Homoplastic versus xenoplastic evolution: exploring the emergence of key intrinsic and extrinsic traits in the montane genus *Soldanella* (Primulaceae)

Ivan Rurik[1,†], Andrea Melichárková[2,†], Eliška Gbúrová Štubová[2,3], Jaromír Kučera[2], Judita Kochjarová[4], Ovidiu Paun[5] ID, Peter Vďačný[1] and Marek Slovák[2,6,*] ID

[1]*Department of Zoology, Comenius University Bratislava, Ilkovičova 6, 842 15 Bratislava, Slovak Republic,*
[2]*Institute of Botany, Plant Science and Biodiversity Centre, Slovak Academy of Sciences, Dúbravská cesta 9, 845 23 Bratislava, Slovak Republic,*
[3]*Slovak National Museum, Natural History Museum, Vajanského nábrežie 2, 810 06 Bratislava, Slovak Republic,*
[4]*Department of Phytology, Faculty of Forestry, Technical University Zvolen, Masarykova 24, 960 53 Zvolen, Slovak Republic,*
[5]*Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, 1030 Vienna, Austria, and*
[6]*Department of Botany, Charles University, Benátská 2, 128 01 Prague, Czech Republic*

SUMMARY

Specific ecological conditions in the high mountain environment exert a selective pressure that often leads to convergent trait evolution. Reticulations induced by incomplete lineage sorting and introgression can lead to discordant trait patterns among gene and species trees (hemiplasy/xenoplasy), providing a false illusion that the traits under study are homoplastic. Using phylogenetic species networks, we explored the effect of gene exchange on trait evolution in *Soldanella*, a genus profoundly influenced by historical introgression. At least three features evolved independently multiple times: the single-flowered dwarf phenotype, dysploid cytotype, and ecological generalism. The present analyses also indicated that the recurring occurrence of stoloniferous growth might have been prompted by an introgression event between an ancestral lineage and a still extant species, although its emergence via convergent evolution cannot be completely ruled out. Phylogenetic regression suggested that the independent evolution of larger genomes in snowbells is most likely a result of the interplay between hybridization events of dysploid and euploid taxa and hostile environments at the range margins of the genus. The emergence of key intrinsic and extrinsic traits in snowbells has been significantly impacted not only by convergent evolution but also by historical and recent introgression events.

Keywords: phylogenetic networks, trait reconstruction, convergent evolution, phylogenetic regression, genome size, *Soldanella*, European alpine system.

INTRODUCTION

The significance of hybridization in plant evolution and diversity has been extensively researched and disputed over the past few decades (Soltis & Soltis, 2009; Stull et al., 2023). Nevertheless, disentangling phylogenetic relationships and trait evolution in phylogenies heavily influenced by historical hybridization and introgression can be a challenging task, requiring continuous and thorough research (e.g., Loiseau et al., 2021), even with the recent success rendered by novel phylogenetic approaches (Bastide et al., 2018; Hibbins et al., 2023; Karimi et al., 2020; Solís-Lemus et al., 2017; Solís-Lemus & Ané, 2016; Teo et al., 2023; Wang et al., 2021). Extensive hybridization and introgression obscure evolutionary history *per se*, providing an inaccurate perspective on trait evolution (Degnan & Rosenberg, 2009; Hibbins et al., 2023; Maddison, 1997; Maddison & Knowles, 2006; Mallet et al., 2016; Wang et al., 2021). Trait reconstruction may be especially difficult in lineages influenced by ancient hybridization, which typically happens between the progenitors of modern lineages and species (Stull et al., 2023). Furthermore, topological discordance between species and gene trees primarily caused by idiosyncratic lineage sorting (ILS), hybridization, gene duplication, or horizontal gene transfer might, in the

case of ancestral trait reconstruction, result in character state patterns similar to those inferred for homoplastic traits, a phenomenon known as hemiplasy (Avise & Robinson, 2008). Hemiplasy is a concept initially linked to ILS (Avise & Robinson, 2008) and thus, more recently, the term 'xenoplasy' has been introduced to characterize cases where a trait pattern arises from inheritance across species boundaries through hybridization or introgression (Wang et al., 2021). Hemiplastic and xenoplastic traits thus may provide an illusion of independent and recurring origins of characters, leading to false inferences of convergent and parallel evolution (Hibbins et al., 2020; Mendes et al., 2018; Mendes & Hahn, 2016; Wang et al., 2021).

Due to historically oscillating conditions, radiated lineages from high mountain environments represent important models for investigating the impact of gene flow on evolutionary history and trait evolution. High mountain lineages are capable of rapidly adapting to a wide range of environments, resulting in adaptive phenotypic evolution (e.g., Pouchon et al., 2018). This is especially noticeable in the dynamic and challenging habitats of the arctic–alpine environment, which allows for polytopic convergent adaptation to similar ecological conditions even at the species level. As a result, independent biological entities respond to essentially similar selection pressures during parallel ecotypic divergence by the emergence of similar phenotypic changes (e.g., Brandrud et al., 2017; Knotek et al., 2020; Szukala et al., 2022; Wos et al., 2022). These ecotypes are considered early stages of speciation and serve as a starting point for further diversification. Moreover, if gene exchange events occur in such high mountain plant lineages, convergence in certain features may be caused by introgression among them. Overlooking this fact may lead to erroneous conclusions, as it can be challenging to distinguish between genuine convergencies and the consequences of introgression events.

In the present study, we examined the impact of gene flow on trait evolution within the mountain-dwelling genus *Soldanella* (snowbells, Primulaceae), which has been considerably influenced by historical introgression (Slovák et al., 2023). *Soldanella* includes up to 25 taxa occupying mountain habitats in the European Alpine System (Bellino et al., 2015; Niederle, 2003, 2016; Ozenda, 1985; Zhang et al., 2001; Zhang & Kadereit, 2002). Snowbells are long-lived perennials with caespitose or stoloniferous growth and showy flowers appearing immediately after snow melts. They form two phenotypic groups corresponding to the two previously recognized but genetically unsupported sections, namely 'Tubiflores' and 'Soldanella' (Figure S1; Slovák et al., 2023; Zhang et al., 2001). The 'Tubiflores' group includes *S. minima* and *S. pusilla*. They are characterized by overall dwarfism, single white to pink bell-shaped flowers, and capsules opening with five teeth. The

dwarf phenotype with a reduced flower number and floral morphology of *S. minima* and *S. pusilla* is hypothesized to have evolved twice in parallel as an adaptation to the arctic–alpine environment (Steffen & Kadereit, 2014; Zhang et al., 2001). In contrast, the 'Soldanella' group comprises taxa that bear several violet to violet-blue funnel-shaped flowers per stem and capsules that open with 10 teeth. They predominantly inhabit humid forest habitats, often close to water bodies (Zhang et al., 2001; Zhang & Kadereit, 2002). However, many species from the 'Soldanella' group are able to colonize habitats in the alpine zone as well (Kliment & Valachovič, 2007; Slovák et al., 2023; Štubňová et al., 2017; Valachovič et al., 2019). Even though the alpine-dwelling populations (ecotypes) are remarkably smaller than their forest counterparts in the 'Soldanella' group, their flower numbers and floral morphologies are not as conspicuously reduced as in the 'Tubiflores' group. Thus, it appears that this ecological setup has evolved independently for phylogenetically distant species inhabiting different mountain ranges. From a karyological perspective, the genus is homoploid with rather low genome size variation, ranging from 2C = 2.95–3.99 pg (Štubňová et al., 2017). The existence of two major cytotypes has been proven, namely euploid with $2n = 40$ and dysploid with $2n = 38$ chromosomes (Figure S1). The dysploid cytotype seems to have evolved in four mostly unrelated species, and so it is thought to be another example of convergent evolution in this genus (Slovák et al., 2023; Štubňová et al., 2017). Although only modest within-genus variation in genome size was uncovered, a geographically associated increase in genome size of some species living along the limits of the genus distribution range is well noticeable. Genome expansions might have evolved in response to a more challenging environment, particularly in more southern regions where climatic factors are not optimal for such moisture- and cold-adapted plants. Alternatively, the increase in genome size could also be attributable to hybridization and introgression events. Nonetheless, all assumptions on karyological and cytogenetic evolution within the genus have not been formally tested using rigorous data and analytical approaches hitherto.

The genus evolved in the Miocene epoch, but most speciation events are dated to the Quaternary period (Boucher, Casazza, et al., 2016; Boucher, Niklaus, et al., 2016; de Vos et al., 2014; Kadereit et al., 2004; Slovák et al., 2023; Zhang et al., 2001). It has been hypothesized that the ancestors of the genus first occupied mountain forests (Zhang et al., 2001). The forest *S. villosa* and the alpine species from the 'Tubiflores' group were suggested to be more ancestral, whereas other forest taxa inhabiting the eastern part of the distribution range (the Carpathians as well as the S and E Balkans) seemed to be more recent and less divergent (cf. Zhang et al., 2001). Both hypotheses

have recently received significant support from phylogenomic analyses, which have also revealed widespread introgression and hybridization impacting the majority of the currently recognized snowbell species. Gene exchange events happened frequently during all time horizons regardless of phenotype or ecology (Slovák et al., 2023). The study further indicated that Balkan species (*S. chrysosticta*, *S. pindicola*, and *S. rhodopaea*) are of hybridogeneous origin. Moreover, extensive introgression was detected both among high alpine species, as well as among Carpathian taxa with large ecological amplitude.

