Received: 6 December 2023 | Accepted: 16 April 2024

DOI: 10.1002/ajb2.16361

RESEARCH ARTICLE

American Journal of Botany

Challenge accepted: Evolutionary lineages versus taxonomic classification of North American shrub willows (*Salix*)

Pia Marinček[1] | **Étienne Léveillé-Bourret**[2] | **Ferris Heiduk**[1] | **Jing Leong**[3,4] | **Stéphane M. Bailleul**[5] | **Martin Volf**[3,4] | **Natascha D. Wagner**[1]

[1]Department of Systematics, Biodiversity and Evolution of Plants (with Herbarium), University of Goettingen, Untere Karspüle 2, D-37073, Göttingen, Germany

[2]Institut de recherche en biologie végétale (IRBV), Département de sciences biologiques, Université de Montréal, 4101 Sherbrooke est, Montréal, QC, H1X 2B2, Canada

[3]Biology Centre of the Czech Academy of Sciences, Branisovska 31, 37005 Ceske Budejovice, Czech Republic

[4]Faculty of Science, University of South Bohemia, Branisovska 31, 37005 Ceske Budejovice, Czech Republic

[5]Division recherche et développement scientifique, Jardin botanique de Montréal, 4101 Sherbrooke est, Montréal, QC, H1X 2B2, Canada

Correspondence

Natascha D. Wagner, Department of Systematics, Biodiversity and Evolution of Plants (with Herbarium), University of Goettingen, Untere Karspüle 2, D-37073 Göttingen, Germany.
Email: natascha.wagner@uni-goettingen.de

Abstract

Premise: The huge diversity of *Salix* subgenus *Chamaetia/Vetrix* clade in North America and the lack of phylogenetic resolution within this clade has presented a difficult but fascinating challenge for taxonomists to resolve. Here we tested the existing taxonomic classification with molecular tools.

Methods: In this study, 132 samples representing 46 species from 22 described sections of shrub willows from the United States and Canada were analyzed and combined with 67 samples from Eurasia. The ploidy levels of the samples were determined using flow cytometry and nQuire. Sequences were produced using a RAD sequencing approach and subsequently analyzed with ipyrad, then used for phylogenetic reconstructions (RAxML, SplitsTree), dating analyses (BEAST, SNAPPER), and character evolution analyses of 14 selected morphological traits (Mesquite).

Results: The RAD sequencing approach allowed the production of a well-resolved phylogeny of shrub willows. The resulting tree showed an exclusively North American (NA) clade in sister position to a Eurasian clade, which included some North American endemics. The NA clade began to diversify in the Miocene. Polyploid species appeared in each observed clade. Character evolution analyses revealed that adaptive traits such as habit and adaxial nectaries evolved multiple times independently.

Conclusions: The diversity in shrub willows was shaped by an evolutionary radiation in North America. Most species were monophyletic, but the existing sectional classification could not be supported by molecular data. Nevertheless, monophyletic lineages share several morphological characters, which might be useful in the revision of the taxonomic classification of shrub willows.

KEYWORDS

divergence time, morphological character evolution, phylogenomics, ploidy estimation, RAD sequencing, radiation, Salicaceae

The genus *Salix* L. (Salicaceae) comprises about 450 species of trees and shrubs that are mainly distributed in the northern hemisphere. About two thirds of the species occur in Eurasia, and approximately 140 species occur in North America. Argus (2010) divided the genus into five subgenera, *Salix* subg. *Longifoliae* (Andersson) Argus, *Salix* subg. *Protitea* Kimura, *Salix* subg. *Salix*, *Salix* subg. *Vetrix* Dumort., and *Salix* subg. *Chamaetia* Nasarow. The first three, *Salix* subg. *Longifoliae*, *Salix* subg. *Protitea,* and *Salix* subg. *Salix* will be referred to as the "tree willows" and be abbreviated as *Salix* subg. *Salix* s.l. As for *Salix* subg. *Vetrix* and *Salix* subg. *Chamaetia*, these together will be referred to as the "shrub willows". The majority of species belong to the shrub willows (subg. *Chamaetia/Vetrix* clade) of which about 100 species occur in North America (Argus, 2010). Species diversity in this clade range from

creeping arctic-alpine dwarf shrubs to medium-sized trees occuring from low to high latitudes and from coastal lowlands to high mountain tops (Argus, 1997; Wagner et al., 2018). The species provide vital habitat and food sources for a diversity of wildlife, making them important components of natural communities (Skvortsov et al., 1999). Thus, shrub willows play an ecologically significant role in ecosystems and provide a taxonomically rich system to study.

Classical taxonomy and systematics in *Salix* have proven to be extremely difficult because of dioecy, simple reduced flowers, common natural formation of hybrids, high intraspecific phenotypic variation, and the presence of polyploid species (Skvortsov et al., 1999; Hörandl et al., 2012; Cronk et al., 2015). Despite these difficulties, in his monograph on North American *Salix*, Argus (1997) classified the species mainly based on morphological characters. His latest treatment of *Salix* in the *Flora of North America* (Argus, 2010) recognizes 19 sections and 63 species for subgenus *Vetrix* and 27 species organized in eight sections for subgenus *Chamaetia*. Informative morphological characters are important for species identification but might not reflect evolutionary lineages. The existing taxonomic classification, however, was never comprehensively assessed with molecular markers because of a lack of informative marker systems. Previous phylogenetic analyses using the traditional Sanger sequencing method were able to separate subgenus *Salix* from a large clade uniting the two subgenera *Chamaetia* and *Vetrix* but lacked any infrageneric resolution for *Vetrix* (Chen et al., 2010; Percy et al., 2014; Lauron-Moreau et al., 2015; Wu et al., 2015). Additionally, plastome variation within the *Chamaetia/Vetrix* clade is impressively low given the morphological and geographical diversity (Percy et al., 2014; Wagner et al., 2021b). More recently, DNA re-sequencing and target enrichment approaches were used for *Salix*; however, the studies either focused on higher taxonomic levels (i.e., family or genus level) or were based on a small and/or geographically restricted sampling (Sanderson et al., 2020; Gulyaev et al., 2022). In a recent *Populus* L. and *Salix* study employing a target capture data set on a broad sampling, Sanderson et al. (2023) revealed that high amounts of ancestral hybridization resulted in conflicting phylogenetic signals, especially in genus *Salix*. RAD sequencing (Baird et al., 2008) was recently used to overcome the lack of phylogenetic information within the *Chamaetia/Vetrix* clade (Gramlich et al., 2018; He et al., 2021; Wagner et al., 2018, 2020, 2023). Using this method, Wagner et al. (2018) published the first well-resolved phylogeny of diploid European members of the *Chamaetia/Vetrix* clade. Followed by an enlarged study on the origin of polyploidy within Eurasian shrub willows (Wagner et al., 2020) and a study on hexaploid shrub willows of the European Alps (Wagner et al., 2023). RAD-sequencing data was also used to analyze the species radiation in the Hengduan Mountain System (He et al., 2021). In *Salix*, RAD loci almost exclusively represent the nuclear genome and provide tens of thousands of SNPs from both conservative and rapidly evolving noncoding genomic regions (Gramlich et al., 2018). Hence, this method has the potential to resolve the complex phylogenetic relationships of North American willows.

This study is part of a project to build a world-wide backbone phylogeny of shrub willows. It is based on a representative sampling of North American species that will be analyzed using a RAD-sequencing protocol. The main aims are (1) to establish a phylogenomic framework of the *Chamaetia/Vetrix* clade, (2) to provide additional morphological and cytological characters and to use them in ancestral character state analyses, and (3) to combine this information to draw conclusions on the existing sectional classification of Argus (2010).

MATERIALS AND METHODS

Sampling

For this study, we sampled 132 individuals representing 46 species from North America spread across 22 of the 27 sections sensu Argus (2010) of the *Chamaetia/Vetrix* clade. Three individuals of North American species of *Salix* subg. *Salix* s.l., as well as two previously published European *S. triandra* L. samples (Wagner et al., 2018) were included to serve as outgroups (Wagner et al., 2020). Additionally, data from 67 samples (61 previously published [Wagner et al., 2018, 2020, 2023b] and 6 new) of 33 Eurasian species of shrub willows were included in the study, resulting in a total of 202 individuals representing 78 *Salix* species (Appendix S1). For some species, we collected material in North America and Eurasia (e.g., *S. reticulate* L., *S. glauca* L.). The samples were collected in the Montreal Botanical Garden and on field trips to California (2018, 2019, 2022), Minnesota (2021), Washington (2022), and Alaska (2022). Individual plants were identified to the species level using the key of Argus (2010), his regional *Salix* identification guides (Argus, 2004, 2012) and other identification literature (Viereck and Little, 1972; Collet, 2002; Hoag, 2005). We aimed to sample representatives of all sections. Since the Eurasian sectional classification is not part of this study, the sampling for this region was reduced. Studies with a focus on Eurasian *Salix* species were and will be presented elsewhere (Wagner et al., 2018, 2020, 2021a, 2023).

DNA extraction and RAD sequencing

DNA of all samples was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and the manufacturer's instructions with slight modifications as described by Marinček et al. (2023). After quality check with Nanodrop 2000 (Thermo Fischer Scientific, Waltham, MA, USA) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific), the DNA extracts were sent to Floragenex (Portland, OR, USA) where the sequencing library was prepared as described by Baird

et al. (2008) using the restriction enzyme PstI, size selected for 300–500 bp with Pippin Prep (Sage Science, Beverly, MA, USA) and barcoded by individual. Multiplexed libraries were sequenced on Illumina sequencing machines. The quality of the resulting single-end 100-bp long sequence reads was checked using FastQC v.0.10.1 (Andrews, 2010). We demultiplexed the reads with ipyrad v.0.9.52 (Eaton and Overcast, 2020) and used this pipeline to run the assembly for 202 individuals using the high-performance cluster of the Gesellschaft für Wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG HPC, Göttingen, Germany). After initial investigation of optimal assembly parameters, clustering threshold was kept at 85% as done in previous studies of the *Salix Chametia/Vetrix* clade (e.g., Wagner et al., 2018, 2020). A maximum of four alleles per site was allowed in the final consensus sequence. Two assemblies were run; the first included all 202 sequenced individuals and the second with 80 individuals, where a single individual per (monophyletic) species was selected randomly. We refer to the latter as the "singletons assembly". For initial tests, we used different settings for the parameter "minimum number of samples sharing a locus (m)" on the full data set, namely m15, m40, m100, and m195, and compared the results by constructing a phylogeny. For final phylogenetic analyses, we used the filtering of m15 and m6 minimum samples per locus for the two assemblies respectively, which filtered out loci shared by less than approximately 8% of individuals. All alignments used for subsequent bioinformatic analyses are available from https://github.com/NataschaWagner28/NorthAmerican_Salix.

