



# Comparative Phylogeography of *Veronica spicata* and *V. longifolia* (Plantaginaceae) Across Europe: Integrating Hybridization and Polyploidy in Phylogeography

Daniele Buono<sup>1</sup>, Gulzar Khan<sup>1</sup>, Klaus Bernhard von Hagen<sup>1</sup>, Petr A. Kosachev<sup>2</sup>, Eike Mayland-Quellhorst<sup>1</sup>, Sergei L. Mosyakin<sup>3</sup> and Dirk C. Albach<sup>1\*</sup>

<sup>1</sup> Institute for Biology and Environmental Sciences, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany,

<sup>2</sup> Faculty of Biology, Altai State University, Barnaul, Russia, <sup>3</sup> M.G. Kholodny Institute of Botany, National Academy of Sciences of Ukraine, Kyiv, Ukraine

## OPEN ACCESS

### Edited by:

Karol Marhold,  
Institute of Botany, Slovak Academy  
of Sciences, Slovakia

### Reviewed by:

Gonzalo Nieto Feliner,  
Real Jardín Botánico (RJB, CSIC),  
Spain

Ines Carrasquer,  
Natural History Museum of Geneva,  
Switzerland

### \*Correspondence:

Dirk C. Albach  
dirk.albach@uni-oldenburg.de

### Specialty section:

This article was submitted to  
Plant Systematics and Evolution,  
a section of the journal  
Frontiers in Plant Science

**Received:** 28 July 2020

**Accepted:** 19 November 2020

**Published:** 01 February 2021

### Citation:

Buono D, Khan G, von Hagen KB,  
Kosachev PA, Mayland-Quellhorst E,  
Mosyakin SL and Albach DC (2021)  
Comparative Phylogeography  
of *Veronica spicata* and *V. longifolia*  
(Plantaginaceae) Across Europe:  
Integrating Hybridization  
and Polyploidy in Phylogeography.  
*Front. Plant Sci.* 11:588354.  
doi: 10.3389/fpls.2020.588354

Climatic fluctuations in the Pleistocene caused glacial expansion-contraction cycles in Eurasia and other parts of the world. Consequences of these cycles, such as population expansion and subsequent subdivision, have been studied in many taxa at intraspecific population level across much of the Northern Hemisphere. However, the consequences for the potential of hybridization and polyploidization are poorly understood. Here, we investigated the phylogeographic structure of two widespread, closely related species, *Veronica spicata* and *Veronica longifolia*, across their European distribution ranges. We assessed the extent and the geographic pattern of polyploidization in both species and hybridization between them. We used genome-scale SNP data to clarify phylogenetic relationships and detect possible hybridization/introgression events. In addition, crossing experiments were performed in different combination between *V. spicata* and *V. longifolia* individuals of two ploidy levels and of different geographic origins. Finally, we employed ecological niche modeling to infer macroclimatic differences between both species and both ploidy levels. We found a clear genetic structure reflecting the geographical distribution patterns in both species, with *V. spicata* showing higher genetic differentiation than *V. longifolia*. We retrieved significant signals of hybridization and introgression in natural populations from the genetic data and corroborated this with crossing experiments. However, there were no clear phylogeographic patterns and unequivocal macroclimatic niche differences between diploid and tetraploid lineages. This favors the hypothesis, that autopolyploidization has happened frequently and in different regions. The crossing experiments produced viable hybrids when the crosses were made between plants of the same ploidy levels but not in the interploidy crosses. The results suggest that hybridization occurs across the overlapping areas of natural distribution ranges of both species, with apparently directional introgression from *V. spicata* to *V. longifolia*. Nevertheless, the two species maintain their species-level separation due to their adaptation to different habitats and spatial isolation rather than reproductive isolation.

**Keywords:** hybridization, polyploidy, comparative phylogeography, steppe, Europe, macroclimatic niche

## INTRODUCTION

The past 30 years have seen an enormous increase in knowledge of how species migrated and evolved, thanks to the analyses of intraspecific genetic variation and progress in the field of phylogeography (Avice, 2000, 2009). Pleistocene refugia and migration corridors have been inferred (e.g., Hewitt, 2000; Harrison and Noss, 2017) and phylogeographic patterns among species have been compared to infer commonalities in dispersal history and differences, indicative of changes in community composition over time (e.g., Brunsfeld et al., 2001; Muellner-Riehl, 2019). Based on such information, hypotheses about the effect of climate change on biodiversity decline have been established (Keppel et al., 2012; Carvalho et al., 2019). A special situation occurs when species interact, i.e., they hybridize or form polyploids. In most cases, such studies examined the origin of hybrids or polyploids on a narrow spatial scale (e.g., Schönswetter et al., 2007; Balao et al., 2015). Undetected hybrids or structure caused by intraspecific lineages with different ploidy levels can potentially mislead analyses to infer genetically distinct groups or can inflate genetic diversity of particular groups of populations. Further, hybridization and polyploidy have potential large effects on a variety of processes shaping the distribution range of a species (Nieto Feliner, 2014; Arrigo et al., 2016; Paule et al., 2017), for example allowing species to exploit new ecological niches. Likewise, the advantages of polyploids over diploid progenitors are the increased number of alleles to mask the deleterious recessive mutations and lead to heterosis in allopolyploids (Gu et al., 2003), and neo-functionalization or sub-functionalization of duplicated genes copies. These factors are considered important for ecological niche expansion and provision of high flexibility toward environmental challenges (Adams and Wendel, 2005; Birchler et al., 2010), facilitating colonization of new and extreme habitats. In consequence, populations of a species affected by hybridization might have survived in places where the non-hybrid populations of that species normally would not be able to occur (Kadereit, 2015). These factors may be especially prevalent in phylogeographic analyses, in which both diploids and polyploids are widespread and sympatric. Here, we compared phylogeographic patterns of two species able to hybridize and assessed how hybridization and polyploidization might have affected these patterns considering past climatic changes. We have studied two species of the genus *Veronica* L., *Veronica spicata* L., and *Veronica longifolia* L. from *Veronica* subgenus *Pseudolysimachium* (W.D.J.Koch) Buchenau, both are widely sympatric across Eurasia and known to have two different ploidy levels, the polyploids (in our case, tetraploids) being derived by autopolyploidy (Bardy et al., 2011; Albach, unpubl.). The two species, however, differ in ecology. *Veronica spicata* occurs on nutrient-poor, open and sunny dry meadows, dunes, or rocks. It is considered to have been a member of the late-glacial steppe-tundra-community and occurs nowadays in the remnants of this vegetation (Pigott and Walters, 1954). It is highly intolerant of shade and competition (Pigott and Walters, 1954), which makes it endangered in large parts of its western distribution area. In contrast, *V. longifolia* is a species inhabiting nutrient-rich, mesic

to moist habitats along rivers, lakes or even ditches or in swamps (Winter et al., 2008).

*Veronica* subgenus *Pseudolysimachium* includes ~30 species distributed across Eurasia, with several of them used in horticulture, especially cultivars of *V. spicata* and *V. longifolia* (Albach et al., 2004; Kosachev et al., 2019). These are also the most common and widespread species of the group. The Altai region, with 13 reported species, is considered the center of its diversity; also, seven hybrids have been described there based on morphology (Kosachev et al., 2015 and 2016). The northern Balkan Peninsula is the second center of diversity of the group, with six species (Albach and Fischer, 2003). The origin of the subgenus has been estimated to have occurred 12.5 Mya ( $\pm 4$  My), likely caused by a polyploidization event (Meudt et al., 2015), thus what we consider diploids in this study are in fact ancient tetraploids. Nevertheless, more recent polyploidization events have occurred within the last 3 My (confidence interval 4–0.5 Mya) within the subgenus (Meudt et al., 2015). Besides polyploidization, introgressive hybridization events between more or less diverged lineages in the subgenus may have occurred. Hybridization has been reported repeatedly in the subgenus. For example, Kosachev et al. (2016), studied 18 species of *V. subg. Pseudolysimachium*, reported an unresolved phylogenetic structure based on cpDNA and ITS1, with most species being polyphyletic, and multiple samples of *V. spicata* and *V. longifolia* showing signs of introgression from the other species. Nevertheless, the ITS1 region at least resolved alleles into six groups with most alleles from *V. longifolia* in one group and those of *V. spicata* in another. Kosachev et al. (2016) also suggested an East Asian origin followed by repeated westwards range expansions across the Pleistocene Eurasian steppes with frequent interspecific hybridization events. Gene flow between *Veronica* species (*V. spicata*, *V. orchidea* Crantz, and *V. barrelieri* H.Schott) has been shown to be common in the Balkan Peninsula, with genetic isolation restricted to few geographic regions (Bardy et al., 2011). Potential hybridization between species in the subgenus has also been detected in the Altai Mountains using SRAP (Sequence-related amplified polymorphism) markers, albeit with weak support (Kosachev et al., 2019). Bardy et al. (2011) suggested that complete homogenization of the identified groups is prevented by ecological divergence, ploidy differences, and geographic isolation. While not being responsible for speciation, hybridization may have caused the present wide morphological variation by generating several variants, which are morphologically intermediate between the recognized species.

Hybridization in the subgenus, though not specifically between *V. spicata* and *V. longifolia*, has been reported in a number of places across Eurasia, such as in the Balkan Mountains, Eastern Alps, Southern Carpathians (Bardy et al., 2011), Altai Mountains (Kosachev et al., 2019; Khan et al., submitted), and possibly other regions (Kosachev et al., 2016). However, it is still unclear whether hybridization occurred more widely across the Eurasian steppes and between *V. spicata* and *V. longifolia* or not. The intermediate plants between both species inferred to be hybrids have taxonomically been called *V. × media* Schrad. Based on crossing experiments between several species of the subgenus, Härle (1932) and Graze (1933)

reported that hybridization between plants with the same ploidy level is possible and the fertile hybrids are able to backcross with their parents. Both species, *V. spicata* and *V. longifolia*, have diploid ( $2n = 34$ ) and tetraploid ( $2n = 68$ ) populations (Albach et al., 2008). The two different ploidy levels found in *V. spicata* and *V. longifolia* ( $2\times$  and  $4\times$ ), however, do not seem to be related to differences in either geographical distribution, ecological requirements, or morphology (Trávníček et al., 2004). However, in earlier publications, Raitanen (1967) reported that morphologically determined samples of *V. spicata* and *V. longifolia* in Fennoscandia seem to hybridize despite differences in ploidy levels (with *V. spicata* occurring in the region only as  $4\times$  and *V. longifolia* mainly as  $2\times$ ). According to Härle (1932), *V. spicata* and *V. longifolia* are self-incompatible, but Wilson et al. (2000) reported *V. spicata* as self-compatible. Similarly, Scalone et al. (2013) suggested that both species are facultatively outcrossing. Pollination in the subgenus is mostly performed by insects (entomophily) and is not species-specific (i.e., performed by non-specialized and mainly generalist pollinators), providing the possibility of hybridization in natural habitats. Bees (Hymenoptera: Apoidea, mainly small- to medium sized species) seem to be the most frequent visitors of the species, but Lepidoptera (Müller, 1881; Knuth, 1909; Ellis and Ellis-Adam, 1994; Albach and Kosachev, unpublished), flies (Diptera: small-sized Syrphidae and various other groups) and small-sized beetles (Coleoptera: Mordellidae, Nitidulidae, and some other non-specialized anthophilous groups) are also observed as occasional flower visitors (Kampny, 1995; Mosyakin, unpublished). It looks like the guilds of pollinators of *V. spicata* and *V. longifolia* are recruited from among the local faunas of non-specialist (both nectar-collecting and pollen-eating) anthophilous insects and thus may differ considerably in various regions. At the same time, such pollinator flexibility allows opportunities for cross-pollination within and between the two plant species, when they occur in close proximity. In addition, both species have rhizomes and can reproduce vegetatively by underground rhizome systems (Wilson et al., 2000), a trait considered important for the establishment and local-scale spread of hybrid and polyploid populations (Herben et al., 2017).

Here, we investigated the phylogeographic structure of *V. spicata* and *V. longifolia* species in parallel and estimated the extent of hybridization between both species. We employed genome wide SNPs (single nucleotide polymorphism) to analyze genetic diversity and hybridization/introgression in both *V. spicata* and *V. longifolia*. Furthermore, we conducted crossing experiments to test for the possibility of generating viable offspring from crosses between *V. spicata* and *V. longifolia* individuals of the same as well as different ploidy levels and geographic origins. Finally, we assessed the geographic distribution of ploidy levels in both species. We based our initial hypotheses on the works of Härle (1932) and Graze (1933), and the observation in the wild by Borsos (1967) and Trávníček et al. (2004). Our initial hypotheses were, first, that hybridization between *V. longifolia* and *V. spicata* is restricted to individuals having the same ploidy levels [e.g., *V. longifolia* ( $2\times$ )  $\times$  *V. spicata* ( $2\times$ ) and *V. longifolia* ( $4\times$ )  $\times$  *V. spicata* ( $4\times$ )], while different ploidy levels [*V. longifolia* ( $2\times$ )  $\times$  *V. spicata* ( $4\times$ )

