1   **Backcrossing to different parents produced two distinct hybrid species**

2   Donglei Wang[1], Jiwang[1], Hao Bi[1], Jianquan Liu[1,2]*, Dafu Ru[2,1]*

3

4[1]Key Laboratory for Bio-resource and Eco-environment of Ministry of Education, College of Life
5Sciences, Sichuan University, Chengdu, PR China
6[2]State Key Laboratory of Grassland Agro-ecosystem, Institute of Innovation Ecology and College of
7Life Sciences, Lanzhou University, Lanzhou, 730000, China

8

9*Corresponding authors: liujq@nwipb.ac.cn; rudf@lzu.edu.cn

10

11**Abstract** Repeated homoploid hybrid speciation (HHS) events with the same parental 12species have rarely been reported. In this study, we used population transcriptome 13data to test paraphyly and HHS events in one conifer *Picea brachytyla*. All analyses 14identified and supported non-sister relationships for the two lineages of *P. brachytyla*. 15The southern lineage was placed within the re-circumscribed *P. likiangensis* species 16complex (PLSC) while *P. brachytyla sensu stricto* (s.s.), comprising only the northern 17lineage, parallels both PLSC and the closely related *P. wilsonii*. In addition, both 18phylogenetic and coalescent analyses suggested that *P. brachytyla* s.s. arose from 19homoploid hybrid speciation between the ancestor of the PLSC before its 20diversification (into the current varieties or species), and *P. wilsonii*, through an 21intermediate hybrid lineage at an early stage and backcrossing to the ancestral PLSC. 22These two parental ancestors also produced another homoploid hybrid species, *P.* 23*purpurea*, in the same way but at a later stage, through the same extinct lineage but 24backcrossing to the other parent, *P. wilsonii*. We reveal the first case that 25backcrossing to different parents of the same extinct hybrid lineage produced two 26different hybrid species. Our results highlight the existence of more reticulate 27evolution during species diversification in the spruce genus and more complex 28homoploid hybrid events than have previously been identified.

29**Key words:** homoploid hybrid speciation, *Picea brachytyla,* population transcriptome 30data, polyphyly, extinct hybrid lineage

31**Introduction**

32Hybridization between two distinct species may generate a new species without any
33change in ploidy level (Rieseber*g et* al., 2003; Nolte and Tautz 2010). This
34phenomenon, homoploid hybrid speciation (HHS), seems to occur more commonly
35than was previously assumed (e.g. Rieseberg 1997; Gross and Rieseberg 2005;
36Mavare*z et* al., 2006; Hermanse*n et* al., 2011; Nieto Feline*r et* al., 2017; Lamichhane*y*
37*et* al., 2018), although only a few of these cases fulfill the strict criterion of Schumer
38*et al*., (2014) that hybridization created reproductive isolation (RI) of the hybrid
39species from its parents. However, such RI may occur through inheriting parental
40genetic incompatibilities (Schumer *et al*., 2015; Brennan *et al*., 2019), although this is
41difficult to prove, especially in trees, by QTL or other methods. Revealing more
42candidate HHS cases is therefore very useful in increasing our understanding of
43reticulate species diversification even when there is no direct evidence of RI caused
44by hybridization (Nieto Feliner *et al*., 2017).

45   Here we aim to examine the homoploid hybrid origin of *Picea brachytyla sensu*
46*stricto* (s.s.), from a genus in which all the recognized species or varieties are diploids
47(Wright 1955). The traditional circumscription of this species comprises two non-
48sister lineages, the southern lineage in Yunnan province and the northern one in
49Sichuan and Chongqing province (Ru *et al*., 2016; Lyu *et al*., 2020). The southern
50lineage (*P. brachytyla*-southern lineage) shares a recent ancestor with three varieties
51(vars. *likiangensis*, *rubescens* and *linzhiensis*) of *P. likiangensis* as a species complex
52(PLSC), and it may have originated from a common radiation with strong gene flow
53and historical hybridization (Sun *et al*., 2018). The northern lineage, which contains
54the population where the type specimen was collected, comprises *P. brachytyla* s.s.,
55paralleling PLSC and *P. wilsonii* (Lyu *et al*., 2020). However, phylogenetic analyses
56of a few individuals of *P. brachytyla* s.s. suggested that it is more closely related to *P.*
57*likiangensis* (species complex), based on nuclear loci, or to *P. wilsonii,* on the basis of
58chloroplast DNA sequences (Ran *et al*., 2015; Shen *et al*., 2019; Shao *et al*., 2019).
59Further population genetic analyses based on more individuals supported these

60inferences (Lyu *et al*., 2020). In addition, *P. brachytyla* combines morphological traits

61of *P. likiangensis* and *P. wilsonii* (Fu *et al*., 1999; Lyu *et al*., 2020). Taken together,

62the conflicting phylogenetic relationships and intermediate morphological traits

63suggest a likely hybrid origin of *P. brachytyla* s.s. from two other species (or species

64complexes) (Shen *et al*., 2019; Shao *et al*., 2019; Lyu *et al*., 2020). *P. brachytyla* s.s.

65occurs in low-altitude humid valleys while the PLSC is distributed in high-altitude

66mountains and *P. wilsonii* is found in northern low-altitude but relatively dry

67mountains (Fu *et al*., 1999; Lyu *et al*., 2020). It should be noted that another species,

68*P. purpurea*, was shown to have also originated through hybridization between *P.*

69*likiangensis* and *P. wilsonii* (Sun *et al*., 2014; Ru *et al*., 2018). Thus, similar to HHS

70in sunflowers (Rieseberg *et al*., 1997), more than one diploid hybrid species may have

71originated from the same parents in the group. In addition, var. *rubescens* of *P.*

72*likiangensis* was found to have species-specific nuclear ancestry from *P. purpurea*

73based on population genetic structure analyses (Ru *et al*., 2018), large-scale shared

74genetic variations from introgression (Sun *et al*., 2018) and inconsistent phylogenetic

75relationships based on different DNA sequences (Shen *et al*., 2019; Shao *et al*., 2019).

76This variety may be better excluded from the PLCS while the *P. brachytyla*-southern

77lineage should be added to the PLCS when modelling likely HHS events betweem

78PLSC and *P. wilsonii*. All these findings suggest that the hybridization history and

79circumscription of these lineages and their ancestors are more complex than expected

80and that further detailed studies are needed.

81 In the present study, we used population transcriptomic data to trace the

82evolutionary origin of *P. brachytyla* s.s. We present, for the first time, transcriptome

83data for 78 individuals of *P. brachytyla* and related species. We added previously

84published transcriptome data from more 114 individuals of all related species to carry

85out comprehensive analyses. We aimed to address the following questions: (1) Do

86population genomic data support polyphyly of *P. brachytyla*? How should the PLSC

87best be defined? (2) Did *P. brachytyla* s.s. originate through homoploid hybridization

88between the common ancestor of the PLCS and *P. wilsonii*, or bifurcate from one of

89these with further gene flow from the other? (3) If it originated through HHS, did *P.*

90*brachytyla* s.s. originate earlier or later than *P. purpurea* as a result of hybridization

91between the same parents?

92Material and methods

93Material and RNA sequencing

94We collected and re-sequenced transcriptomes of 78 individuals (Table S1) for both
95lineages of *P. brachytyla* and related species or taxa (*P. farreri*, *P. wilsonii*, and *P.*
96*purpurea*) following the methods of Ru *et al*., (2018). We used an Illumina HiSeq
972500 platform to generate 150 bp paired-end raw reads and deposited data sets for all
98individuals in BioSample (average number of raw bases > 6 Gb; Table S1). We
99further downloaded transcriptomes of 108 individuals of this species and others
100published earlier (three varieties of *P. likiangensis*, vars. *likiangensis*, *linzhiensis* and
101*rubescens*, *P. wilsonii* and *P. purpurea*; Ru *et al*., 2016, 2018; Feng *et al*., 2019; Shen
102*et al*., 2019; Shao *et al*., 2019). A total of 186 individuals from the core distribution of
103each taxon were used for our subsequent analyses (Fig. 1 and Table 1). For population
104sampling, the individuals collected were spaced at least 500 m apart. Individuals from
105the regions of contact of any two taxa likely, according to the literature, to have
106undergone high levels of gene flow in the recent past were excluded. For all
107phylogenetic analyses, *P. breweriana* was used as outgroup.

