1 **The challenge of species delimitation in the diploid-polyploid complex *Veronica* subsection**

2 ***Pentasepalae.***

3 Nélida Padilla-García[a,b,c,*], Blanca M. Rojas-Andrés[a,b,1], Noemí López-González[a,b], Mariana

4 Castro[d], Sílvia Castro[d], João Loureiro[d], Dirk C. Albach[e], Nathalie Machon[c], M. Montserrat

5 Martínez-Ortega[a,b]

6 [a] *Departamento de Botánica, University of Salamanca, E-37007, Salamanca, Spain*

7 [b] *Biobanco de ADN Vegetal, University of Salamanca, Edificio Multiusos I+D+i, Calle Espejo*

8 *s/n, 37007 Salamanca, Spain*

9 [c] *Centre d'Ecologie et des Sciences de la Conservation (CESCO, UMR7204), Sorbonne*

10 *University, MNHN, CNRS, UPMC, CP51, Paris, France*

11 [d] *CFE, Centre for Functional Ecology, Department of Life Sciences, University of Coimbra,*

12 *Calçada Martim de Freitas s/n, 3000-456 Coimbra, Portugal*

13 [e] *Department of Biology and Environmental Sciences (IBU), AG Biodiversity and Evolution of*

14 *Plants, Carl von Ossietzky Universität Oldenburg, Oldenburg 26111, Germany*

15

16 * Corresponding author at*:* Departamento de Botánica, University of Salamanca, E-37007,

17 Salamanca, Spain

18 *E-mail address:* nelidapg@gmail.com (N. Padilla-García)

19

20 **Abstract** A reliable taxonomic framework and the identification of evolutionary lineages are

21 essential for effective decisions in conservation biodiversity programs. However, phylogenetic

22 reconstruction becomes extremely difficult when polyploidy and hybridization are involved.

23 *Veronica* subsection *Pentasepalae* is a diploid-polyploid complex of ca. 20 species with ploidy

24 levels ranging from $2x$ to $10x$. Here, DNA-ploidy level estimations and AFLP fingerprinting

25 were used to determine the evolutionary history, and species boundaries were reviewed in an

26 integrated approach including also previous data (mainly morphology and sequence-based

[1] *Present address: Department of Molecular Evolution and Plant Systematics & Herbarium (LZ), Institute of Biology, Leipzig University, Leipzig, Germany*

phylogenetic reconstructions). Molecular analyses were performed for 243 individuals from 95

populations, including for the first time all taxa currently recognized within the subsection.

Individuals of uncertain identity and from mixed-ploidy populations were also analyzed.

Phylogenetic reconstruction identified four main groups corresponding almost completely to the

four clusters identified by genetic structure analyses. Multiple autopolyploidization events have

occurred in the tetraploid *V. satureiifolia* giving rise to octoploid entities in central Europe and

north of Spain, whereas hybridization between a newly described diploid species, *V. dalmatica*

sp. nov. and diploid *V. orbiculata* seems to have occurred in several populations from the

Balkan Peninsula. Furthermore, homoploid hybridization events are reported for the first time

for the subsection. Our study has established the taxonomic status of taxa, for the most part

recovered as monophyletic. Monophyly of *V. linearis* has been confirmed and a clear amphi-

Adriatic distribution of *V. kindlii* has been demonstrated. In addition, cryptic taxa within the

group have been identified. Our study allows the recognition of *V. crinita* and *V. thracica* at the

specific rank and a new species, *Veronica dalmatica*, is fully described. This study highlights

the implications of polyploidy in species delimitation, and illustrates the importance to conserve

polyploid populations as potential sources of diversification due to evolutionary significance of

genome duplications in plant evolution.

Keywords AFLPs, autopolyploidy, hybridization, *Pentasepalae*, phylogeny, *Veronica*

dalmatica

1. Introduction

        The delimitation of species boundaries is a classic problem for biologists. Until about

seventy years ago, taxonomists have focused on morphological differences between species

rather than on coherence with the evolutionary history of the species for their circumscription.

However, since the 1940's, a wider interest in the evolutionary history of organisms arose

(Huxley, 1940). In 1950, Hennig published his theory of phylogenetic systematics giving rise to the origin of cladistics, which revolutionized the field of taxonomy (Hennig, 1950). Despite originally considered for the analysis of morphological characters, it is equally suitable for other types of characters that have been used with taxonomic purposes during the last decades. Currently, molecular phylogenies, complementing morphological characters, are the key instruments for biologists and biosystematists who try to understand the evolutionary processes that shape the history of species. Nevertheless, evolutionary histories involving radiations or complex processes such as hybridization, introgression and/or polyploidization, complicate phylogenetic reconstruction (Naciri and Linder, 2015). This, together with a lack of morphological differences and uncertainties over reproductive isolation among polyploids and their diploid progenitors, has resulted in taxonomic biases within polyploid complexes (Soltis et al., 2007, Barley et al., 2013). Here, the importance of the species concept is fundamental. The biological concept of species proposed by Mayr (1942) is difficult to apply when working with closely related species in which hybridization and introgression are common. Most plant taxonomists have traditionally relied on morphology to delimit species boundaries (i.e., morphological species concept), whereas others adopted in the last decades a concept based on genetic differences and monophyly (i.e., genetic and phylogenetic species concepts). However, in species groups with frequent hybridization and polyploidization, the general lineage concept (de Queiroz 2005, 2007) may be more appropriate. According to this concept, species are defined as separately evolving metapopulation lineages, which can be identified by different properties that species accumulate during the process of speciation (e.g., reproductive isolation, morphological or genetic differences, monophyly, etc.). This general lineage concept has been broadly adopted and promoted the development of an integrative taxonomic approach in which multiple and complementary methods are integrated to delimit species boundaries (Dayrat, 2005). This approach argues that taxonomic inference should be based on congruence across different types of characters and analyses. When results from different sources of data are

1   incongruent, caution to delimit species is preferable since taxonomic conclusions may have

2   significant implications (Carstens et al., 2013). For example, regarding conservation and

3   sustainable use of natural resources, in accordance with the Convention on Biological Diversity

4   (https://www.cbd.int/gti/importance.shtml), taxonomy is necessary for effective decision

5   making, because it provides basic understanding about the components of biodiversity. In a

6   world where wild species are increasingly under threat, the conservation status of a taxon can

7   only be correctly evaluated under the light of a clear taxonomic framework (Mace, 2004). The

8   identification and preservation of evolutionary processes is also essential in conservation

9   programs, especially for endangered, rare and endemic species.

10   In the present study, complementary methodologies are used to address the taxonomic

11   challenges of a study group with a complex evolutionary history, *Veronica* subsection

12   *Pentasepalae* Benth. This subsection is a monophyletic lineage within *Veronica* subgenus

13   *Pentasepalae* (Benth.) M.M. Mart. Ort., Albach & M.A. Fisch. (Rojas-Andrés et al., 2015). It

14   has a recent origin (mean crown age 2.8 Mya., Meudt et al., 2015) and is composed of ca. 20

15   closely related perennial species distributed in Eurasia and North Africa. Interestingly, the group

16   comprises five species and three subspecies endemic to single countries or sometimes only a

17   small area within one country. Some of them are included in regional, national and/or

18   international Red Lists (Peñas de Giles et al., 2004; Cabezudo et al., 2005; Alcántara de la

19   Fuente et al., 2007; Petrova and Vladimirov, 2009; Bilz, 2011; International Union for

20   Conservation of Nature, 2016), although there is a clear lack of information for numerous

21   species that have not yet been carefully evaluated. The most important diversification center of

22   *V.* subsection *Pentasepalae* is the Balkan Peninsula. The group is characterized by the presence

23   of a pentapartite calyx with the fifth sepal being significantly smaller, by a capsule usually

24   rounded at the base, and a base chromosome number of $x = 8$. However, although the group is

25   well defined within *Veronica* (Albach et al., 2008), the existence of morphologically

26   intermediate forms within the subsection due to overlapping morphological character states

makes *V.* subsection *Pentasepalae* one of the taxonomically most complicated groups within the genus (Albach and Meudt, 2010). Since Bentham described *V.* subsection *Pentasepalae* in 1846, numerous taxonomic treatments have been proposed (for a historical review of monographs and Floras, see Rojas-Andrés and Martínez-Ortega, 2016), and several studies based on morphological, karyological or molecular data have tried to elucidate the evolutionary history of the group (e.g., Martínez-Ortega et al., 2000, 2004, 2009; Andrés-Sánchez et al., 2009). In the most recent molecular study, Rojas-Andrés et al. (2015) used nuclear and plastid DNA sequence data to perform a phylogenetic analysis of the subsection. Despite the contributions of that study to the understanding of the evolutionary history of the group, a high degree of incongruence was found between the ITS and plastid DNA trees, probably caused by hybridization and incomplete lineage sorting (ILS). Hence, some questions about the monophyly and the relationships among species remained unresolved. Such questions are unlikely to be answered using a few loci alone, especially considering the prevalence of hybridization and polyploidization in the genus (Albach and Chase, 2004).

The variety of ploidy levels in the subsection, ranging from diploid to decaploid (data previous to 2008 summarized by Albach et al., 2008; Rojas-Andrés et al., 2015), indicate that polyploidy has been a crucial process in the evolution of the group. Polyploidy or whole-genome duplication (WGD) is a frequent mechanism of evolution and speciation in flowering plants (Stebbins, 1950; Grant, 1971; Soltis et al., 2004, 2009, 2015; Mayrose et al., 2011; Kellog, 2016). Despite ongoing research regarding the distinction between the types of polyploids (Levin, 2002; Soltis et al., 2010; Husband et al., 2013; Doyle and Sherman-Broyles, 2017), two main categories are generally recognized based on their origin: i) autopolyploids that arise within a species, via intraspecific hybridization and duplication of similar genomes (homologous) and ii) allopolyploids formed by interspecific hybridization and chromosome doubling of genomes that are more or less divergent (homeologous). The prevalence of different types of polyploids in nature has been intensively discussed (Müntzing, 1936; Stebbins, 1947;

Lewis, 1980; Parisod et al., 2010), and recent studies suggest a parity in the incidence of autopolyploidy and allopolyploidy (Barker et al., 2016 but see Doyle and Sherman-Broyles, 2017). The differentiation between these processes is fundamental to evaluate the importance of polyploidization and hybridization in plant evolution. In this context, the diploid-polyploid complex *Veronica* subsection *Pentasepalae* is an excellent model to gain deeper insights into the contribution of these mechanisms to the evolutionary history of angiosperms.

The aim of this study is to clarify the phylogenetic relationships of *V.* subsection *Pentasepalae* by analyzing the nuclear genome using Amplified Fragment Length Polymorphism (AFLP). In addition to its use in phylogeographic studies, the AFLP technique is now widely used to infer phylogenetic relationships and to identify hybridization and polyploid events in recently evolved polyploid non-model groups (Meudt, 2011; Reberning et al., 2012; Himmelreich et al., 2014; Zozomová et al., 2014). Compared to the previous study by Rojas-Andrés et al., (2015), in addition to using AFLPs, a molecular tool for which markers are distributed throughout the genome, we expand the study to include for the first time all taxa currently recognized within the subsection. Also, we added individuals that are difficult to identify to species and of different ploidy level from mixed-ploidy populations. AFLPs were generated to address the following specific points: **i)** The role of auto- and allopolyploidization processes in the evolutionary history of *V.* subsection *Pentasepalae*; **ii)** Implications of these processes in species delimitation and classification; **iii)** A review of the taxonomic status of *V.* subsection *Pentasepalae* taxa.