Ploidy determination using flow cytometry and molecular data

The ploidy level for silica-gel-dried leaf material of all 132 North American samples was determined using flow cytometry (FC) and the protocol workflow of Doležel et al. (2007). Each sample (~1 cm^2) was reduced to small pieces using TissueLyser II (Qiagen; 3 s at 30 Hz), 200 µL of 1% PVP Otto I extraction buffer was added, then the sample was gently inverted for 1 min. Samples were filtered through a CellTrics filter (30 µm mesh, Sysmex Partec GmbH, Görlitz, Germany) into a FC sample tube, then 800 µL of DAPI-containing Otto II buffer (Otto, 1990) was then added to stain the DNA. Samples were then analyzed in a CyFlow Ploidy Analyzer (Sysmex, Norderstedt, Germany) with CyFlow Cube v.1.6.4.14 (Sysmex) to control the instrument and view the histograms of the DNA content. Diploid *S. caprea*, a species from the *Chamaetia/Vetrix* clade, collected in Göttingen, Germany, was used as an external standard for each measurement cycle. The analysis of each sample was run until a distinct peak of DNA could be seen on the plot. The mean position of the peak was collected from the software to calculate the sample ploidy level using the formula of Doležel et al. (2007), which was adapted for a diploid reference: Sample ploidy = 2(Mean position of sample peak)/Mean position of reference peak. We also evaluated ploidy level of the sequenced reads using the software nQuire (Weiß et al., 2018). We mapped the raw reads of all samples to the available reference genome of *S. purpurea* L. (NCBI accession GCA_027405865.1) using the bwa mem command of the bwa software (Li and Durbin, 2009). We ran the create command in nQuire with quality filtering of bases below the score of 30 and coverage of less than 20. Next, we examined the histogram plots of the SNP frequencies distribution of the created. bin files (command histo) and assessed ploidy levels with the statistical model approach used in nQuire (commands lrdmodel, modeltest and histotest). For samples with histograms showing higher ploidy levels, we ran an additional nQuire pipeline with a lower cutoff of minimum base frequency (10%) for better visualization. Those samples could not be assessed with the nQuire model testing framework (appropriate only for di-, tri-, and tetraploids), and we noted them as "higher ploidy" (>4*x*). Finally, ploidy was determined based on both estimation approaches, where possible, or from the single successful estimation approach. Staining of nuclei in FC might be hindered by cytosolic compounds, such as phenols (Loureiro et al., 2006; Doležel et al., 2007). In addition, genome downsizing is a known process associated with polyploids (Leitch and Bennett, 2004). Thus, we tried to find the correct estimation of uncertain values by averaging FC estimates of multiple accessions per species and considered the already existing chromosome count data from literature, if applicable. Unsuccessful FC measurements or ploidy assessment with nQuire (no signal detected in histogram plots) and inconclusive results are marked in Appendix S1.

Phylogenetic inference of species relationships

Phylogenetic relationships were inferred based on the concatenated alignment of the complete data set using the maximum likelihood (ML) analysis in RAxML v.8.2.4 with GTR + Γ nucleotide substitute model and rapid bootstrap analysis with 100 replicates. Then, the quartet sampling (QS; Pease et al., 2018) pipeline with 100 replicates was applied to test the statistical support of the given topology. The phylogenies of the full sampling of 202 individuals and the singletons assembly of 80 individuals did not show identical topologies regarding species relationships. Therefore, we followed the workflow presented by Hipp et al. (2020) and pruned the topology of the complete sampling tree to single individual per monophyletic species clade. This pruned tree was then used as a phylogenetic constraint when building the singletons assembly tree in RAxML, resulting in identical toplogies for the singletons and the full sampling data set. All analyses were run on the GWDG HPC cluster.

We utilized SplitsTree4 (Huson and Bryant, 2006) to reconstruct possible network-like evolutionary relationships among the North American species. Based on unlinked SNP

data resulting from the m15 data set, we generated the split network by implementing a NeighborNet (Bryant and Moulton, 2002) analysis with variance of ordinary least squares complemented by a bootstrapping with 1000 replicates to test for statistical support. Missing data were treated as unknown.

Divergence time estimation

To estimate divergence time, we used BEAST 2 v.2.7.5 (Bouckaert et al., 2014) on the concatenated alignment of unlinked SNPs of the singletons assembly. This alignment is one of the outputs of the ipyrad pipeline and contained one SNP per locus. We applied the optimized relaxed clock approach with GTR + Γ substitution model with four rate categories and Yule model for speciation. We used uniform priors for time calibration, which estimate equal probability across all ages between the defined lower and upper bounds. This was the most appropriate approach for our sampling, the available and reliable fossil data, and previously estimated time divergence studies of Salicaceae. We defined the maximum age constraint of the root of our phylogeny with from a secondary calibration based on the results of De Mestier et al. (2022) and Liu et al. (2022). Their studies arrived at similar dates for the crown node of Salicaceae (59 Ma in the former and 69 Ma in the latter) using very different approaches and fossil calibration points. These ages are also close to the age of the oldest presumed fossil evidence of Salicaceae (*Populus elatior* and *Salix meeki* from the late Cretaceous; Guo, 1975). The minimum age constraint was set using the earliest reliable fossil of the *Salix* subg. *Vetrix* (Wolfe, 1987; Collinson, 1992), with an estimated age of 23 Ma, which has frequently been used in time divergence analyses of genus *Salix* (Wu et al., 2015; He et al., 2021, 2023; Kikuchi et al., 2023). We therefore imposed a maximum and minimum age on the root of our phylogeny, in the form of a uniform prior of 23 to 69 Ma. We also set a uniform prior with the same age constraints to calibrate the stem lineage of the *Chamaetia/Vetrix* clade, which is nested within the *Salix* and to which both of the constraints apply. We opted for the use of two uniform priors to reasonably bracket the age of both calibration nodes. The effects of the priors were also assessed by running the analysis without the molecular data first. We used the constrained ML topology of the singletons assembly (see section Phylogenetic inference of species relationships) as a starting tree and a multiple monophyletic constraint prior. For the tree to fit the time calibration priors and the analysis being able to start, we created an ultrametric tree with the chronos command from the R (R Core Team 2023) package ape v.5.7.1 (Paradis and Schliep, 2019), calibrating it on the same two nodes with the same maximum and minimum constraint as in the BEAST prior settings. Tree building parameters (subtree slide, narrow exchange, wide exchange, and Wilson-Balding global rearrangement) were disabled by setting them to 0 because we were only interested in molecular time inference. Then, two independent MCMC analyses were run with 30 million generations, and the output log files were investigated with Tracer v.1.7.2 (Rambaut et al., 2018) to assess convergence. The log files and the trees files of the two separate runs were combined with LogCombiner from the BEAST package, discarding the first 25% of the logs as burn-in. With TreeAnnotator (Heled and Bouckaert, 2013), the combined trees file was used to generate a maximum clade credibility tree with common ancestor heights.

The described BEAST methodology modelled the data of the concatenated unlinked SNPs alignment as a single gene under a single substitution model, which might methodologically not be entirely appropriate, however, has been used before (e.g., Cavender-Bares et al., 2015; He et al., 2021). The time divergence analysis in the SNAPPER (Stoltz et al., 2021) plugin from the BEAST package uses a diffusion model approach and is thus suitable for analyses of unlinked SNP data sets. We used the same two age calibration priors and starting tree as in the BEAST analysis for SNAPPER. The tree building parameters were disabled and the analysis was run for 2 million generations. We combined two separate runs, discarded the first 25% as burn-in and generated a maximum clade credibility tree (see above). SNAPPER could only be applied to biallelic SNPs. In addition, the snapper_prep.rb script (https://github.com/mmatschiner/snapp_prep) used to prepare the data set retained only biallelic SNPs that contained a minimum amount of missing data.

We also analyzed changes in the evolutionary rate of the North American shrub willows. Therefore, we used BAMM v.2.5.0 (Rabosky et al., 2014; Rabosky, 2014) on the calculated ultrametric tree from BEAST with a speciation-extinction model and 50 million generations to ensure that the effective sample size was well above 200. To account for the incomplete sampling, we defined the sampling fractions for two groups: the *Chamaetia/Vetrix* clade (23% of the 337 species worldwide (according to Argus [2010]) and *Salix* subg. *Salix* s.l. (2% of the 126 species worldwide (according to Argus [2010]). The outputs were examined to find the best estimate of the rate shift for the dated tree and visualized using the R packages BAMMtools v.2.1.10 (Rabosky et al., 2014), ape and coda v.0.19.4 (Plummer, Best, Cowles and Vines., 2006). The same workflow was applied to the ultrametric tree from the SNAPPER analysis.

Character evolution analyses

We investigated a total of 49 characters for all included species based on descriptions in the literature (Lautenschlager-Fleury and Lautenschlager-Fleury, 1994; Skvortsov et al., 1999; Argus, 1997, 2010; Hörandl et al., 2012), complemented with observations of herbarium specimens and individual photos of living individuals. To circumvent differences between floras and field guides, we simplified some characters. The character table and ML topology (singletons tree) were analyzed in

Mesquite v.3.81 (Maddison and Maddison, 2023) with parsimony ancestral states reconstruction method. Among the investigated characters, we selected 14 that were informative for our study. The selected characters and definitions of the specific character states are listed in Appendix S2.

RESULTS

The comparison of different thresholds for "m" is presented in Appendix S3. The greatest shift in topology was observed at the threshold m40 (10,792 RAD loci) to m100 (3440 RAD loci). The optimal trade-off of informativeness, revealed number of SNPs, and good statistical support were observed for the m15 data set. For the complete sampling and final filtering settings, the ipyrad pipeline assembled 34,953 RAD loci, concatenated to 4,226,091-bp alignment containing 458,198 SNPs and 79.73% missing sites. For the singletons assembly, ipyrad assembled 57,856 loci, concatenated to 6,755,758-bp alignment containing 704,152 SNPs and 71.86% missing sites.

Phylogenetic relationships and divergence time estimation

All presented phylogenies based on the selected settings had identical topologies due to the use of a constraint tree (see Materials and Methods). A simplified phylogenetic tree based on the complete sampling assembly with summarized clades in relation to their sectional classifications can be found in Figure 1. A detailed version of the same RAxML phylogeny displaying all tips is in Appendix S4, and a version of the same tree including quartet-sampling scores is illustrated in Appendix S5. The results of the BEAST analysis based on the singletons assembly (one sample per species) and the complete alignment of 56,711 unlinked SNPs are shown in Figure 2. The ultrametric tree based on 560 biallelic SNPs as estimated with the SNAPPER plugin is shown in Appendix S6. The subg. *Chamaetia/Vetrix* clade (excluding *S. setchelliana* C.R. Ball) was well supported (BS 100, QS 0.97/0/1; Figure 1; Appendices S4, S5). Its divergence time was estimated at 23.06 (12.60–37.89) Myr ago (Ma) with the BEAST analysis and at 13.87 (6.50–22.72) Ma with the SNAPPER analysis (Figure 2; Appendix S6). *Salix setchelliana* was in a sister position to all remaining shrub willows (BS 100) (Figure 1) with similar divergence estimates in both dating analyses—26.46 (14.75–43.68) Ma with BEAST and 25.54 (11.62–42.04) Ma with SNAPPER (Figure 2; Appendix S6). The next branching included *S. reticulata* and *S. vestita* Pursh (both section *Chamaetia*) that were in a sister position to the rest of the *Chamaetia–Vetrix* species. The phylogeny showed two major subclades with a well-supported sister relationship (BS 100, QS 0.69/0.06/0.98, Figure 1; Appendix S5) with an estimated divergence age of 19.84 (11.02–32.83) Ma with BEAST (Figures 2) and 11.51 (5.68–18.50) Ma with SNAPPER (Appendix S6). The monophyletic North American (NA) clade (BS 99, QS 0.023/0.78/0.97, Figure 1) had an estimated divergence time of 17.49 (9.50–28.88) Ma with BEAST (Figures 2) and 10.03 (4.90–14.90) Ma with SNAPPER (Appendix S6). It included most of the samples collected in North America. Its sister clade (BS 100, Figure 1), hereafter referred to as the Eurasian (EuA) clade, had an estimated divergence time of 19.05 (10.44–31.42) Ma, Figure 2) with BEAST and 11.21 (5.48–22.71) Ma with SNAPPER (Appendix S6). It included all Eurasian samples and samples from NA species: *S. bebbiana* Sarg., *S. fuscescens* Andersson and *S. uva-ursi* Pursh. Each of these three samples grouped with Eurasian species belonging to the same taxonomic section. Samples from the Eurasian species *S. viminalis* L. and *S. gracilistyla* Flod. that were collected in North America, where they considered introduced, were positioned among samples of the same species in the EuA clade (see Figure 1; Appendix S4). The speciation-rate reconstruction analysis with BAMM based on the BEAST tree revealed three possible scenarios, one of them supporting no shift in diversification rate with calculated probability $f = 0.16$. The most likely scenario with a probability of $f = 0.69$ suggested an increased speciation rate at the split of EuA and NA clade (Figure 2). The more recent diversification, however, did not appear to be associated with an increase in speciation rate neither for the NA clade nor the EuA clade (Figure 2). The third scenario ($f = 0.15$) recovered a change in rate at the *Chamaetia/Vetrix* divergence node (excluding *S. setchelliana*). The speciation rate reconstruction based on the SNAPPER tree showed the same two likely shift scenarios toward an increase in diversification rate of the *Chamaetia/Vetrix* clade with nearly equal probabilities ($f = 0.54$, $f = 0.45$) as illustrated in Appendix S6.