108Read mapping and individual variant calling

109We used "mem" in Burrows-Wheeler Aligner (BWA) version 0.7.10 (Li and Durbin
1102009) with default parameters to align the high-quality reads to both the *P. abies*
111reference transcriptome and the chloroplast (cp) reference genome (Nystedt *et al*.,
1122013) following our previous methods (Ru *et al*., 2018). For the cp genome, PCR
113duplicates in the alignments were marked and removed before calling variants using
114the program MarkDuplicates.jar from PICARD ver. 1.129
115(http://broadinstitute.github.io/picard/), followed by local realignment around each
116indel using Genome Analysis Toolkit (GATK) (Danecek *et al*., 2011). Single
117nucleotide polymorphisms (SNPs) were extracted using "mpileup" in SAMTOOLS
118ver. 1.8 (L*i et* al., 2009) based on the uniquely mapped reads for all individuals. We
119set the minimum base quality (-Q) and mapping quality (-q) to 20 and 30 respectively.
120To obtain high quality variants for the two references, we filtered our raw SNPs with
121the following criteria using a custom Perl script: (1) SNPs located within a 5-bp
122window of an InDel; (2) SNPs with a phred-scaled quality score <20; (3) SNPs with
123>20% missing bases within each species. Additionally, bases with depth of coverage
124(DP) <10 were set to be missing for each individual (Chapma*n et* al., 2013; L*i et* al.,
1252013; Wan*g et* al., 2013; L*i et* al., 2014). Finally we filtered variant sites with
126minimum allele frequency <0.01 (to ensure that at least 4 alleles were found in our

127sample set) using VCFTOOLS ver. 0.1.14 (Danecek *et* al., 2011) to reduce the false
128discovery rate. For convenience, we refer to the nuclear transcriptomic sequences and
129cpDNA sequences as the N-RNA-seq dataset and C-RNA-seq dataset respectively.

130**Nucleotide diversity and population Differentiation**

131The nucleotide diversity π (Nei and Li 1979) and the population differentiation index
132F_{ST} (Weir and Cockerham 1984) between populations were calculated using
133VCFTOOLS (Danecek *et* al., 2011) on the N-RNA-seq dataset. For the calculation of
134mean genome-wide F_{ST}, the negative values were reassigned to zero. In addition, we
135calculated the d_{XY} value (Foo*te et* al., 2016) per locus, which is the average number of
136nucleotide substitutions, using a custom Perl script.

137**Phylogenetic Tree Reconstruction**

138For the N-RNA-seq dataset, we constructed a neighbor-joining (NJ) tree using
139TreeBeST (http://treesoft.sourceforge.net/treebest.shtml) with 1,000 bootstrap
140replications performed to assess the branch reliability. For the C-RNA-seq dataset, we
141used only 69 individuals representing all taxa and their distributions for phylogenetic
142analyses in order to reduce the time required to identify cp genome (= plastome)
143variations. A maximum-likelihood (ML) tree was constructed using RAxML ver.
1448.1.20 (Stamatakis 2014) with the GTRCAT model and 200 bootstrap replicates. The
145NJ and ML trees were viewed using TreeView ver. 1.6.6 and FigTree ver. 1.4.0
146(http://tree.bio.ed.ac.uk/software/figtree/) respectively.

147**Population Structure and Admixture Analyses**

148The N-RNA-seq dataset was thinned by LD values to reduce the linkage
149disequilibrium effect using PLINK ver.1.07 (Purcel*l et* al., 2007; Danecek *et* al., 2011)
150with the parameter --indep-pairwise 50 5 0.2, resulting in a set of ~0.21 Mb SNPs for
151population structure and admixture analyses. A principal component analysis (PCA)
152was performed using the smartpca program from the EIGENSOFT package ver. 6.0.1
153(Pric*e et* al., 2006), while eigenvectors were generated with the R function region. A
154Tracy–Widom test was performed in R to determine the significance level of the
155eigenvectors. ADMIXTURE ver. 1.23 (Alexander and Lange 2011) was used to
156perform an unsupervised ancestry component analysis, with the K value (number of
157assumed ancestral components) ranging from 2 to 10. For each K, 200 bootstrap
158replicates were performed to calculate cross-errors (CV). The optimal K value was
159indicated by the lowest CV values among the numbers assumed.

160**Species-level transcriptome assembly**

161We obtained high quality reads for each sample by trimming adapter sequences, Poly-
162N, and low quality bases and discarding reads with fewer than 36 bases after trimming
163from the raw data using Trimmomatic (Bolge*r et* al., 2014) with the following

164parameters: ILLUMINACLIP:adapter.fa:2:30:10 LEADING:3 TRAILING:3
165SLIDINGWINDOW:4:20 MINLEN:36.

166   We assembled the species-level transcriptome of *P. brachytyla* s.s. using 14
167randomly selected individuals from each of the populations examined (marked with a
168dark green color in Table S1) using Trinity ver. 2.6.6 (Grabher*r et* al., 2011) with
169default parameters based on pooled libraries to reduce the gene loss caused by random
170variation in expression as much as possible. Similarly, the species-level transcriptome
171of *P. wilsonii* was obtained by the same method as before (Ru *et al.*, 2018). For the *P.*
172*likiangensis* species complex (PLSC), we selected one individual from each of the
173sampled populations of two varieties of *P. likiangensis* (var. *linzhiensis* and var.
174*likiangensis*), the *P. brachytyla* southern lineage in Yunnan and *P. farreri* (marked
175with a light green color in Table S1). We excluded *P. likiangensis* var. *rubescens*
176from PLSC because our preliminary ADMIXTURE analyses suggested that this
177variety shared many species-specific nuclear elements from another hybrid species, *P.*
178*purpurea*, because of the second hybridization history. To obtain a high-quality
179transcriptome, the following analyses were performed: (1) A set of non-redundant,
180representative sequences for the assembled transcriptome was retained by CD-HIT
181ver. 4.6.1 (Huang *et al.*, 2010) with a threshold value of 0.95; (2) Coding and peptide
182sequences in the open reading frame were predicted by TransDecoder ver. 2.0.1 (Haa*s*
183*et* al., 2013) following the instructions described in the relevant wiki
184(https://github.com/TransDecoder/TransDecoder/wiki). This involved (1) extracting
185the long open reading frames, (2) identifying ORFs with homology to known proteins
186via Blast or Pfam (Fin*n et* al., 2016) searches, and (3) predicting the likely coding
187regions. In addition, the high-quality transcriptome generated from the above steps
188was further processed to remove bacterial contaminants using BLAST ver. 2.2.30+
189(Camacho *et al.*, 2009) and the longest transcripts were extracted with a custom Perl
190script. The completeness of assembled transcriptome was assessed using BUSCO
191analyses with embryophyta database (https://busco.ezlab.org/).

192**Phylogenetic analyses of the orthologous sequences of four species or species complex**

193To identify orthologous genes for phylogenetic analyses, we used OrthoMCL (Li *et*
194*al.*, 2003) to delineate gene families and cluster all genes into paralogous and
195orthologous groups based on species-level transcriptomes of the four species or
196species complex. The results were used to generate a 1:1:1:1 orthologous gene dataset
197for *P. brachytyla* s.s., *P. wilsonii*, PLSC, and *P. breweriana* (outgroup) with a custom
198Perl script.
199   The amino acid sequences for each ortholog group (OG) were aligned with
200MAFFT ver. 7.313 (Katoh and Standley 2013), and trimmed to exclude poorly
201aligned regions using TrimAl v1.2 (Capella-Gutierre*z et* al., 2009) with "-fasta -
202gappyout -colnumbering". The protein-coding nucleotide sequences for each OG were
203aligned based on the corresponding amino acid alignments using PAL2NAL v14
204(Suyam*a et al.*, 2006) to ensure the correct reading frames.

205 We then constructed phylogenies per gene using RAxML's rapid bootstrap 206algorithm under the GTRGAMMA model with 100 bootstrap replicates to find the 207best-scoring ML tree. We restricted this analysis to those groups satisfying the 208following criteria: sequence length great than 300 bp with '-' character excluded. The 209gene trees obtained that had less than 70% bootstrap support were excluded from 210further analysis. A custom R script was used to count the number of resulting 211phylogenies showing different topologies.

212 The 3,305 orthologous gene trees with more than 70% bootstrap support for 213branches were used to infer interspecific relationships with PhyloNet ver. 3.6.1 (Than 214*et al*., 2008; Yu *et al*., 2014). Rooted trees were converted into the required input 215format with a custom Perl script. Maximum likelihood with parametric bootstrap 216networks (using the command InferNetwork_MPL) in a coalescent framework, with 217both incomplete lineage sorting and gene flow taken into account, was inferred using 218PhyloNet allowing 0, 1 and 2 reticulations in 100 runs to return the best network.

219 We applied a K_s-based method to estimate divergence between species pairs. K_s 220values for each species pair were calculated using the ML method implemented in 221codeml of the PAML package (Yang 1997) under the F3 × 4 model (Goldman and 222Yang 1994). All pairs with a K_s value of less than 0.001, which would include 223transcript isoforms as well as recent tandem duplications, were discarded and not 224considered in the time estimation.

