Phylogenomics reveals patterns of ancient hybridization and differential diversification contributing to phylogenetic conflict in *Populus* L. and *Salix* L.

Brian J. Sanderson*[1,8,11], Diksha Ghambir[1], Guanqiao Feng[1], Nan Hu[1], Quentin C. Cronk[2], Diana M. Percy[2], Francisco Molina Freaner[3], Matthew G. Johnson[1], Lawrence B. Smart[4], Ken Keefover-Ring[5], Tongming Yin[6], Tao Ma[7], Stephen P. DiFazio[8], Jianquan Liu[7,9], Matthew S. Olson*[1,10]

[1]Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409-3131 USA

[2]Department of Botany, University of British Columbia, Vancouver, BC, V6T 1Z4 Canada

[3]Universidad Nacional Automoa de Mexico, Hermosilla, Mexico

[4]Horticulture Section, School of Integrative Plant Science, Cornell University, Cornell AgriTech, Geneva, New York 14456 USA

[5]Departments of Botany and Geography, University of Wisconsin-Madison, Madison, WI 53706, USA

[6]Key Laboratory of Tree Genetics and Biotechnology of Jiangsu Province and Education Department of China, Nanjing Forestry University, Nanjing, China

[7]Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education & College of Life Sciences, Sichuan University, Chengdu 610065, China

[8]Department of Biology, West Virginia University, Morgantown, WV, 26506 USA

[9]State Key Laboratory of Grassland Agro-Ecosystem, Institute of Innovation Ecology & College of Life Sciences, Lanzhou University, Lanzhou 730000, China

[10]Author for correspondence: *matt.olson@ttu.edu*

*These authors contributed equally to this work.

[11]Current address: The Jackson Laboratory for Genomic Medicine, 10 Discovery Drive, Farmington, CT 06032 USA

Running head: Phylogenomics of Populus and Salix

Supplementary data files will be provided by request to matt.olson@ttu.edu

Abstract

Despite the economic, ecological, and scientific importance of the genera *Populus* L. (poplars, cottonwoods, and aspens) and *Salix* L. Salicaceae (willows), we know little about the sources of differences in species diversity between the genera and of the phylogenetic conflict that often confounds estimating phylogenetic trees. *Salix* subgenera and sections, in particular, have been difficult to classify, with one recent attempt termed a 'spectacular failure' due to a speculated radiation of the subgenera *Vetrix* and *Chamaetia*. Here we use targeted sequence capture to understand the evolutionary history of this portion of the Salicaceae plant family. Our phylogenetic hypothesis was based on 787 gene regions and identified extensive phylogenetic conflict among genes. Our analysis supported some previously described subgeneric relationships and confirmed polyphyly of others. Using an f_{branch} analysis we identified several cases of hybridization in deep branches of the phylogeny, which likely contributed to discordance among gene trees. In addition, we identified a rapid increase in diversification rate near the origination of the *Vetrix*-*Chamaetia* clade in *Salix*. This region of the tree coincided with several nodes that lacked strong statistical support, indicating a possible increase in incomplete

lineage sorting due to rapid diversification. The extraordinary level of both recent and ancient hybridization in both *Populus* and *Salix* have played important roles in the diversification and diversity in these two genera.

Keywords: Salicaceae**,** hybridization*,* sequence capture, ASTRAL species tree, concatenated tree

This study is dedicated to the memory of George W. Argus (1929-2022) whose lifelong pursuit of understanding diversity in Salix laid the foundation for future salicologists.

Introduction

As methods to assess the congruence among the genealogical histories of genes across species have matured (Degnan et al. 2009; Young et al. 2020), the curious association between phylogenetic conflict and rapid diversification has suggested a link between population genetic and macroevolutionary processes (Parins-Fukuchi et al. 2021). Although most genomic regions are expected to reflect the speciation and diversification history of a taxonomic group (species tree), two primary factors contribute to conflict between gene genealogies and species history. Incomplete lineage sorting (ILS) results from the persistence of polymorphism across multiple diversification events and the subsequent random fixation of polymorphism among different lineages (Wu 1991). The influence of ILS is particularly strong during periods of rapid speciation, when effective population sizes are large and when long-term balancing selection results in persistence of polymorphisms (Edwards 2009; Pease et al. 2013; Wang et al. 2020). Second, interspecific gene flow due to hybridization has the potential to generate discordance among high proportions of gene trees (McVay et al. 2017; Morales-Briones et al. 2020). Unlike

69 ILS, however, the patterns of gene discordance are biased with over-representation of one hybrid
70 topology (Green et al. 2010; Durand et al. 2011; Patterson et al. 2012). When taxonomic groups
71 have hybridized throughout diversification, the effects of hybridization on gene tree conflict can
72 span multiple species within a clade (Malinsky et al. 2018). The impacts of hybridization may be
73 particularly common in plants, where fertility of interspecific crosses may be maintained well
74 after speciation (Grant 1981).

75 Contemporary populations of poplars (*Populus*) and willows (*Salix*) are widely known to
76 hybridize, with important evolutionary and ecological consequences along hybrid zones
77 (Brunsfeld et al. 1992; Hardig et al. 2000; Schweitzer et al. 2004; Lexer et al. 2010; Chhatre et
78 al. 2018; Wang et al. 2020). The impacts of hybridization on the evolution and diversification are
79 evident in both genera, where multiple independent chloroplast capture events have occurred
80 early in their diversification (Smith et al. 1990; Brunsfeld et al. 1992; Liu et al. 2017; Wang et al.
81 2020; Gulyaev et al. 2022). This history of hybridization creates challenges for the
82 reconstruction of phylogenetic histories of poplars and willows (Percy et al. 2014). Recent
83 progress using genome-wide data sets, however, have constructed well-supported taxonomic
84 relationships within both genera (Wagner et al. 2020; Wang et al. 2020; Wagner et al. 2021a;
85 Gulyaev et al. 2022; Wang et al. 2022), but the sources of conflict among gene trees have not
86 been fully investigated, especially in *Salix*.

87 Poplars and willows are integral components of temperate, boreal, and arctic ecosystems
88 throughout the northern hemisphere and many species have significant cultural, medical, and
89 economic importance (Stettler et al. 1996; Argus 1997). Phytochemical diversity in these genera
90 spans an impressive array of secondary metabolites, including the aspirin and its derivatives
91 (Desborough et al. 2017) and defensive chemicals such as phenolic glycosides, condensed

92 tannins and hydroxycinnamate derivatives (Tsai et al. 2006; Philippe et al. 2007; Boeckler et al.
93 2011; Keefover-Ring et al. 2022). Morphological variation in both genera ranges from dwarf
94 creeping *Salix* in alpine and arctic zones that were once categorized as a separate genus (Stettler
95 et al. 1996; Argus 1997) to large *Populus* trees in subtropical zones. Many more species are
96 recognized in *Salix* (approx. 450-520 species; Argus 2010) than in *Populus* (approx. 100 species;
97 Shu 1999), suggesting that either *Salix* began to diversify much earlier than *Populus* or
98 speciation rates have increased in *Salix*.

99 *Populus* and *Salix* are the two largest genera in the Salicaceae and all but one species
100 across both genera are dioecious (Rohwer et al. 1984). The Salicaceae whole genome duplication
101 unites all genera except *Azara*, which lacks the duplication (Cronk et al. 2015). Within the
102 Salicaceae, *Populus* and *Salix* are united by a striking synapomorphy of flowers organized into
103 aments or catkins (Meeuse 1975; Argus 2010; Eckenwalder 2010; Cronk et al. 2015) with seed
104 dispersal via wind. This differs from their closest relative *Idesia*, which produces fleshy animal-
105 dispersed fruits. *Populus* and *Salix* differ in that pollen is dispersed by wind in *Populus* and by
106 insects or by both insects and wind in *Salix* (Sacchi et al. 1988; Tamura et al. 2000; Karrenberg
107 et al. 2002), suggesting that factors underlying pollination mode may drive differences in the
108 diversification rate between the two genera (Friedman et al. 2009; Wessinger 2021). The reduced
109 floral structures in both genera exhibit relatively low variability and are used to discriminate
110 among species only at the broadest taxonomic levels (Eckenwalder 1996; Argus 1997). Thus,
111 plant stature and growth form, leaf morphology, and bud characteristics have been important
112 characters for species identification (Dorn 1976; Argus 2010; Eckenwalder 2010) despite the
113 high intraspecific variability and propensity for plasticity of these traits (Wu et al. 2015).