Most species appeared as monophyletic (Figure 1). Exceptions included a group of Alaskan species containing *S. barclayi* Andersson, *S. commutata* Bebb., *S. myrtillifolia* Andersson, and *S. pseudomyrsinites* Andersson (Alaskan tetraploid clade), and *S. glauca*. *Salix glauca* showed three clades, two in the NA clade and one in the EuA clade. Sections *Chamaetia* Dumort., Eurasian *Vetrix* Dumort., and *Vimen* Dumort. were monophyletic in our phylogeny, but the other taxonomic sections were not. Within the NA clade, the sectional classification was not well supported. Our data revealed that *Salix* sections *Cinerella* Seringe (= *Vetrix*), *Glaucae* (Fries) Andersson, *Hastatae* (Fries) A. Kerner, *Lanatae* (Andersson) Koehne, *Myrtilloides* (Borrer) Andersson, *Phylicifoliae* (Fries) Andersson, and *Villosae* (Andersson) Rouy were polyphyletic. These sections comprise species from North America and Eurasia. Sections *Hastatae*, *Myrtilloides*, and *Phylicifoliae* not only showed a split between Old and New World species, but also were polyphyletic within the NA and EuA clades. In *Salix* section *Myrtosalix* A. Kerner, the majority of species formed a monophyletic group (*S. alpina* Scop., *S. breviserrata* Flod., *S. uva-ursi*, *S. saxatilis* Turcz. ex Ledeb., and *S. myrsinites* L.) within the EuA clade. Despite its taxonomic assignment to section *Myrtosalix*, *S. rotundifolia* Trautv. grouped in the NA clade. Within the NA clade, *Salix* sections *Geyeriana* Argus, *Cinerella* (*Vetrix*), *Sitchenses* (Bebb) C.K. Schneider, and *Diplodictyae* C.K. Schneider were para- or

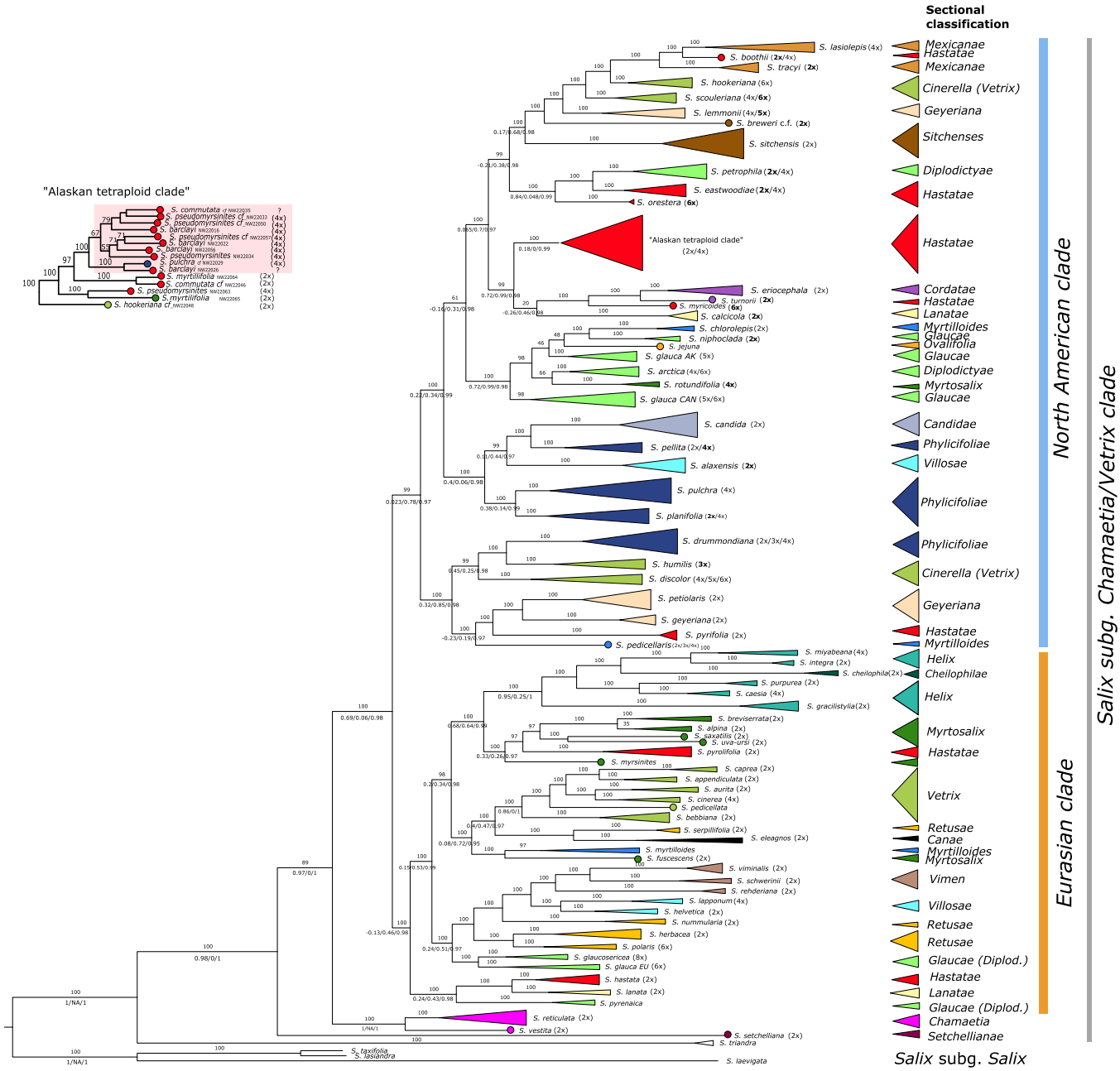

FIGURE 1 RAxML phylogeny based on 34,953 loci and 202 samples representing 78 species. The tree was rooted with three samples of *Salix* subg. *Salix* s.l. Species are color-coded according to their sectional classification Argus (1997) and Skvortsov et al. (1999); sections next to clades follow the same color scheme. The split into the North American (NA) and the Eurasian (EuA) clade is indicated by bars. Ploidy levels indicated in bold refer to new measurements. The Alaskan tetraploid willow clade is presented in detail in the upper left corner. Bootstrap values are above the branches, quartet-sampling values below. The detailed phylogeny with all tips is in Appendices S4, S5.

polyphyletic, respectively. These sections included exclusively North American species. Section "*Cinerella*" (Argus) (partly overlapping with section *Vetrix* sensu Skvortsov, see Discussion) was paraphyletic in the NA clade, separating into two distinct clades, one composed of *S. hookeriana* Barratt ex Hook. and *S. scouleriana* Barratt ex Hook. and the other of *S. humilis* Marsh. and *S. discolor* Muhl.

The NeighborNet network based on SNP data is presented with the RAxML phylogeny for comparison in Appendix S7. We reduced the network to the species in the North American clade to show discordance within this clade. Multiple accessions of one species remained together, except for the species of the Alaskan tetraploid clade. Most clades showed the same composition of species in the NeighborNet analysis and in the RAxML analysis, as indicated by colors (Appendix S7). The NeighborNet indicated a high amount of admixture between the species of the clade containing *S. lasiolepis* Benth., *S. lemmonii*

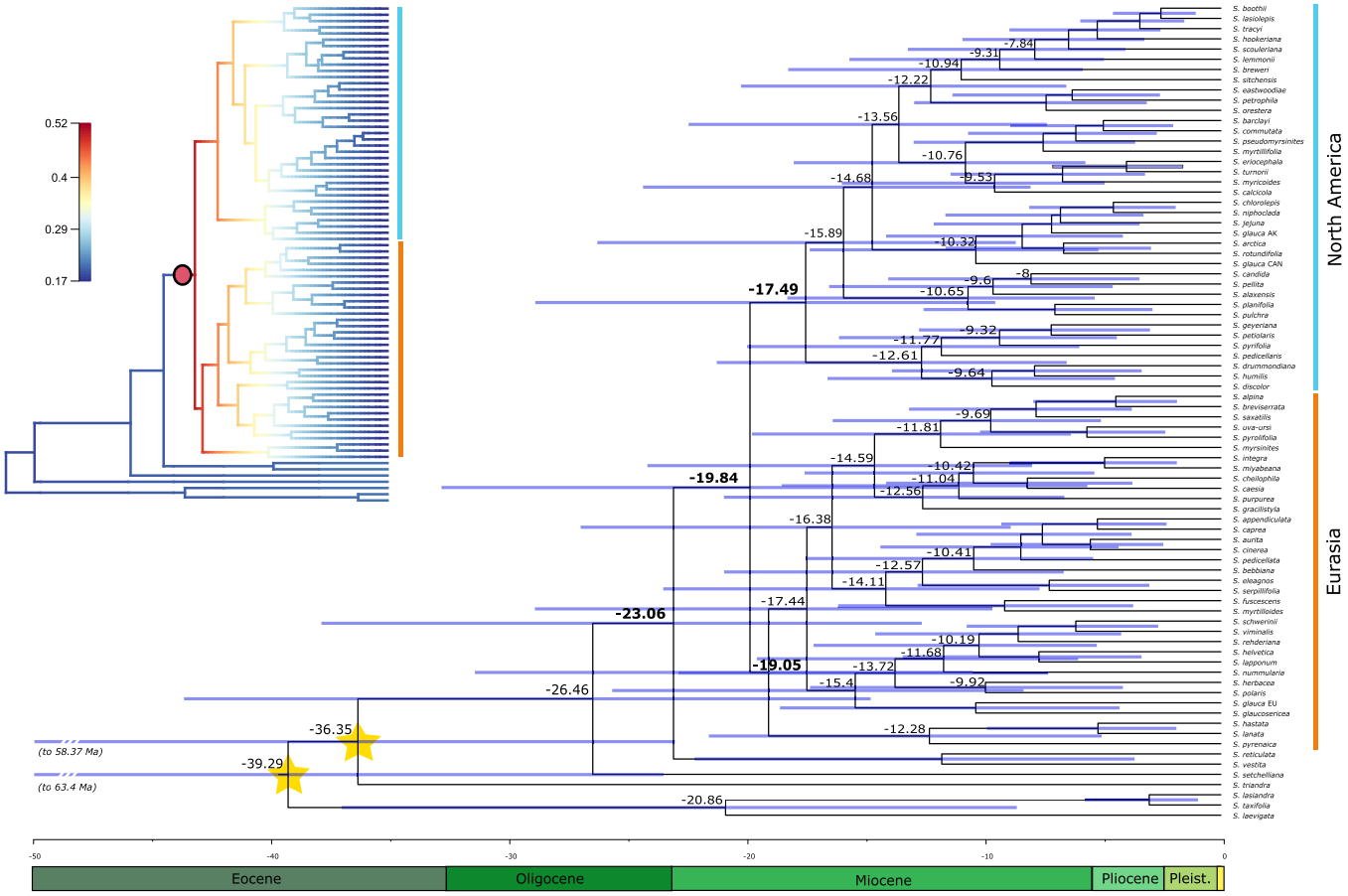

FIGURE 2 Ultrametric tree of the *Chamaetia/Vetrix* clade based on a reduced sampling with one sample per species. Calibration points are indicated with yellow stars (see Materials and Methods). The best estimate of the rate shift ($f = 0.69$) estimated using the BAMM analysis based on the same data set on the left side. The North American (NA) and the Eurasian (EuA) clade are indicated with colored bars.