2. Materials and methods

2.1. Plant material

Samples were collected in the field during 2009–2015 except for one population of *V. satureiifolia* and one population of *V. tenuifolia* subsp. *fontqueri* that were collected in 2002. Localities, initial taxonomic assignment, and further information about samples are given in

1 Table S1. Fresh leaf material was collected and stored in silica gel. For *V. krylovii*, three

2 individuals were included of which two were selected from herbarium material and one from a

3 cultivated specimen in the Botanical Garden of Oldenburg (Germany). *Veronica orientalis,*

4 which belongs to *V.* subsection *Orientales* (Wulff) Stroh of *V.* subgenus *Pentasepalae* (Benth.)

5 M.M. Mart. Ort., Albach & M.A. Fisch. was chosen as outgroup. The complete data-set

6 comprises 243 individuals from 95 populations (outgroup included) covering the geographic

7 distribution of each taxon (Fig. 1). From each location, 2–3 individuals were included, except

8 for populations with mixed-ploidy levels. In these exceptional cases, two individuals per

9 cytotype were analyzed whenever possible. Initial plant identification was based on the most

10 recent taxonomic treatment (Rojas-Andrés and Martínez-Ortega, 2016), with the exception of

11 some taxa that were not recognized by those authors, but whose names are used here to test their

12 status [i.e., *V. crinita* f. *bosniaca* and *V. thracica* were included as synonyms of *V. crinita*, and *V.*

13 *macrodonta* as a synonym of *V. austriaca* subsp. *austriaca* in that taxonomic treatment].

14 Material that was difficult to identify was initially catalogued using morphological affinity to

15 other species (e.g., *V.* affinis *linearis*). Additionally, *V. austriaca* subsp. *jacquinii* / *V. orbiculata*

16 indicates individuals of intermediate morphology between both species. Vouchers are deposited

17 in the herbaria ALTB, GDA, MGC, OLD, SALA, VANF and WU (herbarium acronyms follow

18 Thiers, 2017).

19 *2.2. DNA-ploidy level estimation using flow cytometry*

20 DNA-ploidy levels were estimated by flow cytometry using silica gel dried leaves.

21 Individuals from each sampled population were measured separately. Nuclear suspensions were

22 prepared following the method described by Galbraith et al. (1983) in which the leaf tissue of

23 each individual was chopped together with leaf tissue from an internal standard using a sharp

24 razor blade in a Petri dish containing a buffer solution, namely Woody Plant Buffer

25 (WPB; Loureiro et al., 2007). *Solanum pseudocapsicum* L. (2C = 2.589 pg; Temsch et al.,

26 2010), *Zea mays* L. 'CE-777' (2C = 5.43 pg; Lysak and Dolezel, 1998), *Pisum sativum* L.

‘Ctirad’ (2C = 9.09 pg; Doležel et al., 1998) and *Pisum sativum* L. ‘Kleine Rheinländerin’ (2C =

8.84 pg; Greilhuber and Ebert, 1994) were used as internal standards depending on the C-value

and standard availability. The suspension of isolated nuclei was filtered through a 48 µm nylon

mesh, incubated with RNase to degrade double stranded RNA 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 5,000 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). A proxy of the holoploid

genome size (2C) was calculated as follows: *Veronica* 2C nuclear DNA content (pg) =

(*Veronica* G1 peak mean/internal standard G1 peak mean)*genome size of the internal standard.

The DNA-ploidy level was estimated for each sample based on the values of the genome size

proxy and the available chromosome counts for the studied species (Martínez-Ortega et al.,

2004; Albach et al., 2008).

2.3. DNA extraction and AFLP genotyping

Total genomic DNA was extracted from silica gel dried material following the CTAB

protocol of Doyle and Doyle (1987). 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). All extractions are stored at -80 ºC at the "Biobanco de

ADN Vegetal" (University of Salamanca, Spain). The AFLP procedure followed the method

described by Vos et al. (1995) with slight modifications. Genomic DNA (ca. 100 ng) was

digested with MseI (New England BioLabs) and EcoRI (Fermentas) and ligated to double-

stranded adaptors with T4 DNA-Ligase (Thermo Scientific) in a single restriction-ligation

reaction for 3h at 37 ºC. Products were diluted and pre-amplified using primers EcoRI-A (5'

GAC TGC GTA CCA ATT CA - 3') and MseI-C (5' GAT GAG TCC TGA GTA AC - 3'). *Taq*

DNA Polymerase (BIOTOOLS BandM Labs. S.A) was used in the following PCR conditions: 2

1　min at 72 ºC, 29 cycles of 30 s at 94 ºC, 30 s at 56 ºC and 2 min at 72 ºC, and a final extension

2　of 10 min at 72 ºC. At this step, the pre-selective amplified fragments were visualized on 1%

3　TBE-agarose gel. After dilution of pre-selective products, selective amplifications were

4　performed with the following PCR profile: 10 min at 95 ºC, 13 cycles of 30 s at 94 ºC, 1 min at

5　65 ºC (decreasing 0.7 ºC in each cycle) and 2 min at 72 ºC, followed by 24 cycles of 30 s at 94

6　ºC, 1 min at 56 ºC and 2 min at 72 ºC, with a final extension of 10 min at 72 ºC. All PCR

7　reactions were performed on an Eppendorf-Mastercycler-Pro thermocycler. Twelve individuals

8　from ten different taxa representing the whole diversity of the final dataset were used to screen a

9　total of eight different combinations of selective primers. Four primer combinations were finally

10　selected (Table S2) based on the number and clarity of the peaks, and the polymorphism level

11　observed among individuals, which was checked to be sufficiently variable (i.e., overall genetic

12　similarity among individuals from the same population was higher than that found among

13　individuals from different populations, and much higher than the similarity detected among

14　individuals from different taxa). Final selective PCR products were multiplexed for genotyping

15　using the internal GeneScan 500 LIZ Size Standard (Applied Biosystems) in a multi-capillary

16　sequencer ABI Prism 3730 (Applied Biosystems). Negative controls were run at each step of the

17　process and 4.5% (= 11) of the samples were replicated in each independent run of PCR from

18　the same extracted DNA to assess genotyping errors (Bonin et al., 2004, 2007; Pompanon et al.,

19　2005). Final error rate was estimated after automated scoring according to Bonin et al. (2004)

20　by comparing the 1/0 matrices obtained for the replicated samples. Differences detected here

21　could be due either to technical causes and/or to the automated scoring process. The degree of

22　reproducibility of the data set was also investigated analyzing the placement of replicates in a

23　Neighbor-Joining tree.

24　*2.4. Optimization Procedure of Automated AFLP Scoring*

25　　　　Two different protocols were tested for the optimization of scoring parameters: the

26　protocol of Holland et al. (2008) and the open-source software optiFLP version 1.54 developed

by Arthofer et al. (2011). The results obtained with these methodologies did not show incongruence or significant differences between them. OptiFLP was chosen for our analyses because of its greater flexibility, faster analysis and the possibility to run the program with its "unsupervised mode", which uses phylogenetic tree's robustness to find settings that maximize the differences between groups of profiles. To use the software designed by Arthofer et al. (2011), electropherograms were first visualized in the software PeakScanner v.1.0 (Applied Biosystems) with all default settings except for a "*light smoothing*". Samples with poor quality profiles were discarded and AFLP data were exported to the open-source software optiFLP v.1.54 for the optimization of scoring parameters. Subsequently, fragments were automatically scored with tinyFLP v.1.30 (Arthofer, 2010) using the parameters optimized by optiFLP software (Table S3) and the data matrices from the different markers were concatenated using tinyCAT v.1.2 (Arthofer, 2010). A single scoring procedure was run to create data matrices to be used in subsequent genetic structure analyses.

2.5. Phylogenetic analyses

AFLP data were analyzed in a phenetic framework (i.e., distance based clustering), due to the limitations reported for alternative methods commonly used for phylogenetic reconstruction (Albach, 2007; Himmelreich et al., 2014). With the aim of understanding the phylogenetic relationships between closely related taxa of the complex, and investigating the possible occurrence of hybridization and/or polyploidization processes, a NeighborNet (NNet) was calculated based on Jaccard distances using SplitsTree4 v.4.13.1 (Huson and Bryan, 2006). Additionaly, Neighbor-Joining (NJ) trees based on Jaccard and Nei-Li distances were also built using SplitsTree4 v.4.13.1 and PAUP* 4.0b10 (Swofford, 2003), respectively, to assess the influence of the distance measure on the results. Bootstrap values (1000 replicates) from the NJ tree based on Jaccard distance were transferred to the NeighborNet graph.

2.6. Genetic Structure Analyses

The genetic structure was investigated in the entire AFLP dataset, as well as in data subsets using the same conditions. Data subsets were obtained from the partition of the entire dataset according to the four main K = 4 clusters identified during the initial analysis. Since we are not able to corroborate if the populations under study follow the Hardy-Weinberg model, the genetic structure was initially investigated using two different approaches: non-hierarchical K-means clustering (Hartigan and Wong 1979), which does not assume Hardy-Weinberg equilibrium, and Bayesian clustering analysis based on the MCMC algorithm using Structure v.2.3.4 (Pritchard et al., 2000). Non-hierarchical K-means clustering was performed using the R script of Arrigo et al. (2010). Numbers of K from 1 to 21 were tested and 100,000 independent runs starting from random seeds were performed for each K. To determine the optimal number of genetic clusters, the method of Evanno et al. (2005) was followed as adapted in Arrigo et al. (2010). Bayesian clustering analyses were performed in Structure adopting an admixture model and assuming correlated allele frequencies among populations (Falush et al., 2003) according to a methodology for dominant markers (Falush et al., 2007). Twenty replicates were run for each K from 1 to 21 with a burn-in length of 100,000 generations followed by 1,000,000 additional sampled generations. Structure analyses were run on the computer cluster developed by the UMS 2700 OMSI at the MNHN (Muséum National d'Histoire Naturelle, Paris). The optimal K value was determined using Structure Harvester (Earl and vonHoldt, 2012) following the method of Evanno et al. (2005). The output files were exported to CLUMPP v.1.1.2b (Jakobsson and Rosenberg, 2007) to perform an alignment of cluster assignments across the replicate analyses that we visualized afterwards using Distruct v.1.1 (Rosenberg, 2004). The results obtained with both approaches were independently displayed on a Principal Coordinates Analysis (PCoA) (Krzanowski, 1990) based on the Jaccard distance index using NTSYSpc 2.2 (Exeter Software, Setauket, NY; Rohlf, 2005). The percentages of variance explained by the two different clustering methods were also compared by AMOVA analyses (Excoffier et al., 1992) performed in Arlequin v3.5 (Excoffier et al., 2005, Excoffier and Lischer, 2010).

1   Furthermore, PCO-MC (principal coordinate-modal clustering; Reeves and Richards, 2009) was

2   implemented to test the significance of clusters found in PCoA using PCO-MC software

3   (https://www.ars.usda.gov/plains-area/fort-collins-co/center-for-agricultural-resources-

4   research/plant-germplasm-preservation-research/docs/reeves-pco-mc/). The P-value cutoff was

5   set to 0.9999 and the stability cutoff to 15% to maximize sensitivity to subtle population

6   structure while minimizing type I error (Reeves and Richards, 2009, 2010).