225Testing HHS using coalescent simulations based on population genomic data

226To examine evolutionary relationships among PLSC, *P. brachytyla* s.s. and *P.* 227*wilsonii*, we used fastsimcoal2 ver. 2.6.0.3 (Excoffier *et al*., 2013) to compare 228predefined demographic models using coalescent simulations based on the site 229frequency spectrum of all sampled individuals of these species. We included only the 230four-fold Degenerate Synonymous Sites (4DTV) when constructing two-dimensional 231joint site frequency spectra (2D-SFS) for each pair of species with ngsTools 232(Fumagall*i et* al., 2014). As we did not have information about the ancestral state, we 233treated the transcriptome of *P. abies* as both the reference and the ancestral state. 234After that, we folded all the 2D-SFSs with the 'fold' function implemented in $\partial a \partial i$ 235ver. 1.7.0. (Gutenkunst *et al*., 2009). In total, we used 16 different evolutionary 236models (Fig. S3), of which 11 (model1-model11) represented dichotomous or 237radiative topologies with or without gene flow after divergence, five (model12- 238model14) represented classical models of homoploid hybrid speciation via a single 239hybridization event with or without migration/size-change after divergence, and two 240(model15-model16) represented models of hybrid speciation involving a ghost 241intermediate hybrid lineage in the origin of *P. brachytyla* s.s.

242 For each model, we performed 100,000 coalescent simulations to estimate the 243expected 2D-SFS and computed log-likelihoods based on simulated and observed 2D- 244SFS matrixes. Global maximum likelihood estimates for each model were obtained 245from 50 independent runs, with 30-50 conditional maximization algorithm cycles. The 246relative fit of each of the different demographic models to the data was evaluated

247using the Akaike Information Criterion (AIC), and the model with the minimum AIC
248value was determined as the optimal. We assumed a mutation rate of 4.01×10^{-8} per
249site per generation and a generation time of 50 years (De La Torre *et al*., 2017; Li *et*
250*al*., 2010). A parametric bootstrapping approach was used to construct 95%
251confidence intervals with 50 independent runs for each bootstrap.

252　We used the reduced PLSC (with var. *rubescens* excluded) to examine the HHS
253origin of *P. brachytyla* s.s. because var. *rubescens* contains numerous introgressions
254from *P. purpurea*, which was assumed to originate from the same parents (Ru *et al*.,
2552018). This introgression, and/or the likely hybrid origin through *P. purpurea*, may
256complicate the modelling results. We further examined the diploid hybrid origin of *P.*
257*purpurea* from the reduced PLSC and *P. wilsonii*. We examined four alternative
258speciation models for the origin of *P. purpurea* because we had tested multiple
259models and all models suggested that this species originated through HHS (Ru *et al*.,
2602018) (Fig. S4). We tested which of four HHS models fit the reduced PLSC. Then we
261combined the origins of *P. brachytyla* s.s. and *P. purpurea* together to outline the
262evolutionary relationships among the reduced PLSC, *P. wilsonii*, *P. brachytyla* s.s.
263and *P. purpurea*.

264Ecological niche modelling

265Current potential distributions of the standard PLSC (including var. *rubescens*), *P.*
266*brachytyla* s.s. and *P. wilsonii* were used to examine niche divergences between them
267using the maximum entropy method in MAXENT version 3.4.1 (Phillips, Anderson
268and Schapire 2006; Phillips, Dudík and Schapire 2018) based on 105, 42, and 63
269locations obtained from field observation and herbarium records of, respectively, the
270PLSC, *P. brachytyla* s.s. and *P. wilsonii* (Table S11). The parameters were set as: 20
271replicates, a maximum of 5,000 iterations, 25% random test points, a threshold rule of
27210 percentile training presence applied and a convergence threshold of 0.00001.
273Climate data at a 2.5 arc minute resolution were downloaded from the WorldClim
274database (version 1.4, http://www.worldclim.org). Altitude data were downloaded
275from the SRTM elevation database (https://www2.jpl.nasa.gov/srtm/) and then
276projected to the same resolution as the climate data with ARCGIS. In total, 20
277variables were collected and pairwise Pearson's correlation coefficients (r) (Dormann,
278*et al*., 2013) were calculated with ENMTools version 1.4.4 (Warren *et al*., 2008;
2792010). Over-correlated variables (Pearson's correlation >=0.7) were excluded to avoid
280adverse effects (e.g. bias fitting) on the results.

281　　Niche differences between each pair of species were calculated based on
282Schoener's D (Schoener 1968) and Warren's I statistics (Warren *et al*., 2008), where a
283value of 0 indicates no niche overlap and 1 identical niches.

284Results

285Sampling, Sequencing and Single Nucleotide Polymorphism (SNP) Calling

286In total, we generated transcriptomes for 78 individuals of *P. brachytyla* and related 287species and downloaded the previously published transcriptomes of 108 individuals 288(Table S1, sample IDs starting with SRR) for the present analyses. After quality 289control, we retained an average of 46.37 million (M) reads (50.50 M raw reads) with 2906.18 gillion (G) clean bases per individual (Table S1). The *de novo* transcriptome 291assemblies for *P. brachytyla* s.s. and PLSC produced, respectively, 222,203 and 292313706 transcripts with N50 values of 549 and 667 after redundancy reduction and 293open reading frame (ORF) prediction. The numbers of total assembled bases, total 294Trinity transcripts, and genes, and the average contig length, contig N50, and percent 295GC, are similar to those for transcriptomes of *P. likiangensis*, *P. purpurea* and *P.* 296*wilsonii* from previous work (Ru *et al*., 2018) (Table S2). These assembled 297transcriptomes all with more than 80% BUSCO completeness (Table 2), but they have 298a larger number of contigs with lower N50 than the transcriptome of *P. abies* (Table 299S2) have a larger number of contigs with lower N50 than the transcriptome of *P.* 300*abies*, and we therefore mapped the quality-filtered reads to the revised transcriptome 301of *P. abies* as we did previously (Ru *et al*., 2016) and called SNPs for each individual. 302The average mapping rate for all individuals was 56.6%, with the average coverage of 303the reference transcriptome assembly being 73.9% and a 48.50-fold average effective 304depth (Table S1).

305   A total of 10,237 contigs with 339,165 SNPs were retained using our strict 306criteria after SAMTools calling. *P. likiangensis* (including all of the varieties), *P.* 307*brachytyla* s.s., *P. wilsonii*, *P. farreri*, *P. brachytyla*-southern lineage and *P. purpurea* 308contained 160,394, 180,329, 160,224, 99,459, 116,558, 140,250 SNPs respectively 309(Table S3). Among *P. likiangensis*, *P. brachytyla* s.s. and *P. wilsonii*, 31,319 SNPs 310were specific to *P. brachytyla* s.s.*,* 35,314 SNPs to *P. wilsonii* and 30,505 SNPs to *P.* 311*likiangensis* (Fig. S1A). The number of SNPs shared between *P. brachytyla* s.s. and 312either *P. likiangensis* (119,636) or *P. wilsonii* (114,657) was higher than that 313between *P. likiangensis* and *P. wilsonii* (95,536) (Fig. S1A). Among *P. likiangensis,* 314*P. farreri* and *P. brachytyla*-southern lineage, 41,626 SNPs were specific to *P.* 315*likiangensis*, 2,545 to *P. farreri* and 3,898 to *P. brachytyla*-southern lineage (Fig. 316S1B). About 88.51% SNPs (88,028 of 99,459) of *P. farreri* were shared between *P.* 317*farreri* and *P. brachytyla*-southern lineage but *P. farreri* still shared more with *P.* 318*likiangensis* (94,136 of 99,459) (Fig. S1B).

319**Nucleotide Diversity and interspecific Differentiation**

320Nucleotide diversities (π) of *P. likiangensis*, *P. brachytyla* s.s., *P. wilsonii*, *P.* 321*brachytyla*-southern lineage, *P. farreri* and *P. purpurea*, were all similar (Table S4). 322Mean genome-wide differentiations (F_{ST}) between every pair out of the six taxa were 323greater than 0.05, except for that between *P. farreri* and *P. likiangensis* 324(0.043±0.052). The F_{ST} value between *P. likiangensis* and *P. wilsonii* (0.116±0.093) 325was higher than that between *P. brachytyla* s.s. and either *P. likiangensis* 326(0.089±0.082) or *P. wilsonii* (0.105±0.092) (Table S5 and Fig. 2B and D). Similarly, 327absolute genetic divergence estimated by d_{XY} showed greater divergence between *P.*

328likiangensis and *P. wilsonii* than between *P. brachytyla* s.s. and either *P. likiangensis*
329or *P. wilsonii* (Table S5 and Fig. 2A and C). Both F_{ST} and d_{XY} indicated that *P. wilsonii*
330and *P. farreri* have the greatest divergence among all comparisons, and *P. wilsonii*
331and *P. brachytyla*-southern lineage take second place, while *P. brachytyla*-southern
332lineage is very close to *P. farreri* (F_{ST}: 0.064±0.080; d_{XY}:0.0099±0.013) (Table S5 and
333Fig. 2A and B). These divergences suggested a close relationship among *P.*
334*likiangensis*, *P. farreri* and *P. brachytyla*-southern lineage and we tentatively treated
335them as the *P. likiangensis* species complex (PLSC) for convenience.