Bebb, *S. tracyi* C.R. Ball, *S. hookeriana*, and *S. scouleriana*. A similar pattern was observed for the Alaskan tetraploid clade. For the clade containing *S. calcicola* Fernald & Wiegand, a split of this species and the remaining entities of this clade occurred in the NeighborNet analyses. The analyses revealed many squares between members of clades at an early-branching position in the RAxML analyses. For example, polyploid *S. drummondiana* was separate from *S. humilis* and *S. discolor*, however, showed reticulation with *S. pulchra* Cham. and *S. planifolia* Pursh, which were close to *S. pellita* (Andersson) Bebb and *S. candida* Flüggé ex Willd. in the network.

Estimated ploidy levels

We were able to observe and measure a peak in FC measurements for 117 silica samples, but no peak was obtained for 16 samples. By estimating ploidy levels of molecular data with nQuire, we were able to successfully identify the ploidy of 89 samples and a ploidy level higher than tetraploid for 26 samples, for a total of 115 estimations. For 18 samples, the observed histograms did not show a clear pattern, and model estimation was not possible. The results were reported in Appendix S1.

We reported ploidy levels for 10 species with no previous records for ploidy. While most of our ploidy determinations confirmed results of former studies, for 13 samples representing seven species, the ploidy levels differed from the literature. We came to inconclusive results for nine samples of seven species, hinting, however, toward the possibility of new ploidy levels that we were unable to confirm in this study. We found more than one ploidy level among samples of three species (*S. drummondiana*, *S. glauca*, *S. planifolia*). While polyploid samples were distributed over the entire tree without any specific pattern, we observed a tetraploid shrub willow clade comprising samples of several species collected in Alaska (Figure 1). We also detected a tendency toward similar ploidy levels among several sister species across the phylogeny (e.g., *S. planifolia* and *S. pulchra* [4*x*], *S. petiolaris* Sm. and *S. geyeriana* [2*x*]).

Morphological character coding

The results of 14 selected characters (Figures 3–5; Appendix S8) revealed no exclusively shared character for

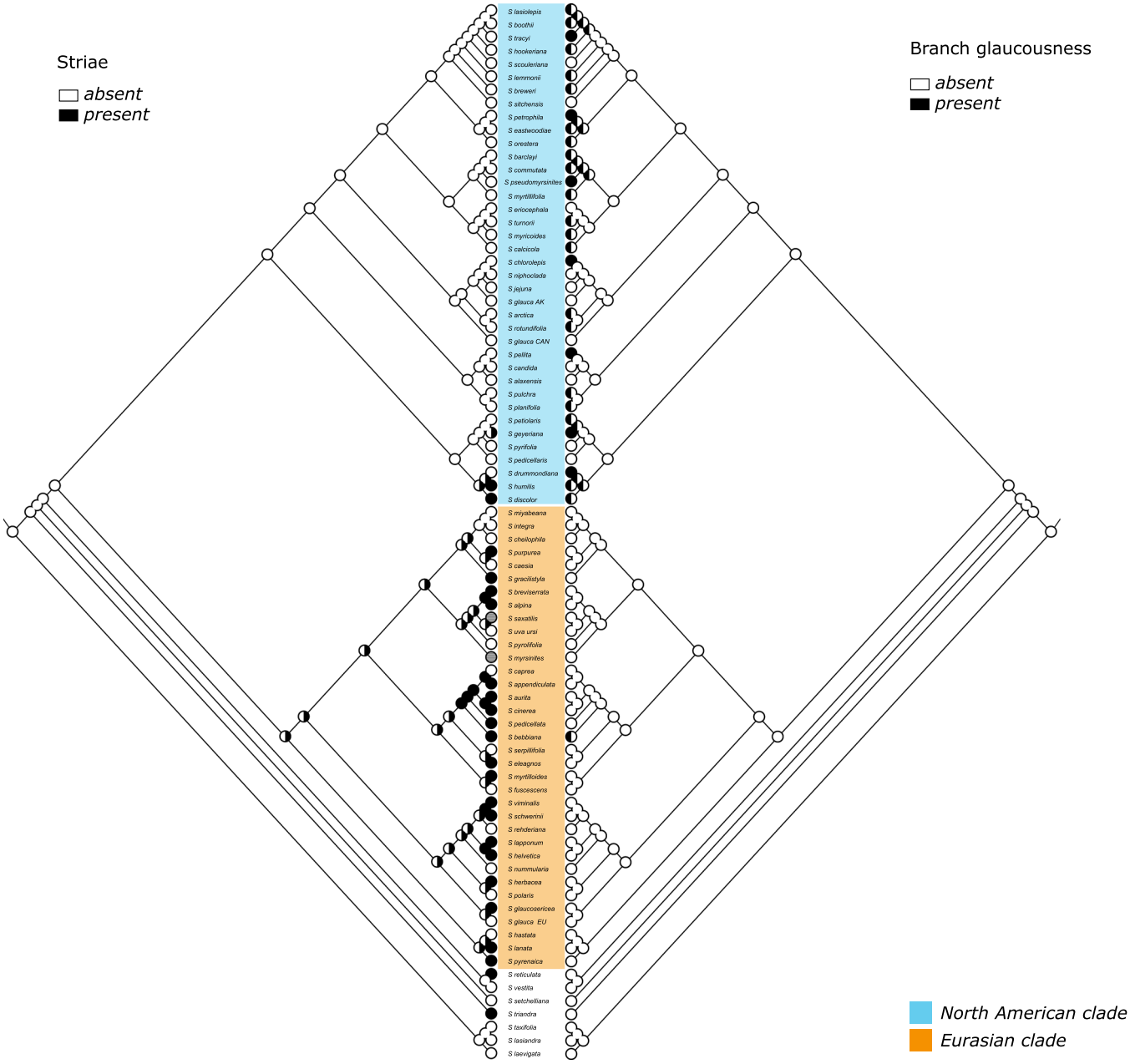

FIGURE 3 Ancestral character state analysis for striae and branch glaucousness on the singletons phylogeny.

the NA nor EuA clade, nor for a single clade or lineage. Nevertheless, some characters, especially those related to reproductive structures, seemed to be informative in regard to certain clades and subgroups.

The dwarf habit evolved several times independently in both major clades (Appendix S8). Branch glaucousness is a character that is present in several clades in the NA clade, but almost completely lacking among the sampled Eurasian species. Striae, vertical marks at the branches below the bark, were present in most included Eurasian species, but almost entirely lacking in the North American clade (Figure 3). The characters bract color and anther color had similar shared patterns in the NA clade: bracts are generally dark, and anthers tend to turn from red or purple to yellow in most samples (Appendix S8). In the Eurasian clade, we found a higher tendency toward light-colored and bicolor bracts and unchanging yellow or red/purple anthers. Traits associated with vegetative leaves were highly variable throughout the shrub willows, though they sometimes showed clade-specific patterns. Those characters were denticulation of margin, glaucousness of the abaxial (lower) leaf blade, and ad-/abaxial leaf indumentum (Appendix S8). Characters associated with inflorescences and flowers, interestingly, often coincided with the phylogenetic clades identified in this study. Those are time of flowering, leaf type on catkin branchlets, presence of staminate abaxial

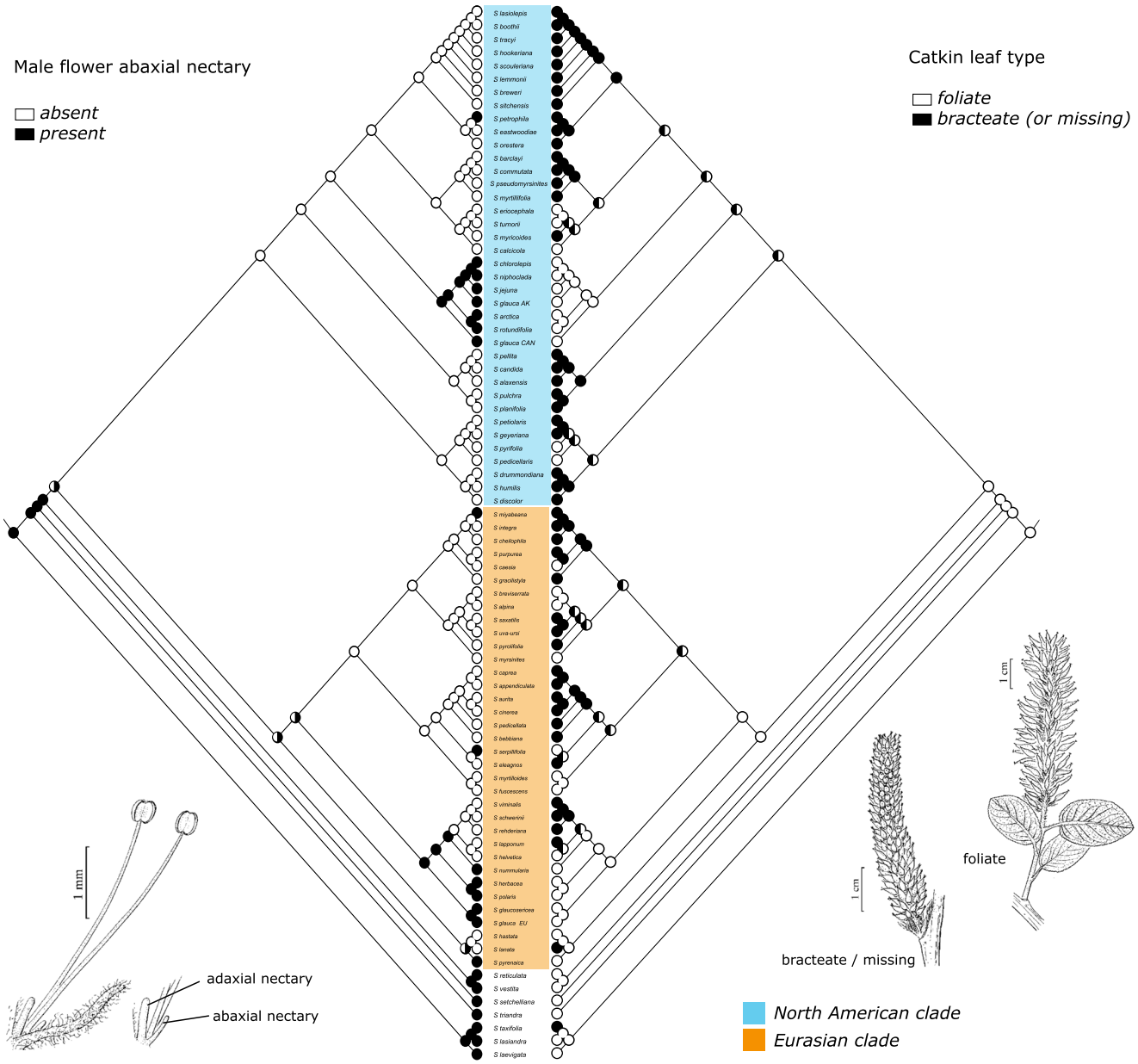

FIGURE 4 Ancestral character state analysis for male flower abaxial nectary and catkin leaf type on the singletons phylogeny. Illustrations of characters abaxial nectary and bracteate/foliate catkins are from Argus (2010, John Myers, illustrator), courtesy of the Flora of North America Association.

nectary, color of floral bracts, size of pistillate adaxial nectary (in relation to stipes), and ovary indumentum (Figures 4, 5; Appendix S8).