114 Here we seek to understand the sources of phylogenetic conflict in both *Populus* and
115 *Salix* by comparing their historical patterns of hybridization and relative rates of diversification,
116 both factors that can contribute to conflict with species trees. Our ASTRAL phylogeny was
117 based on genome-wide sequence capture loci from a large set of mostly North American and
118 Asian diploid *Salix* and *Populus* species. The phylogeny of a subset of these *Salix* samples was
119 previously investigated using DNA barcode markers, but the resulting tree lacked resolution
120 (Percy et al. 2014). Also, a phylogeny of the majority of the *Populus* samples were previously
121 analyzed by Wang et al. (Wang et al. 2020) using a different set of whole genome loci, providing
122 a "positive control" to confirm that our targeted sequence capture array designed for Salicaceae
123 (Sanderson et al. 2020) successfully reconstructs the species tree. We use this tree to compare the
124 impact of hybridization on gene tree discordance and the chronology and rates of diversification
125 across these two genera. The specific goals of our study are to: 1) provide an integrated
126 phylogenetic hypothesis of the sister genera *Populus* and *Salix*, 2) to estimate the timing and
127 rates of diversifications of major clades within each genus, and 3) to assess the association
128 between hybridization history and regions of the phylogeny that exhibit conflict among gene
129 trees. Finally, we discuss the implications of our results in the context of other well-known
130 hybridizing groups of species (syngameons) species such as oaks (Cannon et al. 2020) and pines
131 (Buck et al. 2022).
132

133 **Methods**

134 The 166 samples included in this study were drawn from new collections, older dried
135 herbarium samples, and previously sequenced genomes, and represented all five *Populus*
136 sections, all five *Salix* subgenera, and 25 *Salix* sections (Table S1). All species were considered

137 diploids based on chromosome counts reported in www.tropicos.org except *S. discolor*, which is
138 likely a polyploid, and *S. richardsonii*, for which there was no information concerning
139 chromosome counts. *Salix* species were primarily native to North America, but also from Europe
140 and Asia. Fourteen *Populus* samples (7 species) and 83 *Salix* samples (45 species) were
141 genotyped using a custom sequence target capture kit designed for the Salicaceae (Supplemental
142 Methods; Sanderson et al. 2020). An additional 6 outgroup species (*Azara dentata*, *A.*
143 *integrifolia*, *A. lanceolata*, *A. microphylla*, *Carrierea calycina*, *Idesia polycarpa*, and
144 *Poliothyrsis sinensis*), 54 poplar samples (26 species), and 8 *Salix* samples (7 species) were
145 genotyped using whole genome sequencing. All sequences were assembled into putatively
146 homologous gene sequences using the HybPiper pipeline (Johnson et al. 2016); for the whole
147 genome sequences, this removed all loci except those included in the target capture array. After
148 filtering, removal of alignments with paralogs, and removal of genes with excessively long
149 branches in the gene tree, 787 alignments remained and were used for all downstream
150 phylogenetic analyses (See Supplemental Methods for details; Table S2).
151 Two approaches were used for phylogeny estimation: 1) A two-step approach using
152 ASTRAL that first estimated trees for each gene and then identified the best tree based on
153 minimizing quartet distances among gene trees (Mirarab et al. 2014), and 2) identifying the
154 maximum likelihood tree based on concatenating all genes in our sample. Full details of these
155 analyses are provided in the Supplemental Methods. In brief for the two-step approach, gene
156 trees were estimated using IQTREE 2.0.3 (Nguyen et al. 2015) and an ASTRAL tree of all
157 individuals and the species tree (Rabiee et al. 2019) were inferred using ASTRAL-MP (v 5.12.2
158 (Yin et al. 2019). In our ASTRAL tree of all individuals, four species were represented by
159 intraspecific samples that did not cluster together (Fig. S1): *Salix bebbiana*, *S. eriocephala*, *S.*

160 *pseudomonticola*, and *Populus ningshanica*; these are represented by a postscript 1 & 2 in the
161 figures and Table S1. For this reason, we used two individuals to represent each of these four
162 species in the ASTRAL species tree. Both local posterior probability and 100 multilocus
163 bootstraps were calculated for each ASTRAL tree. For the concatenation approach, a single
164 alignment was constructed by concatenating all 787 genes (1,058,955 sites). IQTREE was used
165 to estimate the most likely concatenated tree using the GTR+F+R10 substitution model. 1000
166 mulitlocus bootstraps and SH-alRT tests (Guindon et al. 2010) also were computed for this tree.
167 A dated species tree was calculated using *BEAST2 (Heled et al. 2010). Because of the
168 long computation times required for sampling, five genes were chosen for dated tree estimation
169 (9,264 sites: SapurV1A.0003s0350, SapurV1A.0045s0240, SapurV1A.0050s0650,
170 SapurV1A.0139s0330, SapurV1A.0260s0050). These genes were first selected for high
171 consistency with species tree topologies, minimized root-to-tip variance, and maximal tree length
172 (Smith et al. 2018) and then screened for consistency of basal nodes with the ASTRAL species
173 tree. Because these five genes represented <1% of the genes used for generating our species tree,
174 we selected an additional five genes using the same criteria to assess the consistency of the
175 results across different gene sets (8,568 sites: SapurV1A.0211s0160, SapurV1A.0789s0070,
176 SapurV1A.0857s0020, SapurV1A.0900s0040, SapurV1A.1178s0060). See the Supplemental
177 methods for additional information on gene selection and for *BEAST2 parameter settings. The
178 calibration date for the root of the tree (divergence between *Azara* and all others) was drawn
179 from a normal distribution with a mean of 65.0 Ma and standard deviation of 1.0 (following
180 Wang et al. 2020), and the calibration of the crown clade of *Populus* + *Salix* was drawn from a
181 normal distribution with a mean of 49 Ma and a standard deviation of 1.0 following Percy et al.
182 (2014) and based on the *Pseudosalix handleyi* fossil (Boucher et al. 2003) from the Eocene

183 Green River formation that has been dated at ca. 49 Ma (Smith et al. 2010). Using a wider
184 standard deviation of 3.0 for the distributions resulted in much larger variance in the estimates of
185 node ages, but only slight changes in the estimates for divergence times (Fig. S4).

186 Patterns of current and historical hybridization within the *Populus* and *Salix* clades were
187 estimated using ABBA-BABA, f_4, and f_{branch} analyses (Patterson et al. 2012; Malinsky et al.
188 2018) calculated using Dsuite (Malinsky et al. 2021). The f_{branch} analysis is heuristic and is
189 designed to account for phylogenetic correlation among f_4-ratio results calculated with
190 phylogenetically correlated samples. The f_{branch} metrics assign significance to internal branches
191 in the phylogeny when excess sharing of alleles that is consistent with hybridization is found
192 across a clade (Malinsky et al. 2018; Malinsky et al. 2021). We generated separate VCF files for
193 *Populus* and *Salix* to calculate f_4-ratio and f_{branch} statistics using Dsuite using GATK v4.2.6.1
194 (McKenna et al. 2010) as detailed in the Supplemental Methods.

195 Two methods were used to address diversification rates across the Salicaceae. First,
196 Bayesian Analysis of Macroevolutionary Mixtures (BAMM; Rabosky 2014) was used to identify
197 credible shifts in the diversification rate across lineages with the *expectedNumberOfShifts* prior
198 set to 1.0. Because our species-level sampling was not uniform across genera, we adjusted for
199 non-random incomplete taxon sampling (Table S3). For each dated tree, two independent
200 MCMC chains using different seeds were run for 10 million generations each resulting in
201 ESS>400, and the 95% credible set of shift configurations was calculated after removing 10%
202 for burn-in. Second, branch-specific diversification rates were estimated using the ClaDS model
203 (Maliet et al. 2019) and calculated using data augmentation (Maliet et al. 2022). Finally, we used
204 the "STructured Rate Permutations on Phylogenies" (STRAPP) test (Rabosky et al. 2015) to

assess the association between categorical traits associated with wind versus insect+wind pollination, wind versus animal seed dispersal, and tree versus shrub-dwarf shrub growth form.