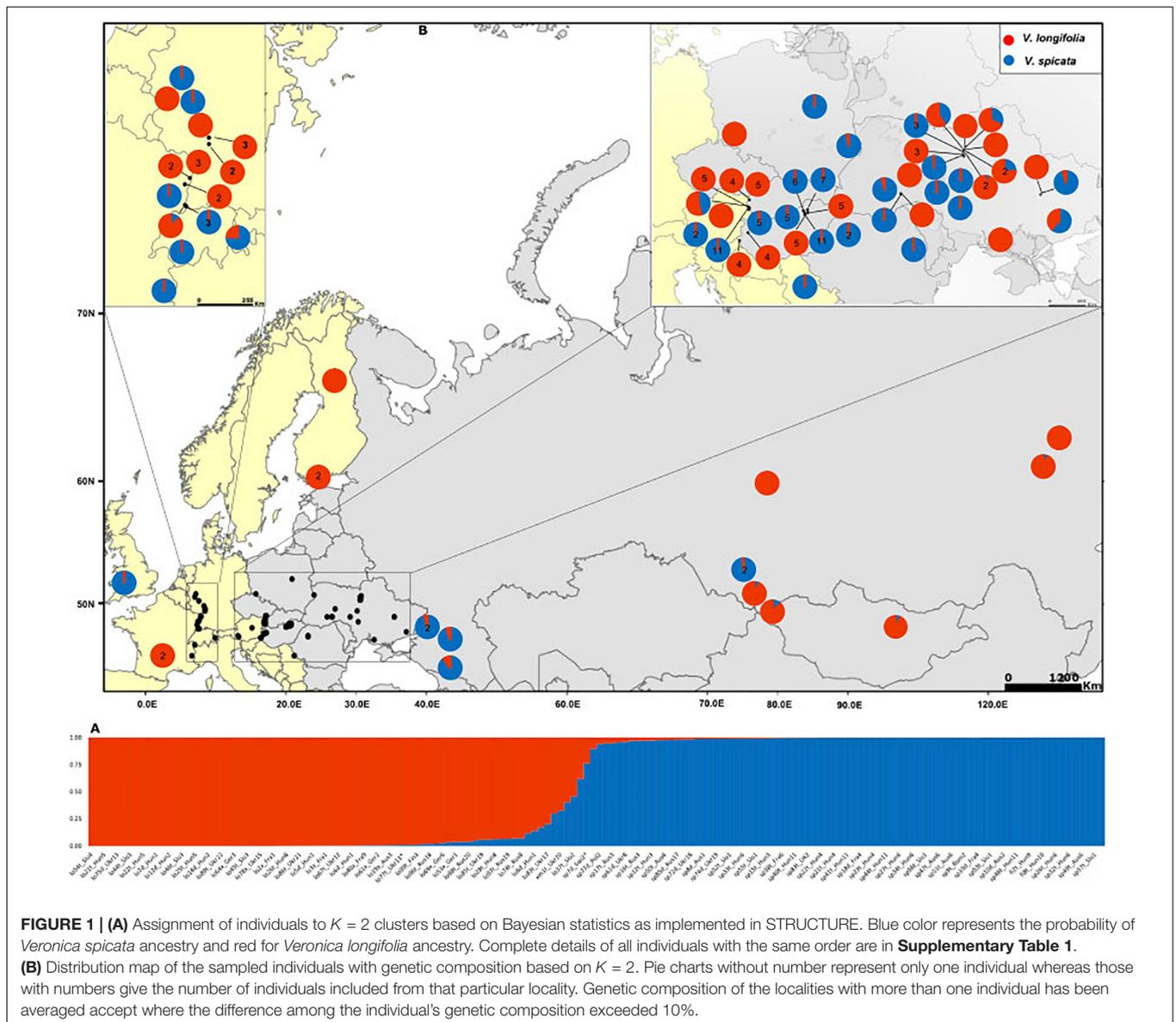
and *V. longifolia* ( $4\times$ )  $\times$  *V. spicata* ( $2\times$ )] represent an important crossing barrier. Second, we also expected a significant genetic intraspecific structure across the native distribution ranges of the species reflecting ecological and spatial isolation in the wild. Lastly, we hypothesized that ecological niches of the two species did not differ based on the wide occurrence but considered it likely that macroclimatic differentiation between diploid and tetraploid individuals at both intra- and interspecific levels occurred in both species.

## MATERIALS AND METHODS

### Sampling

We collected 177 accessions: 81 of *V. spicata*, 74 of *V. longifolia*, one putative hybrid (*V.  $\times$  media*), and *Veronica schmidtiana* as an outgroup for the molecular analyses. *V. schmidtiana* has been used as outgroup since it is the only species unequivocally positioned within the subgenus and sister to the clade containing both ingroup species (Figure 1 and Supplementary Table 1). The accessions were collected on various field excursions during 2000–2019 in different areas across the distribution range of the species (Figure 1B) and identified using the morphological key provided in Albach and Fischer (2003). During the excursion, fresh leaf samples were collected and dried in silica gel (Voucher information; Supplementary Table 1). For crossing experiments, living plants of various accessions have been cultivated in the greenhouse of the Botanical Garden of Carl von Ossietzky-University (see details below). Vouchers for all crosses are deposited in the herbarium of Carl von Ossietzky University of Oldenburg (OLD).

For the analysis of ecological niches, we generated a dataset of 404 specimens with known ploidy levels based on chromosome counting or flow cytometry and sufficiently precise geographical coordinates (about one arc minute resolution or better; Figure 2 and Supplementary Table 2) from literature data and our own estimations. New records of ploidy level directly measured by us constitute 38% of the data. In the first step, we excluded data points with less than 25 km linear distance from another point in each group (Table 1) using the geosphere package (Hijmans et al., 2019) for R (R Core Team, 2019). This reduces the chance that individuals might stem from the same (meta-) population and it is also convenient for statistical purposes when there is only one data point per grid square or less at the same time. This step also reduces a bias in geographical sampling, which can lead to distorted results (Fourcade et al., 2014). After thinning, we used a total of 248 data points. The threshold of using 25 km linear distance (and not an even greater distance) for exclusion was based on the assumption that a lower sample number is negatively affecting the statistical confidence. Therefore, we stuck to our 248 data set. Following our genetic analysis about 4% on average of all existing individuals in both species (or data points in the macroclimatic analysis) might have been compromised by introgression. It will become evident further below that the variation within each group and the overlap between groups was that high that these potential 4% could not have perturbed the results in any way. Plants for which potential



hybrid influence was detected in our genetic analyses were not part of the macroclimatic analyses except for two individuals. One of these individuals was included in tetraploid *V. longifolia* and one in tetraploid *V. spicata* based on their typical morphology and the majority of the genetic markers found. Excluding both would not make any difference because their climatic parameters were typical for their respective groups.

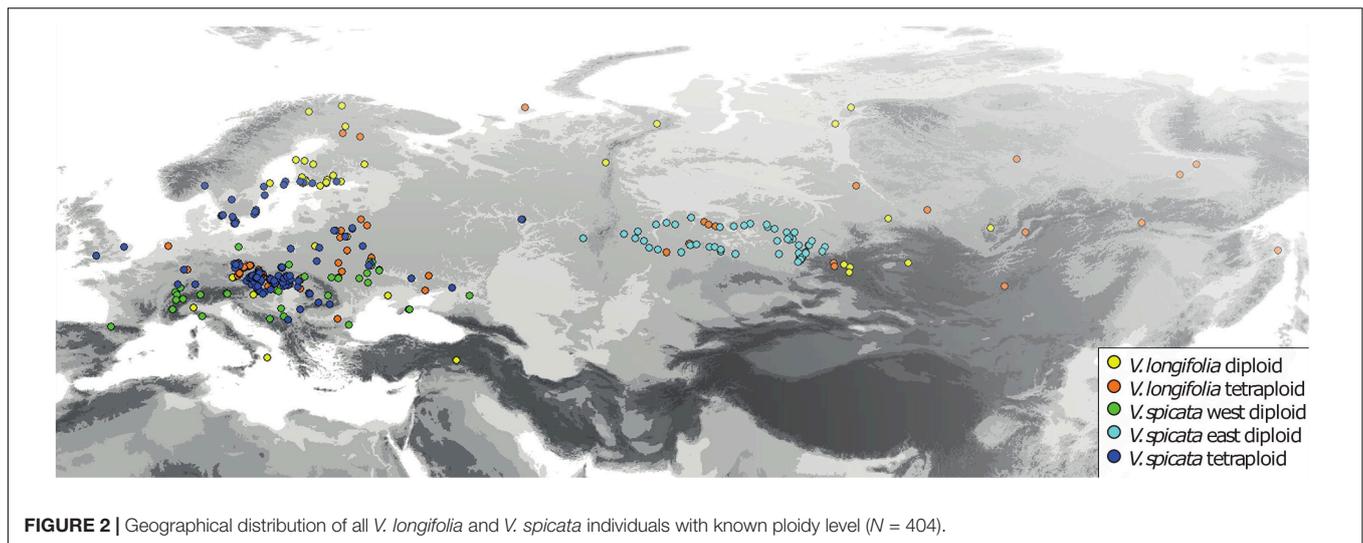
## Flow Cytometry and DNA Extraction

Ploidy information of all samples was assessed by flow cytometry following Meudt et al. (2015) on a CyFlow SL (Partec, Germany). Overall, we report here 156 new measurements of ploidy for the two species. Since intra-population ploidy variation occurs rarely in *V. subg. Pseudolysimachium* (~1.8% of the cases; Bardy et al., 2011), we measured for the majority of populations only one sample after confirming the finding

of Bardy et al. (2011) for 12 populations in which we measured three individuals (**Supplementary Table 2**). DNA was extracted from about 20 mg of dried leaf material using innuPREP Plant DNA Kit (Analytic Jena AG, Jena, Germany) following the manufacturer's instructions. The quality of DNA was checked using gel electrophoresis (1% agarose gel). Photometric DNA concentration and A260/A280 ratio was measured using TECAN infinite F200 Pro (Tecan Group AG, Männedorf, Switzerland).

## Genotyping by Sequencing

Several datasets from different years and sequencing runs were combined for the analysis in this study. Combinability of runs was checked in preliminary analyses including the same individual sequenced in different runs. All GBS (genotype by sequencing) libraries were prepared following Elshire et al.



**FIGURE 2** | Geographical distribution of all *V. longifolia* and *V. spicata* individuals with known ploidy level ( $N = 404$ ).

**TABLE 1** | Number of points for which georeferenced data and ploidy level was known.

<i>Veronica</i> groups	All available	After thinning	Subsets
Diploid <i>V. longifolia</i>	48	40	
Tetraploid <i>V. longifolia</i>	102	65	
Diploid <i>V. spicata</i>	118	74	33 west, 41 east
Tetraploid <i>V. spicata</i>	136	69	
Total	404	248	

In the second column only the points with more than 25 km linear distance from each other were included. Only diploid *V. spicata* was divided in two subsets for some analyses.

(2011), using the methylation-insensitive restriction enzyme *MspI* (5'...CAYNN<sup>+</sup>NNRTG...3') with modifications as described by Siadjeu et al. (2018). Briefly, 200 ng genomic DNA were included in restriction-digestion with one unit *MspI* (New England Biolabs, NEB) in 30  $\mu$ l (1 $\times$  NEB4 buffer) at 37°C for 1 h and heat inactivated at 80°C for 20 min. Adaptor ligation was conducted with 15  $\mu$ l of restriction product and 3  $\mu$ l ligation adapters (Ovation Rapid DR Multiplex System, Nugen Technologies, Leek, Netherlands) in 12  $\mu$ l of master mix (4.6  $\mu$ l D1 water, 6  $\mu$ l L1 ligation buffer mix, and 1.5  $\mu$ l L3 ligation enzyme) at 25°C for 15 min and heat inactivated at 65°C for 10 min. The ligation product was finalized in 20  $\mu$ l of the kits 'final repair' master mix at 72°C for 3 min which was then purified with magnetic beads (homemade Serapure beads). The purified ligated products were amplified with a low number of PCR-cycles (10 cycles) using 4  $\mu$ l of 5 $\times$  MyTaq buffer (Bioline), 0.2  $\mu$ l polymerase and 1  $\mu$ l (10 pmol/ $\mu$ l) of standard Illumina TrueSeq amplification primers. To remove small fragments from the libraries, all amplicons were purified again with magnetic beads. In this way, libraries were prepared for all individuals, normalized, and pooled. An additional step of purification was performed to remove the PCR polymerase through Qiagen MinElute Columns (Hilden, Germany) from the pooled library. The genomic library was sent to LGC genomics (Berlin, Germany) for sequencing.

## RAD-Seq *de novo* Assembly Into Loci and Genotyping of SNPs

Sequencing quality was checked using FASTQC version 0.11.9 (Andrews, 2010) and all the results in HTML format were aggregated in MULTIQC version 1.9 (Ewels et al., 2016). At the end, we included the reads with mean Phred quality score >28, without any adapter contamination and a per base *N* content close to 0. In addition, to ensure and avoid any type of false positive/negative errors in SNP calling, we employed process\_radtags step from the STACKS version 2.4 (Catchen et al., 2013). The settings for process\_radtags were: trimming after 135 bp; remove all uncalled bases with the option -c and low quality scores with the option -q. After confirming upon the quality of the sequencing, we used STACKS version 2.4 (Rochette et al., 2019) pipeline to assemble the raw reads into *de novo* loci and to genotype the SNPs. The pipeline consists of the sequential execution of different tools as (I) USTACKS, (II) CSTACKS, (III) SSTACKS, (IV) TSV2BAM, (V) GSTACKS, and (VI) POPULATIONS. USTACKS was used to construct a set of putative loci and detect SNP at each locus following: maximum distance allowed between stacks = 4; minimum depth of coverage to create a stack = 5; maximum number of stacks at a single *de novo* locus = 5. Some of these settings were dictated by having tetraploid individuals in the dataset. Subsequently, the tool CSTACKS was executed to build a catalog (a set of consensus loci) from the set of processed samples, allowing 4 mis-matches between sampled loci when building the catalog. SSTACKS was then used to search a set of stacks against the catalog produced. We used TSV2BAM to transpose the R2 reads, by orienting it to their corresponding loci instead by sample associated with each single-end locus assembled *de novo* in the previous steps. Following TSV2BAM, we assembled the paired-end reads into a contig, merging the contig with the single-end locus, and aligned reads from individual samples to the locus within GSTACKS. Lastly, the tool POPULATIONS from STACKS was used to filter the data and produce VCF (Variant Call Format, ver. 4.2) files. The minimum percentage of individuals across

populations required to process a locus (option  $-R$ ) was selected as 0.5 after inspection of several phylogenetic trees obtained with various values of  $-R$  option (0.9, 0.75, 0.5; results not shown) since this setting has important effect on the maximum missing data (for example with  $R = 0.5$ , 50% of missing data). At the end, we generated two different VCF files: one with more than one SNP per locus (concatenated), and one with a single (the first) SNP per locus (using the populations option  $-write-single-snp$ ), assuming that loci are independent (e.g., unlinked). In the downstream population genomics, we only used the biallelic SNPs assumed to be unlinked to ensure that allele frequencies are not correlated. We included the post process filtering step to remove the variants with low ( $<8$ ) and high coverages ( $>800$ ). This step is necessary as the variants sequenced at very low coverage may represent only one of four alleles in tetraploids and the one with high coverage may result from repetitive regions. Subsequently, we removed the variants and individuals with more than 50% missing data. This post-process step was performed using the R package *VCFR* (Knaus and Grünwald, 2017) and self-written R script (**Supplementary Material 1**). The tool *PARALOG-FINDER* implementing the method of McKinney et al. (2017) was applied to remove paralogs in our dataset before the downstream analyses, however, no appreciable differences in the results were found and we kept the original datasets. Our final dataset for population genomics included 27% missing data for concatenated loci (290114 SNPs; 153 individuals) and 25.73% missing data on unlinked loci (8723 SNPs; 157 individuals).

The two datasets contain different numbers of individuals as the post process filtering step removed samples using the same limit of missing information admitted, which differs slightly between samples from unlinked and concatenated dataset (since on unlinked dataset missing data is related with only the first SNP per locus; **Supplementary Table 1**).