DISCUSSION

Shrub willows (*Salix* subg. *Chamaetia/Vetrix*), with more than 100 species, have represented a challenging study system in North America for decades due to their diversity and the lack of informative molecular data sets. Existing studies were often based on a sparse, mostly geographically restricted sampling and lacked phylogenetic resolution (Percy et al., 2014; Lauron-Moreau et al., 2015; Wu et al., 2015; Zhang et al., 2018; Sanderson et al., 2020; Gulyaev et al., 2022). Recent studies using modern genomic methods, such as target enrichment data, showed conflicting signals and highlighted ancestral hybridization events (Sanderson et al., 2020, 2023). Our study presents additional insights into the phylogeny of North American shrub willows based on a comprehensive sampling of the *Chamaetia/Vetrix* clade and a good representation of the species' genomes. The comparison of different data sets (Appendix S3) revealed the importance of a sufficient number of informative sites to receive a well-supported phylogeny in closely related groups. The comparison showed that the clades themselves were generally stable. In contrast, the backbone topology showed some incongruencies, which were also reflected by the QS scores observed in the RAxML phylogeny

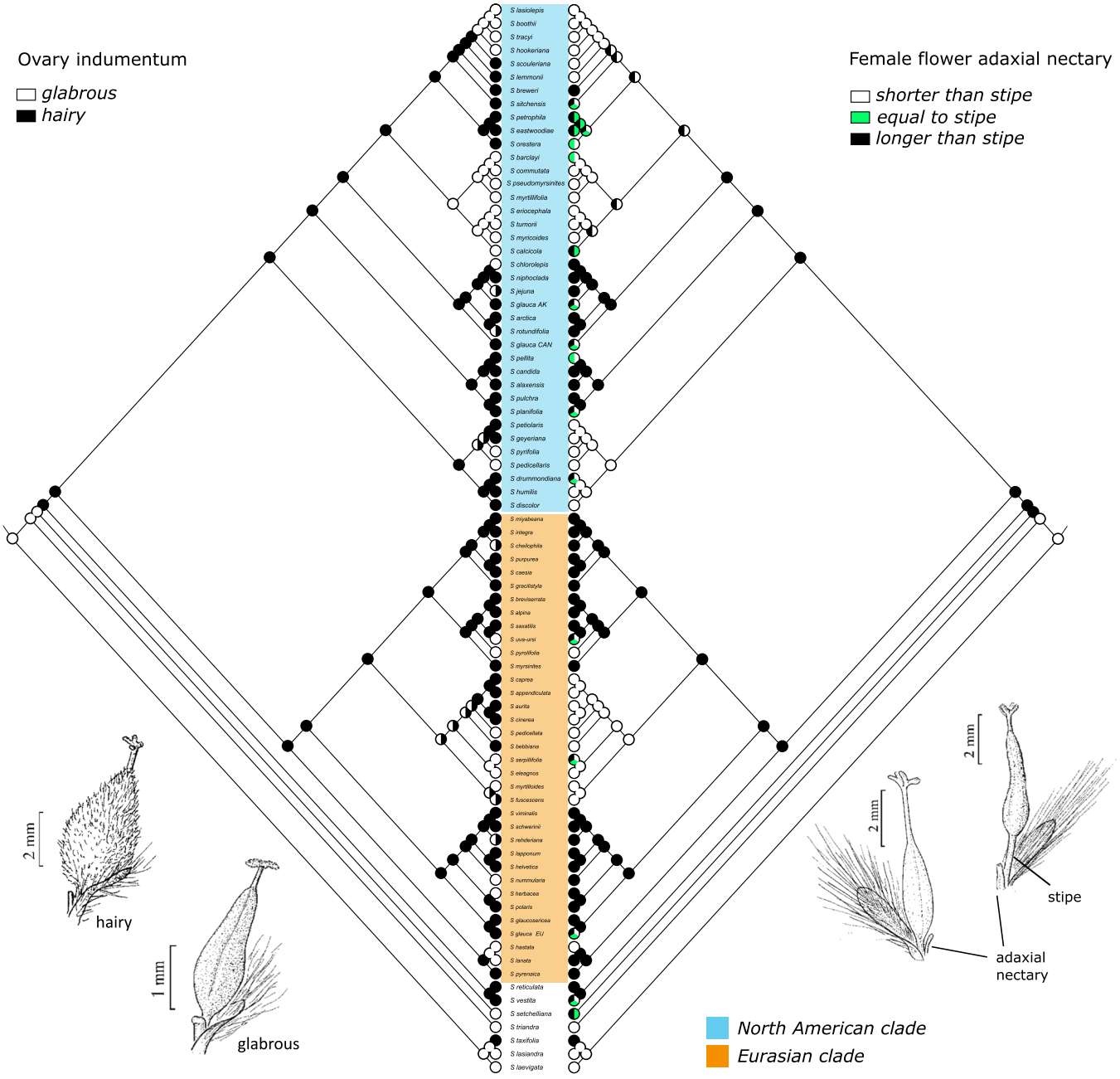

FIGURE 5 Ancestral character state analysis for ovary indumentum and female flower adaxial nectary on the singletons phylogeny. Illustrations of characters ovary indumentum and female flower adaxial nectary are from Argus (2010, John Myers, illustrator), courtesy of the Flora of North America Association.

and the resulting NeighborNet network (Appendices S5, S7). The observed incongruencies in the backbone topology are in accordance with past results (Wagner et al., 2020; Sanderson et al., 2023).

Morphological character state analyses

Traditional sectional classification was based on morphological characters (Argus, 2010). However, in *Salix*, identification at the species level is beset with difficulty due to the phenotypic plasticity that individual willow species present, especially in the face of hybridization. Our study illustrates how these factors pose difficulties for the taxonomic treatment of shrub willows. Although the current taxonomic classification is not supported by our data (Figure 1), we could show that some morphological traits form a pattern within certain clades. Based on our data, the character states related to leaves (form, margin, and indumentum) were highly variable, both intraspecifically and within phylogenetic clades. However, some lineages shared some leaf traits. For example, the clade

containing *S. arctica*, *S. glauca*, and *S. niphoclada* share the character of entire leaf margin (Appendix S8). Thus, leaf-associated characters might be useful in small-scale analyses. In general, we observed a clade-specific pattern of character states that were attached to ecological adaptation (e.g., size, hairiness of leaves and ovaries, bract color). Characters related to reproductive structures (flowers and/or catkins) were more informative for evolutionary lineages as reviewed before for angiosperms in general (Kay et al., 2006). These morphological characters may be most useful when attempting to redefine the sections to accommodate for this new phylogeny.

Comparing the NA and the EuA clade, our data showed that none of the characters tested is exclusive for one clade. However, the data revealed, for example, that striae (Figure 3), an important character used to identify many species in Europe (Skvortsov et al., 1999; Hörandl et al., 2012), seems to be prominent in Eurasia, whereas only a few North American species had this trait. Interestingly, it has never been used for species identification in the *Flora of North America* (Argus, 2010). Next, dark floral bracts are more common in North America and based on our results, is a shared character state for this clade that was lost several times (Appendix S8). Dark bracts might act as an adaptation to cold climate, especially in early spring. The bracts might improve flower development by increasing the flower's temperature, which was shown for example in *Rheum mobile*, a high Himalayan Polygonaceae (Song et al., 2013). In Eurasia, exclusively dark bracts are present mainly in species of high latitudes, such as *S. saxatilis* and *S. lanata* L., while most tested Eurasian species have bicolored and light bracts. Eurasian lowland species that were naturalized in North America, like *S. caprea* L., *S. cinerea* L., and *S. purpurea*, are described with exclusively bicolor bracts in European literature (Lautenschlager-Fleury and Lautenschlager-Fleury, 1994; Hörandl et al., 2012), but with dark bracts in the *Flora of North America* (Argus, 2010). This feature might be a local adaptation to the comparative colder climate at the same latitudinal range in North America or the result of a bottleneck event or local introgression of the introduced species with the native ones. Since we did not include samples of these species originating from North America in our analyses, we cannot test these hypotheses.

On the basis of our data, we could identify some shared morphological traits for certain subclades (Figures 3–5). To present one example in more detail, the clade of arctic-alpine, small to medium-sized shrub willows, composed of *S. chlorolepis* Fernald, *S. niphoclada* Rydb., *S. jejuna* Fernald, *S. glauca*, *S. arctica*, and *S. rotundifolia,* shared the presence of an abaxial nectary at the male flower and a larger adaxial nectary on female flowers. These observations suggest the importance of pollinator attraction for this clade (Figures 4, 5). Abaxial staminate floral nectaries were considered an ancestral character in the genus *Salix* and were traditionally used for subgeneric, as well as sectional classification, linking dwarf shrubs to tree willows (Argus, 1997; Skvortsov et al., 1999). Our study suggests that their presence is more likely the result of an adaptation to arctic-alpine environments, as shown for high-altitude Pan-Himalayan *Salix* species (He et al., 2021) and further confirmed by our study because European alpine species (e.g., *S. serpillifolia* Scop.) also had this character (Figure 4). The clade also shared the presence of foliate, vegetative-like leaves on flowering branchlets, a nonglaucous abaxial leaf surface and an entire leaf margin (only *S. glauca* sometimes with serrulate leaves) (Appendix S8). All these shared morphological traits support a common origin for the members of this clade. However, these shared characters evolved several times elsewhere in the phylogeny. Thus, the evolution of adaptations to arctic-alpine environments happened several times independently, showing the ability of shrub willows to express certain characteristics repeatably. This result is in congruence with former studies, e.g., the independent evolution of dwarf shrubs (Wagner et al., 2018).

Our character coding analyses illustrate two things. First, more than a few characters are needed to classify the species in a section. These difficulties were already presented by Argus (1997) where several New World sections were linked together with a plethora of morphological characteristics. Second, a system such as shrub willows can evolve phenotypic characters multiple times. Convergent evolution could be shown for many characters used for subgeneric and sectional classification, e.g., growth habit (Wagner et al., 2018) and other adaptations to the harsh environment of arctic-alpine habitats (Körner, 2003). The variation in morphology might further be promoted by polyploidy (Comai, 2005; Paun et al., 2009; Soltis and Soltis, 2009), which is present in all major clades (Figure 1; Wagner et al., 2020, 2023). Our results lead to the conclusion that for shrub willows, each evolutionary lineage needs to be analyzed separately when assessing morphological traits. Indeed, we were able to identify several shared characters within biological clades.