7

8   **3. Results**

9   *3.1. DNA-ploidy level determination*

10   DNA-ploidy level estimations according to flow cytometric analyses are shown in Table

11   S1 and Fig. 1. Ploidy was determined for most samples (94%), but for a few (6%) this was

12   problematic likely due to the age of leaf material. In general, our results were in accordance

13   with previous data with the group harboring diploids ($2x$), tetraploids ($4x$), hexaploids ($6x$) and

14   octoploids ($8x$). Heterogeneity in DNA-ploidy level within a taxon were found only for *V.*

15   *austriaca* subsp. *austriaca* ($4x$, $6x$), *V. austriaca* subsp. *jacquinii* ($2x$, $6x$), *V. orbiculata* ($2x$, $4x$)

16   and *V. rosea* ($2x$, $4x$). In most cases, only one DNA-ploidy level was observed per population,

17   with the exception of one population of *V. rosea* in Algeria ($2x$, $4x$), and two populations of *V.*

18   *orbiculata* ($2x$, $4x$) in Bosnia and Herzegovina and Croatia, where two DNA-ploidy levels were

19   observed. All populations initially determined as *V. austriaca* subsp. *jacquinii* / *V. orbiculata*

20   were tetraploid except for one population from Montenegro (pop. 19), which was shown to be

21   diploid and was finally ascribed to a new species which is described here (i. e., *V. dalmatica* sp.

22   nov., see section 5). Similarly, a population from Bosnia and Herzegovina initially identified as

23   *V.* affinis *jacquinii* (finally ascribed to *V. dalmatica*) was confirmed to be diploid (pop. 20), as

24   well as one population of *V.* affinis *linearis* from FYROM (pop. 42, with an *a posteriori*

25   identification as *V. linearis*). Most of the individuals initially catalogued as *V.* affinis *kindlii*

26   (pop. 32, 33, 35) or *V.* affinis *orsiniana* (pop. 51, 52, 53, 54, 55) were diploids, except for two

1   populations labeled as *V*. affinis *kindlii* (from Greece and Montenegro, pop. 31 and 34

2   respectively) that were found to be tetraploids.

3   *3.2. Automated scoring of the AFLP data and degree of reproducibility*

4   A total of 1127 polymorphic fragments were scored with the software tinyFLP (Table

5   S2). The error rate per locus obtained for our final data-set optimized with the optiFLP software

6   was on average 2.55%. In the NJ analyses, six of the eleven replicated samples were placed with

7   their respective original samples, with a bootstrap value > 98%. The other five replicated

8   samples were recovered at least in the same cluster as their respective original samples and

9   others of the same population.

10   *3.3. Phylogenetic reconstruction*

11   Phylogenies reconstructed with different distance methods were congruent with one

12   another and supported the monophyly of most of the previously recognized species by Rojas-

13   Andrés and Martínez-Ortega (2016), with high bootstrap values (Table 1). However, since only

14   one population of *V. rhodopea* was included in our study, monophyly of this particular species

15   could not be confirmed. On the other hand, most of the internal nodes of the NJ trees were not

16   supported by bootstrapping (Fig. S1), and the NeighborNet (Fig. 2) showed a high degree of

17   reticulation existing in the group. Nevertheless, four main groups (I, II, III and IV) were

18   identified according to the placement of individuals in the network and are presented below.

19   Group I. This low-supported group comprised five monophyletic diploid species: *V.*

20   *kindlii* (BS = 100), *V. linearis* (BS = 99.9), *V. orsiniana* (BS = 100), *V. rhodopea* (BS = 100),

21   and *V. teucrioides* (BS = 100). Diploid individuals from the south of Italy (ind. 135–140) of

22   uncertain taxonomic identity, which are morphologically similar to *V. orsiniana* (*V.* affinis

23   *orsiniana*), were recovered as monophyletic together with *V. kindlii*. One tetraploid population

24   initially determined as *V.* affinis *kindlii* (ind. 80–82), was recovered as an independent lineage

25   (BS = 100).

Group II. The monophyly of this group was clearly supported (BS = 99). It comprised three species recovered as monophyletic with bootstrap values of 100%. The diploid *V. tenuifolia* with 3 subspecies [subsp. *tenuifolia*, subsp. *javalambrensis*, subsp. *fontqueri*] and the tetraploid *V. aragonensis* are endemic to the Iberian Peninsula. The third species *V. rosea*, mostly diploid but comprising some tetraploid individuals in a single population, is endemic to North Africa.

Group III. This group included mostly diploid species with highly supported monophyly but there was low support for it as a whole. First, *V. krylovii* (BS = 100), one of the species representing the subsection in Siberia and Kazakhstan, was recovered as a strongly supported clade. Second, *V. prostrata* from central Europe was found to be also well resolved (BS = 100), as well as *V. turrilliana* (BS = 100), an endemic taxon from the border region of Bulgaria and Turkey. Individuals identified as *V. crinita* were recovered in two distinct lineages (BS = 100). Finally, this group comprised diploid individuals of *V. austriaca* subsp. *jacquinii* (BS = 100), and mixed cytotypes (2*x*, 4*x*) of *V. orbiculata* (BS = 71.7), as well as tetraploid individuals of intermediate morphology recorded as *V. austriaca* subsp. *jacquinii* / *V. orbiculata*.

Group IV. Following the initial taxonomic classification, this group included four polyploid taxa: i) tetraploid and hexaploid cytotypes of *V. austriaca* comprising three subspecies: subsp. *austriaca*, subsp. *dentata*, and subsp. *jacquinii*; ii) the tetraploid species *V. satureiifolia*; iii) the octoploid *V. sennenii*, endemic from the north of Spain; and iv) octoploid individuals of *V. teucrium* var. *teucrium*, and var. *angustifolia*. The monophyly of this group was well supported (BS = 93.9) but not the monophyly of most of the species within it. Phylogenetic analysis (Fig. 2) only supported the monophyly of hexaploid individuals of *V. austriaca* subsp. *jacquinii* but with a very low bootstrap value (BS = 59.6). *Veronica satureiifolia* and *V. sennenii* were recovered together with octoploid individuals identified as *V. teucrium* subsp. *angustifolia* (BS = 67.5).

3.4. Genetic Structure

1   Following the method implemented in Structure Harvester (Evanno et al., 2005),

2 Bayesian clustering analysis supported an optimal partition of the subsection in three clusters.

3 On the contrary, non-hierarchical K-means clustering analysis of the same dataset estimated K =

4 2 as the most likely number of genetic clusters. However, PCoA (Fig. 3) and AMOVA analyses

5 (Table 2) demonstrated that the clustering proposed by Structure explained a higher percentage

6 of the variance among groups than K-means (see Table 2). Accordingly, we here focus on

7 results of Bayesian clustering. High levels of admixture were found in Bayesian clustering

8 analyses performed at higher values of K (Fig. S2C). It should be pointed out that most taxa

9 included in our analyses (with the exception of polyploids from group IV and of *V. teucrioides*

10 from group I) were recovered as independent clusters when Structure analyses were performed

11 at K = 20 (Fig. S2C). In addition, an exclusive cluster was found grouping tetraploid

12 populations of *V. austriaca* subsp. *jacquinii* / *V. orbiculata* and *V. orbiculata.*

13   The clusters revealed by Structure at K = 4 (Fig. 2 and Fig. S2A) generally concurred

14 with the groups identified by the NeighborNet with the only exception of *V. orsiniana* (Table 1),

15 and partially corresponded with geographic regions: cluster A included a group of narrow

16 endemics mostly restricted to the south of the Balkan Peninsula; cluster B comprised the three

17 well recognized species from the Iberian Peninsula and North Africa; cluster C recovered most

18 diploids from the Balkan Peninsula, *V. krylovii* from Russia, and *V. orsiniana*; cluster D

19 included all the polyploid taxa mainly from central Europe and north of Spain.

20   Additional Bayesian clustering analyses within clusters A, B, C and D estimated an

21 optimal K = 5, K = 3, K =3 and K = 2, respectively (Fig. S2B). According to the results

22 obtained for cluster A, the four diploid species and the tetraploid population of *V.* affinis *kindlii*

23 from Mt. Vermion (pop. 31) were recovered in independent clusters. In cluster B, the three

24 species from the Ibero-NorthAfrican group (*V. aragonensis*, *V. rosea* and *V. tenuifolia*) were

25 recovered in independent and homogeneous clusters almost without admixture among them.

26 When analyses were performed within cluster C, one cluster grouped a single species (i.e., *V.*

orsiniana) and the other two clusters divided the taxa from group III in two subgroups. One subgroup included diploids of *V. austriaca* subsp. *jacquinii* and *V. orbiculata* and all intermediate populations displaying high levels of admixture [*V. austriaca* subsp. *jacquinii* / *V. orbiculata, V. crinita* (= *V. crinita* f. *bosniaca*) and *V.* affinis *kindlii*]. Another subgroup, also with a certain degree of admixture, comprised the remaining diploid species from group III [*V. crinita*, *V. crinita* (= *V. thracica*), *V. turrilliana*, *V. prostrata*, *V. krylovii*]. Within cluster D, there was an obvious geographic pattern in which polyploids form two separate clusters, however, with many individuals forming a continuous gradation of proportion scores between the two clusters (Fig. 4). Hexaploid individuals classified as *V. austriaca* subsp. *jacquinii* (pop. 11–16) were included in one cluster with a very high proportion score (> 0.99; data not shown). By contrast, tetraploid individuals identified as *V. satureiifolia*, octoploid populations assigned to *V. senneni* and *affinis* individuals, had a proportion score > 0.99 (data not shown) to be defined in a second cluster. Most individuals of *V. teucrium* var. *angustifolia* showed a high genetic affinity to this second cluster with low levels of admixture.

        PCO-MC analysis recovered ten species as significant independent clusters: *V. aragonensis*, *V. dalmatica*, *V. kindlii*, *V. linearis*, *V. orbiculata*, *V. orsiniana*, *V. prostrata*, *V. rosea*, *V. tenuifolia*, and *V. turrilliana* (Table 1).

4. Discussion

4.1. The importance of auto- and allopolyploidization in the subsection and the recurrent formation of polyploids

        The diversity of cytotypes and the existence of mixed-ploidy levels within species and populations in the group reveal that polyploidization has occurred likely continuously since the origin of the subsection ca. 2.8 Mya (Meudt et al., 2015). The pattern of reticulation shown by the NeighborNet (Fig. 2) and the high levels of admixture found by Structure suggest that *V.* subsection *Pentasepalae* is composed of species that are in the initial stages of divergence.

1   Furthermore, based on morphological and (phylo-)genetic intermediacy between potential

2   parental species, hybridization is confirmed in *V.* subsection *Pentasepalae*. In addition, ILS

3   cannot be excluded as a cause of the lack of resolution observed for internal nodes of the

4   NeighborNet and NJ trees. Nevertheless, phylogenetic analyses demonstrate that polyploid taxa

5   distributed mainly in central Europe (group IV) constitute a very well supported group (Fig. 2).