Results

Of the 1216 loci that were targeted, 787 passed filtering criteria with an average of 162 individuals per gene (2.5% & 97.5% quantiles: 144 & 165), 1236 sites per sequence (2.5% & 97.5% quantiles: 673 & 2629), 305.0 parsimony-informative sites (2.5% & 97.5% quantiles: 134.7 & 673.7), 132.3 singletons (2.5% & 97.5% quantiles: 45 & 333.7), and 908.2 constant sites (2.5% & 97.5% quantiles: 452.0 & 1715.6) per gene. Our overall findings support monophyly for both *Salix* and *Populus*,

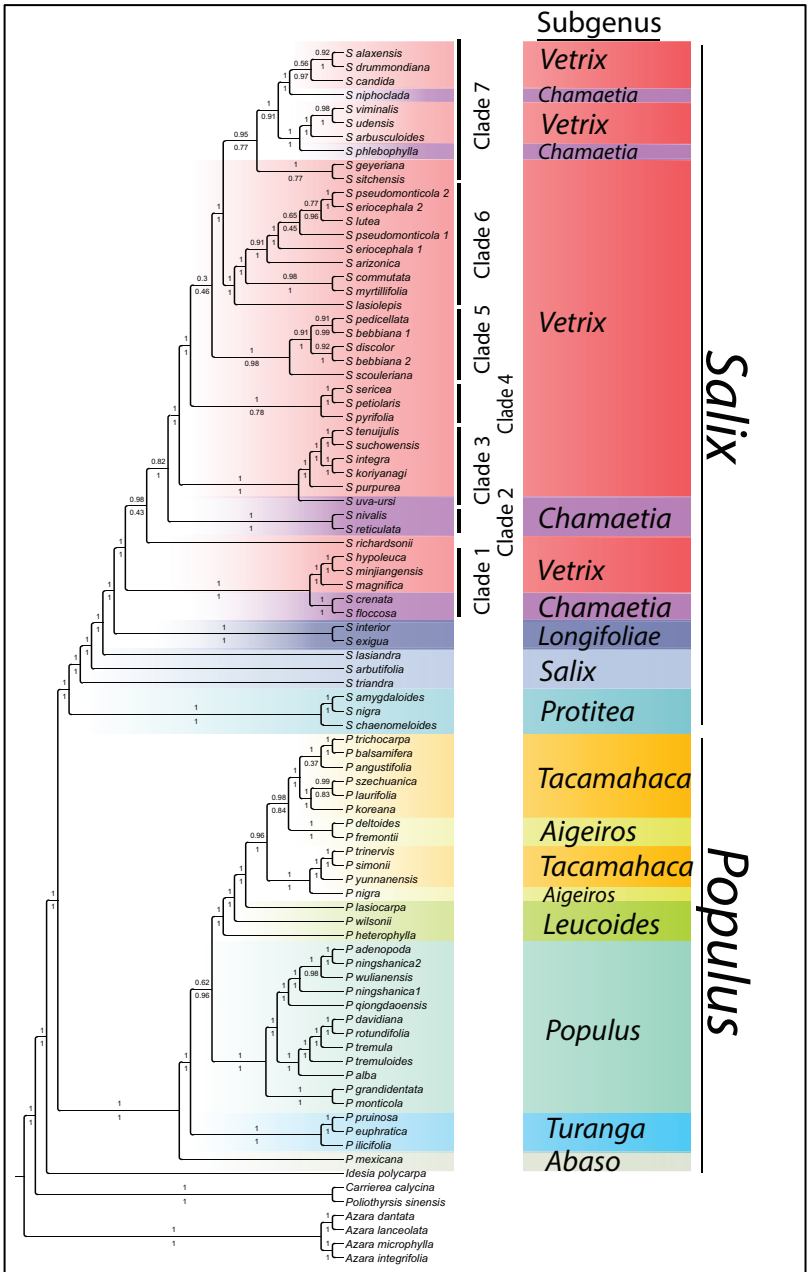

Figure 1. ASTRAL species tree constructed using 787 gene trees. Values above branches indicate posterior probability support and values below branches indicate bootstrap support. Although the most strongly supported relationship among clades is presented, clades with <90% posterior probability or <75% bootstrap support should be viewed with caution. Colors indicate taxon associations with historically identified subgenera.

228 and classification of the basal portions of each genus were consistent with prior morphological
229 and molecular analyses (Figs. 1, S1; Eckenwalder 1996; Wagner et al. 2018; Wang et al. 2020;
230 Gulyaev et al. 2022). Our phylogeny also supports recent discoveries of the polyphyly of the
231 *Tacamahaca* and *Aigeiros* subgenera in *Populus* (Wang et al. 2020) and the polyphyly of the
232 *Vetrix* and *Chamaetia* subgenera in *Salix*. Notably in *Salix*, our study and that of Wager et al.
233 (2018) each identified at least 4 independent evolutionary events leading to dwarf willows
234 (subgenus *Chamaetia*), which are prominent components of northern hemisphere arctic and
235 alpine ecosystems. However, because these studies had little taxonomic overlap, it is difficult to
236 discern whether we identified the same or different events.

237 Most of our samples identified as the same species clustered together in both the
238 ASTRAL tree of individuals and the concatenated tree (Fig. S1). Both the ASTRAL tree and the
239 concatenated tree exhibited high support values for many nodes (especially the concatenated
240 tree). In general, the ASTRAL tree supports the taxonomy of clades that diversified near the root
241 of each genus and provides less support for recently differentiated clades (Fig. S2). The
242 concatenated tree and ASTRAL tree differed in key topological components (Fig. S1) including
243 the placement of *Populus* subgenus *Turanga*. Morphological analyses (Eckenwalder 1996) and a
244 larger set of genes previously analyzed from many of the same *Populus* individuals using only
245 first and second codon positions (Wang et al. 2020), however, generally support a similar
246 topology as our ASTRAL tree (see Discussion). Within *Salix*, the concatenated and ASTRAL
247 trees concurred for the relative placement of subgenera *Protitea*, *Salix,* and *Longifoliae*, but
248 differed in the placement of groups within *Vetrix* and *Chamaetia*. Within subgenus *Salix* the
249 concatenated and species trees differed in the relative placement of *S. lasiandra*, *S. triandra*, and
250 *S. arbutifolia*. Finally, there was generally consistent support between the concatenated and

251 species tree for the monophyly of clades within the *Salix*+*Chamaetia* group with the exceptions
252 of the placement of *S. scouleriana*, *S. lasiolepis, S. richardsonii, S. geyeriana,* and *S. sitchensis*.
253 Importantly, the ASTRAL tree generally exhibited less bootstrap and SH-alRT support for nodes
254 than the concatenated tree, perhaps indicating that the confidence placed on relationships in the
255 concatenated tree is overestimated. Clades in *Populus* were generally more strongly supported by
256 consistency among gene trees than clades in *Salix* (Figs. 1, S3); this lack of support was
257 especially apparent in the backbone of the *Chamaetia*+*Vertix* clade**.** Because conflict among
258 gene trees likely results from either ancient or ongoing gene flow among taxonomic groups
259 (hybridization) or incomplete lineage sorting (ILS), we tested for patterns of historical gene flow
260 and rapid diversification in our tree.

261 Levels of biased intraspecific gene flow indicative of hybridization were approximately
262 the same in *Populus* and *Salix* (52% of D_{tree} and 60% of D_{min} were significant after Benjamini-
263 Hochberg adjustment in *Salix* vs. 64% of D_{tree} and 58% of D_{min} in *Populus* (Tables S4, S5, S6,
264 S7). Because D and f_4 statistics across a clade are phylogenetically correlated, we used the
265 heuristic f_{branch} analysis based on the ASTRAL species tree to assess the history and timing of
266 hybridization during the diversification of *Populus* and *Salix* (Fig. 2, Table S8, S9). The number
267 of f_{branch} statistics that indicated >5% gene flow across species due to hybridization were similar
268 in *Salix* (6.6% of f_{branch} values above 5%) and *Populus* (7.0% of f_{branch} values above 5%) but were
269 more commonly associated with deep internal branches of *Salix* than *Populus* (Tables S8, S9).
270 Up to three ancient hybridization events appear to have influenced interspecific gene flow in
271 *Salix*. First, the f_{branch} metrics indicated evidence for gene flow between the ancestors of both *S.*
272 *triandra* and *S. arbutifolia* and ancestors of the lineage leading to subgenera *Longifoliae*, *Vetrix*
273 and *Chamaetia* (f_{branch} = 0.088 & 0.113; arrow 1 in Fig. 2; Table S8). Second, ancient geneflow