Divergence time and speciation rates

The divergence time for the *Chamaetia/Vetrix* clade as estimated in the BEAST analysis (excluding *S. setchelliana*) in our study is in accordance with some past findings (Wu et al., 2015; He et al., 2021). Both studies used the 23-Myr-old *Salix* subg. *Vetrix* fossil for their calibration, but used it in different prior settings compared to our study. Kikuchi et al. (2023) used this fossil to calibrate the root node of *Chamaetia/Vetrix*, which resulted in a divergence time about 10 Myr earlier than in our study. The divergence of the *Chamaetia/Vetrix* clade as estimated in the SNAPPER analysis, however, suggested a younger divergence time, about 14 Ma, which agrees with the results of Percy et al. (2014) and Zhang et al. (2018). Time divergence analyses always present a special challenge because they require extensive care in the selection of settings, especially when it comes to the use of information for time calibration and settings of the priors. Contrasting results of different studies

likely come from the use of different calibration strategies rather than from the data or the phylogenetic and dating algorithms used. In our study, however, we used the same calibration settings in two methodologically different analyses and a different data set and obtained different estimations for the nodes associated with the *Chamaetia/Vetrix* clade. We were unable to evaluate the suitability of both time divergence analyses. The methodologically less-appropriate BEAST analysis was based on the full data set of unlinked SNPs, and the analysis with the SNAPPER plugin could only analyze 1% of all the unlinked SNPs, but utilized a more appropriate model for our data set. Nonetheless, the estimated diversification time of the NA clade falls in the Miocene (~23 to 5 Ma) for both estimation approaches. This time period is associated with a warmer climate, peaking in the middle Miocene (17 to 15 Ma) followed by a gradual cooling and reestablishment of a major ice sheet on Antarctica by 10 Ma (Zachos et al., 2001). It is interesting to note that our different estimated times of diversification scenarios of the North Americas shrub willows, 17 and 10 Ma, correspond to the climatic extremes of the Miocene. However, a more detailed investigation would be needed to correlate divergence with either, or both, climatic extremes. Our results also indicate that the lineages of the *Chamaetia/Vetrix* clade pre-dated the climatic oscillations of the Quaternary, events that acted as a booster for speciation in many plant lineages (summarized by Kadereit and Abbott, 2021). We did not observe increased speciation rates and rapid radiation of the North American clade of shrub willows during that period, which is in contrast to findings for many other woody lineages in North America (e.g., Hipp et al., 2020). Instead, we observed an increase in speciation rate for the whole *Chamaetia/Vetrix* clade, which was also observed in a previous study on *Salix* (Sanderson et al., 2023). Anyway, the spatiotemporal evolution of willows of North America needs more research.

Polyploidization and hybridization

The observed diversity in ploidy levels in North American willows, sometimes within species, has also been shown previously (Argus, 2010). The slightly lower number of polyploids in the EuA clade is biased by our reduced sampling for this clade with a focus on diploids (but see Wagner et al., 2020, 2023). Presumably, 40% of willow species are polyploid (Suda and Argus, 1968) and have a base chromosome number of $x = 19$ (Argus, 1997). For seven species, our ploidy determinations differed from published records. In addition, inconclusive results for six species show that existing ploidy reports might be incomplete and need further assessment. Usually, ploidy records in the literature were based on observations of a single or a few individuals (Suda and Argus, 1968). In addition, difficulties in plant identification might also explain some of the discrepancies with our results. Hybridization of species with different ploidy levels might cause unexpected ploidy determinations for single individuals that were morphologically identified as a specific species (e.g., in the case of pentaploid *S. lemmonii*, Figure 1; Appendix S1). However, multiple ploidy levels within a single *Salix* species have been reported in the past (Argus, 1965, 2010; Kosiński et al., 2019), e.g., for *S. drummondiana* and *S. glauca* (see Appendix S1). Auto-polyploidy is associated with evolutionary advantages facilitating niche expansion (Spoelhof et al., 2017), which can lead to spatial isolation that has been associated with establishment of autopolyploid *Salix* lineages, as shown for the Asian *S. polyclona* complex (He et al., 2023). We did not test specifically the genomic composition of species that had more than one ploidy level. Nevertheless, the location in the same clade (e.g., 2*x* and 4*x* for *S. planifolia*, Figure 1) might indicate an autopolyploid origin.

A more detailed analysis beyond ploidy screening was out of the scope of this study; therefore, we did not test for allopolyploid origin. Even so, our data showed the frequent and independent occurrence of polyploids in different lineages in the NA clade (see Figure 1). The low QS values for internal nodes (Figure 1; Appendix S5) and the observed pattern in the SplitsTree NeighborNet network (Appendix S7) are an indicator for conflicting signal in the NA clade and thus support our assumption of recent and ancient polyploid formation. Allopolyploids tend to be more variable in their morphological and ecological plasticity and can rapidly establish reproductive barriers with the parent species. These barriers may facilitate the evolution of independent lineages consequently having a huge impact on speciation rates and diversity of certain plant groups (Soltis and Soltis, 2009; Barker et al., 2016). We observed a tetraploid clade comprising four sympatric shrub species from Alaska. These four species also shared some common morphological characteristics and were located in sister position of diploid species. However, most of the observed polyploids were scattered over the tree, and no clade was exclusively diploid (Figure 1). Interestingly, all samples of native species collected in North America that grouped outside the NA clade were diploids: *S. bebbiana*, *S. herbacea*, *S. lanata*, *S. uva-ursi*, *S. reticulata*, and *S. vestita*. This result might suggest that the formation of North American polyploids happened more recently and within North America. We are aware that our methods presented here are not well suited for analyzing polyploid relationships (discussed elsewhere). Nevertheless, our results give some valuable support to the assumption that polyploidy played a major role in the speciation and diversity of willows (Wagner et al., 2020, 2023). More advanced and integrative research methods will hopefully help to better evaluate the role of polyploidy in the evolution of shrub willows in the future.

In addition to polyploidy, hybridization between willow species is frequent and a well-known phenomenon that was previously studied by several authors (e.g., Neumann, 1981; Hardig et al., 2000; Oberprieler et al., 2013; Gramlich et al., 2018; Marinček et al., 2023). Hybridization can boost

genetic variance, and consequently, hybrids might be able to inhabit different ecological niches than their parental species (Mallet, 2007). However, not all hybrid events result in speciation (Gramlich et al., 2018). Due to the high amount of hybridization in willows and the conflicting signals observed in their target enrichment data set, Sanderson et al. (2023) described *Salix* as a syngameon, that is, a species complex of related taxa with frequent hybridization and partial reproductive isolation. This pattern has also been shown in other willow studies (Hardig et al., 2000; Murphy et al., 2022) and has been well studied in oaks (Hipp et al., 2020). In this study, we did not focus on hybrid evolution in willows, but were aiming to provide a robust backbone phylogeny. Thus, we avoided hybrids from our analyses that were identified by uncertain morphology. Nevertheless, we could specifically show netlike relationships for some of the Alaskan shrub willows, i.e., *S. barclayi*, *S. commutata*, *S. pseudomyrsinites*, and *S. myrtillifolia*. These four species grow in sympatry in glacier forefields or wet bogs, and they also show morphological continuums as well as a shift from diploidy to tetraploidy (Alaskan tetraploid clade, Figure 1; Appendix S4). This pattern is a common observation for species complexes (e.g., *Ranunculus auricomus* L. complex; Karbstein et al., 2022) and was reported in *Salix* before (He et al., 2023). We suggest that this complex presents an example of a syngameon of not-yet distinct lineages with high amounts of hybridization followed by allopolyploidization and thus needs further phylogenetic and taxonomic genetic investigation with a more comprehensive sampling.

About the sectional classification of North American shrub willows

The observed phylogeny confirmed previous findings regarding the split into two major clades representing tree willows (*Salix* subg. *Salix* s.l.) and shrub willows (*Salix* subg. *Chamaetia/Vetrix* clade). Within the shrub willows, two clades were observed: a purely NA clade, comprising most of the samples originating from North America, and a EuA clade. Our results contrast with those of Sanderson et al. (2023) and Gulyaev et al. (2022); however, our taxon sampling differed from their studies. The split into a New World and an Old World clade was previously shown only for tree willows (Chen et al., 2010). In accordance with other studies, our results illustrated that there is no support for the separation of the two subgenera *Chamaetia* and *Vetrix* (Wu et al., 2015; Wagner et al., 2018, Sanderson et al., 2023). The composition of the observed clades is partly in accordance with previously published studies, for example, the sister relationships of *S. viminalis*, *S. udensis* Trautv. et Mey., and *S. arbusculoides* Andersson or the close relationship of *S. commutata* and *S. myrtillifolia*, as shown by Sanderson et al. (2023). For many clades, it is impossible to draw more detailed conclusions here since the taxon sampling differs between our study and past studies.

Argus's (1997) sectional classification of shrub willows in North America included the classification of species into sections that contain Eurasian species and sections that contain exclusively North American species. Most of the included species were monophyletic and well supported. However, his sectional classification is only partly supported by our phylogeny (Figure 1). We tried to include Eurasian members of all sections that also contained species in North America. In several polyphyletic sections, we observed a split between Old World and New World taxa of the respective section. Additionally, within each of these major clades, the sections were not monophyletic, e.g., *Glaucae* and *Phylicifoliae*. The polyphyly of sections is in accordance with previous studies on Eurasian taxa (Wagner et al., 2023). Argus (1997) stated that sections *Hastatae*, *Myrtilloides*, and *Myrtosalix* are labile in both species composition and placement. Our molecular data support this view, especially for section *Hastatae*, and suggest that the variation within these sections is based on at least six independent evolutionary lineages (Figure 1).

Salix setchelliana from western North America was sister to the remaining species of the *Chamaetia/Vetrix* clade and was also supported by the results of Lauron-Moreau et al. (2015). We assume this special willow species is a vestige of a very old lineage of shrub willows. The isolated position of *S. reticulata*/*S. vestita* (sect. *Chamaetia*) on a long branch as sister to the remaining clade had already been suggested by other authors (Skvortsov et al., 1999; Wagner et al., 2018, 2020; Sanderson et al., 2023; Volf et al., 2023) and is now confirmed by our data.

The members of section *Cinerella* include species that were assigned to section *Vetrix* by Skvortsov et al. (1999). Therefore, we decided to use the same color code in Figure 1 for both sections. For section *Cinerella* (*Vetrix*), not only did we observe a split between the New World and Old World, but also the section was polyphyletic within the NA clade (Figure 1). The split within the NA clade into two distant positions in the phylogeny may reflect different possible allopolyploid origins. Indeed, the four included species are all polyploid (4*x*, 6*x*). In contrast, the European members of section *Vetrix* in the Eurasian clade formed a well-supported clade including *S. bebbiana*, composed of five diploid and one polyploid species. The close relationships of *S. caprea*, *S. bebbiana*, and *S. aurita* L. (synonymous to our Eurasian "*Vetrix*" clade) were already presented by Lauron-Moreau et al. (2015) based on two plastid and one nuclear marker.

Salix sect. *Phylicifoliae* was polyphyletic. Although we did not include members of this section from Eurasia, a recently published study showed polyphyly of sect. *Phylicifoliae* in Europe (Wagner et al., 2023). However, except for *S. drummondiana*, all the other species of this section grouped together in one clade and were tetraploid, just as the European members of this section were also polyploids (Figure 1; Wagner et al., 2023). The position of *S. alaxensis* (Andersson) Coville (sect. *Villosae*) within this clade is a bit surprising; however, the sister species *S. pellita* and *S.*

candida do share some morphological leaf traits with *S. alaxensis*, such as adaxial silky hairs and abaxial dark green and wrinkled surfaces (Argus, 2010; Appendix S8). The position of this species was also supported by a recent study of Sanderson et al. (2023) based on 787 genes in which *S. alaxensis* appeared in sister relationship to *S. drummondiana* and *S. candida*.