Despite previous investigations (Slovák et al., 2023; Steffen & Kadereit, 2014), the role of convergent evolution and introgression in the evolution of intrinsic ('Tubiflores' morphology, dysploid cytotype, stoloniferous growth, and genome size) and extrinsic (alpine habitat) traits within the genus *Soldanella* remain unclear. We generated here novel data matrices from the RAD-seq data obtained in a previous study (Slovák et al., 2023) and utilized a comparative phylogenetic network approach to determine the extent to which historical introgression impacted trait evolution in the *Soldanella* genus. Thus, we formally tested whether selected intrinsic and extrinsic traits evolved independently multiple times in various species or have been introgressed from specific relatives.

RESULTS

Phylogenomic analyses

Because species tree analyses require *a priori* grouping of accessions into species (lineages), species delimitations were tested for all *Soldanella* taxa but especially for those shown to be genetically non-monophyletic and/or whose morphological boundaries are known to be ambiguous. In Bayesian species delimitation analyses, scenarios based on the current taxonomy of the genus *Soldanella* received the highest marginal likelihood (Table S1) except for *S. angusta*, *S. rugosa*, and *S. marmarossiensis*, which were found to be non-monophyletic and could not be considered distinct taxa. Consequently, they were grouped into two genetically homogeneous entities referred to here as MAR A and MAR B.

The SNAPP analysis generated a topology (Figure S2) that was in overall agreement with the maximum likelihood tree obtained in our previous study (Slovák et al., 2023), albeit with four differing nodes. *Soldanella villosa* was depicted as the sister taxon to the remaining snowbells with a posterior probability (PP) of 1.00. A clade formed by the snow bed specialists *S. minima* and *S. alpina*, which co-clustered with *S. pusilla* in a sister position with full support, branched off first. However, statistical support for the *S. minima-S. alpina-S. pusilla* clade was very weak (PP 0.63). The remaining taxa from the Carpathians, Balkans, and the southern Apennines formed a

core cluster referred to as the 'CBA clade', with robust statistical support (PP 1.00). Within the CBA clade, *S. pindicola* from the Balkans diverged first, but in contrast to the maximum likelihood phylogeny, it was followed by the Apennine *S. calabrella-S. sacra* clade instead of *S. carpatica* from the Western Carpathians. The remaining Carpathian and East Balkan species formed a fully supported and well-resolved cluster but two terminal nodes that showed only low posterior probability values and involved *S. montana*, MAR A, and MAR B (Figure S2).

Reconstruction of discrete ancestral traits

Since the evolutionary history of the genus *Soldanella* is profoundly influenced by historical introgression, comparative phylogenetic network approaches, as implemented in the Julia package PhyloNetworks (Karimi et al., 2020), were used to reconstruct the evolution of four discrete characters (floral morphology, cytotype, growth form, and habitat). The TREEMIX network obtained from our previous study (Slovák et al., 2023) served as a scaffold for all reconstruction analyses. First, the major tree of the TREEMIX network was ultrametrized using all compatible nodes of the SNAPP species trees sampled from the posterior distribution of MCMC simulations of the SNAPP analysis as calibration points. Because the SNAPP consensus species tree differed from the major tree extracted from the TREEMIX network by four nodes, reticulations from the TREEMIX network were introduced to the SNAPP species tree as well. To account for the uncertainty in phylogeny inference, all reconstruction analyses were conducted both on the ultrametrized TREEMIX network (Figure 1a–d) and the reticulated SNAPP species tree (Figure S3a–d).

The equal rates substitution model was selected as the best fit model for all four discrete characters studied. Combining the results of all reconstruction analyses, the most recent common ancestor (MRCA) of *Soldanella* was most likely stoloniferous, had the typical 'Soldanella'-like phenotype with multiple violet to blue-violet flowers with internal scales, and euploid cytotype with $2n = 40$ chromosomes. It could live both at the alpine level and in forests, which indicates a large ecological amplitude for the MRCA of *Soldanella*. The MRCA of the CBA clade had the typical 'Soldanella'-like phenotype and the euploid cytotype, but was caespitose and most probably lived in a forest environment (Figure 1a–d; Figure S3a–d).

The evolution of the four analyzed traits could be summarized as follows: (i) The dwarfic 'Tubiflores'-like phenotype evolved at least twice independently, specifically in the *S. pusilla* and *S. minima* clades, both traditionally classified in the 'Tubiflores' section. This convergent evolution was linked with their alpine habitat, whose preference also emerged independently in their ancestors (Figure 1a,c; Figure S3a,c). (ii) The dysploid chromosome
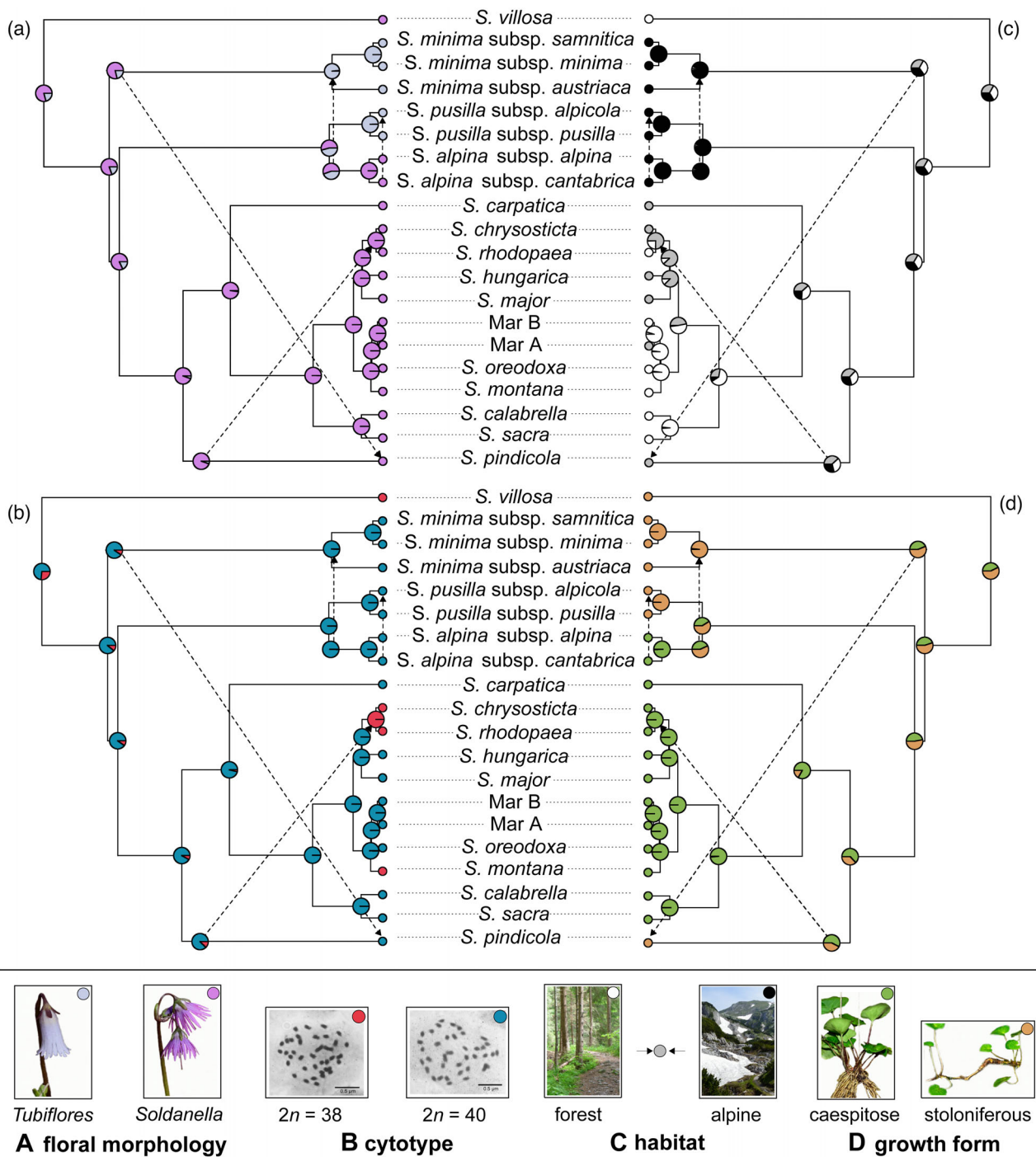
Figure 1. Ancestral state reconstruction of floral morphology (a), cytotype (b), habitat (c), and growth form (d) of snowbells on the ultrametrized TREEMIX network under the best fit equal rates model using the Julia package PhyloNetworks.