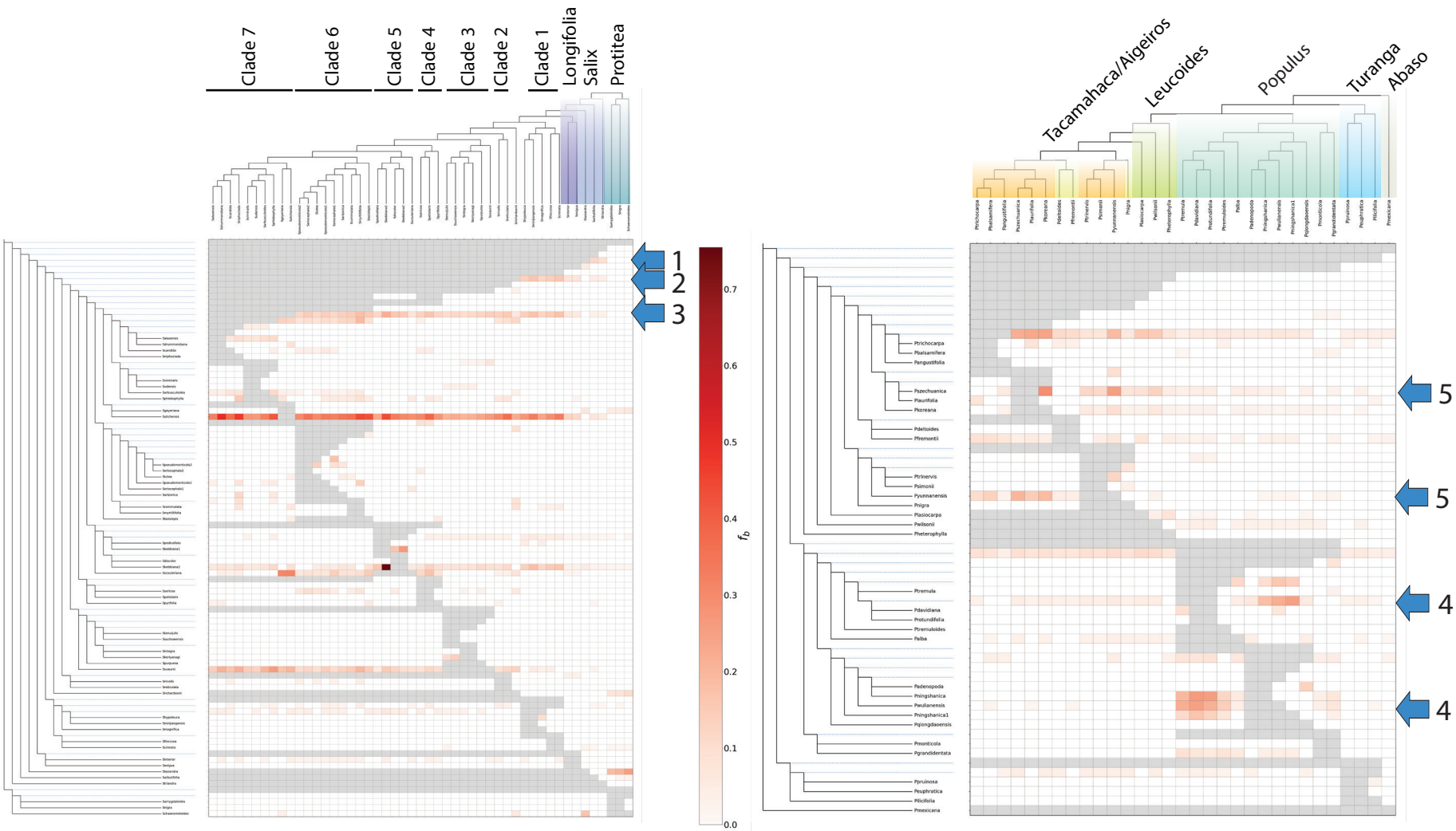

Figure 2. Heuristic f-branch analysis of ongoing and ancient hybridization in Salix and Populus. Arrows indicate gene flow events that are discussed in the text. Color scale in boxes represents the estimated fbranch value, which is indicative of the proportion of the genome that is affected by interspecific gene flow due to hybridization.

274 (f_{branch} > 0.05%) occurred between the ancestors of species in clade 1 with ancestors of clade 2-7
275 (f_{branch} = 0.126 to 0.164; arrow 2 in Fig. 2; Table S8). Finally, the f_{branch} analysis indicates ancient
276 hybridization among ancestors of clades 6 and 7 (f_{branch} = 0.078 to 0.186; arrow 3 in Fig. 2; Table
277 S8), as well as significant hybridization between ancestors of clades 6+7 and clades 1-5 (f_{branch} =
278 0.077 to 0.218; arrow 3 in Fig. 2; Table S8). The lack of evidence of hybridization in *S. geyriana*
279 and strong evidence for hybridization between *S. sitchensis* and an ancestor of clades1-7 suggests
280 that either *S. sitchensis* or *S. geyriana* may be misplaced on the phylogeny and the multiple
281 signals of hybridization result from a single ancient event. In *Populus* the f_{branch} analysis based on
282 the ASTRAL species tree indicated widespread hybridization among ancestors or extant
283 members of the two major clades comprising the *Populus* subgenus (f_{branch} = 0.120 to 0.259;
284 arrows labelled 4 in Fig. 2; Table S9) as well as signals of ancient or ongoing hybridization
285 within the *Tacamahaca/Aigeiros* clade and the *Leucoides* grade (arrows labelled 5 in Fig. 2;
286 Table S9).
287 The two data sets used to estimate the *BEAST2 dated species trees were both large
288 (>8500 sites) and drawn from genes in the top 7% of our objective criteria. Dating of many
289 nodes was within the overlap of 95% highest posterior density for the two estimates (HPD; Fig.
290 3; Tables S10, S11). For instance, the divergence between *Populus* and *Salix* was ~35Mya in
291 both data sets (node 'a' in Fig. 3), the divergence of *P. mexicana* from the remainder of *Populus*
292 was ~15Mya in both data sets (node 'd'), and *Salix* subgenera *Vetrix* and *Chamaetia* began to
293 differentiate ~5 mya in both data sets (node 'c'). Nonetheless, a few estimates of divergence
294 times differed substantially between the data sets with no overlap in 95% HPD. In *Populus* these
295 included the timing of divergence of subgenera *Turanga* (node 'e') and *Populus* (node 'f') from
296 the remainders of the genus, and in *Salix* these included the timing of divergence of subgenera

Protitea (node ‘b’) from the remainder of the genus. This inconsistency reflects variance among the gene histories within *Populus* and *Salix* and identify regions of the tree where interpreting the ages of divergence require elevated caution . We also note here that our HPD estimates were constrained by our selection of the standard deviation of the a priori distributions of calibration dates to σ=1. When σ was set to 3, the uncertainty for node dates increased dramatically (Fig. S4; Tables S10, S12).

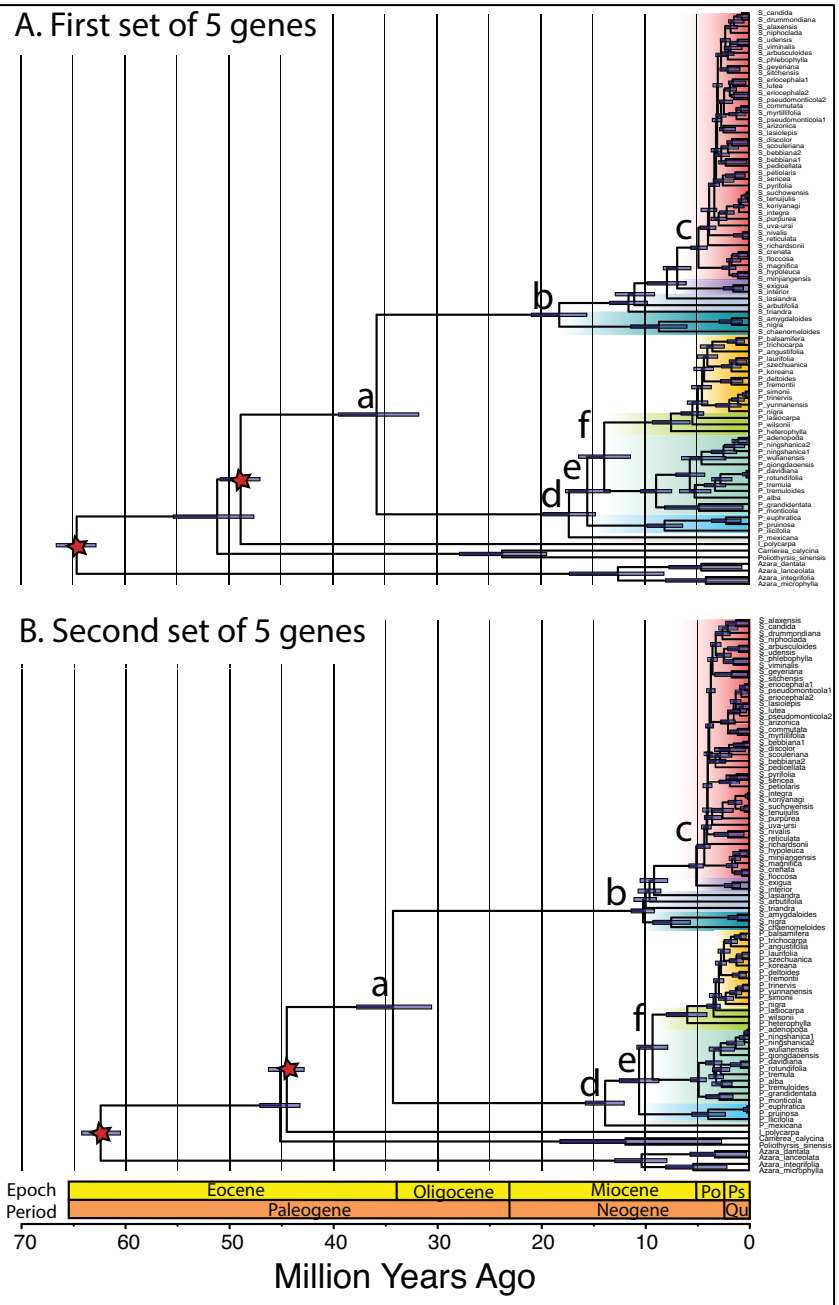

Figure 3. Comparison of dated trees estimated from two different sets of 5 genes using by *BEAST2. A) first set of 5 genes, which was the best set according to criteria described in the methods. B) second set of 5 genes, which was the second-best set. Stars indicate calibration nodes. Node bars represent 95% highest posterior density in node height. Subgenera are colored as in Fig. 1. Letters at nodes are discussed in the text.