336   With the transcriptome of *P. breweriana* used as outgroup, we reconstructed
337genealogies for all 186 individuals using NJ and ML methods based on N-RNA-seq
338and C-RNA-seq respectively. From the NJ tree based on N-RNA-seq, seven lineage-
339specific clusters were identified and the samples of each taxon clustered together with
340the exception of *P. brachytyla*, for which two clusters, *P. brachytyla*-southern lineage
341and *P. brachytyla* s.s. were recovered (Fig. 3A). *P. brachytyla*-southern lineage was
342closely related to var. *likiangensis* and *P. farreri*, and together they were related to
343var. *linzhiensis* and var. *rubescens*. These lineages comprise a monophyletic PLSC
344clade, paralleling *P. brachytyla* s.s., *P. purpurea* and *P. wilsonii*. *P. purpurea* was
345more closely related to *P. wilsonii* than to the others while *P. brachytyla* s.s. was
346sister to the PLSC.

347   ML phylogenetic analyses based on cpDNA variations, however, recovered only
348three well-supported clades: PLSC, *P. brachytyla* s.s. and *P. wilsonii - P. purpurea*.
349The latter two clades were sister to each other while no clear delimitation was found
350between *P. wilsonii* and *P. purpurea* or between different taxa of the PLSC. These
351delimitations and phylogenetic relationships are consistent with previous results
352(Lookwood *et al*., 2013; Sun *et al*., 2014; Ru *et al*. 2018; Lyu *et al*., 2020) (Fig. 3B).
353It was clearly noticeable that the phylogenetic relationships of *P. brachytyla* s.s. were
354discordant between nuclear genomic and plastome trees.

355**Population Structure and ADMIXTURE Analyses**

356Analyses using ADMIXTURE and PCA clustering revealed similar results to the N-
357RNA-seq phylogenetic trees apart from var. *rubescens*. For all individuals, in the
358ADMIXTURE analysis, as the *K* value increased from 2 to 4, four taxa of the PLSC,
359var. *likiangensis*, var. *linzhiensis*, *P. farreri* and *P. brachytyla*-southern lineage,
360shared the same genetic composition (Fig. 3C) while var. *rubescens* shared genetic
361ancestry with the reduced PLSC, *P. wilsonii* and *P. purpurea* when *K* = 2 to 4. *P.*
362*brachytyla* s.s. comprised a separate cluster when *K* = 3 while it exhibited mixed
363ancestry from the reduced PLSC and *P. wilsonii* when *K* = 2. Only when *K* = 4 did *P.*
364*purpurea* stand as a separate cluster, while it exhibited mixed genetic ancestry from
365both the reduced PLSC and *P. wilsonii* when *K* = 2 and 3. If both *P. brachytyla* s.s.
366and *P. purpurea* originated from the same parents, the reduced PLSC and *P. wilsonii*,
367these ADMIXTURE analyses suggested that *P. purpurea* originated later than *P.*
368*brachytyla* s.s.. Within the reduced PLSC, var. *linzhiensis* separated from the other
369three early while *P. farreri* contained a mixture of genetic elements from var.

370linzhiensis and var. *likiangensis* + *P. brachytyla*-southern lineage. Both var.
371*likiangensis* + *P. brachytyla*-southern lineage always belonged to the same genetic
372pool without clear separation when $K = 5$ or 6, even when var. *rubescens* had
373separated as an independent cluster. Further ADMIXTURE analyses of the PLSC
374individuals produced similar results (Fig. S2).

375   In the PCA analysis, the first two components (PC1 and PC2) explained 12.43%
376and 9.10% of the total variance respectively (Fig. 3D and Table S6), and distinguished
377four clear clusters, *P. brachytyla* s.s., *P. wilsonii*, *P. purpurea* and the PLSC. Within
378the PLSC, all individuals of var. *rubescens* comprised a separate cluster from the
379other three.

380PhyloNet test of HHS

381A total of 6,471 orthologous gene groups across four taxa were identified, and after
382filtering, 6,226 of these were used to generate gene trees. A total of 3,305 gene trees
383with ≥70% bootstrap support for all branches were subjected to PhyloNet testing. Of
384these trees, 1,129 (34.16%) clustered the reduced PLSC with *P. brachytyla* s.s.
385(topo1), 959 (29.02%) showed *P. brachytyla* s.s. as an isolated clade with *P. wilsonii*
386and the reduced PLSC clustered together (topo2), and 1,217 (36.82%) clustered *P.*
387*wilsonii* and *P. brachytyla* s.s. together (topo3) (Fig. 4A and B). The resulting
388phylogenetic network inferred by PhyloNet with an assumption of one past
389hybridization event (Fig. 4C) indicated a hybrid origin for *P. brachytyla* s.s.. The
390contributions from the reduced PLSC (~80%) were more than those from *P. wilsonii*
391(~20%) (Fig. 4C).

392Relative divergence time and K_s test of HHS

393We tested whether the divergence times between *P. brachytyla* s.s. and its two
394expected parents (*P. wilsonii* and PLSC) were close to each other based on a K_s-based
395method, as the hybrid species separated from the parents at almost the same time. As
396expected, the K_s value between *P. brachytyla* s.s. and PLSC was similar to the value
397between *P. brachytyla* s.s. and *P. wilsonii*, while all these values were smaller than
398that between PLSC and *P. wilsonii* (Fig. 4D).

399Coalescent analysis of alternative speciation patterns

400The best-fitting model for the origin of *P. brachytyla* s.s. (with the lowest AIC value,
401Table S7) was one involving hybridization and backcrossing (model 15) rather than
402bifurcation followed by introgression. The model indicates that *P. brachytyla* s.s.
403originated through backcrossing between the reduced PLSC and a hybrid extinct
404'ghost' lineage which was initially formed through hybridization between PLSC
405and *P. wilsonii* (Fig. 5A). We further estimated the effective population size of each
406lineage, divergence of the two assumed parents and the timescale of the homoploid
407hybridization events (Table 3 and Fig. 5A). Similarly, we found that *P. purpurea*
408originated through backcrossing, but between *P. wilsonii* and a hybrid 'ghost' lineage

409which was initially also formed through hybridization between the reduced PLSC 410and *P. wilsonii* because this model (model4) had the lowest AIC value (Fig. 5B and 411Fig. S4 and Table S8). The estimated effective population sizes for both parents and 412the timescale of the origin of the extinct hybrid lineage were similar to those for an 413HHS origin of *P. brachytyla* s.s. (Table 3 and Table S9).

414   We therefore assumed that two homoploid hybrid species, *P. brachytyla* s.s. and 415*P. purpurea*, originated through the same extinct intermediate hybrid lineage but by 416backcrossing to different parents. We combined four taxa together to examine their 417evolution by estimating parental contributions and hybrid speciation event timescales. 418We found that the extinct hybrid lineage from the reduced 419PLCS and *P. wilsonii* originated ~9.3 Ma (95%HPDI: 5.5-12.3 Ma) while *P.* 420*brachytyla* s.s. and *P. purpurea* originated ~1.1 Ma (95%HPDI: 0.8-5.2 Ma) and 421~0.50 Ma (95%HPDI: 0.4-3.4 Ma) through further backcrossing to the reduced PLSC 422or *P. wilsonii* (Fig. 5C). The current effective population sizes (Ne) of the reduced 423PLSC, *P. wilsonii* and *P. brachytyla* s.s. were estimated to be 15016 (95%HPDI: 42411106-51868), 27194 (95%HPDI: 18162-89603) and 16340 (95%HPDI: 11700-42557543) respectively, and gene flow from *P. brachytyla* s.s. to PLSC was estimated to 426be greater than that in the opposite direction, while gene flow from *P. brachytyla* s.s. 427to *P. wilsonii* was estimated to be less than that in the opposite direction, and both 428values were greater than that between PLSC and *P. wilsonii* (Table 3 and Table S10).

429   The estimated timescale of the extinct hybrid lineage was a little earlier (9.3 Ma 430versus 7.3 Ma or 7.9 Ma) than that when only one hybrid species was involved in 431calculations. However, the hybrid origin times for both hybrid species were estimated 432to be later than those estimated when only one hybrid species was involved (for *P.* 433*brachytyla* s.s., 1 Ma versus 1.2 Ma while for *P. purpurea*, the estimates were 0.58 434versus 0.49 Ma) . This may be a consequence of the changes in the effective 435population sizes of the four lineages when all of them were involved in the 436estimations.

437**Ecological niche differences between species**

438Environmental niche modeling was carried out to predict the current potential 439distributions of the PLSC, *P. brachytyla* s.s. and *P. wilsonii* groups (Fig. S5A). Eight 440bioclimatic variables (alt: altitude, bio2: mean diurnal range, bio4: temperature 441seasonality, bio8: mean temperature of wettest quarter, bio12: annual precipitation, 442bio14: precipitation of driest month, bio15: precipitation seasonality, bio19: 443precipitation of coldest quarter) were retained in our analysis. Mean area under the 444receiver operator curve (AUC) values were 0.989 for the PLSC, 0.984 for *P.* 445*brachytyla* s.s. and 0.998 for *P. wilsonii*, indicating that all models had high predictive 446ability. Identity tests (D and I) for the comparison between *P. brachytyla* s.s. and 447either PLSC or *P. wilsonii* all rejected the null hypothesis, indicating that *P.* 448*brachytyla* s.s. has extinct niche differentiation from both the PLSC and *P. wilsonii* (*P* 449< 0.001) (Fig. S5B).