number ($2n = 38$) evolved as many as three times independently, namely in *S. villosa*, *S. montana*, and the *S. chrysosticta*-*S. rhodopaea* clade (Figure 1b; Figure S3b). (iii) Since reconstruction analyses revealed the presence of polymorphism in the ecological requirements only at more basal nodes, MAR A and the MRCA of the *S. chrysosticta*,

S. rhodopaea, *S. major*, and *S. hungarica* clade had to independently return to the ancestral ecological polymorphism of the MRCA of the *Soldanella* genus (Figure 1c; Figure S3c). (iv) Regarding the growth form, *S. villosa* as well as the *S. minima* and *S. pusilla* clades most likely maintained the ancestral stoloniferous growth. On the
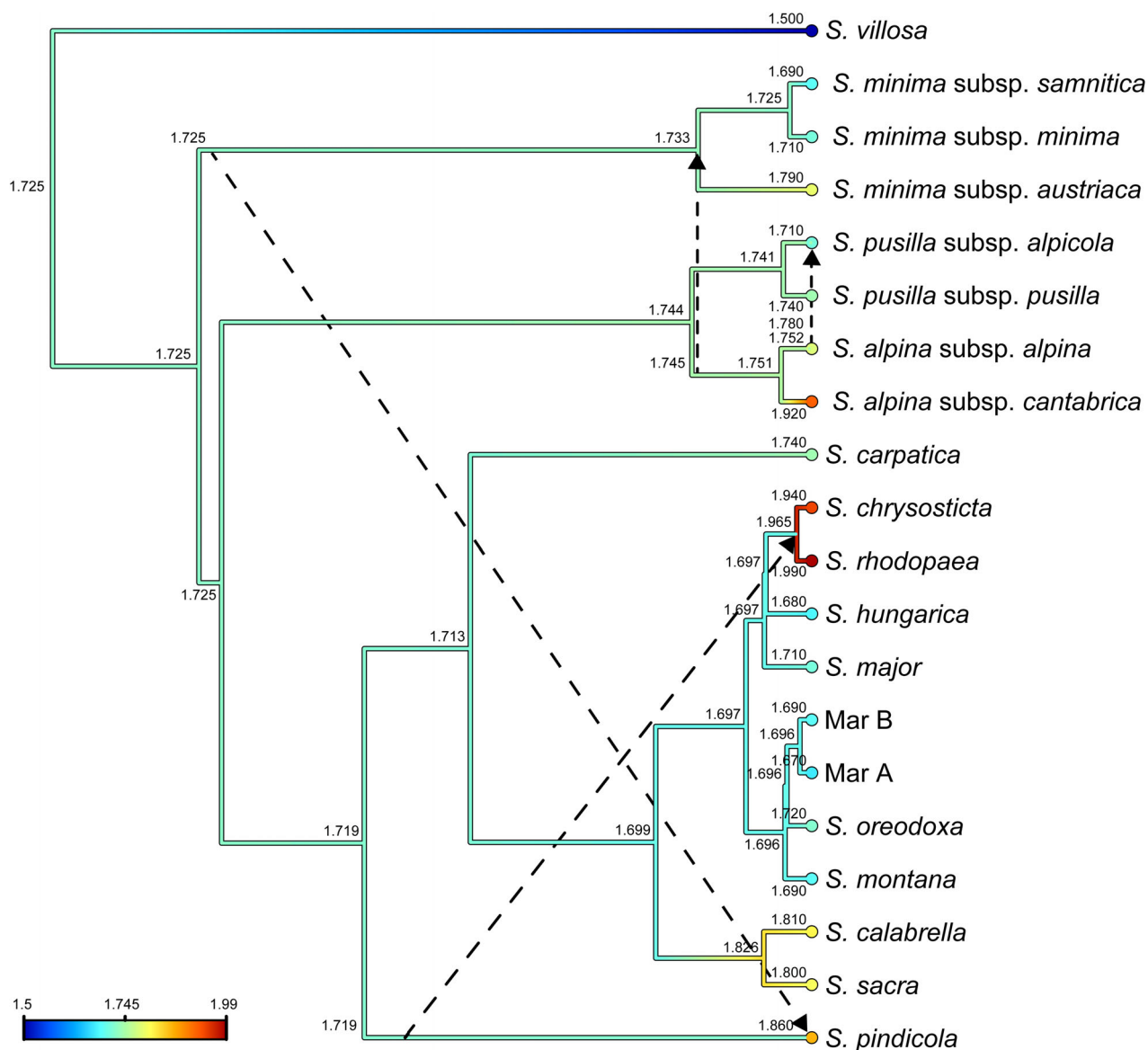
Figure 2. Ancestral state reconstruction of genome size of snowbells on the ultrametrized TREEMIX network under the best fit phylogenetic regression model using the Julia package PhyloNetworks.

other hand, this growth form might be a result of introgression in *S. pindicola* (Figure 1d; Figure S3d).

Reconstruction of genome size and phylogenetic regression

The reconstruction of the snowbell genome size revealed an intriguing fact: while MRCAs of most clades had smaller genomes than extant taxa, *S. villosa*, which branched off first, possesses the smallest genome of all congeners. The highest monoploid genome size values are found in *S. pindicola*, *S. alpina* subsp. *cantabrica*, and the *S. chrysosticta-S. rhodopaea* clade. On the other hand, *S. minima* subsp. *austriaca*, *S. alpina* subsp. *alpina*, as well as the South

Italian species *S. calabrella* and *S. sacra* show only a small increase in genome size (Figure 2; Figure S4).

Shifts in genome size were modeled on the ultrametrized TREEMIX network (Figure 2) and the reticulated SNAPP species tree (Figure S4), using phylogenetic regression with four factors (marginal area, habitat, hybridization event, and cytotype) and their interactions. The outstanding negative genome size shift of *S. villosa* from the MRCA of *Soldanella* makes it, apparently, an outlier (Figure 2; Figure S4). Since such a paradox was observed only in a single species, any explanation would be inherently overfitted. To further prevent overfitting, the significance of regression coefficients was tested also by cross-validation

in that model coefficients were estimated repeatedly, each time excluding (weight = 0) a different species. Phylogenetic regression brought the same best model for the ultrametrized TREEMIX network (log likelihood = 37.3096, AICc = −60.7692) and the reticulated SNAPP species tree (log likelihood = 36.8051, AICc = −59.7796) after *a priori* exclusion of *S. villosa*. AICc scores of all other tested models exceeded the overall best fit model by more than 2.00 and hence were not further considered. The best fit model was Pagel's lambda with λ set to 0.5, which increased individual species variance at the cost of reduced influence of phylogeny. It was inferred with the ridge regression hyperparameter λ set to 0.1. The best model explained the genome size shifts with only two predictors: marginal area and dysploid with hybridization event (Table 1).

When *S. villosa* was included in phylogenetic regression, the overall best model was almost identical to the previous one for both the ultrametrized TREEMIX network (log likelihood = 39.8003, AICc = −61.1414) and the reticulated SNAPP species tree (log likelihood = 39.2352, AICc = −60.0113). The model, however, exploited the fact that *S. villosa* is the only dysploid growing in the marginal area without a recent hybridization event and added this predictor to account for its small genome size. Model coefficients remained significant during cross-validation when the weight of *S. villosa* was at least 0.1. This is not surprising as the added predictor is specific only to *S. villosa*. All other species could be completely removed (weight = 0) during cross-validation without impacting the significance of the predictor coefficients. The best fit model was again Pagel's lambda with λ set to 0.5. It was inferred with the ridge regression hyperparameter λ set to 0.0001. The best model explained genome size shifts using three predictors: marginal area, dysploid in marginal area, and dysploid in marginal area with hybridization event. For all species except for *S. villosa*, the latter two predictors can be summed together having a cumulated effect of +0.1293 in case of the ultrametrized TREEMIX network and +0.1248 in case of the reticulated SNAPP species tree. Noteworthy, these values conform to the 'dysploid with hybridization event' coefficient in models inferred *a priori* without *S. villosa* (Table 1).