## Phylogeographic Structure and Admixture

To assess the genetic structure involving all individuals of both species sampled across the sampling area, we firstly adopted a non-hierarchical approach in STRUCTURE version 2.3 (Pritchard et al., 2000) using the unlinked dataset. STRUCTURE groups the individuals originating from different genetic clusters and identifies migrants. It is also useful in studies involving hybridization using Bayesian statistics by applying Markov Chain Monte Carlo (MCMC) and multiple repetition (Pritchard et al., 2000; Porras-Hurtado et al., 2013). We included all individuals of *V. longifolia* and *V. spicata* in STRUCTURE analysis to assign them to their respective genetic clusters, find the role of geography in their structuring and investigate admixed individuals resulted from the contribution of alleles from both species. The analysis was run for 100000 MCMC (Markov chain Monte Carlo) burning the initial 10% chains, with the admixture model and correlated allele frequencies between populations, from K one to four each with 20 iterations, considering that migration is still occurring. This analysis included morphologically defined *V. spicata* (81 individuals), *V. longifolia* (74 individuals) and *V. × media*

(one individual). To investigate the role of geography in sub-structuring of both species, we ran an additional second-level hierarchical analysis as well including only samples of *V. spicata* and *V. longifolia* with the same parameters except  $K = 13$ . We used the program CLUMPP version 1.1.2 (Jakobsson and Rosenberg, 2007) to clump and align all the 20 iteration for the optimal  $K$  value following the Evanno approach as employed in STRUCTURE HARVESTER version 0.6.94 (Evanno et al., 2005; Earl and von Holdt, 2012) selecting the Full-Search option. To further explore the role of geography, we used TESS3R package in R (Caye et al., 2016; **Supplementary Material 2**) as well. TESS3R determines genetic variation in natural populations considering both genetic and geographic data simultaneously. We used the output results of the second level hierarchical analyses on *V. spicata* and *V. longifolia* from STRUCTURE (results not shown) and combined them with geographic coordinates of all the included individuals.

To check the relationship among the individuals from different geographical localities and complement the STRUCTURE/TESS3R results, we further constructed the maximum likelihood trees as well using the concatenated SNPs data. For this analysis, we used two different datasets, one including all individuals and another one excluding admixed individuals (based on STRUCTURE results). We considered the individuals admixed which had admixture proportion more than 20%. The best substitution model of evolution was assessed with JMODELTEST-NG (Darriba et al., 2019), which provided GTR + I as best model with both BIC and AIC criteria. The actual analysis was then carried out in RAxML version 8.2.12 (Stamatakis, 2014). To inform RAxML about the nature of the dataset containing only concatenated SNPs without invariable sites, we applied the ASC correction as suggested (Stamatakis, 2014) to minimize the impact on branch lengths. To check the robustness of our results, we ran the analysis with 100 bootstrap replicates. In addition to phylogenetic reconstruction, pairwise  $F_{ST}$  statistics were also calculated (Wright, 1949) following Weir and Cockerham (1984). This analysis was implemented using the function *stamppFst* of the R package STAMPP (Pembleton et al., 2013; **Supplementary Material 3**) on the concatenated dataset with 100 bootstrap replicates and a confidence interval of  $\alpha = 0.05$ .

## Characterization of Hybrids Classes and Introgression

After getting the admixing proportion through Bayesian approach between *V. longifolia* and *V. spicata* in STRUCTURE, we allocated the samples to distinct hybrid or pure classes, specifically, we tested to identify the presence of parent species individuals, F1 or F2 hybrids and backcrossed individuals to both species [parental *V. longifolia* (A) or *V. spicata* (B)]. We only included the individuals for which no single class had  $Q \geq 0.80$  probabilities to belong to one parent species. We implemented this analysis using the program NEWHYBRID (Anderson and Thompson, 2002). NEWHYBRIDS assign individuals to specific hybrid classes by looking at the patterns of gene inheritance

within each locus. Since NEWHYBRIDS cannot handle large datasets, we created subsets of 200 unlinked most informative and differentiated SNP loci (method = *AvgPIC*) using the function *gl.nhybrids* included in the R package *DARTR* (Gruber and Georges, 2019; **Supplementary Material 3**). This function identifies loci that exhibit a fixed difference by comparing two sets of parental populations and creates an input file for the program *NewHybrids*. To avoid any further bias caused by the heterozygous loci of parental populations, we only included the pure individuals, those showing less than 1% of admixture from *STRUCTURE* analysis. Additionally, to get the robust results, we run the program with no prior information regarding the status or class of individuals. The analysis was run for  $10^5$  MCMC iterations with a burn-in threshold of  $10^4$ . To run the program *NEWHYBRID*, we used the R package *PARALLELNEWHYBRID* (Wringe et al., 2016, 2017).

To check the possibility of introgression between both species, we measured genomic clines in the admixed individuals shared by the parental species as implemented in *INTROGRESS* (Gompert and Buerkle, 2009, 2010, 2012). *INTROGRESS* analyzes introgression of genotypes between divergent lineages and hybridizing species. In addition, *INTROGRESS* estimates genomic clines from co-dominant, dominant, and haploid marker data, without any requirement of fixed allelic differences between parental populations for the sampled genetic markers. The package tests for deviations of loci from neutral expectations as well. *INTROGRESS* is ideal in cases with high number of informative markers in which the potentially admixed individuals cover the full range of variation between parental species (Gompert and Buerkle, 2010). Since our data is based on continuous sampling, without any spatial correspondence, *INTROGRESS* suits well to our model. Briefly, we divided the data into three groups, two homozygotes (pure parental) and one heterozygote genotypes (admixed individuals), based on the *NEWHYBRID* and *STRUCTURE* results. These groups included 63 individuals of *V. longifolia*, 79 individuals of *V. spicata* and 13 admixed individuals ( $Q \geq 0.2$ ). In the actual analysis, firstly, we estimated hybrid indices and then genomic clines using multinomial regression. We measured introgression for each individual locus in admixed individuals relative to clines using their genotype frequency along an admixture gradient (Buerkle, 2005; Gompert and Buerkle, 2010) by comparing likelihood of regression models to that of a neutral model. The significance of likelihood was accessed using 1000 permutations. Lastly, we counted the introgressed loci from the observed hybrid genotypes by comparing with the probability densities of homozygotes and heterozygotes genotypes as suggested (Nolte et al., 2009). We adjusted the neutral expectation of genomic clines through multiple comparisons using the false discovery rate (FDR) as suggested (Benjamin and Hochberg, 1995; examples in Harrison et al., 2017; Khan et al., 2020).

## Crossing and Germination Experiments

We conducted crossing experiments between individuals of *V. spicata* and *V. longifolia* of both ploidy levels and geographic origins to assess experimentally reproductive barriers between the two species. The experiments were

conducted in the greenhouses of the Carl von Ossietzky-University (Oldenburg, Germany) in spring and summer 2019. When enough inflorescences were available on a single plant, three treatments were applied for each replicate beside the controls (Ctr): H (hybridization), C (negative control) and CS (cage + selfing). To avoid any unwanted pollination from insects, all floral spikes were bagged using tea-bag filters. In treatment H self-pollination was tried to avoid by emasculation of the flowers on daily basis before anthesis. When the flowers opened, the freshly dehiscent anthers of the donor species were removed with clean tweezers and their pollens were transferred by direct contact with stigma on the receptor species. Negative controls (treatment C) were used to determine the minimum amount of pollen contamination acceptable after applying only emasculation. Negative controls were performed in parallel with H on the same plant. Spikes for treatment CS were bagged before anthesis till fruit maturation. Further, the controls (Ctr) were used to determine the germination rate of seeds pollinated with pollen of the same species. They were handled in a similar way to H with the exception that the pollen donor was from the same species and ploidy level as the control.

Seeds obtained from the crossing experiment were sterilized following the protocol suggested by Pickens et al. (2003): (1) 70% ethanol for 2 min; (2) rinse in distilled water for 5 min; (3) 2.6% NaOCl + Tween 20 for 40 min; (4) rinse in sterile water twice for 5 min each. After sterilization, seeds were placed in Petri dishes on filter paper, adding sterilized distilled water. A climate chamber was used to ensure controlled temperature and light conditions, with the settings: light 12/12 h; and temperature 21°C. The position of the single Petri dishes was randomized inside the chamber in order to account for small variation of temperature and light related to the position. Germination rate was defined as the number of seeds germinated after 30 days divided by the total number of seeds plated.

## Polyploidy and Ecological Differentiation

All niche analyses are based on the WorldClim data set (Hijmans et al., 2005) with a resolution of 2.5 min. The 19 bioclim variables were attached to the occurrence data of 248 individuals (see sampling) using DIVA-GIS (Hijmans et al., 2012).

In a preliminary principal component analysis (PCA) including all variables, we found that many of the variables were highly correlated with each other. To avoid collinearity and type II errors caused by high correlation, we applied pairwise Pearson product correlations in R (R Core Team, 2019). This Pearson correlations test retrieved bioclim variables with  $\text{mod} \geq 0.8$  such as: bio4 and bio7; bio5 and bio10; bio1, bio6, bio9, and bio11; bio12, bio13, bio14, bio16, bio17, bio18, and bio19, and the remaining four variables, i.e., bio2, bio3, bio8, and bio15 were not correlated. In the actual analysis we used eight variables, excluding the remaining 11 variables. The choice of the four variables that showed correlation with other variables was based on criteria such as longer time period over shorter period in the definition of the variables and more or less subjective hypotheses about the biological importance in our species. The eight variables included were: bio2 mean diurnal range, bio3 isothermality, bio7 temperature

annual range, bio8 mean temperature of wettest quarter, bio9 mean temperature of driest quarter, bio10 mean temperature of warmest quarter, bio15 precipitation seasonality, and bio17 precipitation of driest quarter.

The reduced data set with 248 specimens and eight bioclimatic variables included was subjected to a PCA using NTSYSPC version 2.2 (Rohlf, 2009). It turned out that the Siberian samples of diploid *V. spicata* (>50° East) grow under rather distinctive climatic conditions. Tetraploid *V. spicata* do not occur there. To accommodate for this, we repeated several of our analyses separating the Siberian samples (>50° East) from diploid *V. spicata* from the West (Table 1). Such a problem did not arise in *V. longifolia*, because its diploid and tetraploid groups occur in the same general geographical regions or there were only very few data points in the one group when far distant from the other group.

To qualitatively compare climatic niche models (or in other words: the macroclimatic envelopes of a group projected on a geographical map) between conspecific diploid and tetraploid groups, we used MAXENT version 3.4.1 (Phillips et al., 2006, 2019) on the data set described above (248 data points divided in four or in five groups). For the application of MAXENT, we followed Merow et al. (2013) and Fourcade et al. (2014), who recommended minimizing the correlation among predictors. Therefore, we again used only the eight bioclim variables as described above. For purposes as in our study, they recommended to use the relevant and complete geographical space for background selection. We therefore first limited the bioclim layers in DIVA-GIS (Hijmans et al., 2012) to 5°W, 135°E, 35°N, and 75°N (Figure 2) and used this restricted rectangle of the eight climatic layers in MAXENT. It was not possible to account for sampling effort in our data, because the data was opportunistically gathered from different times and sources. However, a facultative exclusion of the densely sampled Siberian *V. spicata* and the thinning of all data to 25 km linear distance follow the suggestions of Fourcade et al. (2014) to minimize geographically biased sampling. We tested different output options in MAXENT and chose the logistic output because the results seemed less diffuse (steeper and therefore clearer gradients) than with the other options.

We investigated differentiation in macroclimatic niches between diploids and tetraploids among individual variables using the analysis of variance (ANOVA) tool in EXCEL 2016 (Microsoft Corp., Redmond; significance considered  $p < 0.05$ ). All variables were tested for both taxon combinations and this might lead to statistical type I errors. Therefore, we additionally applied Bonferroni corrections for all 19 parallel comparisons, which leads to  $p < 0.003$  as a much more conservative threshold for significance. Since the differences found between European and Siberian samples of diploid *V. spicata* could be misleading (see below), we repeated the ANOVA for *V. spicata* including only the western samples (west of 50° East).

Finally, to get a quantitative idea about climatic niche differences between the ploidy levels or between species, we calculated Schoener's *D*, a frequently used index for niche overlap, which ranges between 0 and 1 (the latter stands for complete niche overlap) from the original MAXENT results. We used

the precompiled version 1.3 for windows of ENMtools (Warren et al., 2008, 2010) for this and also for the following analysis. To test whether the MAXENT scores exhibit statistically significant ecological differences between several groups, we performed a niche identity test using Schoener's *D* as a test variable. In this test the empirical occurrence points were pooled, their identity randomized, subjected to a MAXENT analysis, and Schoener's *D* was calculated. We did 100 pseudo-replicates per comparison. The resulting distribution of Schoener's *D* represents the null hypotheses that niches differ by chance only and it can be compared with the *D* score of the original data to judge about the probability that niches are identical.