6   It should be pointed out that the higher number of AFLP fragments present in polyploid

7   individuals may produce a bias in posterior analyses towards the apparent monophyly of most

8   of the polyploids. However, this artificial grouping due to a higher number of AFLP fragments

9   in polyploids can be discarded in our dataset because other polyploid species of the subsection

10   are not recovered within group IV (e.g., *V. aragonensis* and *V. orbiculata*). Thus, our study

11   could indicate a common origin of polyploid entities from group IV at the tetraploid level. This

12   hypothesis was previously rejected by Rojas-Andrés et al. (2015) due to the existence of

13   morphological variation among polyploids and to the fact that most polyploid species have a

14   polytopic origin (Soltis & Soltis, 1999). However, a possible explanation is that hexa- and

15   octoploids have emerged (probably several times independently within each lineage) after a

16   previous differentiation of lineages at the tetraploid level. Furthermore, the extinction of diploid

17   or some of the tetraploid ancestors within group IV is also likely, which together with the

18   limitations of the available methodologies hampers the obtention of ancestor-derivative patterns

19   within group IV (Stebbins, 1971; McDade, 1992; Buggs et al., 2014).

20   Our results suggest that polyploid species in the subsection may have emerged by

21   different processes. Whereas autopolyploidization appears to be the main evolutionary force for

22   some taxa, allopolyploidization also seems to be common. Evidence for autopolyploidization is

23   found, for example, in group II. Tetraploid individuals have been found in a population of *V.*

24   *rosea* from Algeria (Table S1, pop. 63), which cluster together with the rest of diploid

25   individuals of the species (Fig. 2, Fig. S2), thus suggesting a recent autopolyploid event

26   occurring within the population.

1   Analyses further point to an autopolyploid origin of the octoploid *V. sennenii* from

2   tetraploid *V. satureiifolia.* The individuals belonging to *V. satureiifolia* and *V. sennenii* are

3   recovered in the same group in the NeighborNet without any clear separation between

4   individuals of both species, and they form a homogeneous cluster in the Bayesian clustering

5   analyses (Fig. 4). Flow cytometric analyses (Table S1) have confirmed that both species (and

6   consequently, both ploidy levels) grow in sympatry in the province of Huesca, in the north of

7   Spain (pop. 69, 4x; and pop. 73, 8x). On the basis of all these results, this is another example of

8   autopolyploid speciation in natural populations (Soltis et al., 2007). Likewise, according to the

9   NeighborNet, most octoploid individuals determined as *V. teucrium* var. *angustifolia* are nested

10   with *V. satureiifolia* and *V. sennenii* and have very similar genetic composition in clustering

11   analyses (Fig. 4). Furthermore, one population of *V. satureiifolia* (pop. 70, 4x) and one of *V.*

12   *teucrium* var. *angustifolia* (pop. 87, 8x) have been found in close proximity (about 400 m) in the

13   region of Île-de-France. *Veronica satureiifolia* and *V. teucrium* var. *angustifolia* were also

14   shown to share the same cpDNA haplotype (Rojas-Andrés et al. 2015). Consequently, a

15   plausible interpretation is that multiple autopolyploidization events might have occurred in the

16   tetraploid *V. satureiifolia* giving rise to octoploids that have been identified as *V. sennenii* in the

17   Iberian Peninsula and *V. teucrium* var. *angustifolia* in France. Alternatively, a past continuous

18   distribution area of the octoploid entity and a subsequent fragmentation scenario cannot be

19   discarded, although it seems unlikely considering the distance of 500 km between the French

20   southernmost and the Spanish northernmost populations, and the existence of the Pyrenees in

21   between.

22   Our results also reveal that at least two episodes of polyploidy have occurred within *V.*

23   *orbiculata* (Fig. 2; Group III), although the processes seem to be more complex. Autopolyploid

24   formation has been detected in one population from Bosnia and Herzegovina (Table S1; pop. 43)

25   and might be occurring in other, not surveyed, populations. Tetraploid individuals found in this

26   population (ind. 107, 108) are nested within the diploid individuals (ind. 109–116) with a BS

1   value of 100% (Fig. 2). Moreover, individuals of both cytotypes are recovered together as a

2   significant cluster in PCO-MC and Structure analyses. In contrast, tetraploid individuals (ind.

3   117–119) from another population in Croatia (Table S1; pop. 45) are recovered well separated

4   from the diploids of the same population (ind. 114–116; BS = 99.5) and are not included

5   together within any significant cluster in PCO-MC analyses. Moreover, an exclusive cluster is

6   found in Structure analyses (Fig. S2C) that groups these tetraploid individuals with tetraploid

7   populations recorded as *V. austriaca* subsp. *jacquinii* / *V. orbiculata*. Thus, these tetraploids are

8   probably the result of an allopolyploidization event. Consequently, *V. orbiculata* is a further

9   example of a diploid-polyploid species with numerous independent origins of polyploid entities

10   as shown in other species (Soltis and Soltis, 1999; Bardy et al., 2010, 2011).

11   There are other strong arguments of recurrent formation of allopolyploids within group

12   III. Specifically, tetraploid individuals from Bosnia and Herzegovina (Table S1; ind. 42–47)

13   recorded as *V. austriaca* subsp. *jacquinii* / *V. orbiculata* due to their transitional morphology,

14   are recovered by the NJ in an intermediate position between these species (Fig. S1). Thus, a

15   hybrid origin of these populations is suggested with diploid *V. austriaca* subsp. *jacquinii* and *V.*

16   *orbiculata* as putative parental species. In addition, diploid individuals located in Montenegro

17   labeled as *V. affinis kindlii* (ind. 83–84), are also recovered in an intermediate position together

18   with a population of *V. crinita*, putatively belonging to f. *bosniaca* (ind. 64–66). Furthermore,

19   the position of individual 85 ($2x$) in the NeighborNet suggests that homoploid hybridization

20   could be also an important evolutionary process occurring in the group, which is here

21   demonstrated for the first time and of which *V.* x *gundisalvi* may represent an additional

22   example (Martínez-Ortega et al., 2004).

23   Last, hybridization and/or introgression events may have affected to the tetraploid

24   population of *V.* aff. *kindlii* located in Mt. Vermion (Greece; Fig. 2, group I, pop. 31). Bayesian

25   clustering analyses show high levels of admixture with the polyploid group IV (Fig. S2A).

26   *Veronica austriaca* subsp. *jacquinii* is the only species from group IV distributed in this

southern area of the Balkan Peninsula. Thus, the position of population 31 in the NeighborNet could be influenced by hybridization and/or introgression processes involving this species and representatives from group I or its ancestors.

4.2. The challenging task of delimitating species within a recently diverged diploid-polyploid complex

Species delimitation within recently diverged plant complexes is currently a major challenge for systematists. In general, at this level of lineage separation, phenotypic differences among species may not be evident and a clear phylogenetic signal is not always obtained (Federici et al., 2013). Consequently, other characteristics (e.g., ploidy levels, differences in habitat, pollinators, phenology, etc.) are important lines of evidence when delimiting species in this recently diverged, phenotypically, and phylogenetically complex groups. This situation requires the adoption of the general lineage concept of species in which different species properties (that have been used as criteria under rival species concepts), serve as lines of evidence to assess lineage delimitation (de Queiroz, 2007).

Identifying biological diversity at the species level is even more challenging when processes such polyploidy are involved in the evolution of a group. Polyploidy has long been considered a mechanism of direct sympatric speciation (Otto and Whitton, 2000; Schemske, 2000; Rieseberg and Willis, 2007). However, recent studies suggest that polyploid speciation is not necessarily an instantaneous process (Husband et al., 2013). The formation of unreduced gametes and other biological traits are fundamental in initial stages of polyploid emergence and establishment (Rieseberg and Willis, 2007; Fowler and Levin, 2016). Both, the rates of production of unreduced gametes and the successful long-term establishment and spread of new polyploid individuals are affected by genetic and environmental factors (Ramsey and Schemske, 1998; Comai, 2005; Lafon-Placette et al., 2016). Regardless of the timing of the process, it has been estimated that 15% of angiosperm speciation events, and even more in *Veronica*, are associated with a ploidy increase (Albach et al., 2008; Wood et al., 2009). Indeed, it has been

corroborated that the effect of genome duplication on countless features (e.g., reproductive biology, phenotype, physiology, geographical and environmental distributions of cytotypes, genetic, epigenetic and genomic consequences, and so forth; reviewed in Ramsey and Ramsey, 2014). Whether caused by ploidy *per se*, adaptation or founder effects and genetic drift, these changes may maintain polyploids as separately evolving metapopulation lineages from their parental taxa, which justify their treatment as separate taxonomic species (de Queiroz, 2007).

Our study has demonstrated that most of the species of *V.* subsection *Pentasepalae* are still in the initial stages of divergence. Moreover, auto- and allopolyploids have been identified within the group. In this situation, the taxonomic status of diploid and polyploid taxa within *V.* subsection *Pentasepalae* is reviewed adopting the general species concept of de Queiroz (2007) and making use of an integrative taxonomic approach. In our case study, no significant differences in habitat preference are observed, experimental data on reproductive biology are not available, and probably many species share pollinators to a great extent. Thus, we have based the decisions of species delimitation on ploidy level, phylogeny and genetic divergence, but also on information from morphology, distribution, ecology, etc., available in Rojas-Andrés and Martínez Ortega (2016). Furthermore, a conservative approach to taxonomy is preferable when incongruences among different lines of evidence are found (Carstens et al., 2013). When such situation was encountered, we maintained the last taxonomic treatment of Rojas-Andrés and Martínez-Ortega (2016).

Moreover, we think that populations identified in this study, for which we have not obtained sufficient evidence to be delimited as species (e.g., tetraploid hybrid populations catalogued as *V. austriaca* subsp. *jacquinii* / *V. orbiculata* that cannot be identified as a different species but could potentially evolve as a distinct lineage) would have to be considered as functional units of biological diversity. In these cases, this recognition would help to address future ecological, evolutionary and taxonomic questions (Ramsey and Ramsey, 2014; Laport and Ng, 2017).

1   Additionally, we consider that further molecular tools (i.e., molecular studies using

2 neutral markers as SSRs; López-González et al., *in prep*), morphological data (e.g., traits with

3 potential impact on individual fitness), and deeper ecological and biological information (e.g.,

4 environmental distribution analyses among cytotypes, crossing experiments to understand

5 reproductive interactions) are needed to re-evaluate whether the species rank is appropriate for

6 some of these taxa (e.g., polyploids from group IV).

7 *4.3. Taxonomic considerations*

8   This study provides new insights into the systematics of the polyploid complex

9 *Veronica* subsection *Pentasepalae*. Our analyses support 20 distinct species in the group. The

10 most recent taxonomic treatment available (Rojas-Andrés and Martínez-Ortega, 2016) has been

11 revised and updated (changes summarized in Table S1).

12   First, the individuals of *V. austriaca* subsp. *jacquinii*, included in this study are placed

13 in two separate phylogenetic lineages differentiated by their ploidy levels (diploids vs.

14 hexaploids) (Fig. 2). Monophyly of the diploid individuals, which represent an example of

15 cryptic species in the subsection, is well supported by phylogenetic reconstruction (BS = 99.8),

16 whereas hexaploids are recovered with a low bootstrap value (BS = 59.6). Additionally, PCO-

17 MC and Bayesian clustering analysis recover these populations as significant and independent

18 clusters (Table 1 and Fig. S2C). Indeed, after an exhaustive revision of herbarium specimens,

19 morphological characters corresponding to each of these species have been found (see section 5).