450Discussion

451Our population genomic analyses confirmed the polyphyly of the previously 452circumscribed *P. brachytyla* (Ru *et al.*, 2016; Lyu *et al.*, 2020) and two non-sister 453lineages were identified: the *P. brachytyla*-southern lineage and *P. brachytyla* s.s.. Our 454further multiple analyses of population genomic data suggested an HHS origin for *P.* 455*brachytyla* s.s. through an extinct intermediate hybrid lineage and backcrossing to one 456parent. Interestingly, we found that the same parents may have given rise to another 457homoploid hybrid species, *P. purpurea*, through the same extinct hybrid lineage, but 458backcrossing to the other parent. To our knowledge, this is the first case illustrating 459that backcrossing to two parents produced two different homoploid species. Our 460findings recovered the high HHS complexity in *Picea*.

461Polyphyly of *P. brachytyla*

462Our population genomic data clearly suggest the polyphyly of the previously 463circumscribed *P. brachytyla*, consistent with previous studies based on a few 464individuals (Lockwood *et al.*, 2013; Zou *et al.*, 2016; Ru *et al.*, 2016; Shao *et al.*, 4652019; Shen *et al.*, 2019) or population genetic data from a few loci (Lyu *et al.*, 2020). 466Phylogenetic and population genetic analyses (PCA and ADMIXTURE structure) 467identified two groups: one of them (Fig. 3), the *P. brachytyla*-southern lineage, is 468closely related to var. *likiangensis* and *P. farreri* while the other, *P. brachytyla* s.s., 469comprises a separate lineage that is independent of all taxa sampled. The *P.* 470*brachytyla*-southern lineage comprises the PLSC together with another three or four 471taxa. Admixture analyses suggested that *P. brachytyla* s.s. originated as a separate 472lineage earlier ($K = 3$) than the *P. brachytyla*-southern lineage because it could not be 473distinguished from var. *likiangensis* even when $K = 7$. The two lineages of *P.* 474*brachytyla* have distinct distributional disjunction in northern and southern regions. In 475addition, the stomatal line is almost absent from the abaxial surface of the leaf in the 476*P. brachytyla*-southern lineage while one or more lines are found for *P. brachytyla* 477s.s., similar to *P. likiangensis*. It remains unknown why both *P. brachytyla*-southern

478lineage and *P. brachytyla* s.s. have two white or pale bands of stomatal lines on the 479flat adaxial leaf surface (Ru *et al*., 2016).

480Species distinctness and homoploid hybrid origin of *P. brachytyla* s.s.

481*P. brachytyla* s.s. is genetically and ecologically delimited from both *P. wilsonii* and 482the PLSC. Niche modelling showed that *P. brachytyla* s.s. occupies an ecological 483niche distinct from those of both the PLSC and *P. wilsonii* (Fig. S5). It should be 484noted that our niche modelling added var. *rubescens* to the PLSC. This taxon, whose 485evolutionary origin is unclear, occurs in a distribution close to that of *P. brachytyla* 486s.s.. Exclusion of this taxon would increase the ecological niche differentiation 487between *P. brachytyla* s.s. and the PLSC. In our population genomic studies, PCA, 488ADMIXTURE and phylogenetic analysis results are highly consistent with one 489another and with the niche modelling results, all of which suggests that *P. brachytyla* 490s.s. comprises a distinct genetic group (Fig. 3) with around 31,319 species-specific 491SNPs (Fig. S1).

492　　However, the phylogenetic relationships of *P. brachytyla* s.s. with the PLSC and 493*P. wilsonii* remain inconsistent. For example, although phylogenetic analyses based on 494nuclear-genome SNPs suggested a close relationship to the PLSC, plastome 495phylogeny clustered all plastome haplotypes of *P. brachytyla* s.s. together with those 496of *P. wilsonii* as a separate clade from that of the PLSC. Phylogenetic analyses based 497on orthologous genes showed a similar pattern: ~63% of the gene trees showed a 498close relationship between *P. brachytyla* s.s. and PLSC while ~37% indicated that it 499was closer to *P. wilsonii* (Fig. 4). In fact, such an inconsistency has also been 500previously reported based on sequencing a few loci from a few individuals or 501populations (Lockwood *et al*., 2013; Ran *et al*., 2015; Shen *et al*., 2019; Zou *et al*., 5022016; Shao *et al*., 2019; Lyu *et al*., 2020). In addition, the specific morphological 503traits of the PLSC and *P. wilsonii* could be seen together in *P. brachytyla* s.s. (Fu *et* 504*al*., 1999; Lyu *et al*., 2020). All of these analyses suggested that *P. brachytyla* s.s. 505might have originated from one or more than one hybridization event between PLSC 506and *P. wilsonii*. We further examined this possibility by means of three analyses. First, 507K_s analysis based on the shared orthologous genes revealed a similar divergence time 508between *P. brachytyla* s.s. and either *P. likiangensis* or *P. wilsonii* while this 509divergence is more recent than that between PLSC and *P. wilsonii* (Fig. 2). Second, 510PhyloNet analysis based on the orthologous gene trees revealed a hybrid origin for *P.* 511*brachytyla* s.s. as expected, with ~81.19% of its nuclear composition derived from 512PLCS and ~19.81% from *P. wilsonii* (Fig. 4). Finally, we modeled alternative 513speciation events and the coalescent simulation analyses strongly supported a hybrid 514origin for *P. brachytyla* s.s. (Fig. 5A). According to the best-fitting model, *P.* 515*brachytyla* s.s. originated in two steps with the first intermediate hybrid lineage being 516formed between the reduced PLSC and *P. wilsonii* followed by backcrossing of this 517extinct intermediate lineage with the PLSC (Table 3). This two-step HHS is very 518similar to that found for another closely related species, *P. purpura* (R*u et* al., 2018).

519All of these analyses consistently support a HHS origin for *P. brachytyla* s.s. from the
520PLSC (before its further diversification) and *P. wilsonii*.

521Two homoploid hybrid species originated from the same parents through a
522'ghost' hybrid lineage to backcross different parents

523In our previous HHS study of *P. purpurea*, we used three varieties of *P. likiangensis* to
524represent the PLSC (Sun *et al.*, 2014; Ru *et al.*, 2018). Later studies (Sun *et al.*, 2018;
525Lyu *et al.*, 2020) and our present analyses suggested that both *P. farreri* and *P.*
526*brachytyla*-southern lineage should be included in the PLSC. In addition, var.
527*rubescens* contained genetic introgression from *P. purpurea*, according to our
528population structure analyses (Fig. 3C) and previous studies (Sun *et al.*, 2018; Shen *et*
529*al.*, 2019; Shao *et al.*, 2019). This introgression may complicate the modelling of
530hypotheses for the alternative speciation events that may have produced *P. purpurea*.
531We therefore excluded var. *rubescens* from the PLSC and used the reduced PLSC and
532*P. wilsonii* to model HHS versus an alternative, bifurcating, origin for *P. purpurea*.
533Similar to previous findings (Ru *et al.*, 2018), *P. purpurea* originated through HHS
534between the reduced PLSC and *P. wilsonii* by a two-step process that formed an
535intermediate hybrid lineage, which further backcrossed with *P. wilsonii* to produce *P.*
536*purpurea* (Fig. 5B and Fig. S4 and Table S8). Thus both *P. purpurea* and *P. brachytyla*
537s.s.. originated by HHS from the same parents, PLSC and *P. wilsonii*. However, *P.*
538*purpurea* originated later than *P. brachytyla* s.s., accumulating fewer species-specific
539mutations and retaining more parental ancestry. The mosaic of parental ancestry in *P.*
540*purpurea* was always obvious when ADMIXTURE structure analyses were applied to
541population genomic data (Fig. 3C). However, for *P. brachytyla* s.s., more species-
542specific mutations blurred the evidence of a mixed ancestry, although inconsistent
543phylogenies (Fig. 3A and B), PhyloNet suggestions (Fig. 4) and coalescent tests (Fig.
5445) together support a HHS origin for this species.
545   In addition, we assumed that the two homoploid hybrid species originated
546through the same extinct intermediate hybrid lineage but that it backcrossed with
547different parents for two reasons. First, this is the most parsimonious hypothesis.