Despite slightly different estimates of node ages in our two estimates, both the BAMM and ClaDS analyses indicated strong support for similar patterns of shifts in the diversification rates within both of our dated trees (Fig. 4; Table S13). In both dated trees diversification rate increased near the *Populus*-*Salix* split and a second increase occurred near the origination of the *Vetrix*-*Chamaetia* clade in Salix (Fig. 4). For the first dated tree, the BAMM analysis identified two branches with high marginal shift probabilities (Fig. 4A), and these shifts were largely supported in the ClaDS analysis (Fig. 4C). The BAMM analysis also identified substantial support for an increase in diversification rate near the *Populus*-*Salix* split for the second dated tree, with the exception that the model with most support indicated that *Populus* subgenus *Abaso* (with only *P. mexicana*)

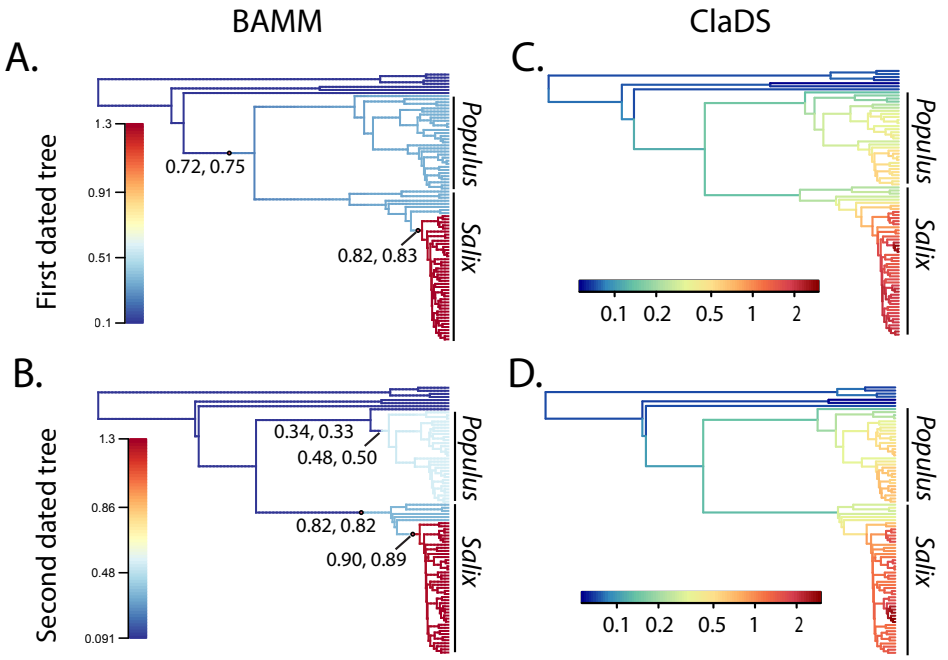

Figure 4. Diversification rate changes across the Salicaceae. Panels A and B show the best estimates of the rate shift set estimated using BAMM for the first and second dated trees, respectively. 60% and 61% (replicate 1 & 2) of samples in the posterior were assigned the the shift configuration shown in A, whereas 40% and 42% (replicate 1 & 2) of samples in the posterior were assigned the the shift configuration shown in B. The 95% credible shift sets for both trees and replicates are presented in Figures S5 and S6. Colors in A and B represent averaged lineage diversification rates according to the scales associated with each panel. Numbers at nodes represent the marginal odds ratios for a shift along branches associated with rate shifts; estimates from the two replicates are presented separated by a comma. Panels C and D show the ClaDS branch-specific rate estimates for the first and second dated trees, respectively. The color scale represents estimate branch specific speciation rates.

342 retained the same diversification rate at the outgroups (Fig. 4B). It is notable that the marginal
343 odds ratio supporting the shift in diversification at the branch that included all of *Populus* was
344 nearly as high (0.33-0.34) as the branch that did not include *P. mexicana* (0.48-0.50), indicating
345 that the shift was likely near the base of the *Populus* + *Salix* clade, but may not have included
346 *Populus* subgenus *Abaso*. The ClaDS analysis also supported a rate shift in this general region of
347 the phylogeny. The BAMM analysis also identified in the second dated tree substantial support
348 for an increase in diversification rate in clade that includes the subgenus *Longifoliae* along with
349 the *Vetrix*-*Chamaetia* clade in *Salix* (Fig. 4B). Again, this pattern was supported in the ClaDS
350 analysis (Fig. 4D).

351 We noted that these transitions in diversification rates were associated with shifts in
352 pollination mode (insect in outgroups and *Salix*; wind in *Populus*), seed dispersal mode (animals
353 in outgroups, wind in *Populus* and *Salix*), and growth form (trees in outgroup, *Populus*, and *Salix*
354 subgenus *Protitea*; shrubs and dwarf shrubs in remaining *Salix* subgenera), and a shift from XY
355 to ZW sex determination systems in *Salix* (Sanderson et al. 2021; Gulyaev et al. 2022; Hu et al.
356 2022). However, there are too few phylogenetically independent trait transitions on our
357 Salicaceae phylogeny to provide a statistically robust test of the associations between traits and
358 diversification rates. STRAPP analyses resulted in no statistically significant associations
359 (P>0.30 for all tests).

360

361 **Discussion**

362 The scale of hybridization on *Populus* and *Salix* is remarkable in its taxonomic, genomic,
363 and chronological extent. Signals of ancient hybridization and introgression across diverged
364 lineages have persisted in descendant lineages, are particularly evident in *Salix* and may better

365 account for gene tree discordance than contemporary hybridization. Our f_{branch} analysis estimated
366 that approximately 10-20% of the *Salix* genomes in the *Vetrix-Chamaetia* clade were affected by
367 the persistence of genes from ancient hybridization events, which based on our dated trees, may
368 have begun as early as 5 Mya. This pattern is consistent with the low levels of chloroplast
369 genomic diversity across the *Vetrix-Chamaetia* subgenera which was hypothesized to have been
370 partially influenced by high levels of hybridization and introgression (Wagner et al. 2021b).
371 Earlier hybridization events affecting genomic variation in subgenera *Salix* and *Longifolae* likely
372 occurred even earlier, perhaps nearly 10Mya. These hybridization events have likely contributed
373 to the difficulties in reconstructing relationships within *Salix* in the present and previous studies
374 (Leskinen et al. 1999; Barkalov et al. 2014; Percy et al. 2014; Lauron-Moreau et al. 2015; Wu et
375 al. 2015; Liu et al. 2016). In *Populus*, signals of hybridization between subgenus *Turanga* and
376 several members of *Aigieros-Tacamahaca* likely contributed to the low support for the position
377 of subgenus *Turanga* clade and may underlie incongruence between the ASTRAL and
378 concatenated trees (Fig. S1). Nonetheless, caution is required when interpreting the f_{branch}
379 analysis because it is dependent on the correct topology of the framework phylogeny. An
380 example of this impact is evident in the high level of hybridization indicated between *S.*
381 *sitchensis* and most other species in the *Vetrix-Chamaetia* clade, yet the lack of hybridization
382 between the putative close relative *S. geyeriana* and these same taxa. This pattern was unlikely to
383 have resulted from recent hybridization between *S. sitchensis* and each one of these species but
384 more likely resulted from the misplacement of *S. sitchensis* and/or *S. geyeriana* in the phylogeny.
385 Previous studies have reported signals of ancient hybridization in both genera. Two
386 ancient chloroplast capture events have been previously identified in *Populus*, one in which an
387 ancestor of *P. heterophylla* captured the chloroplast of *P. mexicana* ancestors (Liu et al. 2017;

388 Wang et al. 2020) and a second in which an ancestor of *P. nigra* captured the chloroplast of *P.*
389 *alba* ancestors (Smith et al. 1990). Our data, however, found no evidence of ancient
390 hybridization among these lineages in the nuclear genomes. This lack of evidence may result
391 from the limited numbers of genes that we sampled, or these ancient hybridization events may
392 have only influenced the chloroplast genomes (Tsitrone et al. 2003). Chloroplast capture
393 reflecting hybridization of ancient lineages also has been reported in *Salix* (Brunsfeld et al. 1992;
394 Hardig et al. 2000; Gulyaev et al. 2022). An ancestral member of subgenus *Longifoliae* captured
395 a chloroplast of a member of subgenus *Protitea* (Gulyaev et al. 2022) and five cases of arguably
396 more recent chloroplast capture within *Salix* section *Longifoliae* were reported by Brunsfeld et
397 al. (1992). Notably, crossing studies indicate that extant members of subgenera *Longifoliae* and
398 *Protitea* are reproductively isolated (Mosseler 1990), so the hybridization generating the former
399 chloroplast capture event likely occurred before reproductive isolation evolved. The deep history
400 of hybridization in *Salix* also is reflected in the evidence for gene flow among ancestral lineages
401 in our study and indicates a long history of concomitant hybridization and speciation in the
402 genus. As others have argued for *Populus* (Cronk et al. 2018), this pattern of ongoing speciation
403 with hybridization may be best represented as a syngameon (Lotsy 1925) and may exhibit
404 emergent properties such as the ability to draw on elevated levels of standing variation for
405 adaptive evolution (Cannon et al. 2020; Cannon 2021). Envisioning *Salix* as a syngameon
406 redefines evolutionary units as larger combinations of hybridizing species and our data and
407 previous results (Hardig et al. 2000; Murphy et al. 2022) suggest that the *Salix* syngameon
408 exhibits a complex web of ongoing hybridization and partial reproductive isolation that has
409 persisted for millions of years.

