi.org/10.1093/bioinformatics/btt763.
Correggiari, A., Roveri, M., Trincardi, F., 1996. Late Pleistocene and Holocene evolution of the North Adriatic Sea. Il Quaternario: Italian Journal of Quaternary Sciences 9, 697–704.
Crowl, A.A., Myers, C., Cellinese, N., 2017. Embracing discordance: Phylogenomic analyses provide evidence for allopolyploidy leading to cryptic diversity in a Mediterranean *Campanula* (Campanulaceae) clade. Evolution 71, 913–922. https://doi.org/10.1111/evo.13203.
Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85, 407–415. https://doi.org/10.1111/j.1095-8312.2005.00503.x.
De Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879–886. https://doi.org/10.1080/10635150701701083.
Delgado L., Rojas-Andrés B.M., López-González N., Padilla-García N., Martínez-Ortega M.M., 2018. Plantaginaceae. In: Marhold, K (ed.), IAPT/IOPB chromosome data 28. Taxon, 67: 1235–1236. https://doi.org/10.12705/676.39.
Diallo, A.M., Nielsen, L.R., Kjær, E.D., Petersen, K.K., Ræbild, A., 2016. Polyploidy can confer superiority to West African *Acacia senegal* (L.) Willd. trees. *Frontiers*. Plant Sci. 7, 821. https://doi.org/10.3389/fpls.2016.00821.
Doyle, J.J., Doyle, J.L., 1987. CTAB DNA extraction in plants. Phytochemical Bulletin 19, 11–15.
Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7.
Elith, J., 2002. Quantitative methods for modeling species habitat: Comparative performance and an application to Australian plants. In: Ferson, S., Burgham, M. (Eds.), Quantitative Methods for Conservation Biology. Springer, New York, pp. 39–58. https://doi.org/10.1007/b97704.
Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.
Fagundes, N.J.R., Ray, N., Beaumont, M., Neuenschwander, S., Salzano, F.M., Bonatto, S. L., Excoffier, L., 2007. Statistical evaluation of alternative models of human evolution. PNAS 104, 17614–17619. https://doi.org/10.1073/pnas.0708280104.
Fedorov, A.A., Kovanda, M., 1976. Plantaginaceae to Compositae (and Rubiaceae). In: Tutin, T.G., Heywood, V.H., Burges, N.A. (Eds.), Flora Europaea 4. Cambridge University Press, Cambridge, pp. 74–93.
Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package) version 3.698. Distributed by the author. Department of Genome Sciences. University of Washington, Seattle (WA).
Fourcade, Y., Engler, J.O., Rödder, D., Secondi, J., 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. PLoS ONE 9, e97122. https://doi.org/10.1371/journal.pone.0097122.
Fox, J., Weisberg, S., 2011. An R Companion to Applied Regression, second ed. Sage Publications, Thousand Oaks (CA). http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.
Frajman, B., Rešetnik, I., Niketić, M., Ehrendorfer, F., Schönswetter, P., 2016. Patterns of rapid diversification in heteroploid *Knautia* sect. Trichera (Caprifoliaceae, Dipsacoideae), one of the most intricate taxa of the European flora. BMC Evol. Biol. 16, 204. https://doi.org/10.1186/s12862-016-0773-2.
Friar, E.A., Prince, L.M., Cruse-Sanders, J.M., McGlaughlin, M.E., Butterworth, C.A., Baldwin, B.G., 2008. Hybrid origin and genomic mosaicism of *Dubautia scabra* (Hawaiian silversword alliance; Asteraceae, Madiinae). Syst. Bot. 33, 589–597. https://www.nature.com/articles/s41467-019-09454-w.
Galbraith, D.W., Harkins, K.R., Maddox, J.M., Ayres, N.M., Sharma, D.P., Firoozabady, E., 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220, 1049–1051. https://doi.org/10.1126/science.220.4601.1049.
Gent, P., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z.L., Zhang, M., 2011. The community climate system model version 4. J. Clim. 24, 4973–4991. https://doi.org/10.1175/2011JCLI4083.1.