Salix section *Glaucae*, especially the species *S. glauca*, show a high amount of morphological variation, and thus the section was the subject of taxonomic debate (Argus, 1965, 1997). Up to five varieties were described; four of them can be found in North America (Argus, 1997, 2010). We did not include many specimens and thus did not try to identify them at the variety level. However, our data presented here showed that the specimens we identified as "*S. glauca*" belong to at least three independent lineages (Figure 1): (1) a European *S. glauca* lineage, which is in sister position to the alpine species *S. glaucosericea* Flod. (*S. glauca* EU, Figure 1), (2) a lineage represented here by shrubby accessions from Alaska (*S. glauca* AK, Figure 1), and (3) a lineage represented by samples from Canada (*S. glauca* CAN, Figure 1). The results suggest that in *Salix*, species may sometimes comprise several independent lineages (possibly representing distinct species). However, independent allopolyploid origin or hybridization with many other species, as reported for *S. glauca*, might influence the topology. Further studies with more diverse samples from specimens across the range and covering varieties of *S. glauca* are necessary to understand the phylogenetic relationships and evolutionary origin of these morphologically similar specimens of *S. glauca*. Skvortsov et al. (1999) placed *S. arctica* within section *Glaucae*, while Argus (1997) placed it in section *Diplodictyae*. Our data placed *S. arctica* in a clade of arctic-alpine shrub and dwarf shrub species that belong to five different *Salix* sections [*Diplodictyae*, *Myrtosalix*, *Myrtilloides*, *Glaucae*, and *Ovalifoliae* (Rydb.) C.K. Schneider]. This clade included diploid, tetraploid, and hexaploid species, and the good support implies that all of the included species originated from the same lineage. Thus, the unification of this clade within one section should be considered.

Salix orestera (6*x*), *S. eastwoodiae* Cockerell ex A. Heller (2*x*) (reported as tetraploid in the literature, but determined as diploid in our study; see Appendix S1), both from sect. *Hastatae*, and *S. petrophila* Rydb. (2*x*) from sect. *Diplodictyae* were all collected in the Sierra Nevada. All three species have overlapping distribution ranges and formed a well-supported clade in our phylogeny. For *S. orestera*, Argus (1997) already speculated that this species might be of hybrid/allopolyploid origin, possibly involving *S. eastwoodiae* or *S. lemmonii*. Finally, it should be noted that our positioning of *S. petrophila*, a small alpine dwarf shrub, is in strong contrast to the findings of Lauron-Moreau et al. (2015) who found this species situated in sister position to the tree willows (*Salix* subg. *Salix* s.l.) based on ITS and plastid markers. We consider this placement as rather unlikely. The morphologically closest species, *S. arctica* and *S. uva-ursi*, were situated in distant lineages (Figure 1), confirming convergent evolution to arctic-alpine environment of shrub willows.

In summary, the sectional taxonomic treatment of Argus (1997), which was based on classical morphological data, was not supported by our molecular data. Nevertheless, our results showed some patterns that indicate close relationships of species that belong to the same section and share some morphological characters. Our study presents a good starting point for a new taxonomic treatment of North American willows. Here, we tried to collect representative species of all sections; however, we only collected about half of the species diversity of the NA shrub willows. Since the sections were not monophyletic, a more complete species set is needed to identify all evolutionary lineages in shrub willows.

CONCLUSIONS

We presented a comprehensive and well-resolved phylogeny of North American shrub willows. The included species were mainly monophyletic; however, the existing sectional classification was not supported by our data. Nevertheless, some shared morphological characters serve to circumscribe the relationships within a monophyletic lineage. Our study shows the need to review the existing traditional classification in the *Salix* subgenus *Chamaetia/Vetrix* clade. The diversity of shrub willows was shaped by a radiation starting in the Miocene within North America. The morphological diversity was further shaped by hybridization and polyploidization events, which is reflected by the presence of polyploid species in each observed clade. Future studies on the spatiotemporal evolution of shrub willows and on their hybridization and polyploidization will shed further light on the mystery surrounding the diversity of these species in North America.

AUTHOR CONTRIBUTIONS

P.M.: Formal analysis; methodology; software; visualization; writing—original draft; writing—review and editing. E.L-.B.: Formal analysis; methodology; writing—review and editing. F.H.: Methodology; resources; writing—review and editing. J.L.: Methodology. S.M.B.: Resources; writing—review and editing. M.V.: Data curation; funding acquisition; methodology; writing—review and editing. N.D.W.: Conceptualization; Funding acquisition; investigation; methodology; project administration; resources; supervision; validation; visualization; writing—original draft; writing—review and editing.

ACKNOWLEDGEMENTS

This paper is dedicated to George W. Argus. We thank Justin Fulkerson (AK), Bruce Baldwin (CAL), Maria Tomoshevich (RU), John Bradican and Elvira Hörandl (GOET), Jeannine Cavender-Bares (MN), Petr Kozel (MN), and Lucy Shroeder

(MN) for help with plant collections and support for collection permits. We thank Marc Appelhans and Gabriele Liessmann for support with mounting and digitization of herbarium vouchers. M.V. and J.V.L. acknowledges funding by The Czech Science Foundation (GACR 23-06855 L). N.W. acknowledges funding by the priority program Taxon-Omics SPP1991 of the DFG (Wa3684/2-1 to N.W.). Sampling in Minnesota was supported by NSF ASCEND BII (DBI 2021898) and Cedar Creek LTER (DEB 1831944) grants. We thank anonymous reviewers for their useful comments on this manuscript. Open Access funding enabled and organized by Projekt DEAL

DATA AVAILABILITY STATEMENT

All demultiplexed raw read data resulting from RAD-seq were submitted to the National Center for Biotechnology Information (NCBI) in the Sequence Read Archive (SRA) under the BioProject ID PRJNA433286. Details are provided in Appendix S1. Filtered and assembled alignments were provided at https://github.com/NataschaWagner28/NorthAmerican_Salix.

ORCID

Pia Marinček http://orcid.org/0000-0003-1620-6516
Étienne Léveillé-Bourret http://orcid.org/0000-0002-0069-0430
Ferris Heiduk http://orcid.org/0009-0009-0246-729X
Jing Leong http://orcid.org/0000-0002-4698-5613
Martin Volf http://orcid.org/0000-0003-4126-3897
Natascha D. Wagner http://orcid.org/0000-0001-6623-7623

REFERENCES

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data, version 0.12.1 for Windows. Computer program and documentation distribted by the author. Website: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [accessed 06 December 2023].

Argus, G. W. 1965. The taxonomy of the *Salix glauca* complex in North America. *Contributions from the Gray Herbarium of Harvard University* 196: 1–142.

Argus, G. W. 1997. Infrageneric classification of *Salix* (Salicaceae) in the New World. *Systematic Botany Monographs* 52: 1–121.

Argus, G. W. 2004. A guide to the identification of *Salix* (willows) in Alaska, the Yukon Territory and adjacent regions. Self-published for July 2004 workshop on willow identification. Website: https://www.naturebob.com/sites/default/files/GuideSalixAK-YT11May05.pdf [accessed 06 December 2023].

Argus, G. W. 2010. *Salix*. *In* Flora of North America Editorial Committee [eds.], Flora of North America north of Mexico [online], Vol. 7, 23–51. Website: http://floranorthamerica.org/Salix. [accessed 06 December 2023].

Argus, G. W. 2012. *Salix*. *In* Jepson Flora Project [eds.], Jepson eFLora. Website: https://ucjeps.berkeley.edu/eflora/eflora_display.php?tid=9887 [accessed 06 December 2023].

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: 1–7.

Barker, M. S., B. C. Husband, and J. C. Pires. 2016. Spreading Winge and flying high: The evolutionary importance of polyploidy after a century of study. *American Journal of Botany* 103: 1139–1145.

Bouckaert, R., T. Vaughan, C. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10(4): e1003537.

Bryant, D., and V. Moulton. 2002. NeighborNet: an agglomerative method for the construction of planar phylogenetic networks. *In* R. Guigó and D. Gusfield [eds.], Proceedings of Second International Workshop on Algorithms in Bioinformatics, WABI 2002, Rome, Italy, 2002. Lecture notes in computer science, vol. 2452, 375–391. Springer, Berlin, Germany.

Cavender-Bares, J., A. González-Rodríguez, D. A. Eaton, A. A. Hipp, A. Beulke, and P. S. Manos. 2015. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. *Molecular Ecology*, 24: 3668–3687.

Chen, J. H., H. Sun, J. Wen, and Y. P. Yang, 2010. Molecular phylogeny of *Salix* L. (Salicaceae) inferred from three chloroplast datasets and its systematic implications. *Taxon* 59: 29–37.

Collet, D. M. 2002. Willows of Southcentral Alaska. Kenai Watershed Forum, Soldotna, AK, USA.

Collinson, M. E. 1992. The early fossil history of Salicaceae: a brief review. *Proceedings of the Royal Society of Edinburgh, B, Biological Sciences* 98: 155–167.

Comai, L. 2005. The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* 6: 836–846.

Cronk, Q., E. Ruzzier, I. Belyaeva, and D. Percy. 2015. *Salix* transect of Europe: latitudinal patterns in willow diversity from Greece to arctic Norway. *Biodiversity Data Journal* 3: e6258.

De Mestier, A., G. Brokamp, M. Celis, B. Falcón-Hidalgo, J. Gutiérrez, and T. Borsch. 2022. Character evolution and biogeography of *Casearia* (Salicaceae): evidence for the South American origin of a pantropical genus and for multiple migrations to the Caribbean islands. *Taxon* 71: 321–347.

Doležel, J., Greilhuber, J., and Suda, J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2: 2233–2244.

Eaton, D. A. R., and I. Overcast. 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36: 2592–2594.

Gramlich, S., N. D. Wagner, and E. Hörandl. 2018. RAD-seq reveals genetic structure of the F2-generation of natural willow hybrids (*Salix* L.) and a great potential for interspecific introgression. *BMC Plant Biology* 18: 317.

Gulyaev, S., X. J. Cai, F. Y. Guo, S. Kikuchi, W. L. Applequist, Z. X. Zhang, E. Hörandl, and L. He. 2022. The phylogeny of *Salix* revealed by whole genome re-sequencing suggests different sex-determination systems in major groups of the genus. *Annals of Botany* 129: 485–498.

Guo, S. X. 1975. The plant of the Xigaze Group from Mount Jolmo Lungma Region. *In* Tibet Scientific Expedition Team, Academia Sinica [eds.], Report of scientific expedition to Mt. Jolmo Lungma Region, 411–425. Science Press, Beijing [in Chinese].

Hardig, T. M., S. J. Brunsfeld, R. S. Fritz, M. Morgan, and C. M. Orians. 2000. Morphological and molecular evidence for hybridization and introgression in a willow (*Salix*) hybrid zone. *Molecular Ecology* 9: 9–24.

He, L., N. D. Wagner, and E. Hörandl. 2021. Restriction-site associated DNA sequencing data reveal a radiation of willow species (*Salix* L., Salicaceae) in the Hengduan Mountains and adjacent areas. *Journal of Systematics and Evolution* 59: 44–57.

He, L., F. Guo, X. Cai, H. Chen, C. Lian, Y. Wang, C. Shang, et al. 2023. Evolutionary origin and establishment of a dioecious diploid-tetraploid complex. *Molecular Ecology* 32: 2732–2749.