548Although it is likely that two ancient hybrid lineages may have originated between 549PLSC and *P. wilsonii*, it is less likely that both of them became extinct. Second, the 550time of origin of the extinct hybrid lineage was estimated to be similar (7.3 Ma or 7.9 551Ma) when only two parents and one assumed hybrid species, *P. brachytyla* s.s. or *P.* 552*purpurea*, was involved (Fig. 5). When all four taxa were involved, the origin of the 553extinct hybrid lineage was estimated to be earlier, around 9.3 Ma (Fig. 5), while the 554two hybrid species, *P. purpurea* and *P. brachytyla* s.s., originated 1 Ma or 0.49 Ma 555(Fig. 5). These speciation events occurred from the late Miocene to the Quaternary 556when the Qinghai-Tibet Plateau (QTP), where the four taxa focused on in this study 557mainly occur, experienced extensive geological and climatic oscillations (Deng and 558Ding 2015; Mulch and Chamberlain 2006). During these oscillations numerous new 559species originated and interspecific hybridizations occurred (Liu *et al*., 2013; Du *et* 560*al*., 2017; Ma *et al*., 2019). These extensive changes may therefore have helped the 561two ancestral parents come into direct contact, resulting in inter-lineage hybridizations 562and backcrosses, which led to the origin of the intermediate hybrid lineage and the 563two extant hybrid species. It remains unknown how the intermediate hybrid lineage 564was extinguished. It is likely that environmental changes or adaptive advantages of 565the newly formed hybrid species resulted in the extinction or replacement of the 566intermediate hybrid lineage (Ru *et al*., 2018).

567   To our knowledge, this is the second reported case in which two parents produced 568more than one homoploid hybrid species. Two sunflower species hybridized to give 569rise to three homoploid hybrid species occurring in different extreme habitats 570(Rieseberg *et al*., 1997, 2003). In our case, one extinct 'ghost' hybrid lineage 571underwent backcrossing events with each of its parents to produce the two extant 572hybrid species (Fig. 5). This finding supports the hypothesis that in the spruce genus, 573reticulate species diversification through hybridization ranther than non-bifurcating 574divergence seems to be more frequent than was previously assumed (Feng *et al*., 5752019).

576

583

584**AUTHOR CONTRIBUTIONS**

585J.L. and D.R. planned and designed the research. D.W. and D.R. conducted fieldwork,

586performed experiments and analysed data etc. D. R., J.L., and D.W. wrote the

587manuscript.

588**DATA ACCESSIBILITY**

589The sequencing data have been deposited in National Genomics Data Center under

590the **BioProject ID**: PRJCA003239.

591ORCID

592*Jianquan Liu* http://orcid.org/0000-0002-4237-7418

593**References**

594Alexander DH, Lange K. (2011). Enhancements to the ADMIXTURE algorithm for
595   individual ancestry estimation. *BMC Bioinformatics*, 12, 246.
596Bolger AM, Lohse M, Usadel B. (2014). Trimmomatic: a flexible trimmer for
597   Illumina sequence data. *Bioinformatics*, 30, 2114-2120.
598Brennan AC, Hiscock SJ, Abbott RJ. (2019). Completing the hybridization triangle:
599   the inheritance of genetic incompatibilities during homoploid hybrid speciation
600   in ragworts (*Senecio*). *AoB Plants*, 11, ply078.
601Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.
602   (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, 10, 421.
603Capella-Gutiérrez S, Silla-Martinez JM,Gabaldon T. (2009). TrimAl: a tool for
604   automated alignment trimming in large-scale phylogenetic analyses.
605   *Bioinformatics*, 25, 1972-1973.
606Chapman MA, Hiscock SJ, Filatov DA. (2013). Genomic divergence during
607   speciation driven by adaptation to altitude. *Molecular Biology and Evolution*, 30:
608   2553-2567.
609Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE,
610   Lunter G, Marth GT, Sherry ST *et al.* (2011). The variant call format and
611   VCFtools. *Bioinformatics*, 27, 2156-2158.

612 De La Torre AR, Li Z, Van de Peer Y, Ingvarsson PK. (2017). Contrasting rates of
613   molecular evolution and patterns of selection among gymnosperms and
614   flowering plants. *Molecular Biology and Evolution*, 34, 1363-1377.

615 Deng T, Ding L. (2015). Paleoaltimetry reconstructions of the Tibetan Plateau,
616   progress and contradictions. *National Science Review*, 2, 417-437.

617 Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber
618   B, Lafourcade B, Leitão PJ *et al.* (2013). Collinearity: a review of methods to
619   deal with it and a simulation study evaluating their performance. *Ecography*, 36,
620   27-46.

621 Du FK, Hou M, Wang W, Mao K, Hampe A. (2017). Phylogeography of *Quercus*
622   aquifolioides provides novel insights into the Neogene history of a major global
623   hotspot of plant diversity in south-west China. *Journal of Biogeography*, 44,
624   294-307.

625 Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. (2013). Robust
626   demographic inference from genomic and SNP data. *PLoS Genetics*, 9,
627   e1003905.

628 Feng S, Ru D, Sun Y, Mao K, Milne R, Liu J. (2019). Trans−lineage polymorphism
629   and nonbifurcating diversification of the genus *Picea*. *New Phytologist*, 222, 576.

630 Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta
631   M, Qureshi M, Sangrador-Vegas A *et al.* (2016). The Pfam protein families
632   database: towards a more sustainable future. *Nucleic Acids Research*, 44, D279-
633   285.

634 Foote AD, Vijay N, Avila-Arcos MC, Baird RW, Durban JW, Fumagalli M, Gibbs RA,
635   Hanson MB, Korneliussen TS, Martin MD *et al.* (2016). Genome-culture
636   coevolution promotes rapid divergence of killer whale ecotypes. *Nature
637   Communications*, 7, 11693.

638 Fu L, Li N, Mills R. (1999). *Pinaceae, flora of China.* Beijing, China: Science Press.

639 Fumagalli M, Vieira FG, Linderoth T, Nielsen R. (2014). ngsTools: methods for
640   population genetics analyses from next-generation sequencing data.
641   *Bioinformatics*, 30, 1486-1487.

642 Goldman N, Yang Z. (1994). A codon-based model of nucleotide substitution for
643   protein-coding DNA sequences. *Molecular Biology and Evolution*, 11, 725-736.

644 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X,
645   Fan L, Raychowdhury R, Zeng Q *et al.* (2011). Full-length transcriptome
646   assembly from RNA-Seq data without a reference genome. *Nature
647   Biotechnology*, 29, 644-652.

648 Gross BL, Rieseberg LH. (2005). The ecological genetics of homoploid hybrid
649   speciation. *Journal of Heredity*, 96, 241-252.

650 Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. (2009). Inferring
651   the joint demographic history of multiple populations from multidimensional
652   SNP frequency data. *PLoS Genetics*, 5, e1000695.

653 Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB,
654   Eccles D, Li B, Lieber M *et al.* (2013). *De novo* transcript sequence
655   reconstruction from RNA-seq using the Trinity platform for reference generation

656   and analysis. *Nature Protocols*, 8, 1494-1512.

657 Hermansen JS, Sætre SA, Elgvin TO, Borge T, Hjelle E, Sætre GP. (2011). Hybrid
658   speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers
659   to gene flow. *Molecular Ecology*, 20, 3812-3822.

660 Huang Y, Niu B, Gao Y, Fu L, Li W. (2010). CD-HIT suite: a web server for
661   clustering and comparing biological sequences. *Bioinformatics*, 26, 680.

662 Katoh K, Standley DM. (2013). MAFFT multiple sequence alignment software
663   version 7: improvements in performance and usability. *Molecular Biology and
664   Evolution*, 30, 772-780.

665 Lamichhaney S, Han F, Webster MT, Andersson L, Grant BR, Grant PR. (2018).
666   Rapid hybrid speciation in Darwin's finches. *Science*, 359, 224-228.

667 Li H, Durbin R. (2009). Fast and accurate short read alignment with Burrows-
668   Wheeler transform. *Bioinformatics*, 25, 1754-1760.

669 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,
670   Durbin R *et al.* (2009). The sequence alignment/map format and SAMtools.
671   *Bioinformatics*, 25, 2078-2079.

672 Li L, Stoeckert Jr. C, Roos DS. (2003). OrthoMCL: identification of ortholog groups
673   for eukaryotic genomes. *Genome Research*, 13, 2178.

674 Li L, Abbott RJ, Liu B, Sun Y, Li L, Zou J, Wang X, Miehe G, Liu J. (2013). Pliocene
675   intraspecific divergence and Plio-Pleistocene range expansions within *Picea
676   likiangensis* (Lijiang spruce), a dominant forest tree of the Qinghai-Tibet Plateau.
677   *Molecular Ecology*, 22, 5237-5255.

678 Li M, Tian S, Yeung CK, Meng X, Tang Q, Niu L, Wang X, Jin L, Ma J, Long K *et al.*
679   (2014). Whole-genome sequencing of Berkshire (European native pig) provides
680   insights into its origin and domestication. *Scientific Reports*, 4, 4678.

681 Li Y, Stocks M, Hemmila S, Kallman T, Zhu H, Zhou Y, Chen J, Liu J, Lascoux M.
682   (2010). Demographic histories of four spruce (*Picea*) species of the Qinghai-
683   Tibetan Plateau and neighboring areas inferred from multiple nuclear loci.
684   *Molecular Biology and Evolution*, 27, 1001-1014.