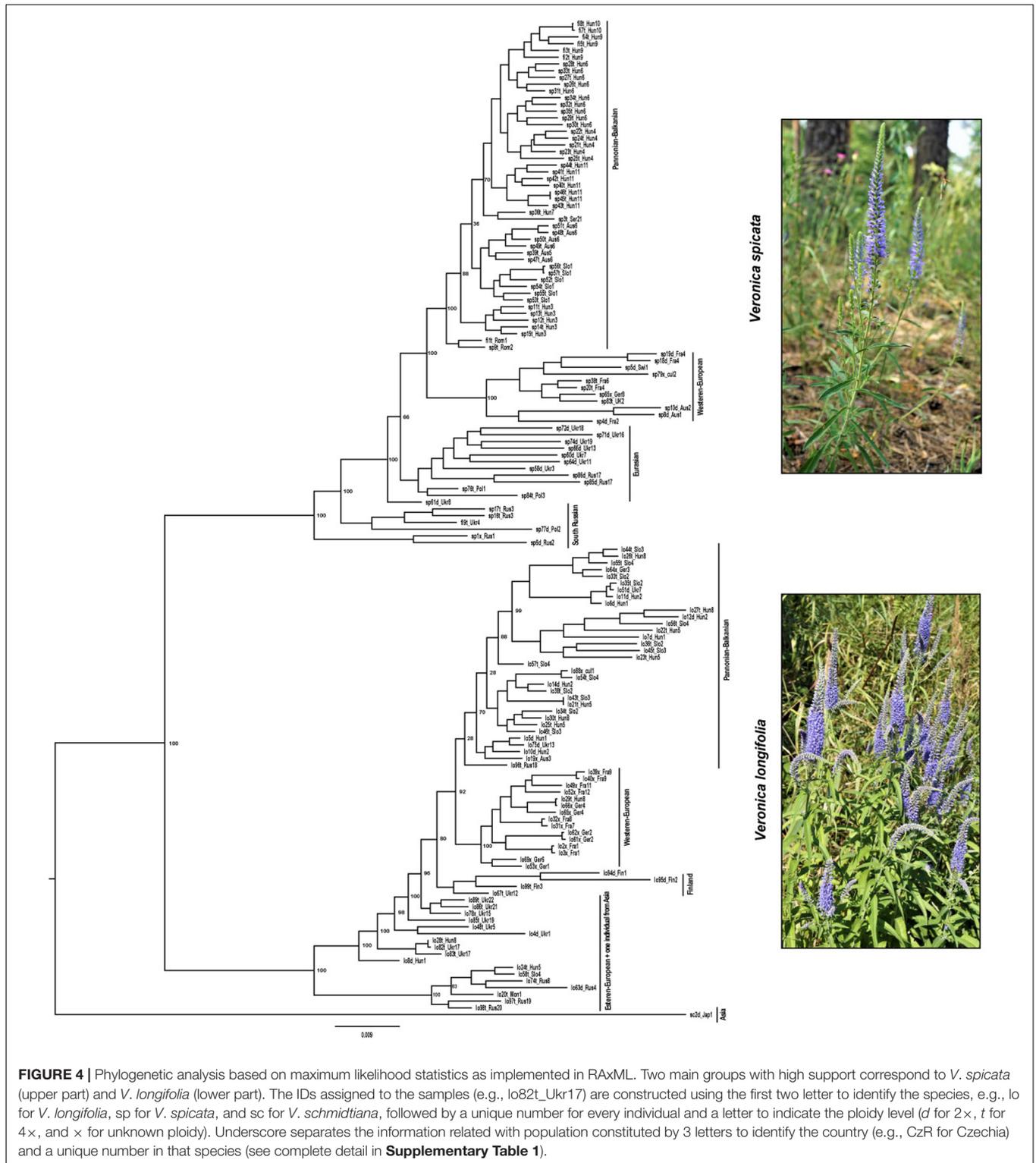
## RESULTS

### Phylogeographic Structure and Admixture

The Bayesian modeling to assess the genetic structure clustered all individuals in two groups (Evanno method;  $K = 2$ ; Figure 1A) corresponding to *V. spicata* and *V. longifolia*. This analysis clearly showed that the genetic structure of both species is in accordance with their morphology-based classification. However, our results also recovered individuals with admixture higher than 20% (Figure 1A and Supplementary Table 1). The threshold of 20% was subsequently used to infer hybrids since individuals with more than 20% did not group according to geography in other analyses. The second level hierarchical clustering on species alone found optimal intraspecific clustering in *V. spicata* at  $K = 2$  (result not shown). However, after averaging the results of STRUCTURE between individuals in the populations and using the R package TESS3R to plot those averages on a map, we recovered a reasonable geographical pattern until  $K = 5$  (Figure 3). The Evanno method indicated an optimal  $K = 2$  for the second level hierarchical clustering analysis using the dataset of *V. longifolia* in STRUCTURE and TESS3R was used to plot the results. At  $K = 2$ , the individuals were structured into one group composed of individuals from Asia and the other including those from Europe (Figure 3) except for one individual from Siberia [Rus18 (lo96t\_Rus18)] grouping with European individuals. Admixture between the two species based on the STRUCTURE analysis, considering individuals with  $Q \geq 0.20$  for the minor ancestry as admixed, occurred in seven individuals. Of these, four individuals were *a priori* identified as *V. longifolia*, two as *V. spicata*, and one considered a hybrid. Admixture occurred in both ploidy levels, six in tetraploids and one in a diploid. It was found widely from Switzerland to Ukraine. The divergence between *V. spicata* and *V. longifolia* was lower, with an  $F_{ST}$  value of  $0.296 \pm 0.002$ , than between *V. longifolia* and *V. schmidtiana* ( $F_{ST} = 0.625$ ), and between *V. spicata* and *V. schmidtiana* ( $F_{ST} = 0.632$ ; Supplementary Table 3), while highly significant for all ( $P = \sim 0.0$ ).

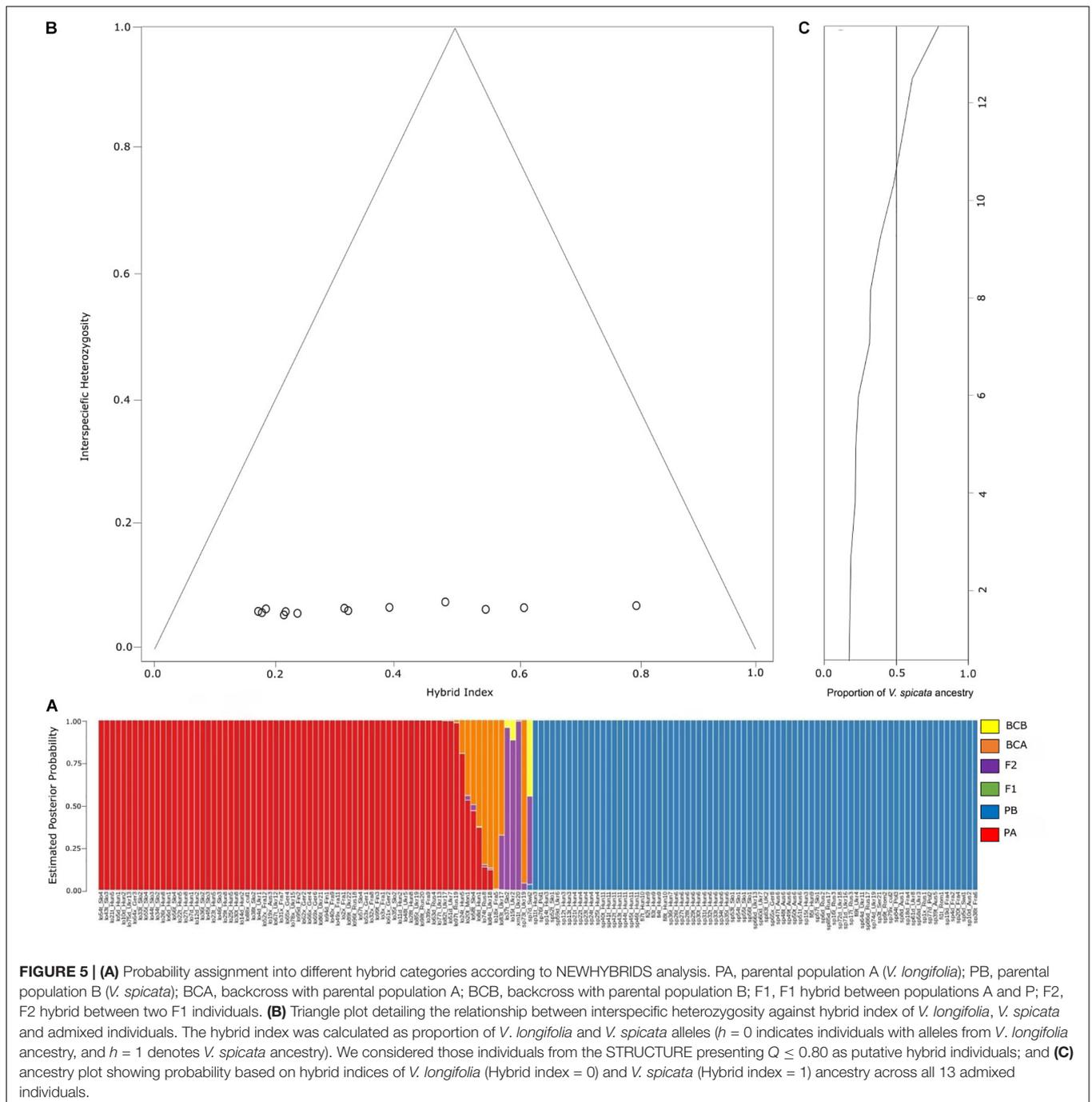
Similarly, the phylogenetic analysis based on maximum likelihood also clustered all individuals into two main groups with high support corresponding to *V. spicata* and *V. longifolia*. Since individuals admixed according to STRUCTURE appeared basally branching and not in geographical meaningful clusters,





in admixed individuals than expected under neutral expectations. The homozygosity inferred was 41% from *V. longifolia* and 51% from *V. spicata* (Figure 5C). We recovered only one heterozygous locus in all admixed individuals, reflecting high homozygosity.

The results infer a higher tendency of homozygotes clines from *V. spicata* with 216 more SNPs than *V. longifolia*. This tendency of 18% more homozygotes from *V. spicata* suggests asymmetrical introgression. For this analysis we only used the non-neutral loci,



showing deviation from a model of neutral introgression based on genome-wide admixture ( $P \leq 0.025$ ; significance following FDR correction; **Table 2**).

## Crossing and Germination Experiments

We have used 32 individuals during the crossing experiment (11 of *V. spicata* and 21 from *V. longifolia*). Overall, 24 crossings were performed in which 13 were based on the same ploidy and 11 on different ploidy levels (**Table 3**). In almost all cases, the negative controls ( $n = 20$ ) resulted in the production of zero

seeds except in two cases, in which the number of seeds was 18 and 4. In successful crosses, we got usually more than 100 seeds per spike. We considered the crossing as successful only when it led to production of more than 20 seeds per spike, attributing the eventual production of small amounts of seeds to contamination by small arthropods, which were able to reach the spikes through the lower opening of the bag. Crossings performed between plants of the same ploidy level ( $2\times \times 2\times$  and  $4\times \times 4\times$ ; **Table 3**), even from very distant geographical origin, resulted in production of viable seeds in all cases, with

**TABLE 2** | Patterns of introgression based on all admixed individuals.

Patterns of Introgression relative to neutral expectations	No of Loci
<b>Non-neutral</b>	1793
<i>V. longifolia</i> (Parent A) homozygotes to admixed individuals	472
<i>V. spicata</i> (Parent B) homozygotes to admixed individuals	688
Heterozygous excess in admixed individuals	1
Heterozygous deficiency in admixed individuals	1603
Homozygous excess in admixed individuals	1160
Homozygous deficiency in admixed individuals	2
<b>Neutral</b>	6930

The results are based on pure parental individuals of *V. longifolia* (63 individuals with  $0.1 \geq Q \geq 0.9$ ), *V. spicata* (79 individuals with  $0.1 \geq Q \geq 0.9$ ), and admixed individuals (13 with  $Q \geq 0.2$ ) following the NEWHYBRID, and STRUCTURE results. The total number of SNPs used were 8,723 with 25.73% missing data. Prediction of introgression at each locus was compared through likelihood of regression model to that of a neutral model and its significance by using 1,000 permutations. Non-neutral loci were selected by adjusting the neutral expectation of genomic clines through multiple comparisons ( $P < 0.025$ ; significance following FDR correction).

germination rates ranging from 18 to 100%. However, the crossings between different ploidy levels ( $2 \times \times 4 \times$  and  $4 \times \times 2 \times$ ; **Table 3**), produced almost no seed in crosses with the tetraploid mother, and in four out of seven crosses when the mother was diploid (**Table 3**). Plants of crosses, which produced less than 10 seeds, did not grow to fully developed plants, except for a single plant. This plant was determined to be tetraploid (Albach S372  $\times$  Albach S634). All five plants of the cross Albach 1450 and Albach S854 analyzed flow cytometrically were tetraploids. Six of the plants from the cross between Albach 1450 and Albach S485 were analyzed with flow cytometry and all were triploid. The controls between plants of the same species and ploidy level showed a high production of seeds and germination rates (>93%) except for the crossing of diploid *V. spicata*, which could be due to sterility or self-incompatibility reactions.

## Polyploidy and Ecological Differentiation

Based on our literature search and our own results, we demonstrate that diploids and tetraploids are about equally common in *V. spicata* (115 vs. 138 measurements, respectively), whereas diploids are less common in *V. longifolia* (49 vs. 102 measurements, respectively). However, these numbers are biased by our sampling and the accessibility of populations.

The ANOVA results for the diploid/tetraploid comparison of eight climatic variables are given in **Table 4** (results for all 19 variables are given in **Supplementary Table 4**). When Siberian diploid samples of *V. spicata* were excluded, mainly the variables dealing somehow with precipitation (e.g., bio17 in **Table 4**) were significantly different between diploids and tetraploids. However, when all *V. spicata* diploids were included, variables somehow dealing with temperature or seasonality (e.g., bio2, 3, 7 in **Table 4**) were significant. This discrepancy is a clear indication and consequence of the distinctly continental climate in Siberia in contrast to the sub-oceanic climate in western and central Europe. The difference between diploids and tetraploids within *V. longifolia* seemed to be less pronounced and was indicated by five significant variables (three after Bonferroni correction).

These variables (e.g., bio8 and 10 in **Table 4**) basically refer to different summer temperatures.

The first four axes of the PCA incorporated 52, 23, 11, and 7% of the total variance. The first axis was mainly determined by variables based on seasonality (bio2, bio3, bio7, and bio15) and precipitation (bio9 and bio17), while the second axis was mainly determined by temperature variables (bio8 and bio10). The 41 eastern samples of diploid *V. spicata* (cyan blue triangle in **Figure 6**) were separate from all other *V. spicata* data points in the PCA. However, the western remainder of diploid *V. spicata* was largely overlapping with tetraploid *V. spicata*, and the same was true for most of the diploid and tetraploid individuals of *V. longifolia*, respectively.

The variables that contributed more than 10% to MAXENT models were: for diploid *V. longifolia* bio7 (43%) and bio17 (28%); for tetraploid *V. longifolia* bio8 (48%), bio 17 (21%), and bio7 (11%); for diploid *V. spicata* (all data) bio17 (41%), bio10 (22%), and bio9 (12%); for diploid *V. spicata* west bio17 (63%) and bio3 (21%), for diploid *V. spicata* east bio7 (34%), bio10 (25%), bio17 (20%), and bio8 (16%); for tetraploid *V. spicata* bio7 (36%) and bio17 (34%). The maps produced by the logistic output of MAXENT revealed slight differences between the conspecific diploid and tetraploid groups of each species (**Supplementary Figure 2**). The main difference within *V. longifolia* was that the plot for diploids indicated suitable climate also in northern Europe (Scandinavia and northeast of the Baltic Sea), but less so for the tetraploids. Within *V. spicata*, the areas suitable for tetraploids and for the western diploids are very similar. This changed when eastern individuals were also included, adding a large corridor from eastern Europe to western Asia to the potentially suitable conditions of the diploids.