20 In addition, the distribution area of the diploid cytotypes is restricted to the Adriatic coast of

21 Albania, Bosnia and Herzegovina, Croatia and Montenegro. Based on all these lines of evidence,

22 we consider that diploid entities of *V. austriaca* subsp. *jacquinii* should be recognized at the

23 specific rank as *V. dalmatica* N.Pad.Gar., Rojas-Andrés, López-González and M.M.Mart.Ort.

24 (see section 5).

25   Second, the analyses presented here allow the recognition of *V. thracica* at the species

26 level. *Veronica thracica* was described by Velenovsky (1893) to differentiate plants mainly

occurring in Bulgaria characterized by white hairy stems and oval-obovate, deeply cordate, almost amplexicaulus leaves. This name was later combined under *V. teucrium* as subspecies (Velenovsky, 1898) or variety (Maly, 1908) and has been related to *V. crinita* by other authors (e.g., Watzl, 1910; Peev, 1995). These individuals were considered within the variation of *V. crinita* in the most recent taxonomic treatment due to their morphological similarities (Rojas-Andrés & Martínez-Ortega, 2016). Genetic data now provide evidence that these populations constitute an independent evolutionary lineage differentiated from typical *V. crinita* described from Hungary (Fig. 2 and Fig. S2C) and it represents an additional example of a cryptic species within the subsection (Martínez-Ortega et al., 2004). Furthermore, after examining the morphological characters of this material, we found that *V. thracica* has dense tomentose indument on leaves and stems, formed by patent to slightly incurvate hairs that confer a whitish (light green *in vivo*) color to the plant. In contrast, *V. crinita* has villous indument on leaves and stems, which is constituted by crooked, generally interwoven hairs that confer a brownish green color to the plant. The leaves are concolor (i.e., upper and lower leaf sides of the same color) and more densely pilose in *V. thracica*, while they are slightly bicolor (i.e., dark / dive green color of the upper leaf side vs. green color of the underside of the leaf) and comparatively not so densely pilose in *V. crinita*. We consider that there is enough evidence to recognize these lineages as separate species, and consequently, the recognition at the specific rank of the Bulgarian populations is proposed.

Finally, several taxa have been described under *V. teucrium*, but only two varieties have been recognized in the last taxonomic treatment of the subsection (Rojas-Andrés and Martínez-Ortega, 2016): *V. teucrium* var. *teucrium* and *V. teucrium* var. *angustifolia.* These two mostly allopatric octaploid entities are morphologically distinct. *Veronica teucrium* var. *teucrium* is distributed in Germany, Austria, and Bulgaria, while *V. teucrium* var. *angustifolia* occurs in France extending towards western Switzerland and northern Italy. Moreover, both varieties are recovered in different subclusters in our molecular analyses (Fig. 4). Based on all available data,

23

1   we consider that both entities should be recognized at the specific level as *V. teucrium* L. and *V.*

2   *angustifolia* (Vahl) Bernh., respectively. Additionally, apart from geographic differentiation,

3   individuals of *V. teucrium* var. *angustifolia* are very similar to *V. sennenii* in size and

4   appearance. Taken together these data would suggest that *V. sennenii* and *V. teucrium* var.

5   *angustifolia* have the same parental origin, although they could have arisen from the same or

6   different autopolyploid events. If they were considered synonyms, the name *V. angustifolia*

7   (Vahl) Bernh. would prevail at the specific level, according to the principle of priority. But this

8   taxonomic decision should not be firmly adopted until additional exhaustive analyses including

9   more populations of these taxa are performed.

10   Additionally, the importance of some populations from the south of Italy identified in

11   *Flora d'Italia* as *V. austriaca* and their relationship with those from the Balkan Peninsula have

12   been previously highlighted (Fischer, 1982). However, the identity of these plants (ind. 135–140;

13   labeled in this study as *V.* affinis *orsiniana*) has remained unclear for many years. Our analyses

14   confirm the identity of these plants as *V. kindlii* and show that this species should be considered

15   independent from *V. orsiniana* or *V. austriaca* (Table 1, Fig. 2). The name *V. kindlii* has

16   recently been resurrected to designate those populations from the Balkan Peninsula, which were

17   previously known as *V. orsiniana* (Rojas-Andrés et al., 2015). Thus, a clear amphi-Adriatic

18   distribution of *V. kindlii* is now demonstrated. Finally, there is no evidence that the individuals

19   initially identified as *V.* affinis *kindlii* belong to *V. kindlii.* Unfortunately, the taxonomic status

20   of these entities remains unresolved. Additional exhaustive field sampling in order to have a

21   good representation of these unresolved entities and posterior molecular studies could shed

22   some light on the taxonomic identity of these individuals.

23   Another important outcome is the corroboration of the genetic distinctiveness and

24   monophyly of *V. linearis*, a diploid endemic species from FYROM that passed unnoticed for

25   many years. This name did not appear in Floras or monographs of *Veronica*. The plant was

26   initially described as *V. kindlii* var. *linearis* by Bornmüller (1937) and has recently been

1   elevated to the species level based on morphological evidence (Rojas-Andrés and Martínez

2   Ortega, 2016). According to our phylogenetic reconstruction (Fig. 2) and PCO-MC analyses

3   (Table 1), *V. linearis* is recovered as monophyletic within group I and its closest relatives are *V.*

4   *kindlii* and *V. teucrioides* (BS = 74.7 for [*V. linearis + V. kindlii + V. teucrioides*]). In addition,

5   one population with dubious morphological characters labeled as *V.* affinis *linearis* (pop. 42) is

6   recovered within *V. linearis.* Nevertheless, clustering analyses (Fig. S2B) showed introgression

7   with *V. teucrioides*, which is in agreement with the dubious determination based on

8   morphological characters.

9   With regard to the delimitation of varieties and subspecies, the theoretical framework

10   behind their concept is less clear as it is for species. We have attempted to avoid the use of these

11   ranks in the proposed taxonomic changes, but in two cases the data available are not conclusive:

12   i) Within group II, three subspecies are recognized under *V. tenuifolia.* Their different

13   distribution areas and the divergence found in the NeighborNet and NJ trees between

14   populations corresponding to each subspecies could indicate reduced gene flow among them

15   (Fig. 2, Fig. S1), as previously showed by studies based on AFLP and morphology (Martínez-

16   Ortega et al., 2004; Andrés-Sánchez et al., 2009). However, our study does not support the

17   recognition of these taxa at the specific rank (see PCO-MC and genetic structure results in Table

18   1, Fig. S2). Likewise, phylogenetic analyses based on nuclear and plastid DNA sequences did

19   not differentiate among the three subspecies currently recognized (Rojas-Andrés et al. 2015).

20   Due to the incongruences found among different sources of data, we suggest to maintain their

21   current formal rank as subspecies.

22   ii) The subspecific rank has also been retained for some taxonomic entities belonging to

23   group IV (i.e., three subspecies recognized under *V. austriaca*). The lack of resolution in our

24   AFLP analyses (Table 1, Fig. 4) as well as in nuclear and plastid DNA trees (Rojas-Andrés et

25   al., 2015) do not support the recognition of current subspecies as different species, nor the

26   unification in a single species. Unfortunately, the phylogenetic relationships among these taxa

remain unresolved. These subspecies have been described in the last taxonomic treatment given the morphological and chorological differences among them (Rojas-Andrés and Martínez-Ortega, 2016). Consequently, we have retained the subspecies rank within *V. austriaca* (subsp. *austriaca*; subsp. *dentata*, subsp. *jacquinii*), at least until future studies clarify the evolutionary history and taxonomy of these polyploid entities.

5. Conclusions and description of a new species.

The exhaustive sampling of *V.* subsection *Pentasepalae*, and the use of AFLP fingerprinting together with flow cytometry data provided new insights into the evolutionary history and species delimitation of a taxonomically complex plant group, in which auto- and allopolyploidization appear to be active evolutionary processes, even nowadays. Based on all sources of data currently available, *V.* subsection *Pentasepalae* contains at least 20 monophyletic species, five of them narrow endemics. This taxonomic framework is essential to design suitable conservation strategies. Future studies should focus on trying to understand in more detail the role that hybridization has played in the evolution of the subsection, and on ecological factors that make polyploidy so important in plant evolution and speciation.

Veronica dalmatica N.Pad.Gar., Rojas-Andrés, López-González & M.M.Mart.Ort.**, sp. nov**. – Type: Holotype: Bosnia and Herzegovina, Republica Srpska: between Brgat and Trebinje, 42.68289 N, 18.28949 E, 283 m, 10/VI/2015, clearings in a forest of *Carpinus betulus* with *Paliurus spina-christi*. *Leg. M. Martínez Ortega, X. Giráldez, N. Padilla and N. López, MMO6119* (SALA No. 157047!).

Next, we provide a diagnosis between *V. dalmatica* and the morphologically closest taxa, as well as a full description of *V. dalmatica*, which is parallel to the descriptions provided

by Rojas-Andrés and Martínez-Ortega (2016). The indument is described according to Beentje (2010). Two measurements given together refer always to length × width.

Diagnosis: *V. dalmatica* differs from *V. austriaca* subsp. *jacquinii* by its smaller plant size (10–16 vs. 25–50 cm), having shorter stem-hairs (0.3–0.4 vs. 0.8–1.2 mm), smaller leaves (12–16 × 5–10 vs. 20–30 × 10–20 mm), shorter styles (3–5 vs. 4–7 mm) and tiny capsules with a less deep sinus (up to 0.5 vs. 1.0 mm). Attending to chromosome number, *V. dalmatica* is diploid (2n = 16) whereas *V. austriaca* subsp. *jacquinii* is frequently hexaploid [2n = (32), 48, (64), (80)].

Overlapping in ploidy level, *V. dalmatica* (2n = 16) is morphologically differentiated from *V. orbiculata* (2n = 16, 32) by having the eglandular hairs of the stem not arranged in two opposite lines along it. Apical shoot leaves are opposite in *V. dalmatica*, whereas in *V. orbiculata* they can be opposite, alternate or verticillate by three.

Description: *Stems* (6) 10–16 (24) cm long, slightly ascending to erect, covered by eglandular hairs (0.18) 0.30–0.40 (0.81) mm long, incurvate, ± appressed and antrorse, not arranged in 2 opposite lines along the stem; apical shoot bearing (5) 8–11 (16) pairs of leaves. *Leaves* opposite, (8) 12–16 (20) × (3) 5–10 (16) mm; ovate, obovate, or narrowly to widely trullate; more or less rounded or cuneate at the base; pinnatifid to pinnatisect, with linear-lanceolate to narrowly elliptic segments, variable in width, entire, revolute to subrevolute, subglabrous or pilose, covered by hairs (0.06) 0.10–0.18 (0.23) mm long, sessile to shortly petiolate. *Basal leaves* pinnatifid to pinnatisect, segments 0.4–1.0 mm wide; *medium leaves* (i.e., those situated in the central segment of the stem) pinnatipartite to pinnatisect, segments 0.25–1.00 mm wide; *uppermost leaves* pinnatisect, rarely bipinnatifid, segments 0.25–0.60 mm wide. *Leaves of the apical shoot* opposite, linear to lanceolate, narrowly elliptic, entire, dentate-serrate or pinnatifid, revolute to subrevolute. *Racemes* axillary, opposite, exceptionally solitary, bearing (9) 20–40

(48) flowers, loosely to densely arranged; *peduncles* (2.5) 3.0–7.0 (11) cm long, covered by a

non-glandular indument similar to that of the leaves; *bracts* (1.5) 3.0–5.0 (8.0) mm long, linear,

entire, exceptionally pinnatifid to pinnatisect at the base with one or two segments, glabrous or

subglabrous, covered by hairs similar to those covering the leaves; *pedicels* (1.6) 3.0–5.0 (8.5)

mm long. Calyx (0.7) 2.0–4.0 (5.0) mm long, with (4) 5 sepals, linear-lanceolate, usually shorter

than the capsule, glabrous or subglabrous. *Corolla* 9–15 mm in diameter, light or dark blue.