685 Liu J, Moller M, Provan J, Gao LM, Poudel RC, Li DZ. (2013). Geological and
686   ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New
687   Phytologist*, 199, 1093-1108.

688 Lockwood JD, Aleksic JM, Zou J, Wang J, Liu J, Renner SS. (2013). A new
689   phylogeny for the genus *Picea* from plastid, mitochondrial, and nuclear
690   sequences. *Molecular Phylogenetics and Evolution*, 69, 717-727.

691 Lyu LK, Wang DL, Li L, Zhu YY, Jiang DC, Liu JQ, Xu XT. (2020). Polyphyly and
692   species delimitation of *Picea brachytyla* (Pinaceae) based on population genetic
693   data. *Journal of Systematics and Evolution*. doi: https://doi.org/10.1111/jse.12584

694 Ma Y, Wang J, Hu Q, Li J, Sun Y, Zhang L, Abbott RJ, Liu J, Mao K. (2019). Ancient
695   introgression drives adaptation to cooler and drier mountain habitats in a cypress
696   species complex. *Communications Biology*, 2, 213.

697 Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, Linares M. (2006).
698   Speciation by hybridization in *Heliconius* butterflies. *Nature*, 441, 868-871.

699 Mulch A, Chamberlain CP. (2006). Earth science: the rise and growth of Tibet.

700   *Nature*, 439, 670-671.

701 Nei M, Li WH. (1979). Mathematical model for studying genetic variation in terms of
702   restriction endonucleases. *Proceedings of the National Academy of Sciences of*
703   *the United States of America*, 76, 5269-5273.

704 Nieto Feliner G, Álvarez I, Fuertes-Aguilar J, Heuertz M, Marques I, Moharrek F,
705   Pineiro R, Riina R, Rossello JA, Soltis PS *et al.* (2017). Is homoploid hybrid
706   speciation that rare? An empiricist's view. *Heredity*, 118, 513-516.

707 Nolte AW, Tautz D. (2010). Understanding the onset of hybrid speciation. *Trends in*
708   *Genetics*, 26, 54-58.

709 Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F,
710   Delhomme N, Giacomello S, Alexeyenko A *et al.* (2013). The Norway spruce
711   genome sequence and conifer genome evolution. *Nature*, 497, 579-584.

712 Phillips SJ, Anderson RP, Schapire RE. (2006). Maximum entropy modeling of
713   species geographic distributions. *Ecological Modelling*, 190, 231-259.

714 Phillips SJ, Anderson RP, Schapire RE. (2018). *Maxent software for modeling species*
715   *niches   and   distributions   (Version   3.4.1).*   URL
716   http://biodiversityinformatics.amnh.org/open_source/maxent/.   [accessed   8
717   September 2020].

718 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. (2006).
719   Principal components analysis corrects for stratification in genome-wide
720   association studies. *Nature Genetics*, 38, 904-909.

721 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J,
722   Sklar P, de Bakker PI, Daly MJ *et al.* (2007). PLINK: a tool set for whole-
723   genome association and population-based linkage analyses. *American Journal of*
724   *Human Genetics*, 81, 559-575.

725 Ran JH, Shen TT, Liu WJ, Wang PP, Wang XQ. (2015). Mitochondrial introgression
726   and complex biogeographic history of the genus *Picea*. *Molecular Phylogenetics*
727   *and Evolution*, 93, 63-76.

728 Rieseberg LH. (1997). Hybrid origins of plant species. *Annual review of Ecology and*
729   *Systematics*, 28, 359-389.

730 Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T,
731   Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. (2003). Major ecological
732   transitions in wild sunflowers facilitated by hybridization. *Science*, 301, 1211-
733   1216.

734 Ru D, Mao K, Zhang L, Wang X, Lu Z, Sun Y. (2016). Genomic evidence for
735   polyphyletic origins and interlineage gene flow within complex taxa: a case
736   study of *Picea brachytyla* in the Qinghai-Tibet Plateau. *Molecular Ecology*, 25,
737   2373-2386.

738 Ru D, Sun Y, Wang D, Chen Y, Wang T, Hu Q, Abbott RJ, Liu J. (2018). Population
739   genomic analysis reveals that homoploid hybrid speciation can be a lengthy
740   process. *Molecular Ecology*, 27, 4875-4887.

741 Schoener TW. (1968). The anolis lizards of Bimini: resource partitioning in a complex
742   fauna. *Ecological Society of America*, 49, 704-726.

743 Schumer M, Cui R, Rosenthal GG, Andolfatto P. (2015). Reproductive isolation of

744 hybrid populations driven by genetic incompatibilities. *PLoS Genetics*, 11,
745 e1005041.

746 Schumer M, Rosenthal GG, Andolfatto P. (2014). How common is homoploid hybrid
747 speciation? *Evolution*, 68, 1553-1560.

748 Shao CC, Shen TT, Jin WT, Mao HJ, Ran JH, Wang XQ. (2019).
749 Phylotranscriptomics resolves interspecific relationships and indicates multiple
750 historical out-of-North America dispersals through the Bering Land Bridge for
751 the genus *Picea* (*Pinaceae*). *Molecular Phylogenetics and Evolution*, 141,
752 106610.

753 Shen TT, Ran JH, Wang XQ. (2019). Phylogenomics disentangles the evolutionary
754 history of spruces (*Picea*) in the Qinghai-Tibetan Plateau: implications for the
755 design of population genetic studies and species delimitation of conifers.
756 *Molecular Phylogenetics and Evolution*, 141, 106612.

757 Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-
758 analysis of large phylogenies. *Bioinformatics*, 30, 1312-1313.

759 Sun Y, Abbott RJ, Li L, Li L, Zou J, Liu J. (2014). Evolutionary history of purple
760 cone spruce (*Picea purpurea*) in the Qinghai-Tibet Plateau: homoploid hybrid
761 origin and Pleistocene expansion. *Molecular Ecology,* 23, 343-359.

762 Sun Y, Abbott RJ, Lu Z, Mao K, Zhang L, Wang X, Ru D, Liu J. (2018). Reticulate
763 evolution within a spruce (*Picea*) species complex revealed by population
764 genomic analysis. *Evolution*, 72, 2669-2681.

765 Suyama M, Torrents D, Bork P. (2006). PAL2NAL: robust conversion of protein
766 sequence alignments into the corresponding codon alignments. *Nucleic Acids
767 Research*, 34, W609-612.

768 Than C, Ruths D, Nakhleh L. (2008). PhyloNet: a software package for analyzing and
769 reconstructing reticulate evolutionary relationships. *BMC Bioinformatics*, 9, 322.

770 Wang GD, Zhai W, Yang HC, Fan RX, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng
771 LG *et al.* (2013). The genomics of selection in dogs and the parallel evolution
772 between dogs and humans. *Nature Communications*, 4, 1860.

773 Warren DL, Glor RE, Turelli M. (2008). Environmental niche equivalency versus
774 conservatism: quantitative approaches to niche evolution. *Evolution*, 62, 2868-
775 2883.

776 Warren DL, Glor RE, Turelli M. (2010). ENMTools: a toolbox for comparative
777 studies of environmental niche models. *Ecography*, 33, 607-611.

778 Weir BS, Cockerham CC. (1984). Estimating f-statistics for the analysis of
779 population-structure. *Evolution*, 38, 1358-1370.

780 Wright JW. (1955). Species crossability in spruce in relation to distribution and
781 taxonomy. *Forest Science*, 1, 319-349.

782 Yang Z. (1997). PAML: a program package for phylogenetic analysis by maximum
783 likelihood. *Computer Applications in the Biosciences*, 13, 555-556.

784 Yu Y, Dong J, Liu KJ, Nakhleh L. (2014). Maximum likelihood inference of reticulate
785 evolutionary histories. *Proceedings of the National Academy of Sciences of the
786 United States of America*, 111, 16448-16453.

787 Zou JB, Yue W, Li LL, Wang X, Lu J, Duan BB, Liu JQ. (2016). DNA barcoding of
788   recently diversified tree species: a case study on spruces based on 20 DNA
789   fragments from three different genomes. *Trees*, 30, 959-969.

790**Figure legends**

791

792Figure 1. Distributional ranges of *P. wilsonii*, *P. purpurea*, *P. farreri* and *P. likiangensis* var.
793*rubescens*, *P. likiangensis* var. *linzhiensis*, *P. likiangensis* var. *likiangensis*, *P. brachytyla* sensu
794stricto (s.s.) and *P. brachytyla*-southern lineage. Different colored circles indicate locations from
795which samples of each species were collected for RNA-seq.

796

797

798Figure 2. (A, B) Comparison of d_{XY} and F_{ST} among 10 population pairs evaluated for the nuclear
799SNP data set. (C, D) Distribution of genetic differentiation (d_{XY} and F_{ST}) among 3 population pairs
800based on the nuclear SNP data set.