Gilbert, K.J., Andrew, R.L., Bock, D.G., Franklin, M.T., Kane, N.C., Moore, J.S., Moyers, B.T., Renaut, S., Rennison, D.J., Veen, T., Vines, T.H., 2012. Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program structure. Mol. Ecol. 21, 4925–4930. https://doi.org/10.1111/j.1365-294X.2012.05754.x.
Goudet, J., 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Resour. 5, 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x.
Greiner, R., Oberprieler, C., 2012. The role of inter-ploidy block for reproductive isolation of the diploid *Leucanthemum pluriflorum* Pau (Compositae, Anthemideae) and its tetra-and hexaploid relatives. Flora-Morphology, Distribution, Functional Ecology of Plants 207, 629–635. https://doi.org/10.1016/j.flora.2012.07.001.
Griffiths HI, Kryštufek B, Reed JM. 2004. Balkan biodiversity: pattern and process in the European hotspot, first ed. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-2854-0.
Hasumi, H., Emori, S., 2004. K-1 Coupled GCM (MIROC) Description. Center for Climate System Research, University of Tokyo, Japan, Tokyo.
Hardy, O.J., Vekemans, X., 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Resour. 2, 618–620. https://doi.org/10.1046/j.1471-8286.2002.00305.x.
Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A k-means clustering algorithm. Journal of the Royal Statistical Society Series C (Applied Statistics) 28, 100–108. https://doi.org/10.2307/2346830.
Hao, G.Y., Lucero, M.E., Sanderson, S.C., Zacharias, E.H., Holbrook, N.M., 2013. Polyploidy enhances the occupation of heterogeneous environments through hydraulic related trade-offs in *Atriplex canescens* (Chenopodiaceae). New Phytol. 197, 970–978. https://doi.org/10.1111/nph.12051.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. https://doi.org/10.1002/joc.1276.
Hijmans R.J., van Etten J., 2014. raster: Geographic data analysis and modeling. R package version, 2. https://CRAN.R-project. org/package=raster.
Hu, Y.N., Zhao, L., Buggs, R.J., Zhang, X.M., Li, J., Wang, N., 2019. Population structure of *Betula albosinensis* and *Betula platyphylla*: evidence for hybridization and a cryptic lineage. Ann. Bot. 123, 1179–1189. https://doi.org/10.1093/aob/mcz024.
Husband, B.C., Schemske, D.W., Burton, T.L., Goodwillie, C., 2002. Pollen competition as a unilateral reproductive barrier between sympatric diploid and tetraploid *Chamerion angustifolium*. Proc. R. Soc. Lond. B Biol. Sci. 269, 2565–2571. https://doi.org/10.1098/rspb.2002.2196.
Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233.
Janes, J.K., Miller, J.M., Dupuis, J.R., Malenfant, R.M., Gorrell, J.C., Cullingham, C.I., Andrew, R.L., 2017. The K = 2 conundrum. Mol. Ecol. 26, 3594–3602. https://doi.org/10.1111/mec.14187.
Johnston, S.A., Hanneman Jr., R.E., 1982. Manipulations of endosperm balance number overcome crossing barriers between diploid *Solanum* species. Science 217, 446–448.
Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94. https://doi.org/10.1186/1471-2156-11-94.
Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129.
Jørgensen, M.H., Ehrich, D., Schmickl, R., Koch, M.A., Brysting, A.K., 2011. Interspecific and interploidal gene flow in Central European *Arabidopsis* (Brassicaceae). BMC Evol. Biol. 11, 346. https://doi.org/10.1186/1471-2148-11-346.
Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., 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.
Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120. https://doi.org/10.1007/BF01731581.
Kolář, F., Čertner, M., Suda, J., Schönswetter, P., Husband, B.C., 2017. Mixed-ploidy species: progress and opportunities in polyploid research. Trends Plant Sci. 22, 1041–1055. https://doi.org/10.1016/j.tplants.2017.09.011.