Capsule (2.0) 3.0–5.0 (6.0) × (2.0) 3.0–4.5 (5.3) mm, glabrous, widely elliptic or widely

obovate to very widely depressed ovate-obovate, rounded at the base, slightly emarginated or

rounded at apex, sinus up to 0.5 (0.6) mm depth. *Style* (2.8) 3.0–5.0 (6.0) mm long. *Seeds* (0.9)

1.3–1.7 × 1.5–1.8 (2.0) mm, ca. 8 per capsule.

Chromosome Number. – $2n = 16$

Habitat. – Dry and stony meadows, steppes, forest glades and shrublands, rocky slopes; usually

on calcareous soils; (50) 200–1,100 (1,400) m above sea level.

Distribution. – W Balkan Peninsula; Albania, Bosnia and Herzegovina, Croatia, Montenegro.

Etymology. – The epithet indicates geographical distribution of the species. Dalmatia is a

historical region of the Adriatic Sea ranging from Rab (Croatia) to the Bay of Kotor

(Montenegro) including a small area of Bosnia and Herzegovina.

Notes. – The plant is illustrated in Rojas-Andrés and Martínez-Ortega, 2016; Figure 3 (f-i).

Specimens examined. – See Appendix B.

Appendix A. Supplementary Material (Fig. S1, Fig. S2, Table S1, Table S2 and Table S3)

Appendix B.

Specimens examined of *V. dalmatica*. Information listed is country, locality, geographical coordinates, altitude, collection date, habitat, collector names, collector number, and herbarium code (Thiers, 2017).

ALBANIA. **Lezhë:** Lezhë, cerca de Fishte, 41.89112N, 19.67781E, 56 m, 17/VI/2015, pastos secos en flysch, *M. Martínez Ortega, X. Giráldez, N. Padilla & N. López, NPG48* (SALA 157035).

BOSNIA AND HERZEGOVINA. **Republica Srpska**: between Brgat and Trebinje, 42.68289 N, 18.28949 E, 283 m, 10/VI/2015, clearings in a forest of *Carpinus betulus* with *Paliurus spina-christi*, *M. Martínez-Ortega, X. Giráldez, N. Padilla and N. López, MMO6119* (SALA 157047); entre Tjentište y Gacko, 43.18547N, 18.56603E, 1085 m, 13/VII/2010, prados sobre calizas. *S. Andrés, X. Giráldez, M. Martínez Ortega & B. Rojas Andrés, MO5552* (SALA 149274); entre Trebinje y Dubrovnik, 42.68992N, 18.297E, 282 m, 14/VII/2010, sobre rocas calizas en zonas aclaradas, *S. Andrés, X. Giráldez, M. Martínez Ortega & B. Rojas Andrés*, *BR108* (SALA 149284); entre Gaccko y Tjentište, 43.1847N, 18.56578E, 1076 m, 10/VI/2015, prados calizos subalpinos, *M. Martínez-Ortega, X. Giráldez, N. Padilla and N. López*, *MO6123bis* (SALA 157025).

CROATIA. **Dubrovnik-Neretva**: Dubrovnik, entre Sumet y Gornji Brgat, 42.64408N, 18.14644E, 212 m, 15/VII/2010, prados sobre calizas, *S. Andrés, X. Giráldez, M. Martínez Ortega & B. Rojas Andrés, SA384* (SALA 149286); Dubrovnik, Gromača, 42.72444N, 18.01778E, 320 m, 14/VII/2010, prados secos sobre calizas, *S. Andrés, X. Giráldez, M. Martínez Ortega & B. Rojas Andrés, BR112* (SALA 149039).

MONTENEGRO. **Andrijevica:** Andrijevica, a 2 km en dirección Kolasin, 42.73946N, 19.76141E, 989 m, 8/VI/2015, claros de robledal, calizas, *M. Martínez Ortega, X. Giraldez, N. Padilla & N. López, NPG31* (SALA 157015); Andrijevica, a 1 km en dirección Kolasin, pista que sale a la derecha, 42.74523N, 19.77552E, 884 m, 8/VI/2015, prados sobre calizas, *M. Martínez-Ortega, X. Giráldez, N. Padilla & N. López, NPG32* (SALA 157016); **Bar**: entre Sutorman y Karuci, Rumija Planina, 42.16105N, 19.09708E, 738 m, 9/VI/2015, claros de bosque sobre calizas junto a la carretera, *M. Martínez Ortega, X. Giráldez, N. Padilla & N. López, NLG136* (SALA 157017); entre Sutorman y Karuči, Rumija Planina, 42.16219N, 19.09794E, 753 m, 16/VII/2010, prados sobre calizas, *S. Andrés, X. Giráldez, M. Martínez Ortega & B. Rojas Andrés, MO5556* (SALA 149285); **Kotor:** Kotor, Lovcen, 42.41802N, 18.79413E, 904 m, 9/VI/2015, claros sobre calizas, *M. Martínez Ortega, X. Giráldez, N. Padilla & N. López, NLG137* (SALA 157018); **Žabljak:** Meždo, 43.16384N, 19.14908E, 1390 m, 12/VI/2015, prados cortos con enebros, calizas, *M. Martínez Ortega, X. Giráldez, N. Padilla & N. López NLG139* (SALA 157030); Žabljak, cercanías del pueblo, 43.16978N, 19.15008E, 1392 m, 18/VII/2010, prados secos sobre calizas con *Juniperus, S. Andrés, X. Giráldez, M. Martínez Ortega & B. Rojas Andrés, SA392* (SALA 149287).

1   **References**

2   Albach, D.C., 2007. Amplified fragment length polymorphisms and sequence data in the

3   phylogenetic analysis of polyploids: multiple origins of *Veronica cymbalaria*

4   (Plantaginaceae). New Phytol. 176: 481–498. https://doi.org/10.1111/j.1469-

5   8137.2007.02172.x

6   Albach, D.C., Chase, M.W., 2004. Incongruence in Veroniceae (Plantaginaceae): evidence from

7   two plastid and a nuclear ribosomal DNA region. Mol. Phylogenet. Evol. 32: 183–197.

8   https://doi.org/10.1016/j.ympev.2003.12.001

9   Albach, D.C., Meudt, H.M., 2010. Phylogeny of *Veronica* in the Southern and Northern

10   Hemispheres based on plastid, nuclear ribosomal and nuclear low-copy DNA. Mol.

11   Phylogenet. Evol. 54: 457–471. https://doi.org/10.1016/j.ympev.2009.09.030

12   Albach, D.C., Martínez-Ortega, M.M., Delgado, L., Weiss-Schneeweiss, H., Özgökce, F.,

13   Fischer M.A., 2008. Chromosome numbers in Veroniceae (Plantaginaceae): Review and

14   several new counts. Ann. Mo. Bot. Gard. 95: 543–566.

15   http://dx.doi.org/10.3417/2006094

16   Alcántara de la Fuente, M., Goñi, D., Guzmán, D., Puente, J., 2007. *Catálogo de especies*

17   *amenazadas en Aragón. Flora*. Zaragoza: Letras Aragonesas.

18   Andrés-Sánchez, S., Rico, E., Herrero, A., Santos-Vicente, M., Martínez-Ortega, M.M., 2009.

19   Combining traditional morphometrics and molecular markers in cryptic taxa: towards

20   an updated integrative taxonomic treatment for *Veronica* subgenus *Pentasepalae*

21   (Plantaginaceae sensu APG II) in the western Mediterranean. Bot. J. Linn. Soc. 159:

22   68–87. http://dx.doi.org/10.1111/j.1095-8339.2008.00917.x

23   Arrigo, N., Felber, F., Parisod, C., Buerki, S., Alvarez, N., David, J., Guadagnuolo, R., 2010.

24   Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of

wheat. New Phytol.187: 1170–1180. http://dx.doi.org/10.1111/j.1469-
8137.2010.03328.x

Arthofer, W., 2010. tinyFLP and tinyCAT: software for automatic peak selection and scoring of
AFLP data tables. Mol. Ecol. Resour. 10: 385–388. http://dx.doi.org/10.1111/j.1755-
0998.2009.02751.x

Arthofer, W., Schlick-Steiner, B.C., Steiner, F.M., 2011. optiFLP: software for automated
optimization of amplified fragment length polymorphism scoring parameters. Mol. Ecol.
Resour.11: 1113–1118. http://dx.doi.org/10.1111/j.1755-0998.2011.03043.x

Bardy, K.E., Albach, D.C., Schneeweiss, G.M., Fischer, M.A., Schönswetter, P., 2010.
Disentangling phylogeography, polyploid evolution and taxonomy of a woodland herb
(*Veronica chamaedrys* group, Plantaginaceae s.l.) in southeastern Europe. Mol.
Phylogenet. Evol. 57: 771–786. http://dx.doi.org/10.1016/j.ympev.2010.06.025

Bardy, K.E., Schönswetter, P., Schneeweiss, G.M., Fischer, M.A., Albach, D.C., 2011.
Extensive gene flow blurs species boundaries among *Veronica barrelieri*, *V. orchidea*
and *V. spicata* (Plantaginaceae) in southeastern Europe. Taxon, 60: 108–121.

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.
http://dx.doi.org/10.1111/nph.13698

Barley, A.J., White, J., Diesmos, A.C., Brown, R.M., 2013. The challenge of species
delimitation at the extremes: diversification without morhological change in Philippine
sun skinks. Evolution 67: 3556–3572. http://dx.doi.org/10.1111/evo.12219

Beentje, H., 2010. The Kew Plant Glossary: an illustrated dictionary of plant terms. Royal
Botanic Gardens, Kew, Richmond, Surrey.

Bentham, G., 1846. *Veronica*. In: Candolle, A.P. de (Eds.), *Prodromus systematis naturalis regni vegetabilis*, vol. 10. Parisiis [Paris]: sumptibus Victoris Masson, pp. 458–491.

Bilz, M., 2011. *Veronica turrilliana*. The IUCN Red List of Threatened Species 2011: e.T165211A5990780.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., Taberlet, P., 2004. How to track and assess genotyping errors in population genetic studies. Mol. Ecol. 13: 3261–3273. http://dx.doi.org/10.1111/j.1365-294X.2004.02346.x

Bonin, A., Ehrich, D., Manel, S., 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. Mol. Ecol. 16: 3737–3758. http://dx.doi.org/10.1111/j.1365-294X.2007.03435.x

Bornmüller, J., 1937. Zur Flora von Mazedonien. *Repertorium Specierum Novarum Regni Vegetabilis* 42: 126–42.