DISCUSSION

Phylogenetic network analyses enable the reconstruction of trait evolution in snowbells despite historical introgression

Neglecting introgression as a mechanism of trait inheritance could result in inaccurate inferences about the homoplastic nature of trait evolution (e.g., Hibbins et al., 2023; Solís-Lemus & Ané, 2016; Wang et al., 2021). Although comparative methods for phylogenetic networks have been recently established to address the reconstruction of trait evolution in the presence of gene flow (Bastide et al., 2018;

Table 1 Parameters of the best fit phylogenetic regression models for genome size of snowbells inferred from the ultrametrized TREEMIX network and the reticulated SNAPP species tree with Pagel's λ set to 0.5

	Coefficient	95% conf. interval	Std. error	t-test P-value	Coefficient	95% conf. interval	Std. error	t-test P-value
Model without *S. villosa*	Ultrametrized TREEMIX network				Reticulated SNAPP species tree			
Genome size of MRCA of *Soldanella* (intercept)	1.7280	1.6823–1.7738	0.021568	<1e-21	1.7236	1.6760–1.7712	0.022490	<1e-21
Marginal area	0.1271	0.0860–0.1682	0.019400	<1e-05	0.1303	0.0840–0.1765	0.021799	<1e-04
Dysploid with hybridization event	0.1221	0.0632–0.1811	0.027816	0.0005	0.1191	0.0561–0.1822	0.029750	0.0010
Model with *S. villosa*	Ultrametrized TREEMIX network				Reticulated SNAPP species tree			
Genome size of MRCA of *Soldanella* (intercept)	1.7253	1.6795–1.7710	0.021575	<1e-21	1.7206	1.6727–1.7685	0.022591	<1e-21
Marginal area	0.1326	0.0892–0.1761	0.020497	<1e-05	0.1361	0.0868–0.1853	0.023215	<1e-04
Dysploid in marginal area with hybridization event	0.4871	0.3751–0.5990	0.052798	<1e-07	0.4814	0.3663–0.5965	0.054287	<1e-06
Dysploid in marginal area	−0.3578	−0.4637 to −0.2519	0.049939	<1e-05	−0.3566	−0.4656 to −0.2475	0.051442	<1e-05

Dysploids with hybridization events were observed only in marginal area. *Soldanella villosa* is the single dysploid taxon occurring in a marginal area without a recent hybridization event, and, therefore, the latter two coefficients in models with *S. villosa* can be summed together for all taxa but *S. villosa*. They have a cumulative effect of +0.1293 in case of the ultrametrized TREEMIX network and +0.1248 in case of the reticulated SNAPP species tree. These values conform rather well to the 'dysploid with hybridization event' coefficient in models inferred without *S. villosa*.

Hibbins et al., 2023; Karimi et al., 2020; Solís-Lemus et al., 2017; Solís-Lemus & Ané, 2016; Teo et al., 2023; Wang et al., 2021), research in this field is still nascent. In the present study, a phylogenetic network approach, utilizing reticulation patterns obtained from TREEMIX and also mapped onto the SNAPP species tree, was employed to reconstruct the evolution of one quantitative (genome size) and four discrete (floral morphology, growth form, habitat, and cytotype) traits in a phylogeny profoundly influenced by historical introgression (Figures 1a–d and 2; Figures S3a–d and S4). It is important to note that only introgression events identified through TREEMIX analysis were considered. On the other hand, introgression events among the young Carpathian taxa (*S. major*, *S. hungarica*, and the *S. marmarossiensis-S. angusta-S. rugosa* group), detected by the D-suite statistics in our previous study (Slovák et al., 2023), were not considered in the present analyses. Nevertheless, since these taxa were placed in the most terminal clade and shared the same trait states for all studied characters, ignoring those introgression events did not impact the present interpretations.

We demonstrate here that both convergent evolution and historical introgression played roles in shaping trait evolution in snowbells. Taking into account, the intensity and frequency of historical gene flow previously identified (Slovák et al., 2023), one would anticipate a large impact of gene flow on trait evolution. However, this does not seem to be the case. The present analyses revealed that three (floral morphology, cytotype, and habitat) out of the four studied discrete traits underwent convergent evolution with minimal influence from historical introgression. On the other hand, introgression likely played a role in the occurrence of multiple genome size expansions (see below) and the independent emergence of stoloniferous growth.

A proper application of the comparative phylogenetic network approach requires careful consideration of the specific biological aspects of the system studied and the technical nuances of the algorithms utilized. Moreover, network approaches may face challenges in identifying reticulation signals when the level of introgression in a biological system is weak (cf. Wang et al., 2021). Nonetheless, the potential for traits to be introgressed from parental species is very limited in such cases. A converse situation has been observed in snowbells, whereas the presence of introgression signals was identified recurrently through various statistical techniques (Slovák et al., 2023). A clear distinction between modes of trait inheritance may be hindered when traits manifest repeatedly among more distantly related species (e.g., Hibbins et al., 2023; Karimi et al., 2020; Wang et al., 2021). Furthermore, distinguishing among homoplastic, hemiplastic, or xenoplastic trait emergence can be intricate especially when the studied biological system has been subjected to convergent evolution and when ILS and introgression have significantly affected

its evolutionary history (cf. Wang et al., 2021). Building on our previous work about snowbells (Slovák et al., 2023), we believe that the identified reticulation edges do not arise from hemiplasies induced by deep ILS but represent actual introgression events. Therefore, in the present study, the main challenge lies in distinguishing between convergent evolution and trait evolution induced by gene flow. In this context, only the stoloniferous growth form remains somewhat problematic. According to the present trait reconstruction analyses, the MRCA of *Soldanella* had stoloniferous growth and this growth form is still associated with early diverged lineages (Figure 1d; Figure S3d). While clonal growth in *S. villosa* and *S. minima* most likely represents an old plesiomorphy of the genus, the stoloniferous growth of the south Balkan *S. pindicola* seems to be introgressed from an unknown ancestor closely related to *S. minima* (Figure 1d; Figure S3d). Since this introgression event occurred between relatively distant taxa and, in addition, one of them is extinct now. This hybridization event also represents the case of 'ghost introgression' (Ottenburghs, 2020), which further complicates the final inference about the inheritance of stoloniferous growth in *Soldanella*. As evident from the present reconstructions, the MRCA of the CBA clade very likely had a caespitose growth form, favoring the hypothesis of convergent evolution of stoloniferous growth in *S. pindicola*. To sum up, we cannot unequivocally determine whether the emergence of clonal growth in *S. pindicola* is a result of historical introgression, convergent evolution, or even their combination (Figure 1d; Figure S3d). More robust genomic data and a reference genome are required to more reliably discern between these alternative scenarios.

Another factor that could significantly impact comparative phylogenetic network approaches is the outgroup selection. Karimi et al. (2020) demonstrated that the inclusion versus exclusion of a given outgroup might result in different reticulation patterns. In the present study, we could not use any of the closest relative outgroup genera (e.g., *Hottonia*, *Bryocarpum*, or *Omphalogramma*), as they are phylogenetically too distant and their genome architecture (especially the chromosome number) is fairly different from that of *Soldanella* (de Vos et al., 2014). The inclusion of *Hottonia* in phylogenetic inferences would very likely bring strongly biased results due to the very high percentage of missing data caused by the deep genetic divergence between *Hottonia* and *Soldanella* (for further details, see the 'Experimental Procedures' section below). Moreover, technical limitations such as potential errors in gene or species tree inference, analysis of a large number of taxa (Hibbins et al., 2023), or reliance on bi-allelic data exclusively and neglecting potential gene duplication (Wang et al., 2021), could significantly influence conclusions of studies employing the comparative phylonetwork approach and should be taken into account as well.

Homoplastic nature of phenotype, ecology, and cytotype in snowbells

Soldanella is distinguished by its distinctive flower morphology with five laciniate corolla lobes, a feature absent in other genera within the Primulaceae family (Hu & Kelso, 1996; Pax, 1905). Alongside the typical 'Soldanella'-like morphotype, the genus also includes a dwarf phenotype characterized by a reduced floral number and structure, exclusively associated with alpine environments (Figure S1; Zhang et al., 2001; Zhang & Kadereit, 2002). The evolution of the alpine phenotype in *S. minima* and *S. pusilla* was recognized as a striking case of convergence in snowbells (Steffen & Kadereit, 2014; Zhang et al., 2001). Our phylogenetic network reconstruction also strongly supports this view (Figure 1a; Figure S3a). In fact, the phenotypes of *S. minima* and *S. pusilla* conform to the criteria of the 'alpine syndrome', which is characterized by overall dwarfism, reduced biomass production, and larger, more conspicuous flowers, representing a typical phenotypic adaptation to the challenging alpine zone environments (Halbritter et al., 2018). While Zhang et al. (2001) proposed that all characters associated with the alpine syndrome in snowbed species are governed by one or a few major genes with pleiotropic effects, the genetic basis of alpine adaptation is likely to be more complex and polygenic, involving the interplay of multiple gene arrays responsible for physiological processes and phenotypic features (e.g., Bohutínská et al., 2021). Moreover, apart from numerous convergent adaptation genes, a significant proportion of genomic adaptation signatures may be species-specific (Rellstab et al., 2020). Notably, the colonization of alpine environments by snowbells is not limited to *S. minima* and *S. pusilla* (Figure 1c; Figure S3c). Similar to other mountain plants (e.g., Bohutínská et al., 2021; Knotek et al., 2020; Szukala et al., 2022; Wos et al., 2022), alpine adaptation has independently evolved multiple times within the CBA clade (Figure 1c; Figure S3c; cf. Kliment & Valachovič, 2007; Štubňová et al., 2017; Valachovič et al., 2019). Although these species exhibit an overall dwarfism trend with size reductions in all organs, they still maintain the typical 'Soldanella'-like appearance, with no reduction in the flower number or morphology in contrast to the snowbed specialists *S. minima* and *S. pusilla* (cf. Kochjarová et al., 2017; Slovák et al., 2023; Zhang et al., 2001; Zhang & Kadereit, 2002). This suggests that overall dwarfism evolves when a niche shifts from a forest to an alpine zone, but it does not necessarily coincide with a reduction in the flower number and morphology.