Krystufek, B., Buzan, E.V., Hutchinson, W.F., Hänfling, B., 2007. Phylogeography of the rare Balkan endemic Martino's vole, Dinaromys bogdanovi, reveals strong differentiation within the western Balkan Peninsula. Mol. Ecol. 16, 1221–1232. https://doi.org/10.1111/j.1365-294X.2007.03235.x.

Kučera, J., Marhold, K., Lihová, J., 2010. *Cardamine maritima* group (Brassicaceae) in the amphi-Adriatic area: a hotspot of species diversity revealed by DNA sequences and morphological variation. Taxon 59, 148–164. https://doi.org/10.2307/27757059.

Kutnjak, D., Kuttner, M., Niketić, M., Dullinger, S., Schönswetter, P., Frajman, B., 2014. Escaping to the summits: phylogeography and predicted range dynamics of *Cerastium dinaricum*, an endangered high mountain plant endemic to the western Balkan Peninsula. Mol. Phylogenet. Evol. 78, 365–374. https://doi.org/10.1016/j.ympev.2014.05.015.

Lafon-Placette, C., Köhler, C., 2016. Endosperm-based postzygotic hybridization barriers: developmental mechanisms and evolutionary drivers. Mol. Ecol. 25, 2620–2629. https://doi.org/10.1111/mec.13552.

Lafon-Placette, C., Johannessen, I.M., Hornslien, K.S., Ali, M.F., Bjerkan, K.N., Bramsiepe, J., Glöckle, B.M., Rebernig, C.A., Brysting, A.K., Grini, P.E., Köhler, C., 2017. Endosperm-based hybridization barriers explain the pattern of gene flow between Arabidopsis lyrata and Arabidopsis arenosa in Central Europe. Proc. Natl. Acad. Sci. 114, E1027–E1035. https://doi.org/10.1073/pnas.1615123114.

Lakušić, D., Liber, Z., Nikolić, T., Surina, B., Kovačić, S., Bogdanović, S., Stefanović, S., 2013. Molecular phylogeny of the Campanula pyramidalis species complex (Campanulaceae) inferred from chloroplast and nuclear non-coding sequences and its taxonomic implications. Taxon 62, 505–524. https://doi.org/10.12705/623.1.

Levin D.A., 2002. The Role of Chromosomal Change in Plant Evolution, first ed. Oxford University Press, Oxford, UK. https://doi.org/10.1086/425787.

Liepelt, S., Cheddadi, R., de Beaulieu, J.L., Fady, B., Gömöry, D., Hussendörfer, E., Konnert, M., Litt, T., Longauer, R., Terhürne-Berson, R., Ziegenhagen, B., 2009. Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.)—a synthesis from palaeobotanic and genetic data. Rev. Palaeobot. Palynol. 153, 139–149. https://doi.org/10.1016/j.revpalbo.2008.07.007.

Liu, K., Warnow, T.J., Holder, M.T., Nelesen, S.M., Yu, J., Stamatakis, A.P., Linder, C.R., 2011. SATe-II: very fast and accurate simultaneous estimation of multiple sequence alignments and phylogenetic trees. Systematic Biolology 61, 90–106. https://doi.org/10.1093/sysbio/syr095.

López-González, N., Mayland-Quellhorst, E., Pinto-Carrasco, D., Martínez-Ortega, M.M., 2015. Characterization of 12 polymorphic SSR markers in *Veronica* subsect. *Pentasepalae* (Plantaginaceae) and cross-amplification in 10 other subgenera. Appl. Plant Sci. 3 (10) https://doi.org/10.3732/apps.1500059.

López-Jurado, J., Mateos-Naranjo, E., Balao, F., 2019. Niche divergence and limits to expansion in the high polyploid *Dianthus broteri* complex. New Phytol. 222, 1076–1087. https://doi.org/10.1111/nph.15663.

Loureiro, J., Rodriguez, E., Doležel, J., Santos, C., 2007. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. Ann. Bot. 100, 875–888. https://doi.org/10.1093/aob/mcm152.

Lowry, E., Lester, S.E., 2006. The biogeography of plant reproduction: potential determinants of species' range sizes. J. Biogeogr. 33, 1975–1982. https://doi.org/10.1111/j.1365-2699.2006.01562.x.

Madlung, A., 2013. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. Heredity 110, 99–104. https://doi.org/10.1038/hdy.2012.79.

Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latalowa, M., Litt, T., Paule, L., Roure, J.M., Tantau, I., van der Knapp, W.O., Petit, R.J., de Beaulieu, J.L., 2006. A new scenario for the Quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. New Phytol. 171, 199–221. https://doi.org/10.1111/j.1469-8137.2006.01740.x.

Majure, L.C., Judd, W.S., Soltis, P.S., Soltis, D.E., 2012. Cytogeography of the Humifusa clade of *Opuntia* ss Mill. 1754 (Cactaceae, Opuntioideae, Opuntieae): correlations with pleistocene refugia and morphological traits in a polyploid complex. *Comparative*. Cytogenetics 6, 53. https://doi.org/10.3897/CompCytogen.v6i1.2523.