The results of the quantitative niche overlap analysis (Schoener's *D*) and the niche identity test for all pairwise comparisons indicated significant niche differentiation (**Table 5**). These analyses recovered the least differentiation between *V. longifolia* tetraploid and *V. spicata* diploid and highest difference between *V. spicata* diploid and *V. spicata* tetraploid.

## DISCUSSION

In this study, we investigated the phylogeographic structure using genome-wide SNPs considering hybridization, introgression and polyploidization of two congeneric species, *V. spicata* and *V. longifolia* across their vast European distribution ranges. Though *V. longifolia* exhibited lower geographic structure, our results demonstrated clear geographical patterns for both species. Hybridization and introgression through backcrossing and polyploidization was detected in several natural populations across the European distribution of both species, but apparently these cases of hybridization have not lead to large-scale merger of the species, neither such hybridization obscured the species boundaries. Finally, experimental crossings revealed high level of compatibility between the two species in cases in which individuals have the same ploidy levels. We have not recovered any clear geographical or macroclimatic patterns related to polyploidy. Thus, the two species do not form hybrid zones on a

**TABLE 3 |** Crossing experiment results.

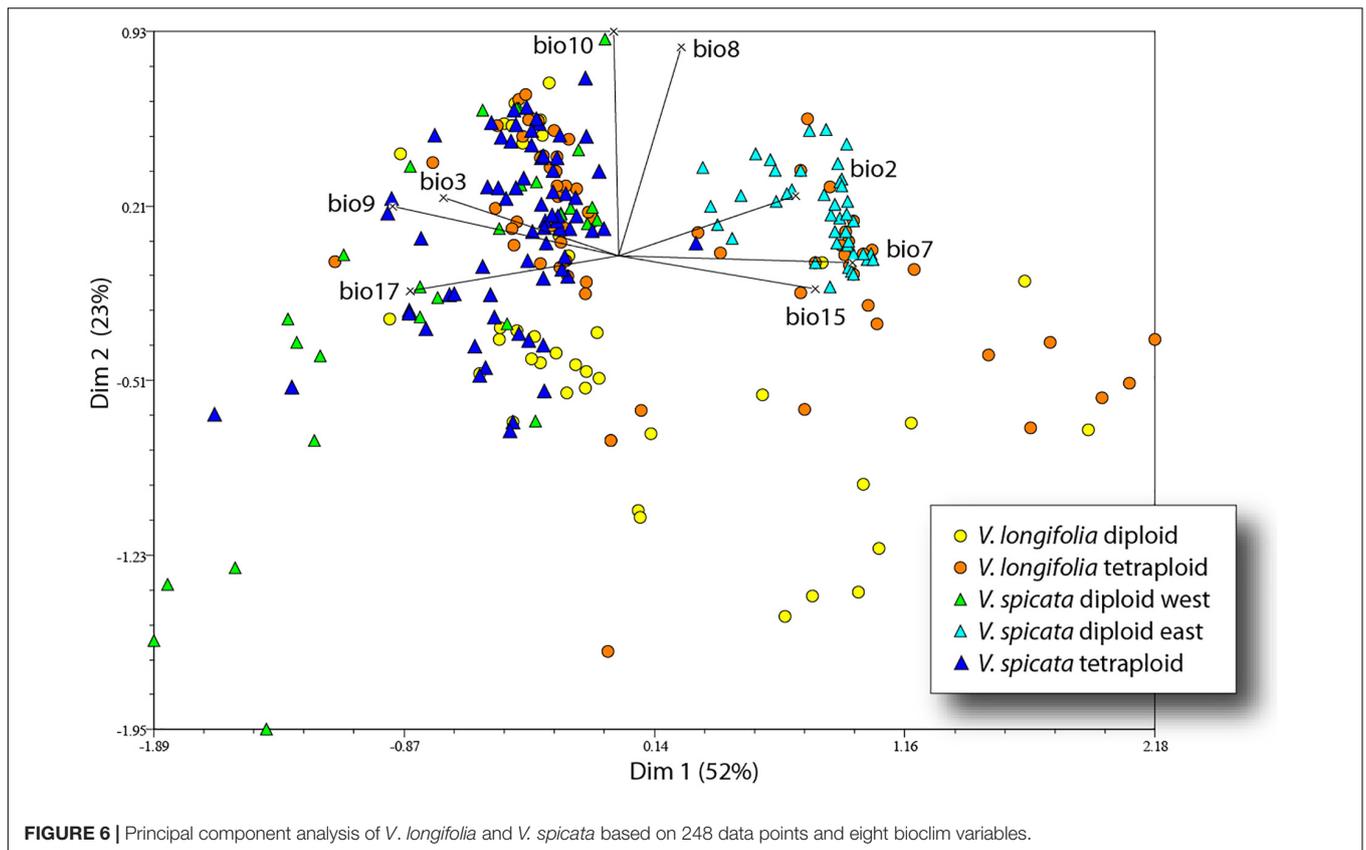
Hybrids (H)										
Ploidy	♂		♀						ID	
(♂ × ♀)	Species	Origin	Species	Origin	Seeds obt.	Seeds plated	Germ.30 d	Germ.%	♂	♀
2x × 2x	<i>V. spicata</i>	Switzerland	<i>V. longifolia</i>	Altai Mount.	>100	50	47	94	Albach S526	Albach S625
	<i>V. longifolia</i>	Altai Mount.	<i>V. spicata</i>	France	>100	50	50	100	Albach S625	Albach S758
	<i>V. longifolia</i>	Altai Mount.	<i>V. spicata</i>	Switzerland	>100	50	48	96	Albach S625	Albach S526
	<i>V. longifolia</i>	Finland	<i>V. spicata</i>	Romania	>300	50	50	100	Albach S848	Albach S376
	<i>V. longifolia</i>	Finland	<i>V. spicata</i>	Hungary	>100	50	46	92	Albach S849	Albach S748
	<i>V. longifolia</i>	Finland	<i>V. spicata</i>	Poland	>100	50	9	18	Albach S848	Albach S485
	<i>V. longifolia</i>	Finland	<i>V. spicata</i>	Poland	>100	50	13	26	Albach S849	Albach S485
	<i>V. spicata</i>	France	<i>V. longifolia</i>	Finland	>50	50	12	24	Albach S758	Albach S849
2x × 4x	<i>V. longifolia</i>	Finland	<i>V. spicata</i>	Hungary	0	0	0	0	Albach S848	Albach S869
	<i>V. spicata</i>	Altai Mount.	<i>V. longifolia</i>	Ukraine	~10	0	0	0	Albach S634	Boiko 020
	<i>V. spicata</i>	Poland	<i>V. longifolia</i>	Czechia	0	0	0	0	Albach S485	Albach S499
4x × 2x	<i>V. longifolia</i>	Yakutia	<i>V. spicata</i>	France	4	4	4	100	Albach S372	Albach S758
	<i>V. longifolia</i>	Yakutia	<i>V. spicata</i>	Altai Mount.	9	9	6	67	Albach S372	Albach S634
	<i>V. longifolia</i>	Ukraine	<i>V. spicata</i>	Poland	>50	28	19	68	Albach 1450	Albach S485
	<i>V. longifolia</i>	Ukraine	<i>V. spicata</i>	Poland	~100	50	1	2	Albach 1450	Albach S485
	<i>V. longifolia</i>	Ukraine	<i>V. spicata</i>	Altai Mount.	0	0	0	0	Boiko 020	Albach S634
4x × 4x	<i>V. longifolia</i>	Ukraine	<i>V. spicata</i>	Hungary	>100	50	49	98	Boiko 020	Albach S869
	<i>V. spicata</i>	Hungary	<i>V. longifolia</i>	Czechia	>100	50	48	96	Albach S869	Albach S499
	<i>V. spicata</i>	Hungary	<i>V. longifolia</i>	Ukraine	>100	50	50	100	Albach S869	Albach 1450
	<i>V. longifolia</i>	Ukraine	<i>V. spicata</i>	Hungary	>100	50	40	80	Boiko 020	Albach S869
Positive Control (Ctr)										
2x × 2x	<i>V. longifolia</i>	Finland	<i>V. longifolia</i>	Finland	~30	33	31	94	Albach S849	Albach S849
	<i>V. spicata</i>	Austria	<i>V. spicata</i>	Austria	>50	50	50	100	Albach S749	Albach S749
4x × 4x	<i>V. spicata</i>	Austria	<i>V. spicata</i>	Austria	>100	50	50	100	Albach S749	Albach S749
	<i>V. longifolia</i>	Czechia	<i>V. longifolia</i>	Czechia	~37	35	34	97	Albach S499	Albach S499
	<i>V. longifolia</i>	Czechia	<i>V. longifolia</i>	Czechia	>50	50	50	100	Albach S499	Albach S499

The number of seeds obtained refers to single floral spikes. Germ. 30d refers to seed germinated 30 days after plating. ID refers to vouchers in herbarium OLD. Hybrids indicate crossings between different species, while positive controls crossings between same species and ploidy level. The IDs assigned to the samples (e.g., sp30aSwi\_d1) were constructed using the first two letter to identify the species (sp for *V. spicata*, and lo for *V. longifolia*), followed by a unique letter e.g., 30a representing an individual of population 30, and 3 letter code of country of origin (e.g., Swi for Switzerland). An underscore separates the information related with polyploidy e.g., t for tetraploid and d for diploid.

**TABLE 4 |** One-way ANOVA calculations comparing bioclim variables (Hijmans et al., 2005) of diploid (2) and tetraploid (4) *Veronica longifolia* and of diploid (2) and tetraploid (4) *Veronica spicata*.

Bioclim variable	<i>V. longifolia</i> 2/4			<i>V. spicata</i> 2 total/4			<i>V. spicata</i> 2 west/4	
	Mean 2 (N = 40)	Mean 4 (N = 65)	p-value	Mean 2 (N = 74)	Mean 4 (N = 69)	p-value	Mean 2 (N = 33)	p-value
bio2	8.99	9.58	<b>0.047</b>	9.65	8.62	<b>&lt;0.001*</b>	8.42	0.431
bio3	25.20	25.61	0.629	25.36	29.00	<b>&lt;0.001*</b>	29.10	0.894
bio7	36.28	38.89	0.185	39.35	29.81	<b>&lt;0.001*</b>	29.17	0.383
bio8	13.39	16.90	<b>&lt;0.001*</b>	16.20	16.32	0.861	13.68	<b>0.006</b>
bio9	-5.18	-7.18	0.267	-7.10	0.11	<b>&lt;0.001*</b>	0.93	0.292
bio10	15.27	17.26	<b>&lt;0.001*</b>	17.72	17.40	0.365	16.92	0.332
bio15	39.48	42.85	0.334	36.48	32.40	<b>0.031</b>	27.24	<b>0.009</b>
bio17	85.88	77.75	0.258	98.34	105.32	0.406	153.31	<b>&lt;0.001*</b>

The latter was tested with and without the diploid Siberian samples (2 total and 2 west, respectively). The *V. spicata* mean 4 column is valid for both comparisons. Only the eight uncorrelated bioclim variables are shown here but the test was done for all 19 variables (see **Supplementary Table 4**). Significant results ( $p < 0.05$ ) are in bold. Significant results after Bonferroni correction for multiple tests ( $p < 0.003$ ) have an asterisk.



**FIGURE 6** | Principal component analysis of *V. longifolia* and *V. spicata* based on 248 data points and eight bioclim variables.

**TABLE 5** | The results of the niche overlap (Schoener's *D*) and niche identity tests.

	Schoener's <i>D</i>	Niche identity test mean <i>D</i> ± <i>SD</i>	<i>p</i> -value
<i>V. longifolia</i> tetraploid vs. diploid	0.61	0.79 ± 0.04	<0.01
<i>V. spicata</i> tetraploid vs. diploid	0.39	0.84 ± 0.03	<0.01
<i>V. spicata</i> tetraploid vs. diploid (west)	0.66	0.81 ± 0.04	<0.01
<i>V. longifolia</i> tetraploid vs. <i>V. spicata</i> diploid	0.69	0.82 ± 0.03	<0.01
<i>V. longifolia</i> diploid vs. <i>V. spicata</i> diploid	0.51	0.81 ± 0.04	<0.01
<i>V. longifolia</i> tetraploid vs. <i>V. spicata</i> tetraploid	0.40	0.81 ± 0.04	<0.01
<i>V. longifolia</i> diploid vs. <i>V. spicata</i> tetraploid	0.41	0.78 ± 0.05	<0.01
<i>V. longifolia</i> diploid vs. <i>V. spicata</i> diploid (west)	0.46	0.78 ± 0.06	<0.01
<i>V. longifolia</i> tetraploid vs. <i>V. spicata</i> diploid (west)	0.40	0.77 ± 0.04	<0.01

The mean, the standard deviation (*SD*) of the *D* value, and the *p*-values were derived from the distribution of *D* values from 100 × pseudoreplicates.

continental scale but likely have formed hybrid zones of different extent and temporal duration on a local scale ever since the two species split about 1.5 million years ago (Khan et al., submitted).

## Phylogeographic Structure of *V. longifolia* and *V. spicata* Across Eurasia

Both the population genetic structure and maximum likelihood tree indicated a basal split between European and Central to East Asian (plus two European) samples of *V. longifolia* after removal of admixed ( $Q > 0.2$  for the minor ancestry) individuals. The higher amount of inferred admixture for these individuals (4–20%; **Supplementary Table 1**) is not considered to be real evidence for introgression but rather an artifact of STRUCTURE based on genetic divergence of these samples from the majority of (European) individuals. Currently, we lack evidence for this hypothesis but this is currently being analyzed in more detail. Such a split between European and Central to East Asian populations has previously been found in *Clausia* (Franzke et al., 2004) with a split between groups found around 70° East and further west along the Ural Mountains in *Camelina microcarpa* (Žerdoner Čalasan et al., 2019). This split along the Ural Mountains is in line with the division of periglacial steppes in two large ecosystems east and west of the Ural Mountains based on paleo-environmental reconstructions (Hurka et al., 2019). Considering the postulated East Asian origin of the subgenus (Kosachev et al., 2016), we hypothesize that *V. longifolia* originated in Eastern Asia and subsequently dispersed westward toward Europe and further differentiated there. A split between the European and Asian populations in *V. longifolia* has already been suggested by Trávníček (2000),

who considered the differences between Asian and European plants substantial enough to treat them as two different species. However, given our sparse sampling among Asian populations and the lack of clear morphological differentiation, we currently refrain from adopting this view.