The entire genus demonstrates a remarkable ecological amplitude (Figure 1c; Figure S3c), which is also observed in its closest relatives, the genera *Bryocarpum* Hook. f. & Thoms. and *Omphalogramma* Franch. from Asia (Huang et al., 2019). Under this assumption, alpine-dwelling populations may repeatedly evolve in most Carpathian species (*S. carpatica*, *S. hungarica*, *S. major*, and the *S. marmarossiensis* group) by tapping into preexisting genetic variation (cf. Bohutínská et al., 2021) and activating adaptations to the alpine environment already present in ancestral snowbell lineages. However, a recent genomic study about the parallel adaptation of *Arabidopsis* lineages to the alpine environment suggested that as the divergence between compared lineages increases, the reuse of genes leading to parallelism decreases (Bohutínská et al., 2021). The evolution of the single-flowered morphotype with reduced floral morphology is limited to species from early branching lineages (Figure 1a; Figure S3a; Slovák et al., 2023). Thus, greater phylogenetic divergence may impose significant constraints on the reuse of genes responsible for the evolution of the single-flowered phenotype in the CBA clade, whereas an opposite scenario is expected for ecological preadaptation to the alpine zone, which is evident across all snowbell lineages. Furthermore, a notable distinction between *S. minima* and *S. pusilla* and the rest of snowbells lies in their ecological differences, as the former two species are exclusive snowbed specialists, fully adapted to thrive in the extremely cold environments of snowbed communities (Körner et al., 2019). It appears that this phenotype has been shaped by long-term intense environmental selective pressure during Quaternary glaciations, closely associated with their ability to survive in highly glaciated areas of European mountains. In contrast, other species possess a broad ecological range and can potentially survive at lower elevations and further southward, reducing the need for severe selection observed in snowbed species. For survival in alpine zone niches, dwarfism alone, without reduction in the flower number and structure has proven to be sufficient. Extensive mutual introgression among species and lineages of the CBA clade has very likely allowed genes responsible for adaptation to challenging environments to be easily exchanged. Gene flow might have thus fueled the convergent emergence of alpine ecotypes within the CBA clade. Conversely, historical gene flow between snowbed specialists and species from the CBA clade was low, greatly limiting the introgression of the single-flowered phenotype into species of the CBA clade (cf. Slovák et al., 2023).

Finally, while we believe that alpine zone-adapted dwarfic populations from the CBA clade represent well-adapted ecotypes, we cannot definitively rule out the possibility that they are to some extent the result of phenotypic plasticity rather than genuine adaptation. Considering the lack of a single-flowered morphotype in the CBA clade, phenotypic plasticity may provide a more parsimonious explanation. Establishing the genuine existence of these ecotypes will require explicit phenotypic investigations and differentiation from phenotypic plasticity through reciprocal

transplantation and cultivation experiments (e.g., Knotek et al., 2020; Walter et al., 2016, 2018; Wos et al., 2022).

Another notable instance of evolutionary convergencies in snowbells is the independent evolution of dysploid cytotypes at least three times independently, namely in *S. villosa*, *S. montana*, and the *S. chrysosticta–S. rhodopaea* clade (Figure 1b; Figure S3b; cf. Slovák et al., 2023; Štubňová et al., 2017). The recurrent emergence of dysploid cytotypes has been considered a major driver of diversification and speciation in various plant groups (e.g., Blöch et al., 2009; Moraes et al., 2016; Singhal et al., 2017). In contrast to numerous plant lineages (Melichárková et al., 2020; Singhal et al., 2017), each emergence of dysploid cytotype in snowbells led to a speciation event. Chorological and ecological observations indicate that the evolution of dysploid cytotype is primarily associated with forest environments and mid-altitudinal zones (Kochjarová et al., 2017; Valachovič et al., 2019; Zhang & Kadereit, 2002). Interestingly, this hypothesis contradicts the findings that the occurrence of meiotic abnormalities, which may result in the formation of aneuploid gametes necessary for dysploid cytotype evolution, is expected to be higher in colder and harsher environments (such as alpine zones and cold deserts) than in more climatically stable forest ecosystems (e.g., Tantray et al., 2021). Additionally, our recent study (Slovák et al., 2023) indicates that species with dysploid cytotype have not been reproductively isolated from euploid populations (see also Mandáková & Lysak, 2018; Melichárková et al., 2020; Winterfeld et al., 2018).

Polytopic, independent genome size expansion is most likely associated with introgression and marginal distribution range

Interestingly, *Soldanella* lineages with larger genome sizes, namely the *S. chrysosticta–S. rhodopaea* clade, *S. pindicola*, *S. alpina* subsp. *cantabrica* and partially also the *S. calabrella–S. sacra* clade, predominantly grow at the southern and western margins of the genus distribution (Slovák et al., 2023; Štubňová et al., 2017). Moreover, most of these taxa were also involved in historical introgression events (Slovák et al., 2023). These findings align with previous hypotheses proposed by Štubňová et al. (2017), suggesting that genome expansion could result from the proliferation of repetitive DNA elements (such as transposable elements and satellite DNA) following hybridization events, or the accumulation of these elements due to stressful environments at the edges of the genus distribution range. In the present study, we specifically examined various scenarios involving hybridization, marginal distribution, ecological preferences, cytotype, and their interactions. Using phylogenetic regression on networks, we revealed that the increase in genome size was driven mainly by two factors, marginal distribution area and a dysploid with hybridization event. Indeed, the most dramatic genome size

expansion occurred when dysploid species inhabiting the edge of their distribution range underwent hybridization, as observed in the Balkan species *S. chrysosticta* and *S. rhodopaea* (Figure 2; Figure S4).

The proliferation of mobile DNA elements may be triggered by a range of biotic and abiotic stresses, with the fusion of genomes from two separate species playing a crucial role (Balao et al., 2017; McLintock, 1984). This hybridization-induced genomic instability has been referred to as 'genome shock' and was previously considered a major driver of transposable element proliferation (McLintock, 1984). However, the concept of genomic shock remains a subject of debate, as it has been supported by some studies (e.g., Parisod et al., 2010; Senerchia et al., 2015; Tusso et al., 2022) but contradicted by others (Eriksson et al., 2022; Göbel et al., 2018; Kawakami et al., 2011; Ungerer & Kawakami, 2013). In conclusion, we propose that the independent evolution of larger genomes in snowbells is most likely a result of the interplay between hybridization events of dysploid taxa and hostile environments at the range boundaries of the genus, possibly in combination with reduced selection efficacy at low effective population sizes (cf. Bui & Grandbastien, 2012; Kumar & Bennetzen, 1999).

EXPERIMENTAL PROCEDURES

Sampling design, RAD-seq library preparation, sequencing, and data processing

The present analysis builds upon the data generated in a recent study by Slovák et al. (2023). Nevertheless, new data matrices were generated to meet the requirements of state reconstruction analyses. Except for *S. chrysosticta* subsp. *serbica*, which failed to be sequenced, all currently recognized *Soldanella* taxa, including samples from the *locus classicus* of each taxon, were analyzed here. Altogether, 119 individuals from 118 sampling locations were studied (Table S2; Slovák et al., 2023). Initially, *Hottonia palustris* was used as the closest available outgroup (de Vos et al., 2014). Unfortunately, preliminary analyses with various SNP filtering settings resulted in a high percentage of missing data (approximately 87%), likely due to deep genetic divergence between *Hottonia* and *Soldanella* as well as due to their different chromosome numbers. Consequently, *Hottonia* accessions had to be excluded from further analyses.

DNA extraction, single digest RAD-seq library preparation, and SNP filtering of raw data are described in detail elsewhere (Slovák et al., 2023). Briefly, SNPs were called using a pipeline that involved mapping data from each individual against a pseudoreference of RAD loci assembled de novo with Stacks (Catchen et al., 2013). Additional SNP filtration was conducted using Vcftools v.0.1.13 (Danecek et al., 2011), with a maximum allowed level of missing data at 5%. For specific information regarding the composition of RAD-seq data matrices, including the number of recovered loci and missing data, see Table S3.