410 Although our recovered topology for *Populus* was similar to Wang et al. (2020), it was
411 not identical. A notable difference between our tree and that of Wang et al. (2020) is that our tree
412 placed *P. angustifolia* as sister to the *P. trichocarpa-P. balsamifera* clade, whereas the Wang et
413 al. (2020) placed *P. angustifolia* at the base of a larger clade including multiple North American
414 and Asian species. Importantly, the placement of Wang et al. (2020) indicates that either *P.*
415 *szechuanica*, *P. laurifolia*, and *P. koreana* or *P. angustifolia* may have speciated due to
416 vicariance or long-distance dispersal from North America to Asia and the divergence of *P.*
417 *angustifolia* predates this event. The difference in the placement of *P. angustifolia* is particularly
418 interesting and worthy of further study because it commonly hybridizes with both *P. trichocarpa*
419 and *P. balsamifera* (Brayshaw 1965; Chhatre et al. 2018), which may influence patterns of gene
420 tree coalescence, and the perceived relationships among species. Interestingly, some foliar
421 pathogens of *P. angustifolia* are related to foliar pathogens of Asian members of *Tacamahaca*
422 and absent in North American members (Busby et al. 2012), lending support to the hypothesis of
423 a recent trans-Beringian migration of *P. angustifolia*.
424 The estimates for divergence times among subgenera within *Populus* and *Salix* presented
425 here must be considered tentative because we did not calibrate internal nodes in each genus.
426 Many fossils of *Populus* and *Salix* from both Asia and North America have been identified
427 (Collinson, M.E. 1992), but it has been difficult to accurately assign them to extant taxonomic
428 groups without robust phylogenies. We chose to rely primarily on the molecular clock to
429 estimate diversification dates within these genera instead of including internal calibrations
430 following Percy et al. (2014) and Zhang et al. (2018). We note that the estimates of
431 diversification dates by Wu et al. (2015) were much earlier than we report here and included an
432 internal *Salix* late-Oligocene (ca. 26Ma) leaf fossil to calibrate of the origination of subgenus

433 *Vetrix* (Wolfe pers. comm. 1991 in Collinson 1992). Also, *Populus* fossils hypothesized to
434 belong to subgenus *Tacamahaca* are described from the Late Oligocene Creede flora of
435 Colorado (Wolfe et al. 1990; Collinson 1992), a date much earlier than our estimate of the
436 diversification of this subgenus. However, our and other recent phyogenies have identified both
437 subgenus *Tacamacaha* and *Vetrix* as polyphyletic (Wagner et al. 2018; Wang et al. 2020),
438 suggesting that the characters used to categorize the fossils should be reexamined. The last
439 thorough review of Salicaceous fossils was published in 1992 (Collinson 1992), and updating of
440 this group in relation to the most up-to-date phylogenies would aid greatly in developing a more
441 accurate estimate of diversification times in the Salicaceae.

442 In *Salix* we identified a burst of diversification near the origin of the *Vetrix-Chamaetia*
443 clade that likely increased the level of incomplete lineage sorting (ILS) and contributed to the
444 lack of support for inferred relationships within this clade (Roch et al. 2015). We also identified
445 a second increase in the diversification rate near the divergence of *Populus* from *Salix*. The
446 mechanisms driving these increases in diversification rate remain speculative. Although these
447 shifts are accompanied with changes in seed dispersal, pollination vectors, and growth form
448 (Argus 2010; Eckenwalder 2010), and even a shift from an XY to a ZW sex determination
449 system in *Salix* (Gulyaev et al. 2022; Hu et al. 2022), each of these shifts occurs only once or
450 twice on the tree, so too few phylogenetically independent events have occurred for powerful
451 statistical tests of association. Nonetheless, these patterns may be useful for meta-analyses
452 including a larger set of taxonomic groups, or future studies may find more detailed patterns
453 associated with genetic or morphological changes that can shed light on the drivers of
454 diversification in these two important genera.

455 The continued development of analytical methods for quantifying ancient interspecific
456 gene flow will permit investigations into the prevalence of ancient hybridization, its impacts on
457 adaptation, and its biological and phylogenetic correlates. Ancient hybridization inferred from
458 chloroplast capture has been identified in large numbers of plants groups (Rieseberg et al. 1991).
459 However, chloroplast capture may result from unique properties of the hybridizing species
460 (Tsitrone et al. 2003) and may not be present in all taxonomic groups affected by ancient
461 hybridization. Interspecific gene flow is usually not categorized as either contemporary or
462 ancient, but the relative influences of ancient versus recent introgression events may provide
463 insight into the sources of different classes of genetic variants under selection (e.g. Menon et al.
464 2021). Recent studies in *Quercus*, *Pinus,* and *Populus* have documented adaptive introgression
465 among hybridizing groups of multiple species (Chhatre et al. 2018; Leroy et al. 2020; Buck et al.
466 2022). Along with *Populus* and *Salix*, ancient gene flow has also been identified in *Quercus*,
467 where phylogenetic conflict in chloroplast genomes suggests hybridization during early
468 diversification (Yang et al. 2021) indicating at least two plant families with a combination of
469 both contemporary and ancient interspecific gene flow. Plant families differ in their propensity to
470 hybridize, and there is a strong phylogenetic signal for hybridization (Whitney et al. 2010). Thus,
471 it is likely that the ancestors of contemporary groups with strong propensities for hybridization
472 also hybridized. Understanding patterns of ancient interspecific gene flow in relation to
473 contemporary hybridization may provide further insight into biological factors correlated with
474 hybridization (Mitchell et al. 2019) and factors associated with the development of syngameons.
475
476
477

Conclusion

Examples of the extended effects of contemporary hybridization in *Salix* and *Populus* are well-known (Hardig et al. 2000; Evans et al. 2008; Lexer et al. 2010). The cumulative effects of persistent hybridization over eons as diversification unfolded, however, may result in different qualities of adaptation and diversification than isolated cases of contemporary hybridization. In *Salix* and *Populus* this history has resulted in a tangle of gene histories within each genus, with some clades having developed monophyly and others that may never be resolved. Hybridization also may have contributed changes in the diversification rate of species in the *Vetrix-Chamaetia* clade of *Salix*. Understanding the characteristics associated with and generated from long-term and persistent interspecific gene flow will elucidate whether these properties confer fundamentally different pattens of adaptation and speciation (Cannon 2021).

Data Accessibility

All alignments and gene trees used in this research are available at:
https://datadryad.org/stash/share/48HL7Z3TAgH1hOQLwWM1EzundQOFxN0fwoRxXMdXi_s
All data not already available at NCBI will be uploaded to the NCBI short read archive prior to publication – we are currently finalizing the metadata.

Acknowledgements

We thank Peter Zhelev and Gancho Slavov for *S. triandra* collections, Andrew Hamstetter for insightful discussions of diversification analyses, and Pascal Title for computational assistance with BAMM. This research was supported by grants from the US National Science Foundation (1542509, 1542599, 1542479, 1542486) and the National Natural Science Foundation of China (31561123001).

Literature Cited

Argus GW. 1997. Infrageneric classification of Salix (Salicaceae) in the new world. Systematics Botany Monographs, 52:1-121.