Marquardt, D.W., 1970. Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. Technometrics 12, 591–612. https://doi.org/10.2307/1267205.

Marques, I., Loureiro, J., Draper, D., Castro, M., Castro, S., 2018. How much do we know about the frequency of hybridisation and polyploidy in the Mediterranean region? Plant Biology 20, 21–37. https://doi.org/10.1111/plb.12639.

Martínez-Ortega, M.M., Sánchez Sánchez, J., Rico, E., 2000. Palynological study of *Veronica* Sect. *Veronica* and Sect. *Veronicastrum* (Scrophulariaceae) and its taxonomic significance. Grana 39, 21–31. https://doi.org/10.1080/00173130150503777.

Martínez-Ortega, M.M., Delgado, L., Albach, D.C., Rico, E., Elena-Rossello, J.A., 2004. Species boundaries and phylogeographic patterns in cryptic taxa inferred from AFLP markers: *Veronica* subgen. *Pentasepalae* (Scrophulariaceae) in the Western Mediterranean. Syst. Bot. 29, 965–986. https://doi.org/10.1600/0363644042451071.

Médail, F., Diadema, K., 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. J. Biogeogr. 36, 1333–1345. https://doi.org/10.1111/j.1365-2699.2008.02051.x.

Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M., Anderson, R.P., 2014. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol. Evol. 5, 1198–1205. https://doi.org/10.1111/2041-210X.12261.

Feliner, G.N., 2014. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. Perspectives in Plant Ecology, Evolution and Systematics 16, 265–278. https://doi.org/10.1016/j.ppees.2014.07.002.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M.A.S.S., 2007. The vegan package. Community ecology package 10, 631–637.

Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Annu. Rev. Genet. 34, 401–437. https://doi.org/10.1146/annurev.genet.34.1.401.

Padilla-García, N., Rojas-Andrés, B.M., López-González, N., Castro, M., Castro, S., Loureiro, J., Albach, D.C., Machon, N., Martínez-Ortega, M.M., 2018. The challenge of species delimitation in the diploid-polyploid complex *Veronica* subsection *Pentasepalae*. Mol. Phylogenet. Evol. 119, 196–209. https://doi.org/10.1016/j.ympev.2017.11.007.

Parrott, W.A., Smith, R.R., 1986. Evidence for the existence of endosperm balance number in the true clovers (*Trifolium* spp.). Can. J. Genet. Cytol. 28, 581–586. https://doi.org/10.1139/g86-085.

Pearson, R.G., Raxworthy, C.J., Nakamura, M., Townsend, Peterson A., 2007. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. J. Biogeogr. 34, 102–117. https://doi.org/10.1111/j.1365-2699.2006.01594.x.

Peev, D.R., 1972. New taxa and ploidy levels of some Bulgarian *Veronica* species. Proceedings of the Bulgarian Academy of Sciences 25, 811–814.

Peng, Y.Y., Wei, Y.M., Baum, B.R., Jiang, Q.T., Lan, X.J., Dai, S.F., Zheng, Y.L., 2010. Phylogenetic investigation of *Avena* diploid species and the maternal genome donor of *Avena* polyploids. Taxon 59, 1472–1482. https://doi.org/10.1002/tax.595012.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Podnar, M., Mayer, W., Tvrtković, N., 2004. Mitochondrial phylogeography of the Dalmatian wall lizard, *Podarcis melisellensis* (Lacertidae). Org. Divers. Evol. 4, 307–317. https://doi.org/10.1016/j.ode.2004.04.004.

Pope, L.C., Domingo-Roura, X., Erven, K., Burke, T., 2006. Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. Mol. Ecol. 15, 371–386. https://doi.org/10.1111/j.1365-294X.2005.02815.x.

Popp, M., Oxelman, B., 2007. Origin and evolution of north american polyploid *Silene* (Caryophyllaceae). Am. J. Bot. 94, 330–349. https://doi.org/10.3732/ajb.94.3.330.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Pritchard, J.K., Wen, W., Falush, D., 2010. Documentation for structure software: version 2.3. University of Chicago, Chicago.

Psonis, N., Antoniou, A., Karameta, E., Leaché, A.D., Kotsakiozi, P., Darriba, D., Kozlov, A., Stamatakis, A., Poursanidis, D., Kukushkin, O., Jablonski, D., Crnobrnja-Isailović, J., Gherghel, I., Lymberakis, P., Poulakakis, N., 2018. Resolving complex phylogeographic patterns in the Balkan Peninsula using closely related wall-lizard species as a model system. Mol. Phylogenet. Evol. 125, 100–115. https://doi.org/10.1016/j.ympev.2018.03.021.