In contrast to *V. longifolia*, the earliest branching samples in *V. spicata* are from the South Russian – North Caucasian region (Figure 4). This argues against the East Asian origin but rather indicates early differentiation along the southern border of the extant distribution, though not necessarily an origin within that region. The only Asian samples of that species included here were nested among samples from Ukraine, which suggests that *V. spicata* migrated eastwards to the Altai Mountains either from Southern Russia or from a widespread Eurasian steppe group. The results of the STRUCTURE analysis support the latter view that the samples of *V. spicata* from the Altai Mountains and Eastern Europe belong to a group of widespread Eurasian steppe genotypes and the Southern Russian populations belong to a separate population group that differentiated early in the evolution of the species. As in *V. longifolia*, these results show higher amounts of admixture, which we again attribute more to intraspecific genetic divergence than to real introgression from *V. longifolia* (see above). In this case, it is noteworthy that sample sp1x\_Rus1 morphologically belongs to *V. spicata* subsp. *transcaucasica* Bordz., which is morphologically differentiated from typical *V. spicata*. The existence of widespread Eurasian steppe genotypes is somehow counter to the more general assumption that many species during glacial phases of the Pleistocene survived in southern refugia, e.g., the three southern European peninsulas with a subsequent spread across Europe and eastwards to Siberia (Taberlet et al., 1998; Hewitt, 2001). This generalized phylogeographic pattern, being in general true for many species, is, however, not considered a universal key suitable for explaining all phylogeographic patterns of European and Eurasian plant species anymore. Several other patterns for taxa of vascular plants have been proposed as well (e.g., Suchan et al., 2019; Friesen et al., 2020; Volkova et al., 2020, and references therein), such as survival in more northern refugia (even cryptic ones: see Stewart and Lister, 2001), ecological shifts with possible survival in habitats and plant communities not now peculiar to the species, etc. (see a paleobotanical overview in Bezusko et al., 2011). These patterns reflect that each species reacted to the Pleistocene and Holocene environmental and climate changes in an idiosyncratic way reflecting that many characters and ecological adaptations are specific to that species. Phylogeographic patterns of Eurasian steppe plants should also be considered against the background of the general development and transformations of the Eurasian steppe zone (see an overview in Hurka et al., 2019, and references therein), especially during the Pleistocene and Holocene climatic and environmental cycles. Thus, an analysis with better sampling across the Eurasian steppe is necessary to identify refugia of these two species of *Veronica*.

The ML phylogram (Figure 4) further suggests that *V. spicata* expanded westwards to Eastern Europe (southern Poland and Ukraine) from the Southern to Central Russian refugium. All individuals west of Ukraine form two geographically distinct groups, one constitutes a Western European (Germany to

United Kingdom) and another a Pannonian-Balkanian group. Our TESS3R (Figure 3A) results detected five genetic clusters in *V. spicata*, which form geographically delimited groups mostly in congruence with the ML phylogram. These are (I) the Western European group with individuals from United Kingdom, France, Germany, Switzerland, and Western Austria; (II) Pannonian-Balkanian group that includes Serbian, Romanian, Hungarian, Slovenian, Eastern Austrian, and Southern Polish individuals; (3) the South Russian group with an outlying individual from Poland; (4) a widespread Eurasian group composed by Ukrainian populations including the samples from the Russian Altai; and (5) a rather isolated Swiss individual. Consequently, our results agree with previous analyses and support the existence of a western, a southeastern European, a Caucasian, and an Eastern refugium (Stewart and Lister, 2001).

The single diploid individual (sp7d\_Swi2) sampled from population Swi2 constitutes a separate Swiss group in the STRUCTURE analysis. STRUCTURE analysis indicates slight admixture for this *V. spicata* individual, while it shows no morphological similarities with *V. longifolia*. Our NewHybrids analysis indicates that sp7d\_Swi2 has a partial probability to be a F<sub>2</sub> hybrid or a backcross with *V. longifolia*, explaining the uniqueness of this Swiss population. Another alternative for this uniqueness might be that sp7d\_Swi2 originates from a separate refugial area. This is in concordance with Trávníček et al. (2004) who mentioned *V. spicata* diploid relicts in Switzerland and other areas of Europe. During the last Pleistocene glaciations, the Alps were covered with glaciers, and many species found refuge in the present-day territory of Italy, which was then free of ice (Lugon-Moulin and Hausser, 2002). After glaciations, many species recolonized the Alps from the South following different migration pathways (e.g., through the Rhodanian pathway; Parisod, 2008). A phylogeographic analysis of *V. spicata* in the Alps seems necessary to resolve this question. However, our study warns that hybridization and introgression from *V. longifolia* needs to be considered in such an analysis.

Intraspecific differentiation as in *V. spicata* is not as clear in *V. longifolia*. This may partly be due to our sampling, but also to the higher connectivity of suitable habitats (*V. longifolia* is considered mainly a river corridor plant; Winter et al., 2008; Figure 3B), which allows high gene flow. Nevertheless, phylogenetic analysis, though hampered by low sampling points in western Asia, indicates that *V. longifolia* spread westward across the Russian steppe, reached first Ukraine (Figure 4), from where it seems to have spread to Fennoscandia independently from the rest of Europe. The remaining samples form two clades, one in Western Europe (France, and Germany) and another from Central to Eastern Europe (from Austria to Ukraine).

## Effects of Polyploidization

The genetic structure of both species seems to be largely unrelated to ploidy levels, since both geographical groups in *V. longifolia* and the European populations in *V. spicata* contain both ploidy levels with some geographically coherent clusters but no clear pattern. Thus, recurrent polyploidization events across the distribution area best explain the pattern, although

we cannot locate specific areas where tetraploids originated. The Balkan group of *V. spicata* is constituted only by 4× individuals and the East European group only by 2× individuals, while the other groups have both ploidy levels present. Diploid *V. spicata* have been reported from Bosnia (Fischer, 1974), Bulgaria (Graze, 1933), Serbia (Pustahija et al., 2013), Hungary (Löve, 1977), and Romania (**Supplementary Table 2**) but they have not been analyzed genetically and could potentially be misidentified or, in the case of our Romanian sample, could be the westernmost extension of the Ukrainian group. It is intriguing that there are no tetraploid individuals of *V. spicata* known in Asia. This can have at least three different explanations somewhat depending on the unknown frequency of polyploidization. First, it might be that tetraploids have not originated there by chance; second, they might not have migrated there by chance; or, third, they might have originated there with similar frequency as in Europe, but the habitats were not suitable. Unfortunately, the ANOVA analysis is inconclusive (**Table 4**), because significant niche differences within Europe seem mainly due to less precipitation in habitats of tetraploids and higher precipitation in diploids but when looking at the total range of the diploids, the significant differences indicate different temperature regime only. The same pattern can be seen in the PCA (**Figure 6**) in which the diploid and tetraploid *V. spicata* groups were largely overlapping and only the Siberian diploids were separate from all other *V. spicata* points. Also, the analysis of niche overlap (**Table 5**) is not a good or valid indicator for putative differences between ploidy levels in our study. Admittedly, there were significant differences between the ploidy levels. This test statistic, however, seems overly sensitive to unequal geographical sampling and we could not adjust our sample to accommodate for this. An example for this problem was, that the measured niche overlap between diploid and tetraploid *V. spicata* was smaller than between any other comparison, even between the different species. In conclusion, Schoener's *D* test statistic may not be suitable for such a wide and therefore partly incomplete geographical sampling as in our study.

Since tetraploid *V. spicata* have been found mainly in the Pannonian-Balkan region and diploids have rarely been found in the region, it should be investigated in more detail whether the tetraploid *V. spicata* from the region constitute a genetically differentiated infraspecific taxon with higher drought resistance (**Table 4**), a trait reported for other autopolyploids, although this is not a universal pattern (e.g., Yang et al., 2014). In that respect, the connection between the Pannonian-Balkan tetraploids and the other tetraploids needs to be studied further. Are these tetraploids of different origin, which could not spread as much as the Balkan populations? Despite this question, tetraploid *V. spicata* fits the common scenario found in many other European polyploids that were better able to colonize North European areas after glacier retreats (see **Figure 2**; blue vs. green points), possibly by being better adapted to cold conditions (Brochmann et al., 2004; but see Tkach et al., 2008), although this is not directly reflected in the comparison of climatic variables shaping ecological niches (e.g., similar bio10 values in **Table 4**). An alternative explanation for the success of polyploids applicable here is that due to fixed heterozygosity

polyploids are able to cope better with inbreeding depression after colonization events in new areas (Birchler, 2012). In contrast, in *V. longifolia* the diploids spread further north, with the tetraploids requiring higher temperature (e.g., bio10 in **Table 4**). Thus, the macroclimatic niche signal we found is weak and partly inconsistent, but if there was any, it seemed that polyploidization had different effects in the two related species. Finally, our genetic data showed that there was and is a recurrent evolution of polyploids in both species, which produces many unrelated and unconnected autopolyploid lineages. Thus, our explorations of climatic niches fit well to the theory, that the *Veronica* polyploids still carry the macroclimatic niche requirements of their individuals and immediate diploid ancestral population rather than a common, "typical" and detectable polyploid niche requirement.

## Hybridization

The genetic divergence between *V. spicata* and *V. longifolia* was found to be low ( $F_{ST} = 0.29$ ), indicating structured populations with connectivity (**Supplementary Table 3**). This suggests that the amount of evolutionary divergence between the parental species is not high enough to cause significant genetic incompatibilities (Moyle and Nakazato, 2010). Further, the possibility for hybridization to occur was revealed in our crossing experiments, which also demonstrated that plants of the two species can form viable seeds independent of their geographic origin, but only when both parents have the same ploidy level (**Table 3**). Third, macroclimatic differences between the two species appear minor (**Table 5**), which is consistent with the coexistence of the two species in close proximity (e.g., Hun6, Ukr13, and Ukr16-20).

Hybridization due to secondary contacts is considered to be a common phenomenon following recolonization after retreat of Pleistocene glaciers (Hewitt, 2001). During these glaciations (glacial phases, in contrast to interglacials), communities of steppe vegetation and its ecological analogs (such as cold and dry periglacial steppes or similar dry grassland plant communities) were favored by predominantly continental climate in most part of Europe, thus possibly providing large and connected open habitats (Mai, 1995; Bezusko et al., 2011; Hurka et al., 2019); such habitats were most probably suitable for *V. spicata* and ensured its survival and dispersal during Pleistocene cycles. On the other hand, interglacial phases of the Pleistocene were normally warmer and more humid in most of Europe and thus promoted such habitats as riparian meadows and floodplain grasslands developing on fluvio-glacial deposits, which were more favorable for survival and dispersal of *V. longifolia*. Thus, most probably these two species reacted in different manners to the environmental changes of the Pleistocene and Holocene but shifts in environmental conditions brought them into frequent contact.

Unfortunately, direct paleobotanical (paleopalynological) data on the history of *Veronica* in the Late Quaternary are limited because pollen grains of those entomophilous species are rare in fossil and subfossil pollen samples (see Bezusko et al., 2011, 2018); also, distinguishing species of *Veronica* based solely on their pollen is problematic (Tsymbalyuk and Mosyakin, 2013). However, some paleobotanical (in fact, paleoecological)

evidence on our two species can be obtained indirectly, through reconstructions of the possible distribution of ancient plant communities suitable for these species. In such assumptions we should, however, take into considerations possible habitat shifts of species in the past, as well as the existence of plant communities of the past that were only roughly corresponding to modern plant communities of the region concerned (Bezusko et al., 2011). In any case, such paleoreconstructions should be done in parallel and in agreement with experimental and molecular results, such as those presented here.

Our results indicate that hybridization events can be found between *V. spicata* and *V. longifolia* across the whole distribution range in Europe, with tetraploid and diploid hybrid individuals identified in populations from Ukraine, Slovakia, and Switzerland and likely happened for a long time. However, reliable identification of hybrids can be difficult. Here, we adopted a genetic definition based on results of the admixture analysis in combination with phylogenetic evidence. Individuals with higher admixture levels branched earlier among the respective species. Thus, the genetic divergence appears continuous from intraspecific variation to hybrids (Supplementary Table 1), and geographic coherence was added to justify the cut-off.

NEWHYBRIDS results are in line with STRUCTURE (Figures 1, 5). We observed that individuals with high probability to be hybrids or back-crosses with one of the two parental species in the STRUCTURE analysis showed a significant introgression probability proportional to the genotype frequency class expected (Anderson and Thompson, 2002). Overall, 13 individuals had a significant posterior probability to be one of the four hybrid classes (Figure 5). Backcrossing was mainly inferred toward *V. longifolia*, with only one exception, suggesting that introgression is asymmetrical from *V. spicata* to *V. longifolia*. Introgression seems to be not uniform inside natural populations, as indicated by the fact that some individuals resulted to be highly introgressed, while other individuals from the same population (e.g., Ukr19) were not admixed. Furthermore, in natural populations hybridization seems to be a phenomenon not restricted only to tetraploid individuals, even though it seems to occur less frequently in diploids (in only one case the identified hybrid was a diploid compared to five cases among tetraploids). Our crossing results indicate that the vitality of diploid and tetraploid hybrid individuals is comparable, while crossings between different ploidy levels produce seeds with a lower seed development, germination success and survivability, suggesting ploidy level as a crossing barrier. In line with this, hybridization in the two species was only observed in the populations in which *V. spicata* and *V. longifolia* had the same ploidy (Ukr17–20) and not in the populations where the two species differed in their ploidy (Hun6, Ukr13, and Ukr16–18). This result is in accordance with Härle (1932), who suggested within-ploidy but not between-ploidy cross-compatibility. Stebbins (1971) already hypothesized that inter-cytotype gene flow is mostly unidirectional from diploids to tetraploids and this has been demonstrated in the last years with molecular tools (e.g., Jørgensen et al., 2011; Zohren et al., 2016). Whereas documented cases of introgression in the wild involving the same ploidy level are frequent, less common examples between different ploidy levels are known. This is due

to the fact that triploid hybrids are not produced (triploid block), and if produced, they are highly sterile (Ramsey and Schemske, 1998). In fact, despite thousands of flow cytometric analyses in *Veronica*, uneven ploidy levels have never been reported in the wild (Bardy et al., 2011, **Supplementary Table 1** for *V. subg. Pseudolysimachium*, but see also Bardy et al., 2010; Padilla-García et al., 2018; Rojas-Andrés et al., 2020; López-González et al., 2020). However, we produced some triploids in our greenhouse from interploidal crosses (Table 3), although we did not yet grow them to maturity to check for fertility. Thus, in some rare cases, triploid individuals may exist in the wild and functional gametes may be produced by triploid hybrids. This can break ploidy barrier and, through backcross with one of the parental species, result in gene transfer (Chapman and Abbott, 2010).