Bayesian species delimitation

In the initial phase, a SNAPP coalescent-based species tree was built from the complete dataset and all taxa. The SNAPP tree

served as a scaffold for subsequent network-based trait reconstruction. Since species tree analyses require *a priori* grouping of accessions into species (lineages), species delimitations were tested for all *Soldanella* taxa but especially for those shown to be genetically non-monophyletic and/or whose morphological boundaries were reported to be ambiguous (cf. Slovák et al., 2023; Zhang et al., 2001; Zhang & Kadereit, 2002). To this end, Bayes factor delimitation (BFD) method of Leaché et al. (2014) was employed. This technique utilizes the SNAPP approach to directly estimate 'species' trees of different species circumscription from unlinked bi-allelic markers without explicitly reconstructing gene trees (Bryant et al., 2012). These coalescent analyses were performed on selected accessions representing all analyzed taxa (Table S1). To expedite computation time, accessions were divided into independent sub-datasets for each of the two BFD analyses (Matrix RAD_1A–D; Table S3). For each sub-dataset, two to six alternative circumscriptions were proposed and compared. The fundamental scenario adhered to the current taxonomy (cf. Zhang et al., 2001; Zhang & Kadereit, 2002), while other scenarios were modified versions in which at least two taxa were combined into a single taxonomic unit (Table S1). Two independent runs were conducted. Subsequently, the validity of the 'species' trees obtained from different scenarios was compared using Bayes factors. Each analysis was run with a gamma prior for the speciation rate (shape parameter = 2, scale parameter = 2), 20 different steps for path sampling, 100 000 generations per step in the MCMC simulations, sampling every 100th generation, and logging every 100th tree. The first 10% of samples were considered as burn-in.

Species tree inference

The Bayesian multispecies coalescence-based method implemented in SNAPP was employed to reconstruct the species tree (Bryant et al., 2012). Due to the high computational demands of Bayesian coalescent analysis, a reduced matrix (Matrix RAD_2) was used to build the species tree (Table S3). This reduced matrix included at least two individuals from each taxon and a single SNP per RAD-seq locus. In our previous study (Slovák et al., 2023), a maximum likelihood-based phylogeny revealed that accessions from *S. montana* subsp. *gubalowkae* and *S. tatricola* are genetically indistinguishable from *S. montana* and *S. marmarossiensis*, respectively. However, only a single accession was available for each of these taxa, preventing their inclusion in the Bayesian species delimitation analyses. Nevertheless, based on the maximum likelihood analyses, they could still be confidently assigned to the two respective taxa mentioned above. Additionally, the non-monophyletic *S. angusta*, *S. marmarossiensis*, and *S. rugosa* were grouped into two tentative taxonomic units, MAR A and MAR B. The SNAPP analysis was run for 25 million generations, sampling every 100th generation and logging every 100th tree. Following the recommendations of Bryant et al. (2012) and Drummond and Bouckaert (2015), a Yule model was used. The remaining priors were set to default values. The convergence of all parameters to the stationary distribution in each run was inspected with Tracer v.1.7 (Rambaut et al., 2018). Effective sample size values were checked to ensure they exceeded 200, with some allowed to be greater than 100. Trees were summarized using TreeAnnotator v.2.4.4 (Bouckaert et al., 2014), with a 25% burn-in.

Reconstruction of discrete ancestral traits

The evolutionary history of one extrinsic and three intrinsic traits of snowbells was reconstructed taking into account historical introgression. The single extrinsic polymorphic character included

habitat (forest versus alpine), while the three intrinsic characters were as follows: floral morphology ('Soldanella' versus 'Tubiflores'), growth form (caespitose versus stoloniferous), and cytotype (euploid with $2n = 40$ chromosomes versus dysploid with $2n = 38$ chromosomes). Character states and their coding are summarized in Table S4.

The TREEMIX network calculated in our previous study (Slovák et al., 2023) served as a scaffold for reconstruction analyses. However, the network had to be ultrametrized before the reconstruction of ancestral traits. First, the major tree was extracted from the TREEMIX network. Then, the posterior distribution of MCMC simulations of the SNAPP analyses (see above) was scanned for species trees that differed from the major tree by a maximum of two nodes. The most compatible trees were retained to calculate a maximum clade credibility consensus tree in TreeAnnotator v.2.4.4 (Bouckaert et al., 2014). The function chronos implemented in the ape v.5.7.1 R package was used to ultrametrize the major tree of the TREEMIX network by applying the correlated model with $\lambda = 0.01$ and node ages of all compatible nodes of the consensus species tree as calibration points. Finally, edge lengths from the ultrametrized major tree were applied to the TREEMIX network. Since the SNAPP tree differed from the major tree extracted from the TREEMIX network by four nodes, reticulations from the TREEMIX network were introduced to the species tree as well to account for the uncertainty in phylogeny inference.

The evolution of four characters of snowbells was reconstructed on the ultrametrized TREEMIX network and the reticulated SNAPP species tree under substitution models for discrete traits as implemented in the Julia package PhyloNetworks v0.16.2 (Karimi et al., 2020). The single extrinsic polymorphic character was fitted with the Equal Rates Substitution Model (ERSM), while the binary intrinsic characters were fitted separately with the ERSM and also with the Binary Trait Substitution Model (BTSM). The likelihood ratio test with a level of significance at 0.05 selected the ERSM as the best substitution model for all characters.

Phylogenetic regression and reconstruction of genome size

Phylogenetic regression on networks (Bastide et al., 2018), as implemented in the Julia library PhyloNetworks v0.16.2 (Karimi et al., 2020), was further customized for the purposes of the present study to support different data whitening methods, ridge regression, and leave one out cross-validation (LOOCV).

Bastide et al. (2018) used Cholesky whitening to remove correlation among species in the dataset. However, Cholesky whitening is not suitable in combination with LOOCV, as whitened data values strongly depend on the order of samples. Therefore, ZCA-cor (ZCA stands for zero-phased components analysis) whitening method (for a review, see Kessy et al., 2018) was applied here. This technique maximizes the correlation between original and whitened data and hence minimizes the impact of sample order during LOOCV. Although both whitening methods provide the same regression outcome on full datasets, ZCA-cor gives much more stable results during cross-validation when regression is performed on reduced datasets.

Ridge regression was performed instead of ordinary least squares regression to reduce model overfitting. Ridge regression adds an L2 regularization penalty for each estimated coefficient to the objective function. The penalty is controlled by the hyperparameter λ, which was applied to the coefficient value divided by its standard error to standardize the regularization pressure. The hyperparameter λ was determined by applying LOOCV and

minimizing the sum of residuals (estimation errors) on unseen data given the model fit calculated on the remaining samples.

Weighted LOOCV was implemented to allow reducing the weight of a sample, instead of its complete removal. Weighted LOOCV reduces the weight of one sample in each step, fits the model, and calculates the estimation error for the sample with reduced weight. The objective is to minimize the sum of estimation error on 'almost unseen' data. When the weight is set to zero, weighted LOOCV becomes ordinary LOOCV.

Both the ultrametrized TREEMIX network and the reticulated SNAPP species tree were used to reconstruct the genome size. The following factors were considered to explain shifts in the genome size of snowbells: hybridization event, marginal area, habitat, and cytotype. Altogether, 24 indicator predictors representing these four factors and their interactions are listed in Tables S5 and S6. All combinations of up to six predictors with an absolute value of cross-correlation less than 0.95 were considered. For each combination of predictors, Brownian motion (BM), Pagel's lambda, and scaling hybrid models, as implemented in PhyloNetworks, were separately fitted. Their variance components were estimated using a maximum likelihood criterion, which is a prerequisite for comparing models with different predictors using a likelihood ratio test. In total, more than 22 000 models were statistically analyzed and tested for the full dataset, while more than 15 000 models were fitted when *S. villosa* was excluded. Only models with significant regression coefficients ($P < 0.05$) were further cross-validated using LOOCV with the following weights: 0.000, 0.001, 0.010, 0.050, and 0.100 for the full dataset and weight = 0.000 for the dataset without *S. villosa*. To prevent model overfitting, coefficients were tested for significance ($P < 0.05$) also during LOOCV and then ridge regression was carried out with the best hyperparameter λ ranging from 0.0001 to 5.0, as inferred from LOOCV.

The best fit model was selected using the three following tests. (i) An *F*-test was used to compare BM models with the same or nested predictors (one of the models has only a subset of predictors) at the significance level of 0.05. (ii) Likelihood ratio test served to compare a BM model with either Pagel's lambda or a scaling hybrid model with the same or nested predictors (BM is a special case with $\lambda = 1$), or Pagel's lambda models with nested predictors or scaling hybrid models with nested predictors. The significance level of 0.05 was used to reject models with significantly worse likelihood. (iii) Corrected Akaike information criterion (AICc) was calculated for each model. Models whose AICc scores exceeded the overall best fit model by 2.0 or more were rejected.

Ancestral genome sizes of snowbells were reconstructed on the ultrametrized TREEMIX network and reticulated SNAPP species tree using PhyloNetworks with (i) the best model selected as described above for the full dataset, (ii) predictor values of internal nodes as described in Table S5, and (iii) predictor values of extant taxa as listed in Table S6.

AUTHOR CONTRIBUTIONS

The conception and design of the study were performed by MS, IR, PV, and OP. JKu, EGŠ, and JKo conducted the data collection and material preparation for genetic analyses. IR, AM, and PV performed the data analyses. The first draft of the manuscript was written by MS, IR, and PV, with all authors providing comments on previous versions of the manuscript. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data underlying this study have been made freely available through the Dryad Digital Repository: https://doi.org/10.5061/dryad.mcvdnck41.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Representative members of the genus *Soldanella*.

Figure S2. The Bayesian coalescent-based species tree, generated by SNAPP using 4007 single nucleotide polymorphisms and 58 *Soldanella* accessions.