Ramsey, J., Schemske, D.W., 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annu. Rev. Ecol. Syst. 29, 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467.

Ramsey, J., Ramsey, T.S., 2014. Ecological studies of polyploidy in the 100 years following its discovery. Philosophical Transactions of the Royal Society B: Biological Sciences 369, 20130352. https://doi.org/10.1098/rstb.2013.0352.

Rešetnik, I., Frajman, B., Schönswetter, P., 2016a. Heteroploid *Knautia drymeia* includes *K. gussonei* and cannot be separated into diagnosable subspecies. Am. J. Bot. 103, 1300–1313. https://doi.org/10.3732/ajb.1500506.

Rešetnik, I., Baričevič, D., Rusu, D.B., Carović-Stanko, K., Chatzopoulou, P., Dajić-Stevanović, Z., Goncearіuc, M., Grdiša, M., Greguraš, D., Ibraliu, A., Krasniqi, E., Liber, Z., Murtić, S., Pećanac, D., Radosavljević, I., Stefkov, G., Stešević, D., Šoštarić, I., Šatović, Z., Jug-Dujaković, M., 2016b. Genetic diversity and demographic history of wild and cultivated/naturalised plant populations: evidence from Dalmatian Sage (*Salvia officinalis* L., Lamiaceae). PLoS ONE 11, e0159545. https://doi.org/10.1371/journal.pone.0159545.

Rojas-Andrés, B.M., Albach, D.C., Martínez-Ortega, M.M., 2015. Exploring the intricate evolutionary history of the diploid–polyploid complex *Veronica* subsection *Pentasepalae* (Plantaginaceae). Bot. J. Linn. Soc. 179, 670–692. https://doi.org/10.1111/boj.12345.

Rojas-Andrés, B.M., Martínez-Ortega, M.M., 2016. Taxonomic revision of *Veronica* subsection *Pentasepalae (Veronic*a, Plantaginaceae sensu APG III). Phytotaxa 285, 1–100. https://doi.org/10.11646/phytotaxa.285.1.1.

Rojas-Andrés, B.M., Padilla-García, N., de Pedro, M., López-González, N., Delgado, L., Albach, D.C., Castro, M., Castro, S., Loureiro, J., Martínez-Ortega, M.M., 2020. Environmental differences are correlated with the distribution pattern of cytotypes in *Veronica* subsection *Pentasepalae* at a broad scale. Ann. Bot. 125, 471–484. https://doi.org/10.1093/aob/mcz182.

Ronfort, J., Jenczewski, E., Bataillon, T., Rousset, F., 1998. Analysis of population structure in autotetraploid species. Genetics 150, 921–930.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., Feldman, M.W., 2002. Genetic structure of human populations. Science 298, 2381–2385. https://doi.org/10.1126/science.1078311.

Sánchez Agudo, J.A., Rico, E., Sánchez Sánchez, J., Martínez-Ortega, M.M., 2009. Pollen morphology in the genus *Veronica* L. (Plantaginaceae) and its systematic significance. Grana 48, 239–257. https://doi.org/10.1080/00173130903364723.

Schinkel, C.C., Kirchheimer, B., Dullinger, S., Geelen, D., De Storme, N., Hörandl, E., 2017. Pathways to polyploidy: indications of a female triploid bridge in the alpine species *Ranunculus kuepferi* (Ranunculaceae). Plant Syst. Evol. 303, 1093–1108. https://doi.org/10.1007/s00606-017-1435-6.

Schuelke, M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 18, 233–234. https://doi.org/10.1038/72708.

Servick, S., Visger, C.J., Gitzendanner, M.A., Soltis, P.S., Soltis, D.E., 2015. Population genetic variation, geographic structure, and multiple origins of autopolyploidy in *Galax urceolata*. Am. J. Bot. 102, 973–982. https://doi.org/10.3732/ajb.1400554.

Slotte, T., Huang, H., Lascoux, M., Ceplitis, A., 2008. Polyploid speciation did not confer instant reproductive isolation in *Capsella* (Brassicaceae). Mol. Biol. Evol. 25, 1472–1481. https://doi.org/10.1093/molbev/msn092.