Thus, as suggested by Bardy et al. (2011), hybridization may have generated a great morphological variation by generating several morphologically intermediates between species. Some hybrids seem to have intermediate morphology, while others seem to be more similar to one of the parents morphologically and/or in term of ecological niche. For example, the individual *V. × media* (xm1t\_Ukr20), classified as F<sub>2</sub> hybrid, showed intermediate morphological characters between *V. spicata* and *V. longifolia*, while the other hybrids detected, even those inferred to be F<sub>2</sub> such as lo37t\_Slo2 and lo15t\_Ukr2, did not. The predominance of backcrossing to *V. longifolia* combined with the majority of hybrids *a priori* identified as *V. longifolia* suggests that there is selective pressure favoring a *V. longifolia*-phenotype for the hybrids and, thus, backcrossing with *V. longifolia*. However, this selection does not occur up to the seedling stage since we did not observe maternal effects in our crossings (Table 3).

## CONCLUSION

The genus *Veronica*, with its high species richness and diversification, as well as a great ecological amplitude, constitutes an important model study system which allows investigating several aspects and factors of plant evolution. In the present study, further evidence is provided to support the hypothesis, already proposed for *V. subg. Pseudolysimachium* in other regions (Balkans – Bardy et al., 2011; Altai – Kosachev et al., 2019), that hybridization between *V. spicata* and *V. longifolia* is common but these two species remain separate and distinct due to their different ecological adaptations and partial geographical or spatial isolation, with hybrids being probably less competitive than their parents. Furthermore, as shown in the crossing experiment, ploidy can constitute a further reproductive barrier, preventing gene flow between natural populations of different ploidy levels. Our analyses further demonstrate that, despite similar distribution ranges and flower biology, the two species differ markedly in their phylogeographic structure. Dry grasslands appear to have occurred more scattered than riparian habitats with their high connectivity, leading to more elaborate population structure in *V. spicata* as compared to that in *V. longifolia*. Increasing the number of populations sampled and performing sampling in a more geographically homogeneous way across Eurasia may provide interesting and more solid data on the

origins and phylogeographic histories of the two species and their evolutionary interactions. Further ecological studies in different regions would be necessary to evaluate the effective ecological competitiveness of the hybrids compared with their parents and to identify which ecological niches they can occupy. The ecology of the hybrids is totally unknown; it can be intermediate between parents, but also one of the two extremes.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: 10.5061/dryad.kkwh70s3.

## AUTHOR CONTRIBUTIONS

DA, PK, SM, and GK conceived the idea. DB and EM-Q performed the experiments. DB, KBH, and GK conducted formal analyses. PK, EM-Q, and DA contributed reagents, materials, and analysis tools. DA, GK, DB, PK, and SM authored the drafts of the article. All authors contributed, made multiple revisions, and approved the final draft.

## ACKNOWLEDGMENTS

We acknowledge the support from the VW-Foundation (grant number 90256), state task (Russian Federation) project number FZMW-2020-0003), and the German Research Foundation (DFG) project AL 632/15-1 within the priority program SPP-1991

## REFERENCES

- Adams, K. L., and Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* 8, 135–141.
- Albach, D., and Fischer, M. (2003). AFLP- and genome size analyses: contribution to the taxonomy of *Veronica* subg. *Pseudolysimachium* sect. *Pseudolysimachion* (*Plantaginaceae*), with a key to the European taxa. *Phyt. Balc.* 9, 401–424.
- Albach, D. C., Martínez-Ortega, M. M., Delgado, L., Weiss-Schneeweiss, H., Özgökçe, F., and Fischer, M. A. (2008). Chromosome numbers in *Veroniceae* (*Plantaginaceae*): review and several new counts. *Ann. Missouri Bot. Gard.* 95, 543–567. doi: 10.3417/2006094
- Albach, D. C., Martínez-Ortega, M. M., Fischer, M. A., and Chase, M. W. (2004). A new classification of the tribe *Veroniceae*-problems and a possible solution. *Taxon* 53, 429–452. doi: 10.2307/4135620
- Anderson, E. C., and Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160, 1217–1229.
- Andrews, S. (2010). *FASTQC, a Quality Control Tool for High Throughput Sequence Data*. Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed February 1, 2020).
- Arrigo, N., de La Harpe, M., Litsios, G., Zozomová-Lihová, J., Španiel, S., Marhol, K., et al. (2016). Is hybridization driving the evolution of climatic niche in *Alyssum montanum*. *Am. J. Bot.* 103, 1348–1357. doi: 10.3732/ajb.1500368
- Avise, J. C. (2000). *Phylogeography: the History and Formation of Species*. Cambridge, MA: Harvard University Press.
- Avise, J. C. (2009). Phylogeography: retrospect and prospect. *J. Biogeogr.* 36, 3–15. doi: 10.1111/j.1365-2699.2008.02032.x

Taxon-Omics. This work benefited from the sharing of expertise within the DFG priority program SPP 1991. Computations were performed at the HPC Cluster HERO (High End Computing Resources Oldenburg) funded by the DFG through its Major Research Instrumentation Program (INST 184/108-1 FUGG) and the Ministry of Science and Culture (MWK) of the state of Lower Saxony. We are grateful for material from France and Germany by Laurent Hardion and from Austria by Monika Kriechbaum. Finally, we thank Silvia Kempen, Imke Notholt, and Sabrina Schöngart for help in the lab and in the greenhouse and Jan Reimuth for help in the niche modeling.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.588354/full#supplementary-material>

**Supplementary Figure 1** | Phylogenetic analysis based on maximum likelihood statistics as implemented in RAXML including hybrids. Coloring and labeling follow **Figure 4**. The IDs assigned to the samples (e.g., lo82t\_Ukr17) are constructed using the first two letter to identify the species, e.g., lo for *Veronica longifolia*, sp for *V. spicata*, and sc for *V. schmidtiana*, followed by a unique number for every individual and a letter to indicate the ploidy level (*d* for 2 $\times$ , *t* for 4 $\times$ , and  $\times$  for unknown ploidy). Underscore separates the information related with population constituted by 3 letters to identify the country (e.g., CzR for Czechia) and a unique number in that species (see complete detail in **Supplementary Table 1**).

**Supplementary Figure 2** | Macroclimatic maps resulting from the logistic output of MAXENT (Phillips et al., 2019). Macroclimatic suitability of regions for the groups analyzed ranges from very high (red) to poor (blue). White squares indicate points used for training the model. **(A)** Diploid *V. longifolia*. **(B)** Tetraploid *V. longifolia*. **(C)** All of diploid *V. spicata*. **(D)** Tetraploid *V. spicata*. **(E)** Only western individuals of diploid *V. spicata*. **(F)** Only eastern individuals of diploid *V. spicata*.

- Balao, F., Casimiro-Soriguer, R., García-Castaño, J. L., Terrab, A., and Talavera, S. (2015). Big thistle eats the little thistle: does unidirectional introgressive hybridization endanger the conservation of *Onopordum hinojense*? *New Phytol.* 206, 448–458. doi: 10.1111/nph.13156
- Bardy, K. E., Albach, D. C., Schneeweiss, G. M., Fischer, M. A., and Schönswetter, P. (2010). Disentangling phylogeography, polyploid evolution and taxonomy of a woodland herb (*Veronica chamaedrys* group, *Plantaginaceae* s.l.) in southeastern Europe. *Mol. Phyl. Evol.* 57, 771–786. doi: 10.1016/j.ympev.2010.06.025
- Bardy, K. E., Schönswetter, P., Schneeweiss, G. M., Fischer, M. A., and 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. doi: 10.1002/tax.601010
- Benjamin, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bezusko, L. G., Mosyakin, S. L., and Bezusko, A. G. (2011). *Patterns and Trends of Development of the Plant Cover of Ukraine in the Late Pleistocene and Holocene*. Kiev: Alterpress.
- Bezusko, L. G., Tsymbalyuk, Z., and Mosyakin, S. (2018). Distribution of some representatives of *Scrophulariaceae sensu lato* on the plain part of Ukraine during the Holocene: paleofloristic and palynomorphological aspects. *Visnyk Lviv. Univ. Ser. Biol.* 78, 96–99. doi: 10.30970/vlubs.2018.78.19
- Birchler, J. A. (2012). “Genetic consequences of polyploidy in plants,” in *Polyploidy and Genome Evolution*, eds P. S. Soltis and D. E. Soltis (Berling: Springer), 21–32. doi: 10.1007/978-3-642-31442-1\_2

- Birchler, J. A., Yao, H., Chudalayandi, S., Vaiman, D., and Veitia, R. A. (2010). Heterosis. *Plant Cell* 22, 2105–2112.
- Borsos, O. (1967). Über einige *Rorippa*- und *Veronica*-Arten. *Acta Bot. Acad. Sci. Hung.* 13, 1–10. doi: 10.1007/978-3-319-23534-9\_1
- Brochmann, C., Brysting, A. K., Alsos, I. G., Borgen, L., Grundt, H. H., Scheen, A. C., et al. (2004). Polyploidy in arctic plants. *Biol. J. Linn. Soc.* 82, 521–536. doi: 10.1111/j.1095-8312.2004.00337.x
- Brunsfeld, S. J., Sullivan, J., Soltis, D. E., and Soltis, P. S. (2001). “A comparative phylogeography of northwestern North America: a synthesis,” in *Integrating Ecology and Evolution in a Spatial Context*, eds J. Silvertown and J. Antonovics (Oxford: Blackwell Science), 319–340.
- Buerkle, C. A. (2005). Maximum-likelihood estimation of a hybrid index based on molecular markers. *Mol. Ecol. Notes* 5, 684–687. doi: 10.1111/j.1471-8286.2005.01011.x
- Carvalho, S. B., Torres, J., Tarroso, P., and Velo-Antón, G. (2019). Genes on the edge: a framework to detect genetic diversity imperiled by climate change. *Glob. Chang. Biol.* 25, 4034–4047. doi: 10.1111/gcb.14740
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. doi: 10.1111/mec.12354
- Caye, K., Deist, T. M., Martins, H., Michel, O., and François, O. (2016). TESS3: fast inference of spatial population structure and genome scans for selection. *Mol. Ecol. Res.* 16, 540–548. doi: 10.1111/1755-0998.12471
- Chapman, M., and Abbott, R. (2010). Introgression of fitness genes across a ploidy barrier. *New Phytol.* 186, 63–71. doi: 10.1111/j.1469-8137.2009.03091.x
- Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., and Flouri, T. (2019). ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* 37, 291–294. doi: 10.1093/molbev/msz189
- Earl, D. A., and von Holdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Res.* 4, 359–361. doi: 10.1007/s12686-011-9548-7
- Ellis, W. N., and Ellis-Adam, A. C. (1994). To make a meadow it takes a clover and a bee: the entomophilous flora of N.W. Europe and its insects. *Bijdragen Dierkunde* 63, 193–220. doi: 10.1163/26660644-06304001
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354
- Fischer, M. (1974). Beitrag zu einer systematischen Neubearbeitung der Gruppe um *Pseudolysimachion spicatum* (L.) Opiz (*Veronica spicata* L.). *Phyton*. 16, 29–47.
- Fourcade, Y., Engler, J. O., Rödder, D., and Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS One* 9:e97122. doi: 10.1371/journal.pone.0097122
- Franzke, A., Hurka, H., Janssen, D., Neuffer, B., Friesen, N., Markov, M., et al. (2004). Molecular signals for late tertiary/early quaternary range splits of an Eurasian steppe plant: *Claudia aprica* (Brassicaceae). *Mol. Ecol.* 13, 2789–2795. doi: 10.1111/j.1365-294x.2004.02272.x
- Friesen, N., Žerdoner Čalasan, A., Neuffer, B., German, D. A., Markov, M., and Hurka, H. (2020). Evolutionary history of the Eurasian steppe plant *Schivereckia podolica* (Brassicaceae) and its close relatives. *Flora* 268:151602. doi: 10.1016/j.flora.2020.151602
- Gompert, Z., and Buerkle, C. A. (2009). A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Mol. Ecol.* 18, 1200–1224. doi: 10.1111/j.1365-294x.2009.04098.x
- Gompert, Z., and Buerkle, C. A. (2010). INTROGRESS, A software package for mapping components of isolation in hybrids. *Mol. Ecol. Res.* 10, 378–384. doi: 10.1111/j.1755-0998.2009.02733.x
- Gompert, Z., and Buerkle, C. A. (2012). BGC: Software for Bayesian estimation of genomic clines. *Mol. Ecol. Res.* 12, 1168–1176.
- Graze, H. (1933). Die chromosomalen verhältnisse in der sektion *Pseudolysimachia* Koch der gattung *Veronica*. *Jahrb. Wiss. Bot.* 77, 507–559.
- Gruber, B., and Georges, A. (2019). *dartR: Importing and Analysing SNP and Silicodart Data Generated by Genome-Wide Restriction Fragment Analysis. R Package Version 1.1.11*. Available online at: <https://CRAN.R-project.org/package=dartR> (accessed February 1, 2020).
- Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., and Li, W. H. (2003). Role of duplicate genes in genetic robustness against null mutations. *Nature* 421, 63–66. doi: 10.1038/nature01198
- Härle, A. (1932). Die arten und formen der *Veronica* - Sektion *Pseudolysimachia* K OCH auf Grund systematischer und experimenteller Untersuchungen. *Biblioth. Bot.* 104, 1–86.
- Harrison, H. B., Berumen, M. L., Saenz-Agudelo, P., Salas, E., Williamson, D. H., and Jones, G. P. (2017). Widespread hybridization and bidirectional introgression in sympatric species of coral reef fish. *Mol. Ecol.* 26, 5692–5704. doi: 10.1111/mec.14279
- Harrison, S., and Noss, R. (2017). Endemism hotspots are linked to stable climatic refugia. *Ann. Bot.* 119, 207–214. doi: 10.1093/aob/mcw248
- Herben, T., Suda, J., and Klimešová, J. (2017). Polyploid species rely on vegetative reproduction more than diploids: a re-examination of the old hypothesis. *Ann. Bot.* 120, 341–349. doi: 10.1093/aob/mcx009
- Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature* 405, 907–913. doi: 10.1038/35016000
- Hewitt, G. M. (2001). Speciation, hybrid zones and phylogeography-or seeing genes in space and time. *Mol. Ecol.* 10, 537–549. doi: 10.1046/j.1365-294x.2001.01202.x
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. doi: 10.1002/joc.1276
- Hijmans, R. J., Guarino, L., and Mathur, P. (2012). *DIVA-GIS Version 7.3.0*.
- Hijmans, R. J., Williams, E., and Vennes, C. (2019). *CRAN. R-Project.org/Package=Geosphere, v. 1.5-10*.
- Hurka, H., Friesen, N., Bernhardt, K. G. G., Neuffer, B., Smirnov, S. V., Shmakov, A. I., et al. (2019). The Eurasian steppe belt. Status quo, origin and evolutionary history. *Turczaninowia* 22, 5–71. doi: 10.14258/turczaninowia.22.3.1
- Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233
- Jørgensen, M. H., Ehrich, D., Schmickl, R., Koch, M. A., and Brysting, A. K. (2011). Interspecific and interploidal gene flow in Central European *Arabidopsis* (Brassicaceae). *BMC Evol. Biol.* 11:346. doi: 10.1186/1471-2148-11-346
- Kadereit, J. W. (2015). The geography of hybrid speciation in plants. *Taxon* 64, 673–687. doi: 10.12705/644.1
- Kampny, C. M. (1995). Pollination and flower diversity in Scrophulariaceae. *Bot. Rev.* 61, 350–366. doi: 10.1007/bf02912622
- Keppel, G., Van Niel, K. P., Wardell-Johnson, G. W., Yates, C. J., Byrne, M., Mucina, L., et al. (2012). Refugia: identifying and understanding safe havens for biodiversity under climate change. *Glob. Ecol. Biogeogr.* 21, 393–404. doi: 10.1111/j.1466-8238.2011.00686.x
- Khan, G., Franco, F. F., Silvae, G. A. R., Bombonato, J. R., Machadoc, M., Afonso, D. P., et al. (2020). Maintaining genetic integrity with high promiscuity: frequent hybridization with low introgression in multiple hybrid zones of *Melocactus* (Cactaceae). *Mol. Phyl. Evol.* 142:106642. doi: 10.1016/j.ympev.2019.106642
- Knaus, B. J., and Grünwald, N. J. (2017). VCFR: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Res.* 17, 44–53. doi: 10.1111/1755-0998.12549
- Knuth, P. (1909). *Handbook of Flower Pollination*, Vol. 3. Oxford: Clarendon Press.
- Kosachev, P. A., Albach, D., and Ebel, A. L. (2015). Check-list of *Veronica* subgen. *Pseudolysimachium* (Plantaginaceae) of siberia. *Turczaninowia* 18, 84–95. doi: 10.14258/turczaninowia.18.3.4
- Kosachev, P. A., Behçet, L., Mayland-Quellhorst, E., and Albach, D. C. (2016). Analyzing reticulate relationships using cpDNA and pyrosequenced ITS1 as exemplified by *Veronica* subgen. *Pseudolysimachium* (Plantaginaceae). *Syst. Bot.* 41, 105–119. doi: 10.1600/036364416x90697
- Kosachev, P. A., Mayland-Quellhorst, E., and Albach, D. C. (2019). Hybridization among species of *Veronica* subgen. *Pseudolysimachium* in the Altai detected by SRAP markers. *Nord. J. Bot.* 37, e02209.