Figure S3. Ancestral state reconstruction of floral morphology (a), cytotype (b), habitat (c), and growth form (d) of snowbells on the reticulated SNAPP species tree under the best fit equal rates model using the Julia package PhyloNetworks.

Figure S4. Ancestral state reconstruction of genome size of snowbells on the reticulated SNAPP species tree under the best fit phylogenetic regression model using the Julia package PhyloNetworks.

Table S1. Bayes factor delimitation (BFD) of *Soldanella* species.

Table S2. List of sample sites for *Soldanella* populations analyzed in the present study.

Table S3. Characterization of RADseq data matrices used in Bayes factor delimitation (BFD) and inference of the Bayesian coalescent-based species tree (SNAPP).

Table S4. List of characters used for ancestral trait reconstructions.

Table S5. Indicator predictors tested in reconstruction of genome sizes using phylogenetic regression.

Table S6. Coding of indicator predictor values of extant taxa used in phylogenetic regression.

REFERENCES

Avise, J.C. & **Robinson, T.J.** (2008) Hemiplasy: a new term in the lexicon of phylogenetics. *Systematic Biology*, **57**, 503–507.

Balao, F., **Trucchi, E.**, **Wolfe, T.M.**, **Hao, B.H.**, **Lorenzo, M.T.**, **Baar, J.** *et al.* (2017) Adaptive sequence evolution is driven by biotic stress in a pair of orchid species (*Dactylorhiza*) with distinct ecological optima. *Molecular Ecology*, **26**, 3649–3662.

Bastide, P., **Solís-Lemus, C.**, **Ané, C.**, **William Sparks, K.** & **Kriebel, R.** (2018) Phylogenetic comparative methods on phylogenetic networks with reticulations. *Systematic Biology*, **67**, 800–820.

Bellino, A., **Bellino, L.**, **Baldantoni, D.** & **Saracino, A.** (2015) Evolution, ecology, and systematics of *Soldanella* (Primulaceae) in the southern Apennines (Italy). *BMC Evolutionary Biology*, **15**, e158.

Blöch, C., **Weiss-Schneeweiss, H.**, **Schneeweiss, G.M.**, **Barfuss, M.H.**, **Rebernig, C.A.**, **Villaseñor, J.L.** *et al.* (2009) Molecular phylogenetic analyses of nuclear and plastid DNA sequences support dysploid and polyploid chromosome number changes and reticulate evolution in the diversification of *Melampodium* (Millerieae, Asteraceae). *Molecular Phylogenetics and Evolution*, **53**, 220–233.

Bohutínská, M., **Vlček, J.**, **Yair, S.**, **Laenen, B.**, **Konečná, M.**, **Fracassetti, M.** *et al.* (2021) Genomic basis of parallel adaptation varies with divergence in *Arabidopsis* and its relatives. *Proceedings of the National Academy of Sciences of the United States of America*, **118**, e2022713118.

Boucher, F.C., **Casazza, G.**, **Szövényi, P.** & **Conti, E.** (2016) Sequence capture using RAD probes clarifies phylogenetic relationships and species boundaries in *Primula* sect. *Auricula*. *Molecular Phylogenetics and Evolution*, **104**, 60–72.

Boucher, F.C., **Niklaus, E.**, **Zimmermann, E.** & **Conti, E.** (2016) Allopatric speciation with little niche divergence is common among alpine Primulaceae. *Journal of Biogeography*, **43**, 591–602.

Bouckaert, R., **Heled, J.**, **Kühnert, D.**, **Vaughan, T.**, **Wu, C.-H.**, **Xie, D.** *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**, e1003537.

Brandrud, M.K., **Paun, O.**, **Lorenzo, M.T.**, **Nordal, I.** & **Brysting, A.K.** (2017) RAD-seq provides evidence for parallel ecotypic divergence in the autotetraploid *Cochlearia officinalis* in Northern Norway. *Scientific Reports*, **7**, e5573.

Bryant, D., **Bouckaert, R.**, **Felsenstein, J.**, **Rosenberg, N.A.** & **Roy Choudhury, A.** (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, **29**, 1917–1932.

Bui, Q.T. & **Grandbastien, M.-A.** (2012) LTR retrotransposons as controlling elements of genome response to stress? In: Grandbastien, M.A. & Casacuberta, J. (Eds.) *Plant transposable elements: impact on genome structure and function*. Berlin: Springer Verlag, pp. 273–296.

Catchen, J., **Hohenlohe, P.A.**, **Bassham, S.**, **Amores, A.** & **Cresko, W.A.** (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Danecek, P., **Auton, A.**, **Abecasis, G.**, **Albers, C.A.**, **Banks, E.**, **DePristo, M.A.** *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

De Vos, J.M., **Hughes, C.E.**, **Schneeweiss, G.M.**, **Moore, B.R.** & **Conti, E.** (2014) Heterostyly accelerates diversification via reduced extinction in primroses. *Proceedings of the Royal Society B: Biological Sciences*, **281**, e20140075.

Degnan, J.H. & **Rosenberg, N.A.** (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, **24**, 332–340.

Drummond, A.J. & **Bouckaert, R.R.** (2015) *Bayesian evolutionary analysis with BEAST*. Cambridge: Cambridge University Press.

Eriksson, M.C., **Mandakova, T.**, **McCann, J.**, **Temsch, E.M.**, **Chase, M.W.**, **Hedren, M.** *et al.* (2022) Repeat dynamics across timescales: a perspective from sibling allotetraploid marsh orchids (*Dactylorhiza majalis* s.l.). *Molecular Biology and Evolution*, **39**, msac167.

Göbel, U., **Arce, A.L.**, **He, F.**, **Rico, A.**, **Schmitz, G.** & **de Meaux, J.** (2018) Robustness of transposable element regulation but no genomic shock observed in interspecific *Arabidopsis* hybrids. *Genome Biology and Evolution*, **10**, 1403–1415.

Halbritter, A.H., **Fior, S.**, **Keller, I.**, **Billeter, R.**, **Edwards, P.J.**, **Holderegger, R.** *et al.* (2018) Trait differentiation and adaptation of plants along elevation gradients. *Journal of Evolutionary Biology*, **31**, 784–800.

Hibbins, M.S., **Breithaupt, L.C.** & **Hahn, M.W.** (2023) Phylogenomic comparative methods: accurate evolutionary inferences in the presence of gene tree discordance. *Proceedings of the National Academy of Sciences of the United States of America*, **120**, e2220389120.

Hibbins, M.S., **Gibson, M.J.** & **Hahn, M.W.** (2020) Determining the probability of hemiplasy in the presence of incomplete lineage sorting and introgression. *eLife*, **9**, e63753.

Hu, C.-M. & **Kelso, S.** (1996) Primulaceae. In: Wu, Z.Y. & Raven, P.H. (Eds.) *Flora of China*, Vol. **15**. Beijing and St. Louis, MI: Science Press and Missouri Botanical Garden, pp. 39–189.

Huang, Y., **Chen, X.**, **Chen, J.** & **Li, Z.** (2019) DNA barcoding and phylogenetic relationships in *Omphalogramma* (Primulaceae) from the Hengduan Mountain region in China. *Phytotaxa*, **402**, 97–106.

Kadereit, J.W., **Griebeler, E.M.** & **Comes, H.P.** (2004) Quaternary diversification in European alpine plants: pattern and process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 265–274.

Karimi, N., **Grover, C.E.**, **Gallagher, J.P.**, **Wendel, J.F.**, **Ané, C.** & **Baum, D.A.** (2020) Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in baobabs (*Adansonia*; Bombacoideae; Malvaceae). *Systematic Biology*, **69**, 462–478.

Kawakami, T., **Dhakal, P.**, **Katterhenry, A.N.**, **Heatherington, C.A.** & **Ungerer, M.C.** (2011) Transposable element proliferation and genome expansion are rare in contemporary sunflower hybrid populations despite widespread transcriptional activity of LTR retrotransposons. *Genome Biology and Evolution*, **3**, 156–167.

Kessy, A., **Lewin, A.** & **Strimmer, K.** (2018) Optimal whitening and decorrelation. *Journal of the American Statistical Association*, **72**, 309–314.

Kliment, J. & **Valachovič, M.** (2007) *Vegetácia Slovenska - Rastlinné spoločenstvá Slovenska. 4. Vysokohorská vegetácia*. Bratislava: Veda.

Knotek, A., **Konečná, V.**, **Wos, G.**, **Požárová, D.**, **Šrámková, G.**, **Bohutínská, M.** *et al.* (2020) Parallel alpine differentiation in *Arabidopsis arenosa*. *Frontiers in Plant Science*, **11**, e561526.

Kochjarová, J., **Kučera, J.**, **Slovák, M.** & **Štubňová, E.** (2017) *Soldanella* L. In: Goliašová, K. & Michalková, E. (Eds.) *Flóra Slovenska VI/4*. Bratislava: Veda, pp. 645–659.