506 Argus GW. 2010. *Salix* L. In: Committee FoNAE editor. Flora of North America North of
507 Mexico [Online]. . New York and Oxford.
508 Barkalov VY, Kozyrenko MM. 2014. Phylogenetic relationships of Salix L. subg. Salix species
509 (Salicaceae) according to sequencing data of intergenic spacers of the chloroplast genome
510 and ITS rDNA. Russian Journal of Genetics, 50:828-837.
511 Boeckler GA, Gershenzon J, Unsicker SB. 2011. Phenolic glycosides of the Salicaceae and their
512 role as anti-herbivore defenses. Phytochemistry, 72:1497-1509.
513 Boucher LD, Manchester SR, Judd WS. 2003. An extinct genus of Salicaceae based on twigs
514 with attached flowers fruits, and foliage from the Eocene Green River Formation of Utah and
515 Colorado, USA. Am J Bot, 90:1389-1399.
516 Brayshaw TC. 1965. Native poplars of southern Alberta and their hybrids. In: Forestry CDo
517 editor. Ottowa, CA, Roger Duhamel , FRSC, Queen's Printer and Controller of Stationary.
518 Brunsfeld SJ, Soltis DE, Soltis PS. 1992. Evolutionary patterns and processes in Salix sect
519 Longifoliae - evidence from chloroplast DNA. Systematic Botany, 17:239-256.
520 Buck R, Ortega-Del Vecchyo D, Gehring C, Michelson R, Flores-Renteria D, Klein B, Whipple
521 AV, Flores-Renteria L. 2022. Sequential hybridization may have facilitated ecological
522 transitions in the Southwestern pinyon pine syngameon. New Phytol.
523 Busby PE, Aime MC, Newcombe G. 2012. Foliar pathogens of Populus angustifolia are
524 consistent with a hypothesis of Beringian migration into North America. Fungal Biology,
525 116:792-801.
526 Cannon CH. 2021. Is speciation an unrelenting march to reproductive isolation? Molecular
527 Ecology, 30:4349-4352.
528 Cannon CH, Petit RJ. 2020. The oak syngameon: more than the sum of its parts. New Phytol,
529 226:978-983.
530 Chhatre VE, Evans LM, DiFazio SP, Keller SR. 2018. Adaptive introgression and maintenance
531 of a trispecies hybrid complex in range-edge populations of Populus. Molecular Ecology,
532 27:4820-4838.
533 Collinson ME. 1992. The early fossil history of Salicaceae - A brief review. Proceedings of the
534 Royal Society of Edinburgh Section B-Biological Sciences, 98:155-167.
535 Cronk QC, Needham I, Rudall PJ. 2015. Evolution of Catkins: Inflorescence Morphology of
536 Selected Salicaceae in an Evolutionary and Developmental Context. Front Plant Sci, 6:1030.
537 Cronk QC, Suarez-Gonzalez A. 2018. The role of interspecific hybridization in adaptive
538 potential at range margins. Molecular Ecology, 27:4653-4656.
539 Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the
540 multispecies coalescent. Trends Ecol Evol, 24:332-340.
541 Desborough MJR, Keeling DM. 2017. The aspirin story - from willow to wonder drug. British
542 Journal of Haematology, 177:674-683.
543 Dorn RD. 1976. A synopsis of American *Salix*. Canadian Journal of Botany, 54:2769-2789.
544 Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for Ancient Admixture between
545 Closely Related Populations. Mol Biol Evol, 28:2239-2252.
546 Eckenwalder JE. 1996. Systematics and Evolution of *Populus* In: Stettler RF, Bradshaw HD,
547 Heilman PE, T.M.Hinckley editors. Biology of *Populus* and its Implications for Management
548 and Conservation. Ottawa, Canada, NRC Research Press, p. 7-56.
549 Eckenwalder JE. 2010. *Populus* L. In: Committee FoNAE editor. Flora of North America North
550 of Mexico [Online], New York and Oxford.

551 Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? Evolution,
552 63:1-19.
553 Evans LM, Allan GJ, Shuster SM, Woolbright SA, Whitham TG. 2008. Tree hybridization and
554 genotypic variation drive cryptic speciation of a specialist mite herbivore. Evolution,
555 62:3027-3040.
556 Friedman J, Barrett SCH. 2009. Wind of change: new insights on the ecology and evolution of
557 pollination and mating in wind-pollinated plants. Annals of Botany, 103:1515-1527.
558 Grant V. 1981. Plant Speciation. New York, Columbia University Press.
559 Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai WW,
560 Fritz MHY*, et al.* 2010. A Draft Sequence of the Neandertal Genome. Science, 328:710-722.
561 Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New
562 Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the
563 Performance of PhyML 3.0. Systematic Biology, 59:307-321.
564 Gulyaev S, Cai XJ, Guo FY, Kikuchi S, Applequist WL, Zhang ZX, Horandl E, He L. 2022. The
565 phylogeny of Salix revealed by whole genome re-sequencing suggests different sex-
566 determination systems in major groups of the genus. Annals of Botany, 129:485-498.
567 Hardig TM, Brunsfeld SJ, Fritz RS, Morgan M, Orians CM. 2000. Morphological and molecular
568 evidence for hybridization and introgression in a willow (Salix) hybrid zone. Molecular
569 Ecology, 9:9-24.
570 Heled J, Drummond AJ. 2010. Bayesian Inference of Species Trees from Multilocus Data. Mol
571 Biol Evol, 27:570-580.
572 Hu N, Sanderson BJ, Guo M, Feng G, Gambhir D, Hale H, Wang D, Hyden B, Liu J, Ma T*, et*
573 *al.* 2022. An unusual origin of a ZW sex chromosome system. PREPRINT (Version 1)
574 available at Research Square.
575 Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ.
576 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-
577 throughput sequencing reads using target enrichment. Applications in Plant Sciences, 4.
578 Karrenberg S, Kollmann J, Edwards PJ. 2002. Pollen vectors and inflorescence morphology in
579 four species of Salix. Plant Systematics and Evolution, 235:181-188.
580 Keefover-Ring K, Carlson CH, Hyden B, Azeem M, Smart LB. 2022. Genetic mapping of
581 sexually dimorphic volatile and non-volatile floral secondary chemistry of a dioecious
582 willow. Journal of Experimental Botany, 73:6352-6366.
583 Lauron-Moreau A, Pitre FE, Argus GW, Labrecque M, Brouillet L. 2015. Phylogenetic
584 Relationships of American Willows (Salix L., Salicaceae). Plos One, 10.
585 Leroy T, Louvet J-M, Lalanne C, Le Provost G, Labadie K, Aury J-M, Delzon S, Plomion C,
586 Kremer A. 2020. Adaptive introgression as a driver of local adaptation to climate in
587 European white oaks. New Phytol, 226:1171-1182.
588 Leskinen E, Alstrom-Rapaport C. 1999. Molecular phylogeny of Salicaceae and closely related
589 Flacourtiaceae: evidence from 5.8 S, ITS 1 and ITS 2 of the rDNA. Plant Systematics And
590 Evolution, 215:209-227.
591 Lexer C, Joseph JA, van Loo M, Barbará T, Heinze B, Bartha D, Castiglione S, Fay MF, Buerkle
592 CA. 2010. Genomic Admixture Analysis in European Populus spp. Reveals Unexpected
593 Patterns of Reproductive Isolation and Mating. Genetics, 186:699-712.
594 Liu X, Wang ZS, Shao WH, Ye ZY, Zhang JG. 2017. Phylogenetic and Taxonomic Status
595 Analyses of the Abaso Section from Multiple Nuclear Genes and Plastid Fragments Reveal

New Insights into the North America Origin of Populus (Salicaceae). Frontiers in Plant Science, 7.
Liu X, Wang ZS, Wang DS, Zhang JG. 2016. Phylogeny of Populus-Salix (Salicaceae) and their relative genera using molecular datasets. Biochemical Systematics and Ecology, 68:210-215.
Lotsy JP. 1925. Species of linneon. Genetica, 7:487-506.
Maliet O, Hartig F, Morlon H. 2019. A model with many small shifts for estimating species-specific diversification rates. Nature Ecology & Evolution, 3:1086-1092.
Maliet O, Morlon H. 2022. Fast and Accurate Estimation of Species-Specific Diversification Rates Using Data Augmentation. Systematic Biology, 71:353-366.
Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D-statistics and related admixture evidence from VCF files. Molecular Ecology Resources, 21:584-595.
Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nature Ecology & Evolution, 2:1940-1955.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M*, et al.* 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20:1297-1303.
McVay JD, Hipp AL, Manos PS. 2017. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. P Roy Soc B-Biol Sci, 284.
Meeuse ADJ. 1975. Taxonomic relationships of Salicaceae and Flacourtiaceae: their bearing on interpretative floral morphology and dilleniid phylogeny. Acta Botanica Neerlandica, 24:437-457.
Menon M, Bagley JC, Page GFM, Whipple AV, Schoettle AW, Still CJ, Wehenkel C, Waring KM, Flores-Renteria L, Cushman SA*, et al.* 2021. Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. Communications Biology, 4:160.
Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics, 30:I541-I548.
Mitchell N, Campbell LG, Ahern JR, Paine KC, Giroldo AB, Whitney KD. 2019. Correlates of hybridization in plants. Evol Lett, 3:570-585.
Morales-Briones DF, Kadereit G, Tefarikis DT, Moore MJ, Smith SA, Brockington SF, Timoneda A, Yim WC, Cushman JC, Yang Y. 2020. Disentangling Sources of Gene Tree Discordance in Phylogenomic Data Sets: Testing Ancient Hybridizations in Amaranthaceae s.l. Systematic Biology, 70:219-235.
Mosseler A. 1990. Hybrid performance and species crossability relationships in willows (Salix). Canadian Journal of Botany-Revue Canadienne De Botanique, 68:2329-2338.
Murphy EK, Cappa EP, Soolanayakanahally RY, El-Kassaby YA, Parkin IAP, Schroeder WR, Mansfield SD. 2022. Unweaving the population structure and genetic diversity of Canadian shrub willow. Sci Rep-Uk, 12.
Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol, 32:268-274.
Parins-Fukuchi C, Stull GW, Smith SA. 2021. Phylogenomic conflict coincides with rapid morphological innovation. Proceedings of the National Academy of Sciences of the United States of America, 118.