Buggs, R.J., Wendel, J.F., Doyle, J.J., Soltis, D.E., Soltis, P.S., Coate, J. E. 2014. The legacy of diploid progenitors in allopolyploid gene expression patterns. Phil. Trans. R. Soc. B, 369: 20130354. http://dx.doi.org/10.1098/rstb.2013.0354

Cabezudo, B., Talavera, S., Blanca, G., Salazar, C., Cueto, M., Valdés, B., Hernández Bermejo, J.E., Herrera, C.M., Rodríguez Hiraldo, C., Navas, D., 2005. *Lista Roja de La Flora Vascular de Andalucía*. Sevilla: Consejería de Medio Ambiente de la Junta de Andalucía.

Carstens, B.C., Pelletier, T.A., Reid, N.M., Satler, J.D., 2013. How to fail at species delimitation. Mol. Ecol. 22: 4369–4383. http://dx.doi.org/10.1111/mec.12413

Comai, L., 2005. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6: 836–846. http://dx.doi.org/10.1038/nrg1711

1 Convention on Biological Diversity. Global Taxonomy Initiative. Why is Taxonomy Important?.

2   https://www.cbd.int/gti/importance.shtml (accessed 19 March 2017).

3 Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85: 407–415.

4   http://dx.doi.org/10.1111/j.1095-8312.2005.00503.x

5 de Queiroz, K., 2005. Different species problems and their resolution. BioEssays, 27: 1263–

6   http://dx.doi.org/1269. 0.1002/bies.20325

7 de Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56: 879–886.

8   http://dx.doi.org/10.1080/10635150701701083

9 Doležel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysák, M.A., Nardi, L., Obermayer, R.,

10   1998. Plant genome size estimation by flow cytometry: inter-laboratory

11   comparison. Ann. Bot. 82: 17–26. https://doi.org/10.1093/oxfordjournals.aob.a010312

12 Doyle, J.J., Doyle J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf

13   tissue. Phytochemical Bulletin, Botanical Society of America 19: 11–15.

14 Doyle, J.J., Sherman-Broyles, S., 2017. Double trouble: taxonomy and definitions of

15   polyploidy. New Phytol. 213: 487–493. https://doi.org/10.1111/nph.14276

16 Earl, Dent A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for

17   visualizing STRUCTURE output and implementing the Evanno method. Conserv.

18   Genet. Resour. 4: 359–361. https://doi.org/10.1007/s12686-011-9548-7

19 Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using

20   the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–

21   2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

1   Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software

2   package for population genetics data analysis. Evol. Bioinform. 1: 47–50.

3   Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform

4   population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10: 564–

5   567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

6   Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from

7   metric distances among DNA haplotypes: application to human mitochondrial DNA

8   restriction data. Genetics 131: 479–491.

9   Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure: extensions to

10   linked loci and correlated allele frequencies. Genetics 164: 1567–1587.

11   Falush, D., Stephens, M., Pritchard, J.K., 2007. Inference of population structure using

12   multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes 7: 574–

13   578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

14   Federici, S., Galimberti, A., Bartolucci, F., Bruni, I., De Mattia, F., Cortis, P., Labra, M., 2013.

15   DNA barcoding to analyse taxonomically complex groups in plants: the case of *Thymus*

16   (Lamiaceae). Bot. J. Linn. Soc. 171: 687–699.

17   http://dx.doi.org/10.3366/anh.2005.32.2.403

18   Fischer, M.A., 1982. *Veronica* L. In: Pignatti, S. (Eds.), *Flora d'Italia*, vol 2. Bologna:

19   Edagricole, pp. 558–573.

20   Fowler, N.L., Levin, D.A., 2016. Critical factors in the establishment of allopolyplois. Am. J.

21   Bot. 103: 1–16.

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. http://dx.doi.org/10.1126/science.220.4601.1049

Grant, V., 1971. *Plant speciation. 1st ed*. New York, NY, USA: Columbia University Press.

Greilhuber, J., Ebert, I., 1994. Genome size variation in *Pisum sativum*. Genome 37: 646–655.

    http://dx.doi.org/10.1139/g94-092

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A k-means clustering algorithm. J.R.

    Stat. Soc. Series C (Applied Statistics), 28: 100–108. http://dx.doi.org/10.2307/2346830

Hennig, W., 1950. *Grundzüge einer Theorie der phylogenetischen Systematik*. Berlin: Deutscher

    Zentralverlag.

Himmelreich, S., Breitwieser, I., Oberprieler, C., 2014. Phylogenetic relationships in the

    extreme polyploidy complex of the New Zealand genus *Leptinella* (Compositae:

    Anthemideae) based on AFLP data. Taxon 63: 883–898.

    https://doi.org/10.12705/634.19

Holland, B., Clarke, A., Meudt, H., 2008. Optimizing automated AFLP scoring parameters to

    improve phylogenetic resolution. Syst. Biol. 57: 347–366.

    https://doi.org/10.1080/10635150802044037

Husband, B.C., Baldwin, S.J., Suda, J., 2013. The incidence of polyploidy in natural plant

    populations: major patterns and evolutionary processes. Pp. 255–276 in Leitch, I.J. et al.

    (eds.), Plant Genome Diversity vol. 2. Vienna: Springer.

Huson, D.H., Bryan, D., 2006. Application of phylogenetic networks in evolutionary studies,

    Mol. Biol. Evol. 23: 254–267. https://doi.org/10.1093/molbev/msj030

Huxley, J., 1940. *The new systematics*. Oxford: Oxford University Press.

1 IUCN 2016. The IUCN Red List of Threatened Species. Version 2016-3. www.iucnredlist.org.

2 (accessed 19 March 2017).

3 Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program

4 for dealing with label switching and multimodality in analysis of population structure.

5 Bioinformatics 23: 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

6 Kellog, E.A., 2016. Has the connection between polyploidy and diversification actually been

7 tested? Curr. Opin. Plant Biol. 30: 25–32. https://doi.org/10.1016/j.pbi.2016.01.002

8 Krzanowski, W.J., 1990. *Principles of multivariate analysis*. Oxford: Clarendon Press.

9 Lafon-Placette, C., Vallejo-Marín, M., Parisod, C., Abbot, R.J., Kôhler, C., 2016. Current plant

10 speciation research: unravelling the processes and mechanisms behind the evolution of

11 reproductive isolation barriers. New Phytol. 209: 29–33.

12 https://doi.org/10.1111/nph.13756

13 Laport, R.G., Ng, J., 2017. Out of one, many: the biodiversity considerations of polyploidy. Am.

14 J. Bot. 104: 1119-1121. https://doi.org/10.3732/ajb.1700190

15 Levin, D.A., 2002. *The role of chromosomal change in plant evolution*. New York, USA:

16 Oxford University Press.

17 Lewis, W.H., 1980. Polyploidy in species populations. In: Lewis W.H., (Eds.) Polyploidy:

18 biological relevance. New York, NY, USA: Plenum Press, pp. 103–144.

19 Loureiro, J., Rodriguez, E., Doležel, J., Santos, C., 2007. Two new nuclear isolation buffers for

20 plant DNA flow cytometry: a test with 37 species. Ann. Bot. 100: 875–888.

21 https://doi.org/10.1093/aob/mcm152

22 Lysak, M.A., Dolezel, J., 1998. Estimation of nuclear DNA content in *Sesleria*

23 (Poaceae). Caryologia 51: 123-132. http://dx.doi.org/10.1080/00087114.1998.10589127

1 Mace, G.M., 2004. The role of taxonomy in species conservation. Philos. T. Roy. Soc. B, 359:
2   711–719.

3 Maly, J.K., 1908. *Beiträge zur Kenntnis der illyrischen Flora*. Magyar Bot. Lapok 7: 229–231.

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

8 Martínez-Ortega, M.M, Sánchez-Agudo, J.A., Rico, E., 2009. *Veronica* L. In: Benedí C, Rico E,
9   Güemes J, Herrero A. (Eds.), *Flora Iberica: plantas vasculares de la Península Ibérica*
10   *e Islas Baleares*, Vol. 13. Plantaginaceae-Scrophulariaceae. Madrid: Real Jardín
11   Botánico, CSIC, pp 360–434.

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

15 Mayr, E., 1942. *Systematics and the origin of species, from the viewpoint of a zoologist*.
16   Harvard University Press.

17 Mayrose, I., Zhan, S.H., Rothfels, C.J., Magnuson-Ford, K., Barker, M.S., Rieseberg, L. H.,
18   Otto, S.P., 2011. Recently formed polyploid plants diversify at lower rates. Science 333:
19   1257–1257. http://dx.doi.org/10.1126/science.1207205

20 McDade, L.A., 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on
21   cladistic analysis. Evolution, 46: 1329–1346. http://dx.doi.org/10.1111/j.1558-
22   5646.1992.tb01127.x

Meudt, H.M., 2011. Amplified fragment length polymorphism data reveal a history of auto- and allopolyploidy in New Zealand endemic species of *Plantago* (Plantaginaceae): new perspectives on a taxonomically challenging group. Int. J. Plant Sci. 172: 220–237.

Meudt, H.M., Rojas-Andrés, B.M., Prebble, J.M., Low, E., Garnock-Jones, P.J., Albach, D.C., 2015. Is genome downsizing associated with diversification in polyploid lineages of *Veronica*? Bot. J. Linn. Soc. 178: 243–266. http://dx.doi.org/10.1111/boj.12276

Müntzing, A., 1936. The evolutionary significance of autopolyploidy. Hereditas 21: 363–378. http://dx.doi.org/10.1111/j.1601-5223.1936.tb03204.x

Naciri, Y., Linder, H.P., 2015. Species delimitation and relationships: The dance of the seven veils. Taxon 64: 3–16. http://dx.doi.org/10.12705/641.24

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

Parisod, C., Holderegger, R., Brochmann, C., 2010. Evolutionary consequences of autopolyploidy. New Phytol. 186: 5–17. https://doi.org/10.1111/j.1469-8137.2009.03142.x

Peev, D. 1995. *Veronica* L. In: Kozuharov, S.I. & Kuzmanov, B.A. (Eds.), *Flora Reipublicae Bulgaricae*, vol. 10. Sofia: Editio Academica "Professor Marin Drinov", pp. 142–189.

Peñas de Giles, J., Martínez-Ortega, M.M., Pérez Latorre, A.V., Cabezudo Artero, B., 2004. *Veronica tenuifolia* subsp. *fontqueri* (Pau) M.M. Mart. Ort., E. Rico. In: Bañares, A., Blanca, G., Güemes, J., Moreno, J. C., Ortiz, S. (Eds.), *Atlas y Libro Rojo de La Flora Vascular Amenazada de España*. Madrid: Dirección General para la Biodiversidad, pp. 564–565.

1   Petrova, A., Vladimirov, V., 2009. Red List of Bulgarian Vascular Plants. Phytol. Balc. 15: 63–

2   94.

3   Pompanon, F., Bonin, A., Bellemain, E., Taberlet, P., 2005. Genotyping errors: causes,

4   consequences and solutions. Nat. Rev. Genet. 6: 847–859.

5   http://dx.doi.org/10.1038/nrg1707

6   Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using

7   multilocus genotype data. Genetics 155: 945–959.

8   Ramsey, J., Ramsey, T.S., 2014. Ecological studies of polyploidy in the 100 years following its

9   discovery. Phil. Trans. R. Soc. B. 369: 1648. http://dx.doi.org/10.1098/rstb.2013.0352

10   Ramsey, J., Schemske, W., 1998. Pathways, mechanisms and rates of polyploidy formation in

11   flowering plants. Annu. Rev. Ecol. Syst. 29: 467–501.