Körner, C., **Riedl, S.**, **Keplinger, T.**, **Richter, A.**, **Wiesenbauer, J.**, **Schweingruber, F.** *et al.* (2019) Life at 0 °C: the biology of the alpine snowbed plant *Soldanella pusilla*. *Alpine Botany*, **129**, 63–80.

Kumar, A. & **Bennetzen, J.L.** (1999) Plant retrotransposons. *Annual Review of Genetics*, **33**, 479–532.

Leaché, A.D., **Harris, R.B.**, **Rannala, B.** & **Yang, Z.** (2014) The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, **63**, 17–30.

Loiseau, O., **Mota Machado, T.**, **Paris, M.**, **Koubínová, D.**, **Dexter, K.G.**, **Versieux, L.M.** *et al.* (2021) Genome skimming reveals widespread hybridization in a neotropical flowering plant radiation. *Frontiers in Ecology and Evolution*, **9**, e668281.

Maddison, W.P. (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523–536.

Maddison, W.P. & **Knowles, L.L.** (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, **55**, 21–30.

Mallet, J., **Besansky, N.** & **Hahn, M.W.** (2016) How reticulated are species? *BioEssays*, **38**, 140–149.

Mandáková, T. & **Lysak, M.A.** (2018) Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology*, **42**, 55–65.

McLintock, B. (1984) The significance of responses of the genome to challenge. *Science*, **226**, 792–801.

Melichárková, A., **Šlenker, M.**, **Zozomová-Lihová, J.**, **Skokanová, K.**, **Šingliarová, B.**, **Kačmárová, T.** *et al.* (2020) So closely related and yet so different: strong contrasts between the evolutionary histories of species of the *Cardamine pratensis* polyploid complex in Central Europe. *Frontiers in Plant Science*, **11**, e588856.

Mendes, F.K., **Fuentes-González, J.A.**, **Schraiber, J.G.** & **Hahn, M.W.** (2018) A multispecies coalescent model for quantitative traits. *eLife*, **7**, e36482.

Mendes, F.K. & **Hahn, M.W.** (2016) Gene tree discordance causes apparent substitution rate variation. *Systematic Biology*, **65**, 711–721.

Moraes, A.P., **Olmos Simões, A.**, **Ojeda Alayon, D.I.**, **de Barros, F.** & **Forni-Martins, E.R.** (2016) Detecting mechanisms of karyotype evolution in *Heterotaxis* (Orchidaceae). *PLoS One*, **11**, e0165960.

Niederle, J. (2003) Remarks on the genus *Soldanella* L. in the West Carpathians. *Acta Musei Richnoviensis, Sectio Natura*, **10**, 171–174.

Niederle, J. (2016) Dřípatky ze Srbska. *Skalničkářův Rok*, **73**, 16.

Ottenburghs, J. (2020) Ghost introgression: spooky gene flow in the distant past. *BioEssays*, **42**, e2000012.

Ozenda, P. (1985) *La végétation de la chaîne alpine dans l'espace montagnard européen*. Paris: Masson.

Parisod, C., Alix, K., Just, J., Petit, M., Sarilar, V., Mhiri, C. et al. (2010) Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytologist*, **186**, 37–45.

Pax, F. (1905) Primulaceae/*Primula*. In: Engler, A. (Ed.) *Das Pflanzenreich, Band IV/237*. Leipzig: H. R. Engelmann, pp. 1–160.

Pouchon, C., Fernández, A., Nassar, J.M., Boyer, F., Aubert, S., Lavergne, S. et al. (2018) Phylogenomic analysis of the explosive adaptive radiation of the *Espeletia* complex (Asteraceae) in the tropical Andes. *Systematic Biology*, **67**, 1041–1060.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**, 901–904.

Rellstab, C., Zoller, S., Sailer, C., Tedder, A., Gugerli, F., Shimizu, K.K. et al. (2020) Genomic signatures of convergent adaptation to Alpine environments in three Brassicaceae species. *Molecular Ecology*, **29**, 4350–4365.

Senerchia, N., Felber, F. & Parisod, C. (2015) Genome reorganization in F1 hybrids uncovers the role of retrotransposons in reproductive isolation. *Proceedings of the Royal Society B: Biological Sciences*, **282**, e20142874.

Singhal, V.K., Kumar, R., Singhal, H., Kumar, P., Kaur, D., Kaur, M. et al. (2017) A profile of male meiosis, chromosomal variation and status in species of *Impatiens* from North-West Himalaya in India. *Caryologia*, **70**, 258–269.

Slovák, M., Melichárková, A., Gbúrová Štubňová, E., Kučera, J., Mandáková, T., Smyčka, J. et al. (2023) Pervasive introgression during rapid diversification of the European mountain genus *Soldanella* (L.) (Primulaceae). *Systematic Biology*, **72**, 491–504.

Solís-Lemus, C. & Ané, C. (2016) Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics*, **12**, e1005896.

Solís-Lemus, C., Bastide, P. & Ané, C. (2017) PhyloNetworks: a package for phylogenetic networks. *Molecular Biology and Evolution*, **34**, 3292–3298.

Soltis, P.S. & Soltis, D.E. (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology*, **60**, 561–588.

Steffen, S. & Kadereit, J.W. (2014) Parallel evolution of flower reduction in two alpine *Soldanella* species (Primulaceae). *Botanical Journal of the Linnean Society*, **175**, 409–422.

Štubňová, E., Hodálová, I., Kučera, J., Mártonfiová, L., Svitok, M. & Slovák, M. (2017) Karyological patterns in the European endemic genus *Soldanella* L.: absolute genome size variation uncorrelated with cytotype chromosome numbers. *American Journal of Botany*, **104**, 1241–1253.

Stull, G.W., Pham, K.K., Soltis, P. & Soltis, D.E. (2023) Deep reticulation: the long legacy of hybridization in vascular plant evolution. *The Plant Journal*, **114**, 743–766.

Szukala, A., Lovegrove-Walsh, J., Luqman, H., Fior, S., Wolfe, T.M., Frajman, B. et al. (2022) Polygenic routes lead to parallel altitudinal adaptation in *Heliosperma pusillum* (Caryophyllaceae). *Molecular Ecology*, **32**, 1832–1847.

Tantray, R., Jan, Y.I., Saleem Wani, M., Kumar Singhal, V. & Gupta, R.C. (2021) Chromosome numbers and meiotic behavior in some species of Asteraceae from high altitudinal regions of Kashmir Himalayas. *Journal of Asia-Pacific Biodiversity*, **14**, 590–606.

Teo, B., Rose, J.P., Bastide, P. & Ané, C. (2023) Accounting for within-species variation in continuous trait evolution on a phylogenetic network. *Bulletin of the Society of Systematic Biologists*, **2**, 1–29.

Tusso, S., Suo, F., Liang, Y., Du, L.-L. & Wolf, J.B.W. (2022) Reactivation of transposable elements following hybridization in fission yeast. *Genome Research*, **32**, 324–336.

Ungerer, M.C. & Kawakami, T. (2013) Transcriptional dynamics of LTR retrotransposons in early generation and ancient sunflower hybrids. *Genome Biology and Evolution*, **5**, 329–337.

Valachovič, M., Štubňová, E., Senko, D., Kochjarová, J. & Coldea, G. (2019) Ecology and species distribution pattern of *Soldanella* sect. *Soldanella* (Primulaceae) within vegetation types in the Carpathians and the adjacent mountains. *Biologia (Bratislava)*, **74**, 733–750.

Walter, G.M., Wilkinson, M.J., Aguirre, J.D., Blows, M.W. & Ortiz-Barrientos, D. (2018) Environmentally induced development costs underlie fitness tradeoffs. *Ecology*, **99**, 1391–1401.

Walter, G.M., Wilkinson, M.J., James, M.E., Richards, T.J., Aguirre, J.D. & Ortiz-Barrientos, D. (2016) Diversification across a heterogeneous landscape. *Evolution*, **70**, 1979–1992.

Wang, Y., Cao, Z., Ogilvie, H.A. & Nakhleh, L. (2021) Phylogenomic assessment of the role of hybridization and introgression in trait evolution. *PLoS Genetics*, **17**, e1009701.

Winterfeld, G., Becher, H., Voshell, S., Hilu, K. & Röser, M. (2018) Karyotype evolution in *Phalaris* (Poaceae): the role of reductional dysploidy, polyploidy and chromosome alteration in a wide-spread and diverse genus. *PLoS One*, **13**, e0192869.

Wos, G., Arc, E., Hülber, K., Konečná, V., Knotek, A., Požárová, D. et al. (2022) Parallel local adaptation to an alpine environment in *Arabidopsis arenosa*. *Journal of Ecology*, **110**, 2448–2461.

Zhang, L.-B., Comes, H.P. & Kadereit, J.W. (2001) Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *American Journal of Botany*, **88**, 2331–2345.

Zhang, L.-B. & Kadereit, J.W. (2002) The systematics of *Soldanella* (Primulaceae) based on morphological and molecular (ITS, AFLPs) evidence. *Nordic Journal of Botany*, **22**, 129–169.