Heled, J., and R. R. Bouckaert. 2013. Looking for trees in the forest: summary tree from posterior samples. *BMC Evolutionary Biology* 13: 1–11.

Hipp, A. L., P. S. Manos, M. Hahn, M. Avishai, C. Bodénès, J. Cavender-Bares, A. A. Crowl, et al. 2020. Genomic landscape of the global oak phylogeny. *New Phytologist* 226: 1198–1212.

Hoag, J. C. 2005. Simple identification key to common willows, cottonwoods, alder, birch and dogwood of the Intermountain West. *Riparian Wetland Project Information Series* 19: 1–16.

Hörandl, E., F. Florineth, and F. Hadacek. 2012. Weiden in Österreich und angrenzenden Gebieten (willows in Austria and adjacent regions), 2nd ed. University of Agriculture of Vienna, Vienna, Austria.

Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.

Kadereit, J. W., and R. J. Abbott. 2021. Plant speciation in the Quaternary. *Plant Ecology and Diversity* 14: 105–142.

Karbstein, K., S. Tomasello, L. Hodač, N. Wagner, P. Marinček, B. H. Barke, C. Paetzold, and E. Hörandl. 2022. Untying Gordian knots: unraveling reticulate polyploid plant evolution by genomic data using the large *Ranunculus auricomus* species complex. *New Phytologist* 235: 2081–2098.

Kay, K. M., C. Voelckel, J. Y. Yang, K. M. Hufford, D. D. Kaska, and S. A. Hodges. 2006. Floral characters and species diversification. *In* L. D. Harder and S. C. H. Barret [eds.], Ecology and evolution of flowers, 311–325. Oxford University Press, Oxford, UK.

Kikuchi, S., S. Setsuko, T. Nagamitsu, and W. Suzuki. 2023. Molecular phylogenetic analyses identify the process of speciation of endemic willow species in the Japanese Archipelago. *Researchsquare*. Website: https://doi.org/10.21203/rs.3.rs-3070336/v1

Körner, C. 2003. Alpine plant life: Functional plant ecology of high mountain ecosystems, 2nd ed. Springer-Verlag, Berlin, Germany.

Kosiński, P., E. Sliwinska, A. Hilpold, and A. Boratyński. 2019. DNA ploidy in *Salix retusa* agg. only partly in line with its morphology and taxonomy. *Nordic Journal of Botany* 37: e02197.

Lauron-Moreau, A., F. E. Pitre, G. W. Argus, M. Labrecque, and L. Brouillet. 2015. Phylogenetic relationships of American willows (*Salix* L., Salicaceae). *PLoS One* 10: e0138963.

Lautenschlager-Fleury, D., and E. Lautenschlager-Fleury. 1994. Die Weiden von Mittel- und Nordeuropa: Bestimmungsschlüssel und Artbeschreibungen für die Gattung Salix L. Springer-Verlag Basel, Basel, Switzerland.

Leitch, I. J., and M. D Bennett. 2004. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* 82: 651–663.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* 25: 1754–1760.

Liu, X., Z. Wang, W. Wang, Q. Huang, Y. Zeng, Y. Jin, H. Li, et al. 2022. Origin and evolutionary history of *Populus* (Salicaceae): further insights based on time divergence and biogeographic analysis. *Frontiers in Plant Science* 13: 1031087.

Loureiro, J., E. Rodriguez, J. Doležel, and C. Santos. 2006. Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. *Annals of Botany* 98: 515–527.

Maddison, W. P., and D. R. Maddison. 2023. Mesquite: a modular system for evolutionary analysis, version 3.81 for Windows. Computer program and documentation distributed by the author. Website: http://www.mesquiteproject.org [accessed 05 September 2023].

Mallet, J. 2007. Hybrid speciation. *Nature* 446: 279–283.

Marinček, P., L. Pittet, N. D. Wagner, and E. Hörandl. 2023. Evolution of a hybrid zone of two willow species (*Salix* L.) in the European Alps analyzed by RAD-seq and morphometrics. *Ecology and Evolution* 13: e9700.

Murphy, E. K., E. P. Cappa, R. Y. Soolanayakanahally, Y. A. El-Kassaby, I. A. P. Parkin, W. R. Schroeder, and S. D. Mansfield. 2022. Unweaving the population structure and genetic diversity of Canadian shrub willow. *Scientific Reports* 12: 17254.

Neumann, A. 1981. Die mitteleuropaeischen *Salix*-Arten. *Mitteilungen der forstlichen Bundes-Versuchanstalt* 134: 1–151.

Oberprieler, C., L. Dietz, C. Harlander, and J. Heilmann. 2013. Molecular and phytochemical evidence for the taxonomic integrity of *Salix alba*, *S. fragilis*, and their hybrid *S.* × *rubens* (Salicaceae) in mixed stands in SE Germany. *Plant Systematics and Evolution* 299: 1107–1118.

Otto, F. 1990. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. *Methods Cell Biology* 33: 105–110.

Paradis, E., and K. Schliep. 2019. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.

Paun, O., F. Forest, M. F. Fay, and M. W. Chase. 2009. Hybrid speciation in angiosperms: parental divergence drives ploidy. *New Phytologist* 182: 507–518.

Pease, J. B., J. W. Brown, J. F. Walker, C. E. Hinchliff, and S. A. Smith. 2018. Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany* 105: 385–403.

Percy, D. M., G. W. Argus, Q. C. Cronk, A. J. Fazekas, P. R. Kesanakurti, K. S. Burgess, B. C. Husband, et al. 2014. Understanding the spectacular failure of DNA barcoding in willows (*Salix*): Does this result from a trans-specific selective sweep? *Molecular Ecology* 23: 4737–4756.

Plummer, M, N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6: 7–11.

R Core Team. 2023. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: https://www.R-project.org/

Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9: e89543.

Rabosky, D. L., M. Grundler, C. Anderson, P. Title, J. J. Shi, J. W. Brown, H. Huang, and J. G. Larson. 2014. BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* 5: 701–707.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.

Sanderson, B. J., S. P. DiFazio, Q. C. Cronk, T. Ma, and M. S. Olson. 2020. A targeted sequence capture array for phylogenetics and population genomics in the Salicaceae. *Applications in Plant Sciences* 8: e11394.

Sanderson, B. J., D. Gambhir, G. Feng, N. Hu, Q. C. Cronk, D. M. Percy, F. M. Freaner, et al. 2023. Phylogenomics reveals patterns of ancient hybridization and differential diversification that contribute to phylogenetic conflict in willows, poplars, and close relatives. *Systematic Biology* 72: 1220–1232.

Skvortsov, A. K. 1999. Willows of Russia and adjacent countries. Taxonomical and geographical revision [I. N. Kadis, translator]. *In* A. G. Zinovjev, G. W. Argus, J. Tahvanainen, and H. Roininen [eds.], Report series 39, Faculty of Mathematics and Natural Sciences, University of Joensuu, Joensuu, Finland.

Spoelhof, J. P., P. S. Soltis, and D. E. Soltis. 2017. Pure polyploidy: Closing the gaps in autopolyploid research. Journal of Systematics and Evolution 55: 340–352.

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

Song, B., J. Stöcklin, Z. Zhang, Y. Yang, and H. Sun. 2013. Seed and microsite limitation in *Rheum nobile*, a rare and endemic plant from the subnival zone of Sino-Himalaya. *Plant Ecology and Diversity* 6: 503–509.

Stoltz, M., B. Baeumer, R. Bouckaert, C. Fox, G. Hiscott, and D. Bryant. 2021. Bayesian inference of species trees using diffusion models. *Systematic Biology* 70: 145–161.

Suda, Y., and G. W. Argus. 1968. Chromosome numbers of some North American *Salix*. *Brittonia* 20: 191–197.

Viereck, L. A., and E. L. Little. 1972. Alaska trees and shrubs. Agriculture Handbook No. 410. U.S.D.A. Forest Service, Washington D.C., Washington, USA.

Volf, M., J. V. Leong, P. de Lima Ferreira, T. Volfová, P. Kozel, P. Matos-Maraví, E. Hörandl, et al. 2023. Contrasting levels of β-diversity and underlying phylogenetic trends indicate different paths to chemical diversity in highland and lowland willow species. *Ecology Letters* 26: 1559–1571.

Wagner, N. D., S. Gramlich, and E. Hörandl. 2018. RAD sequencing resolved phylogenetic relationships in European shrub willows (*Salix* L. subg. *Chamaetia* and subg. *Vetrix*) and revealed multiple evolution of dwarf shrubs. *Ecology and Evolution* 8: 8243–8255.

Wagner, N. D., L. He, and E. Hörandl. 2020. Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Frontiers in Plant Science* 11: 1–38.

Wagner, N. D., L. He, and E. Hörandl. 2021a. The evolutionary history, diversity, and ecology of willows (*Salix* L.) in the European Alps. *Diversity* 13: 1–16.

Wagner, N. D., P. Marinček, L. Pittet, and E. Hörandl. 2023. Insights into the taxonomically challenging hexaploid alpine shrub willows of *Salix* sections *Phylicifoliae* and *Nigricantes* (Salicaceae). *Plants* 12: 1144.

Wagner, N. D., M. Volf, and E. Hörandl. 2021b. Highly diverse shrub willows (*Salix* L.) share highly similar plastomes. *Frontiers in Plant Science* 12: 662715.

Weiß, C. L., M. Pais, L. M. Cano, S. Kamoun, and H. A. Burbano. 2018. nQuire: a statistical framework for ploidy estimation using next generation sequencing. *BMC Bioinformatics* 19: 1–8.

Wolfe, J. A. 1987. An overview of the origins of the modern vegetation and flora of the northern Rocky Mountains. *Annals of the Missouri Botanical Garden* 74: 785–803.

Wu, J., T. Nyman, D. Wang, G. W. Argus, Y. Yang, and J. Chen. 2015. Phylogeny of *Salix* subgenus *Salix* s.l. (Salicaceae): delimitation, biogeography, and reticulate evolution. *BMC Evolutionary Biology* 15: 31.

Zachos, J., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292: 686–693.

Zhang, L., Z. Xi, M. Wang, X. Guo, and T. Ma. 2018. Plastome phylogeny and lineage diversification of Salicaceae with focus on poplars and willows. *International Journal of Business Innovation and Research* 17: 7817–7823.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Sampling table containing detailed information on sample localities, GenBank accession numbers, and results of ploidy determination.

Appendix S2. Character coding table containing all tested character states and definitions thereof.

Appendix S3. Comparison of RAxML analyses based on four different thresholds for loci shared by "m" samples (m15, m40, m100, m195).

Appendix S4. Detailed RAxML phylogeny of 202 taxa based on RAD sequencing data (detailed version of Figure 1).

Appendix S5. Quartet sampling results on RAxML phylogeny of 202 taxa based on RAD sequencing data.

Appendix S6. Ultrametric tree and subsequent BAMM analyses based on 560 filtered biallelic SNPs using SNAPPER.

Appendix S7. Comparison of RAxML and NeighborNet network analyses (only North American clade) using SplitsTree.

Appendix S8. Results of ancestral character state analyses of 14 selected morphological characters.

How to cite this article: Marinček, P., É. Léveillé-Bourret, F. Heiduk, J. Leong, S. M. Bailleul, M. Volf, and N. D. Wagner. 2024. Challenge accepted: Evolutionary lineages versus taxonomic classification of North American shrub willows (*Salix*). *American Journal of Botany* e16361. https://doi.org/10.1002/ajb2.16361