12   http://dx.doi.org/10.1146/annurev.ecolsys.29.1.467

13   Reberning, C.A., Weiss-Schneeweiss, H., Blöch, C., Turner, B., Stuessy, T.F., Obermayer, R.,

14   Villaseñor, J.L., Schneeweiss, G.M., 2012. The evolutionary history of the white-rayed

15   species of *Melampodium* (Asteraceae) involved multiple cycles of hybridization and

16   polyploidization. Am. J. Bot. 99: 1043–1057. http://dx.doi.org/10.3732/ajb.1100539

17   Reeves, P.A., Richards, C.M., 2009. Accurate inference of subtle population structure (and

18   other genetic discontinuities) using principal coordinates. PLoS One, 4: e4269.

19   https://doi.org/10.1371/journal.pone.0004269

20   Reeves, P.A., Richards, C.M., 2010. Species delimitation under the general lineage concept: an

21   empirical example using wild North American hops (Cannabaceae: *Humulus*

22   *lupulu*s). Syst. Biol. 60: 45–59. https://doi.org/10.1093/sysbio/syq056

Rieseberg, L.H., Willis, J.H., 2007. Plant Speciation. Science 317: 910–914.

    http://dx.doi.org/10.1126/science.1137729

Rohlf, F.J., 2005. NTSYS-PC 2.2. Exeter Software, Setauket, New York.

    http://www.exetersoftware.com/cat/ntsyspc/ntsyspc.html

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. 173: 670–692.

    http://dx.doi.org/10.1111/boj.12345

Rojas-Andrés, B.M., Martínez-Ortega, M.M., 2016. Taxonomic revision of *Veronica* subsection
Pentasepalae (*Veronica*, Plantaginaceae sensu APG III). Phytotaxa 285: 1–100.

    http://dx.doi.org/10.11646/phytotaxa.285.1.1

Rojas-Andrés, B.M., Rico, E., Martínez-Ortega, M.M., 2016. A nomenclatural treatment for
Veronica subsect. *Pentasepalae* (Plantaginaceae sensu APG III) and typification of
several names. Taxon 65: 617–627. https://doi.org/10.12705/653.14

Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure.
Mol. Ecol. Notes 4: 137–138. http://dx.doi.org/10.1046/j.1471-8286.2003.00566.x

Schemske, D.W., 2000. Understanding the origin of species. Evolution 54: 1069–1073.
http://dx.doi.org/10.1554/0014-3820(2000)054[1069:UTOOS]2.3.CO;2

Soltis D.E., Albert V.A., Leebens-Mack J., Bell C.D., Paterson A.H., Zheng C., Sankoff D.,
dePamphilis C.W., Wall P.K., Soltis P.S., 2009. Polyploidy and angiosperm
diversification. Am. J. Bot. 96: 336–348. http://dx.doi.org/10.3732/ajb.0800079

Soltis, D.E., Buggs, R.J., Doyle, J.J., Soltis, P.S., 2010. What we still don't know about
polyploidy. Taxon 59: 1387–1403.

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.

Soltis D.E., Soltis P.S., Tate J.A., 2004. Advances in the study of polyploidy since plant speciation. New Phytol. 161: 173–191. http://dx.doi.org/10.1046/j.1469-8137.2003.00948.x

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

Soltis, P.S., Marchant, D.B., Van de Peer, Y., Soltis, D.E., 2015. Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev. 35: 119–125. https://doi.org/10.1016/j.pbi.2005.01.001

Stebbins, G.L.Jr., 1947. Types of polyploids: their classification and significance. Adv. Genet. 1: 403–429. https://doi.org/10.1016/S0065-2660(08)60490-3

Stebbins, G.L.Jr., 1950. *Variation and evolution in plants*. New York, USA: Columbia University Press.

Stebbins G.L.Jr., 1971. Chromosomal evolution in higher plants. London, UK: Edward Arnold.

Swofford, D.L., 2003. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sunderland, MA: Sinauer Associates.

Temsch, E.M., Greilhuber, J., Krisai, R., 2010. Genome size in liverworts. Preslia, 82: 63-80.

Thiers, B., (2017, continuously updated). Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/ih/. (last accessed 15 January 2017).

1   Velenovsky, J., 1893. Dritter Nachtrag zur Flora von Bulgarien. Sitzungsber. Königl. Böhm.

2   Ges. Wiss. Mathe.-Naturwisse. Cl. 37: 1–72

3   Velenovsky, J., 1898. *Flora bulgarica. Descriptio et enumeratio systematica plantarum*

4   *vascularium in principatu Bulgariae sponte nascentium*. Supplementum I. Pragae: F.

5   Rivnác.

6   Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J.,

7   Peleman, J., Kuiper, M., Zabeau, M., 1995. AFLP: a new technique for DNA

8   fingerprinting. Nucleic acids Res. 23: 4407–4414.

9   https://doi.org/10.1093/nar/23.21.4407

10 Watzl, B., 1910. *Veronica prostrata* L., *teucrium* L. und *austriaca* L. Nebst einem Anhang über

11   deren nächste Verwandte. *Abh. K. K. Zool.-Bot. Ges. Wien* 5(5): 1–94.

12   http://bibdigital.rjb.csic.es/ing/Libro.php?Libro=5939

13 Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., Rieseberg, L.H.,

14   2009. The frequency of polyploid speciation in vascular plants. PNAS, 106: 13875-

15   13879. https://doi.org/10.1073/pnas.0811575106

16 Zozomová-Lihová, J., Marhold, K., Spaniel, S., 2014. Taxonomy and evolutionary history of

17   *Alyssum montanum* (Brassicaceae) and related taxa in southwestern Europe and

18   Morocco: diversification driven by polyploidy, geographic and ecological isolation.

19   Taxon 63: 562–591. https://doi.org/10.12705/633.18

Neighbor-Net Groups	Taxa	Ploidy	BS values		Clustering			
			NJ Jaccard	NJ Nei-Li	K-means K=2	Structure K=3	Structure K=4	PCO-MC
GROUP I	*V. kindlii* Adamović	2*x*	100.0	100.0	1/2	A	A	✓
	V. linearis (Bornm.) Rojas-Andrés & M.M. Mart. Ort.	2*x*	100.0	100.0	1	A	A	✓
	V. orsiniana Ten.	2*x*	100.0	100.0	2	A	C	✓
	* *V. rhodopea* (Velen.) Degen. ex Stoj. & Stef.	2*x*	100.0	100.0	2	A	A	—
	V. teucrioides Boiss. & Heldr.	2*x*	99.9	100.0	2	A	A	—
	V. affinis kindlii	4*x*	100.0	100.0	1	A	A	—
GROUP II	*V. aragonensis* Stroh	4*x*	100.0	100.0	1	B	B	✓
	V. rosea Desf.	2*x*, 4*x*	100.0	100.0	2	B	B	✓
	V. tenuifolia subsp. *fontqueri* (Pau) M.M. Mart. Ort. & E. Rico	2*x*	100.0	100.0	2	B	B	—
	V. tenuifolia subsp. *javalambrensis* (Pau) Molero & J. Pujadas	2*x*	100.0	99.0	2	B	B	—
	V. tenuifolia Asso subsp. *tenuifolia*	2*x*	100.0	100.0	2	B	B	—
GROUP III	*V. dalmatica* N. Pad. Gar., Rojas-Andrés, López-González & M.M. Mart. Ort.	2*x*	99.8	100.0	1	A	C	✓
	V. austriaca subsp. *jacquinii* / *V. orbiculata*	4*x*	84.3	83.0	1	A	C	—
	V. crinita Kit	2*x*	100.0	100.0	1	A	C	—
	V. crinita (= *V. crinita* f. *bosniaca*)	2*x*	62.9	61.0	1	A	C	—
	V. thracica Velen.	2*x*	100.0	100.0	1	A	C	—
	V. krylovii Schischk.	2*x*	100.0	100.0	1	A	C	—
	V. orbiculata A. Kern	2*x*, 4*x*	73.0	78.0	2	A	C	✓
	V. prostrata L.	2*x*	100.0	100.0	2	A	C	✓
	V. turrilliana Stoj. & Stef.	2*x*	100.0	100.0	2	A	C	✓
	V. affinis kindlii	2*x*, 4*x*	Not monophyletic	Not monophyletic	1	A	C	—
GROUP IV	*V. austriaca* subsp. *austriaca* L.	6*x*	Not monophyletic	Not monophyletic	1	C	D	—
	* *V. austriaca* subsp. *austriaca* (= *V. macrodonta*)	4*x*	94.4	93.0	1	C	D	—
	V. austriaca subsp. *dentata* (F.W. Schmidt) Watzl	6*x*	Not monophyletic	Not monophyletic	1	C	D	—
	V. austriaca subsp. *jacquinii* (Baumg.) Watzl	6*x*	59.6	56.0	1	C	D	—
	V. satureiifolia Poit. & Turpin	4*x*	63.9	62.0	2	C	D	—
	V. sennenii (Pau) M.M. Mart. Ort. & E. Rico	8*x*	Not monophyletic	Not monophyletic	2	C	D	—
	V. angustifolia Vahl (Bernh.)	8*x*	Not monophyletic	Not monophyletic	2	C	D	—
	V. teucrium L.	8*x*	Not monophyletic	Not monophyletic	2	C	D	—

1 *Only one population sampled for this study

2 **Table 1.** Overview of the boostrap (BS) values detected by different phylogenetic analyses supporting each taxon included in the study. Taxa are shown according to the four groups identified by the placement of taxa

3 in the Neighbor-Net network. Final taxonomic assignments, DNA ploidy level and classification of individuals in clusters using different methodologies are indicated. Species recovered as independent clusters in PCO-

4 MC analyses are indicated by a checkmark (✓).

Clustering approach	K value	Source of variation	Sum of squares	Variance components	Percentage of variation	Statistic
K-means model	**K = 2**	Among clusters	569.761	3.18	**3.10**	Fct= 0.031
		Among populations	16,272.464	52.08	50.75	Fsc= 0.524
		Within populations	7,198.333	47.36	46.15	Fst= 0.538
NeighborNet	**K = 4**	Among clusters	3,310.346	15.78	**15.03**	Fct= 0.150
		Among populations	13,873.399	42.14	40.12	Fsc= 0.472
		Within populations	6,971.500	47.10	44.85	Fst= 0.551
Structure algorithm	**K = 3**	Among clusters	2,589.107	16.12	**14.96**	Fct= 0.150
		Among populations	14,594.638	44.50	41.31	Fsc= 0.486
		Within populations	6,971.500	47.10	43.73	Fst= 0.563
Structure algorithm	**K = 4**	Among clusters	3,398.245	16.60	**15.74**	Fct= 0.157
		Among populations	13,785.499	41.75	39.59	Fsc= 0.470
		Within populations	6,971.500	47.10	44.67	Fst= 0.553
Structure algorithm	**K = 20**	Among clusters	9,018.235	36.00	**34.74**	Fct= 0.347
		Among populations	6,507.650	21.60	20.85	Fsc= 0.319
		Within populations	5,889.167	46. 01	44.41	Fst= 0.556

1

2 **Table 2.** Analysis of molecular variance (AMOVA) performed with different grouping approaches. Percentage of variation explained by different methodological groupings

3 (K-means model, Structure algorithm and NeighborNet) are shown.

4