642 Patterson N, Moorjani P, Luo YT, Mallick S, Rohland N, Zhan YP, Genschoreck T, Webster T,
643 Reich D. 2012. Ancient Admixture in Human History. Genetics, 192:1065-+.
644 Pease JB, Hahn MW. 2013. More accurate phylogenies inferred from low-recombination regions
645 in the presence of incomplete lineage sorting. Evolution, 67:2376-2384.
646 Percy DM, Argus GW, Cronk QC, Fazekas AJ, Kesanakurti PR, Burgess KS, Husband BC,
647 Newmaster SG, Barrett SC, Graham SW. 2014. Understanding the spectacular failure of
648 DNA barcoding in willows (Salix): does this result from a trans-specific selective sweep?
649 Mol Ecol, 23:4737-4756.
650 Philippe RN, Bohlmann J. 2007. Poplar defense against insect herbivores. Canadian Journal of
651 Botany-Revue Canadienne De Botanique, 85:1111-1126.
652 Rabiee M, Sayyari E, Mirarab S. 2019. Multi-allele species reconstruction using ASTRAL.
653 Molecular Phylogenetics and Evolution, 130:286-296.
654 Rabosky DL. 2014. Automatic Detection of Key Innovations, Rate Shifts, and Diversity-
655 Dependence on Phylogenetic Trees. Plos One, 9.
656 Rabosky DL, Huang H. 2015. A Robust Semi-Parametric Test for Detecting Trait-Dependent
657 Diversification. Systematic Biology, 65:181-193.
658 Rieseberg LH, Soltis DE. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants.
659 Evolutionary Trends in Plants, 5:65-84.
660 Roch S, Steel M. 2015. Likelihood-based tree reconstruction on a concatenation of aligned
661 sequence data sets can be statistically inconsistent. Theoretical Population Biology, 100:56-
662 62.
663 Rohwer J, Kubitzki K. 1984. Salix martiana, a regularly hermaphrodite willow. Plant
664 Systematics and Evolution, 144:99-101.
665 Sacchi CF, Price PW. 1988. Pollination of the arroyo willow, Salix-lasiolepis - role of insects
666 and wind. Am J Bot, 75:1387-1393.
667 Sanderson BJ, DiFazio SP, Cronk QCB, Ma T, Olson MS. 2020. A targeted sequence capture
668 array for phylogenetics and population genomics in the Salicaceae. Applications in Plant
669 Sciences, 8.
670 Sanderson BJ, Feng G, Hu N, Carlson CH, Smart LB, Keefover-Ring K, Yin T, Ma T, Liu J,
671 DiFazio SP*, et al.* 2021. Sex determination through X–Y heterogamety in Salix nigra.
672 Heredity.
673 Schweitzer JA, Bailey JK, Rehill BJ, Martinsen GD, Hart SC, Lindroth RL, Keim P, Whitham
674 TG. 2004. Genetically based trait in a dominant tree affects ecosystem processes. Ecology
675 Letters, 7:127-134.
676 Shu Y. 1999. Populus Linnaeus. . In: Song H editor. Flora of China @ eFloras. Missouri
677 Botanical Garden, St. Louis, MO, USA and Harvard UNiversity Herbaria, Cambridge, MA,
678 USA, p. 139-162.
679 Smith ME, Chamberlain KR, Singer BS, Carroll AR. 2010. Eocene clocks agree: Coeval Ar-
680 40/Ar-39, U-Pb, and astronomical ages from the Green River Formation. Geology, 38:527-
681 530.
682 Smith RL, Sytsma KJ. 1990. Evolution of Populus-aigra (sect Aigeiros) - introgressive
683 hybridization and the chloroplast contribution of Populus alba (sect Populus). Am J Bot,
684 77:1176-1187.
685 Smith SA, Brown JW, Walker JF. 2018. So many genes, so little time: A practical approach to
686 divergence-time estimation in the genomic era. Plos One, 13.

687 Stettler RF, Bradshaw HD, Heilman PE, Hinckley TM. 1996. Biology of Populus and its
688 Implications for Management and Conservation. Ottawa, Canada, NRC Research Press.
689 Tamura S, Kudo G. 2000. Wind pollination and insect pollination of two temperate willow
690 species, Salix miyabeana and Salix sachalinensis. Plant Ecology, 147:185-192.
691 Tsai CJ, Harding SA, Tschaplinski TJ, Lindroth RL, Yuan YN. 2006. Genome-wide analysis of
692 the structural genes regulating defense phenylpropanoid metabolism in Populus. New Phytol,
693 172:47-62.
694 Tsitrone A, Kirkpatrick M, Levin DA. 2003. A model for chloroplast capture. Evolution,
695 57:1776-1782.
696 Wagner ND, Gramlich S, Horandl E. 2018. RAD sequencing resolved phylogenetic relationships
697 in European shrub willows (Salix L. subg. Chamaetia and subg. Vetrix) and revealed
698 multiple evolution of dwarf shrubs. Ecology and Evolution, 8:8243-8255.
699 Wagner ND, He L, Horandl E. 2020. Phylogenomic Relationships and Evolution of Polyploid
700 Salix Species Revealed by RAD Sequencing Data. Frontiers in Plant Science, 11.
701 Wagner ND, He L, Horandl E. 2021a. The Evolutionary History, Diversity, and Ecology of
702 Willows (Salix L.) in the European Alps. Diversity-Basel, 13.
703 Wagner ND, Volf M, Hoerandl E. 2021b. Highly Diverse Shrub Willows (Salix L.) Share Highly
704 Similar Plastomes. Frontiers in Plant Science, 12.
705 Wang MC, Zhang L, Zhang ZY, Li MM, Wang DY, Zhang X, Xi ZX, Keefover-Ring K, Smart
706 LB, DiFazio SP*, et al.* 2020. Phylogenomics of the genus Populus reveals extensive
707 interspecific gene flow and balancing selection. New Phytol, 225:1370-1382.
708 Wang YC, Huang J, Li EZ, Xu SJ, Zhan ZF, Zhang XJ, Yang ZQ, Guo FY, Liu KJ, Liu D*, et al.*
709 2022. Phylogenomics and Biogeography of Populus Based on Comprehensive Sampling
710 Reveal Deep-Level Relationships and Multiple Intercontinental Dispersals. Frontiers in Plant
711 Science, 13.
712 Wessinger CA. 2021. From pollen dispersal to plant diversification: genetic consequences of
713 pollination mode. New Phytol, 229.
714 Whitney KD, Ahern JR, Campbell LG, Albert LP, King MS. 2010. Patterns of hybridization in
715 plants. Perspectives in Plant Ecology Evolution and Systematics, 12:175-182.
716 Wolfe JA, Schorn HE. 1990. Taxonomic revision of the Spermatopsida of the Oligocene Creede
717 Flora, Southern Colorado., United States Geological Bulletin, p. 1-40.
718 Wu CI. 1991. Inferences of species phylogeny in relation to segregation of ancient
719 polymorphisms. Genetics, 127:429-435.
720 Wu J, Nyman T, Wang D-C, Argus GW, Yang Y-P, Chen J-H. 2015. Phylogeny of Salix
721 subgenus Salix s.l. (Salicaceae): delimitation, biogeography, and reticulate evolution. BMC
722 evolutionary biology, 15:311-311.
723 Yang YY, Qu XJ, Zhang R, Stull GW, Yi TS. 2021. Plastid phylogenomic analyses of Fagales
724 reveal signatures of conflict and ancient chloroplast capture. Molecular Phylogenetics and
725 Evolution, 163.
726 Yin J, Zhang C, Mirarab S. 2019. ASTRAL-MP: scaling ASTRAL to very large datasets using
727 randomization and parallelization. Bioinformatics, 35:3961-3969.
728 Young AD, Gillung JP. 2020. Phylogenomics — principles, opportunities and pitfalls of big-data
729 phylogenetics. Systematic Entomology, 45:225-247.
730 Zhang L, Xi ZX, Wang MC, Guo XY, Ma T. 2018. Plastome phylogeny and lineage
731 diversification of Salicaceae with focus on poplars and willows. Ecology and Evolution,
732 8:7817-7823.

733