- López-González, N., Bobo-Pinilla, J., Padilla-García, N., Loureiro, J. C. M., Castro, S., Rojas-Andrés, B. M., et al. (2020). Genetic similarities versus morphological resemblance: unraveling a polyploid complex in a Mediterranean biodiversity hotspot. *Molec. Phylogenet. Evol.* doi: 10.1016/j.ympev.2020.107006. [Epub ahead of print].
- Löve, A. (1977). IOPB chromosome number reports LVI. *Taxon* 26, 257–274. doi: 10.1002/j.1996-8175.1977.tb04167.x
- Lugon-Moulin, N., and Hausser, J. (2002). Phylogeographical structure, postglacial recolonization and barriers to gene flow in the distinctive Valais chromosome race of the common shrew (*Sorex araneus*). *Mol. Ecol.* 11, 785–794. doi: 10.1046/j.1365-294x.2002.01469.x
- Mai, D. H. (1995). *Tertiäre Vegetationsgeschichte Europas: Methoden und Ergebnisse*. Heidelberg: Spektrum Akademischer Verlag.
- McKinney, G. J., Waples, R. K., Seeb, L. W., and Seeb, J. E. (2017). Paralogues are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Mol. Ecol. Res.* 17, 656–669. doi: 10.1111/1755-0998.12613
- Merow, C., Smith, J. A., and Silander, J. R. (2013). A practical guide to MAXENT for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36, 1058–1069. doi: 10.1111/j.1600-0587.2013.07872.x
- Meudt, H. M., Rojas-Andrés, B. M., Prebble, J. M., Low, E., Garnock-Jones, P. J., and Albach, D. C. (2015). Is genome downsizing associated with diversification in polyploid lineages of *Veronica*? *Bot. J. Linn. Soc.* 178, 243–266. doi: 10.1111/boj.12276
- Moyle, L. C., and Nakazato, T. (2010). Hybrid incompatibility “snowballs” between *Solanum* species. *Science* 329, 1521–1523. doi: 10.1126/science.1193063
- Muellner-Riehl, A. N. (2019). Mountains as evolutionary arenas: patterns, emerging approaches, paradigm shifts, and their implications for plant phylogeographic research in the Tibeto-Himalayan Region. *Front. Plant Sci.* 10:195. doi: 10.3389/fpls.2019.00195
- Müller, H. (1881). *Alpenblumen, ihre Befruchtung Durch Insekten und ihre Anpassungen an dieselben*. Leipzig: Wilhelm Engelmann.
- Nieto Feliner, G. (2014). Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspect. Plant Ecol. Evol. Syst.* 16, 265–278. doi: 10.1016/j.ppees.2014.07.002
- Nolte, A. W., Gompert, Z., and Buerkle, C. A. (2009). Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Mol. Ecol.* 18, 2615–2627. doi: 10.1111/j.1365-294x.2009.04208.x
- Padilla-García, N., Rojas-Andrés, B. M., López-González, N., Castro, M., Castro, S., Loureiro, J., et al. (2018). The challenge of species delimitation in the diploid-polyploid complex *Veronica* subsection *Pentasepalae*. *Mol. Phyl. Evol.* 119, 196–209. doi: 10.1016/j.ympev.2017.11.007
- Parisod, C. (2008). Postglacial recolonisation of plants in the western Alps of Switzerland. *Bot. Helv.* 118, 1–12. doi: 10.1007/s00035-008-0825-3
- Paule, J., Wagner, N. D., Weising, K., and Zizka, G. (2017). Ecological range shift in the polyploid members of the South American genus *Fosterella* (Bromeliaceae). *Ann. Botany* 120, 233–243. doi: 10.1093/aob/mcw245
- Pembleton, L. W., Cogan, N. O. I., and Forster, J. W. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Res.* 13, 946–952. doi: 10.1111/1755-0998.12129
- Phillips, S. J., Anderson, R., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026
- Phillips, S. J., Dudík, M., and Schapire, R. E. (2019). *Maxent Software for Modeling Species Niches and Distributions (Version 3.4.0)*. Available online at: [http://biodiversityinformatics.amnh.org/open\\_source/maxent/](http://biodiversityinformatics.amnh.org/open_source/maxent/) (accessed February 1, 2020).
- Pickens, K. A., Affolter, J. M., Wetzstein, H. Y., and Wolf, J. H. (2003). Enhanced seed germination and seedling growth of *Tillandsia eizii* in vitro. *Hort. Sci.* 38, 101–104. doi: 10.21273/hortsci.38.1.101
- Pigott, C. D., and Walters, S. M. (1954). On the interpretation of the discontinuous distributions shown by certain British species of open habitats. *J. Ecol.* 42, 95–116. doi: 10.2307/2256981
- Porrás-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, A., and Lareu, M. V. (2013). An overview of STRUCTURE: applications, parameter settings, and supporting software. *Front. Genet.* 29:98. doi: 10.3389/fpls.2019.00098
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Pustahija, F., Brown, S. C., Bogunić, F., Bašić, N., Muratović, E., Ollier, S., et al. (2013). Small genomes dominate in plants growing on serpentine soils in West Balkans, an exhaustive study of 8 habitats covering 308 taxa. *Plant Soil* 373, 427–453. doi: 10.1007/s11104-013-1794-x
- R Core Team (2019). *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Raitanen, P. R. (1967). Taxonomic studies on *Veronica longifolia* and *V. spicata* in Eastern Fennoscandia. *Ann. Bot. Fenn.* 4, 471–485.
- Ramsey, J., and Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Evol. Syst.* 29, 467–501.
- Rochette, N. C., Rivera-Colón, A. G., and Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADSeq-based population genomics. *Mol. Ecol.* 28, 4737–4754. doi: 10.1111/mec.15253
- Rohlf, F. J. (2009). *NTSYSpc: Numerical Taxonomy System. ver. 2.21c*. Setauket, NY: Exeter Software.
- Rojas-Andrés, B. M., Padilla-García, N., Pedro, M. D., López-González, N., Delgado, L., Albach, D. C., et al. (2020). Environmental differences are correlated with the distribution pattern of cytotypes in *Veronica* subsection pentasepalae at a broad scale. *Ann. Bot.* 125, 471–484.
- Scalone, R., Kolf, M., and Albach, D. C. (2013). Mating system variation in veronica (Plantaginaceae): inferences from pollen/ovule ratios and other reproductive traits. *Nord. J. Bot.* 31, 372–384. doi: 10.1111/j.1756-1051.2012.01706.x
- Schönswetter, P., Lachmayer, M., Lettner, C., Prehler, D., Rechner, S., Sonnleitner, M., et al. (2007). Sympatric diploid and hexaploid cytotypes of *Senecio carniolicus* (Asteraceae) in the Eastern Alps are separated along an altitudinal gradient. *J. Pl. Res.* 120, 721–725. doi: 10.1007/s10265-007-0108-x
- Siadjeu, C., Mayland-Quellhorst, E., and Albach, D. C. (2018). Genetic diversity and population structure of trifoliolate yam (*Dioscorea dumetorum* Kunth) in Cameroon revealed by genotyping-by-sequencing (GBS). *BMC Plant Biol.* 18:359. doi: 10.1186/s12870-018-1593-x
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stebbins, G. L. (1971). *Chromosomal Evolution in Higher Plants*. London: Edward Arnold.
- Stewart, J. R., and Lister, A. M. (2001). Cryptic northern refugia and the origins of the modern biota. *Trends Ecol. Evol.* 16, 608–613. doi: 10.1016/s0169-5347(01)02338-2
- Suchan, T., Malicki, M., and Ronikier, M. (2019). Relict populations and Central European glacial refugia: the case of *Rhododendron ferrugineum* (Ericaceae). *J. Biogeogr.* 46, 392–404. doi: 10.1111/jbi.13512
- Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., and Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7, 453–464. doi: 10.1046/j.1365-294x.1998.00289.x
- Tkach, N. V., Hoffmann, M. H., Röser, M., Korobkov, A. A., and von Hagen, K. B. (2008). Parallel evolutionary patterns in multiple lineages of Arctic *Artemisia* L. (Asteraceae). *Evolution* 62, 184–198.
- Trávníček, B. (2000). Notes on the taxonomy of *Pseudolysimachion longifolium* complex (Scrophulariaceae). *Thaiszia* 10, 1–26. doi: 10.1007/978-3-319-23534-9\_1
- Trávníček, B., Lysák, M. A., Čihalíková, J., and Doležel, J. (2004). Karyo-taxonomic study of the genus *Pseudolysimachion* (Scrophulariaceae) in the Czech Republic and Slovakia. *Folia Geobot. Phytotax.* 39, 173–203. doi: 10.1007/bf02805245
- Tsybalyuk, Z. M., and Mosyakin, S. L. (2013). *Atlas of Pollen Grains of Representatives of Plantaginaceae and Scrophulariaceae*. Kyiv: Nash Format Publisher.
- Volkova, P. A., Burlakov, Y. A., and Schanzer, I. A. (2020). Genetic variability of *Prunus padus* (Rosaceae) elaborates “a new Eurasian phylogeographical paradigm”. *Plant Syst. Evol.* 306:1. doi: 10.1007/978-3-319-77088-8\_110-2
- Warren, D. L., Glor, R. E., and Turelli, M. (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62, 2868–2883. doi: 10.1111/j.1558-5646.2008.00482.x
- Warren, D. L., Glor, R. E., and Turelli, M. (2010). ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* 33, 607–611.
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.2307/2408641
- Wilson, G. B., Houston, L., Whittington, W. J., and Humphries, R. N. (2000). *Veronica spicata* L. ssp. *spicata* and ssp. *hybrida* (L.) Gaudin

- (*Pseudolysimachium spicatum* (L.) Opiz). *J. Ecol.* 88, 890–909. doi: 10.1046/j.1365-2745.2000.00501.x
- Winter, C., Lehmann, S., and Diekmann, M. (2008). Determinants of reproductive success: a comparative study of five endangered river corridor plants in fragmented habitats. *Biol. Conserv.* 141, 1095–1104. doi: 10.1016/j.biocon.2008.02.002
- Wright, S. (1949). The genetical structure of populations. *Ann. Eugen.* 15, 323–354. doi: 10.1111/j.1469-1809.1949.tb02451.x
- Wringe, B. F., Stanley, R. R., Jeffery, N. W., Anderson, E. C., and Bradbury, I. R. (2016). *Parallelnewhybrid: An R Package for the Parallelization of Hybrid Detection Using NEWHYBRIDS. R Package Version 0.0.0.9002*. Available online at: <https://github.com/bwringe/parallelnewhybrid> (accessed February 1, 2020).
- Wringe, B. F., Stanley, R. R., Jeffery, N. W., Anderson, E. C., and Bradbury, I. R. (2017). parallelnewhybrid: an R package for the parallelization of hybrid detection using new hybrids. *Mol. Ecol. Res.* 17, 91–95. doi: 10.1111/1755-0998.12597
- Yang, P. M., Huang, Q. C., Qin, G. Y., Zhao, S. P., and Zhou, J. G. (2014). Different drought-stress responses in photosynthesis and reactive oxygen metabolism between autotetraploid and diploid rice. *Photosynthetica* 52, 193–202. doi: 10.1007/s11099-014-0020-2
- Žerdoner Čalasan, A., Seregin, A. P., Hurka, H., Hofford, N. P., and Neuffer, B. (2019). The Eurasian steppe belt in time and space: phylogeny and historical biogeography of the false flax (*Camelina* Crantz, Camelinaeae, Brassicaceae). *Flora* 260:151477. doi: 10.1016/j.flora.2019.151477
- Zohren, J., Wang, N., Kardailsky, I., Borrell, J. S., Joecker, A., Nichols, R. A., et al. (2016). Unidirectional diploid-tetraploid introgression among British birch trees with shifting ranges shown by restriction site-associated markers. *Mol. Ecol.* 25, 2413–2426. doi: 10.1111/mec.13644

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with one of the authors SM.

Copyright © 2021 Buono, Khan, von Hagen, Kosachev, Mayland-Quellhorst, Mosyakin and Albach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.