Soltis, D.E., Soltis, P.S., 1999. Polyploidy: recurrent formation and genome evolution. Trends Ecol. Evol. 14, 348–352. https://doi.org/10.1016/S0169-5347(99)01638-9.

Soltis, D.E., Soltis, P.S., Schemske, D.W., Hancock, J.F., Thompson, J.N., Husband, B.C., Judd, W.S., 2007. Autopolyploidy in angiosperms: have we grossly underestimated the number of species? Taxon 56, 13–30. https://doi.org/10.2307/25065732.

Soltis, P.S., Liu, X., Marchant, D.B., Visger, C.J., Soltis, D.E., 2014. Polyploidy and novelty: Gottlieb's legacy. Philosophical Transactions of the Royal Society B: Biological Sciences 369, 20130351. https://doi.org/10.1098/rstb.2013.0351.

Soltis, D.E., Visger, C.J., Marchant, D.B., Soltis, P.S., 2016. Polyploidy: Pitfalls and paths to a paradigm. Am. J. Bot. 103, 1146–1166. https://doi.org/10.3732/ajb.1500501.

Sonnleitner, M., Weis, B., Flatscher, R., García, P.E., Suda, J., Krejčíková, J., Schneeweiss, G.M., Winkler, M., Schönswetter, P., Hülber, K., 2013. Parental ploidy strongly affects offspring fitness in heteroploid crosses among three cytotypes of autopolyploid *Jacobaea carniolica* (Asteraceae). PLoS ONE 8, e78959. https://doi.org/10.1371/journal.pone.0078959.

Španiel, S., Marhold, K., Zozomová-Lihová, J., 2017. The polyploid *Alyssum montanum-A. repens* complex in the Balkans: a hotspot of species and genetic diversity. 2017. Plant Syst. Evol. 303, 1443–1465. https://doi.org/10.1007/s00606-017-1470-3.

Stebbins, G.L., 1971. Chromosomal evolution in higher plants. Addison Wesley, London, UK.

Steen, S.W., Gielly, L., Taberlet, P., Brochmann, C., 2000. Same parental species, but different taxa: molecular evidence for hybrid origins of the rare endemics *Saxifraga opdalensis* and *S. svalbardensis* (Saxifragaceae). Bot. J. Linn. Soc. 132, 153–164. https://doi.org/10.1111/j.1095-8339.2000.tb01211.x.

Surina, B., Schönswetter, P., Schneeweiss, G.M., 2011. Quaternary range dynamics of ecologically divergent species (*Edraianthus serpyllifolius* and *E. tenuifolius*, Campanulaceae) within the Balkan refugium. J. Biogeogr. 38, 1381–1393. https://doi.org/10.1111/j.1365-2699.2011.02493.x.

Sutherland, B.L., Galloway, L.F., 2017. Postzygotic isolation varies by ploidy level within a polyploid complex. New Phytol. 213, 404–412. https://doi.org/10.1111/nph.14116.

Theodoridis, S., Randin, C., Broennimann, O., Patsiou, T., Conti, E., 2013. Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia*. *Journal of Biogeography* 40, 1278–1289. https://doi.org/10.1111/jbi.12085.

Thiers B., Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. [Continuosly updated]. http://sweetdum.nybg.org/ih/.

Thompson J.D., 2005. Plant evolution in the Mediterranean, first ed. Oxford University Press, Oxford, UK. https://doi.org/10.1093/acprof:oso/9780198515340.001.0001.

Welles, S., Ellstrand, N.C., 2019. Evolution of increased vigor associated with allopolyploidization in the newly formed invasive species *Salsola ryanii*. AoB PLANTS plz039. https://doi.org/10.1093/aobpla/plz039.

Wickham, H., 2010. ggplot2: elegant graphics for data analysis. J. Stat. Softw. 35, 65–88. https://doi.org/10.1007/978-0-387-98141-3.

Williams, E.G., Rouse, J.L., 1990. Relationships of pollen size, pistil length and pollen tube growth rates in *Rhododendron* and their influence on hybridization. Sex. Plant Reprod. 3, 7–17.

Zhan, S.H., Drori, M., Goldberg, E.E., Otto, S.P., Mayrose, I., 2016. Phylogenetic evidence for cladogenetic polyploidization in land plants. Am. J. Bot. 103, 1252–1258. https://doi.org/10.3732/ajb.1600108.