801

802

803Figure 3. (A) Neighbor Joining tree based on all 186 transcriptome sequences with *P. breweriana*
804as the outgroup. (B) Maximum Likelihood (ML) tree based on 69 transcriptome sequences with *P.*
805*breweriana* as the outgroup. (C) Bar plots indicative of assignment probabilities from
806ADMIXTURE analysis of 184 transcriptomes from K =2 to K=6. (D) Principal component
807analysis (PCA) plots showing the first two principal components.

808

809

810Figure 4. (A) Relationships between *P. likiangensis* species complex (PLSC), *P. brachytyla* s.s.
811and *P. wilsonii* (using *P. breweriana* as outgroup) according to ML analysis of 3,305 orthologous
812gene sequences identified by OrthoMCL as having a ratio of 1:1:1:1. (B) The proportion of each
813topology is based on the ortholog groups. (C) ML-bootstrap network for 3,305 orthologous gene
814trees generated by PhyloNet after runs allowing 0, 1 or 2 reticulations. Reticulations are shown in
815blue with inheritance probabilities. Note: Only trees with branch bootstrap values >70% were
816analysed. (D) K_S age distributions for one-to-one orthologs between PLSC and *P. brachytyla* s.s.
817and *P. wilsonii*.

818

819

820Figure 5. (A) Simplified graphical summary of the best-fitting demographic model inferred by
821fastsimcoal2 for the reduced PLSC, *P. brachytyla* s.s., and *P. wilsonii*. (B) Simplified graphical
822summary of the best-fitting demographic model inferred by fastsimcoal2 for the reduced PLSC, *P.*
823*purpurea*, and *P. wilsonii*. (C) Simplified graphical summary of the best-fitting demographic
824model inferred by fastsimcoal2 for the reduced PLSC, *P. brachytyla* s.s., *P. purpurea* and *P.*
825*wilsonii*. The percentages indicate nuclear genomic compositions contributed from parents to the
826hybrid offspring. The right-hand axis indicates the timescale in years before the present.

827
828
829
830
831

832**Table legends**

833

834Table 1 Locations of *Picea* individuals sampled for this study

835

836

837Table 2 BUSCO results for assembly completeness of four spruce transcriptomes

838

839

840Table 3 Inferred demographic parameters of the best-fitting demographic model in Fig. S3

Supporting Information

Figure S1. Venn diagram summarizing the number of shared and exclusive nuclear SNPs called across (A) *P. likiangensis*, *P. brachytyla* s.s. and *P. wilsonii* and (B) *P. likiangensis*, *P. farreri* and *P. brachytyla*-southern lineage.

Figure S2. ADMIXTURE analysis of *P. likiangensis* var. *rubescens*, *P. likiangensis* var. *linzhiensis*, *P. likiangensis* var. *likiangensis*, *P. brachytyla* s.s., *P. brachytyla*-southern lineage and *P. farreri*.

Figure S3. Schematic of 16 different demographic models analysed using fastsimcoal2. Note: the gray bars indicate the joint past population and arrows indicate asymmetric gene flow. The red, green, dark blue and khaki bars represent PLSC, *P. wilsonii*, *P. brachytyla* s.s. and the ghost lineage respectively. Models 1-2, radiative divergences of three species without/with gene flow with T as the radiative time from a common ancestor; Models 3-11: stepwise bifurcating models of three species without/with gene flow with T1 and T2 as the divergence times; models 12-14, hybrid origin of *P. brachytyla* s.s. between the other two without/with gene flow with T1 as the time of divergence between PLSC and *P. wilsonii*, and T2 as the time of origin of *P. brachytyla* s.s.; models 15-16, hybrid origin of *P. brachytyla* s.s. between the other two with an additional ghost lineage (in khaki) and asymmetric gene flow and T1 as the time of divergence between PLSC and *P. wilsonii*, T2 as the time of origin of the ghost lineage and T3 as the time of origin of *P. brachytyla* s.s..

Figure S4. Schematic of 4 different demographic models analysed using fastsimcoal2. Note: the gray bars indicate the joint past population and arrows indicate asymmetric gene flow. The red, green, pink and khaki bars represent PLSC, *P. wilsonii*, *P. purpurea* and ghost lineage respectively. Models 1-2, stepwise bifurcating models of three species with gene flow with T1 and T2 as the divergence times; Model 3, hybrid origin of *P. purpurea* between the other two with gene flow with T1 as the time of divergence time PLSC and *P. wilsonii*, and T2 as the time of origin of *P. purpurea*; Model 4, hybrid origin of *P. purpurea* between the other two with an additional ghost lineage (in khaki) and asymmetric gene flow and T1 as the time of divergence between PLSC and *P. wilsonii*, T2 as the time of origin of the ghost lineage and T3 as the time of origin of *P. purpurea*.

Figure S5. Environmental niche models for three species, and results of identity tests between paired groups. (A) Current potential distributions of PLSC, *P. brachytyla* s.s. and *P. wilsonii* groups, predicted by Maxent. (B) Results of identity tests for three comparisons (PLSC vs. *P. brachytyla* s.s., *P. brachytyla* s.s. vs. *P. wilsonii*, PLSC vs. *P. wilsonii*). The pink bars indicate the null distributions of I, while the light blue bars indicate D. Arrows indicate values of I (pink) and D (light blue) in actual Maxent runs.

Table S1 Summary statistics of Illumina transcriptome data for each sample.

883 Table S2 Transcriptome assembly statistics for the *Picea* species including revised *P. abies*.

884 Table S3 SNPs information for six *Picea* species.

885

886 Table S4 $\pi\pm$SD with *P. abies* as the reference.

887

888 Table S5 Summary statistics for F_{ST} and d_{XY} with *P. abies* as the reference.

889

890 Table S6 Tracy-Widom statistics for the first four eigenvalues in PCA analysis.

891

892 Table S7 Relative likelihoods of the different demographic models shown in Fig. S3.

893

894 Table S8 Relative likelihoods of the different demographic models shown in Fig. S4.

895

896 Table S9 Inferred demographic parameters of the best-fitting demographic model in Fig. S4.

897

898 Table S10 Inferred demographic parameters with PLSC, *P. wilsonii*, *P. brachytyla* s.s. and *P.*

899 *purpurea* included in the analysis.

900

901 Table S11 Species occurrences collected for ecological niche modeling analysis.

902

903
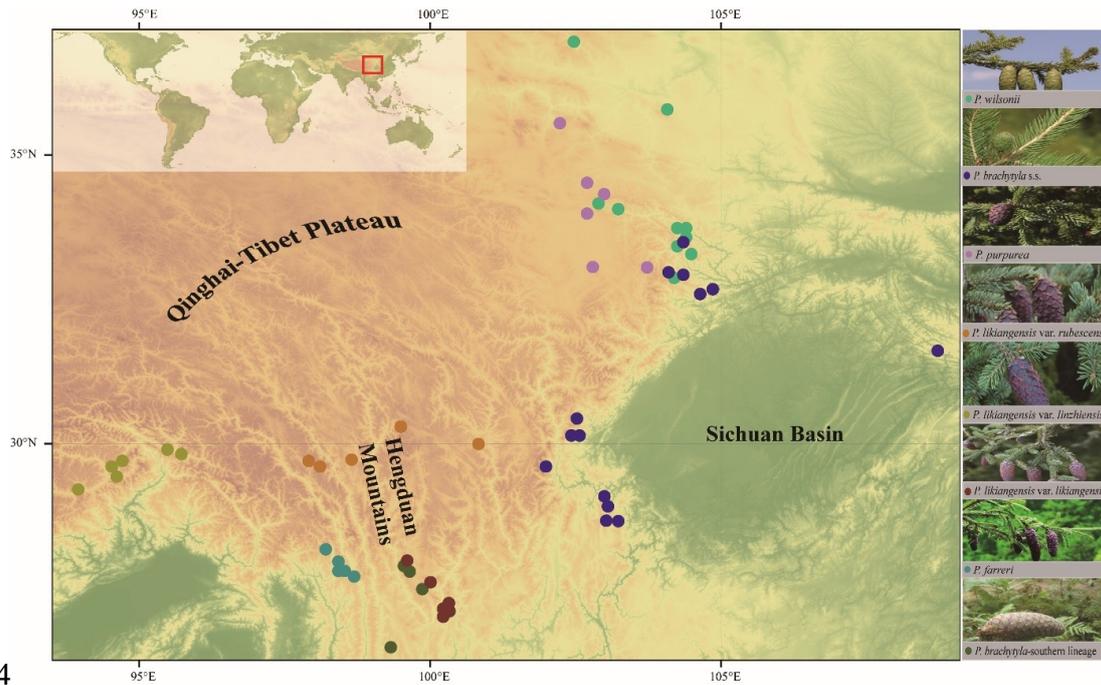
904

905Figure 1. Distributional ranges of *P. wilsonii*, *P. purpurea*, *P. farreri* and *P. likiangensis* var.
906*rubescens*, *P. likiangensis* var. *linzhiensis*, *P. likiangensis* var. *likiangensis*, *P. brachytyla* sensu
907stricto (s.s.) and *P. brachytyla*-southern lineage. Different colored circles indicate locations from
908which samples of each species were collected for RNA-seq.
909
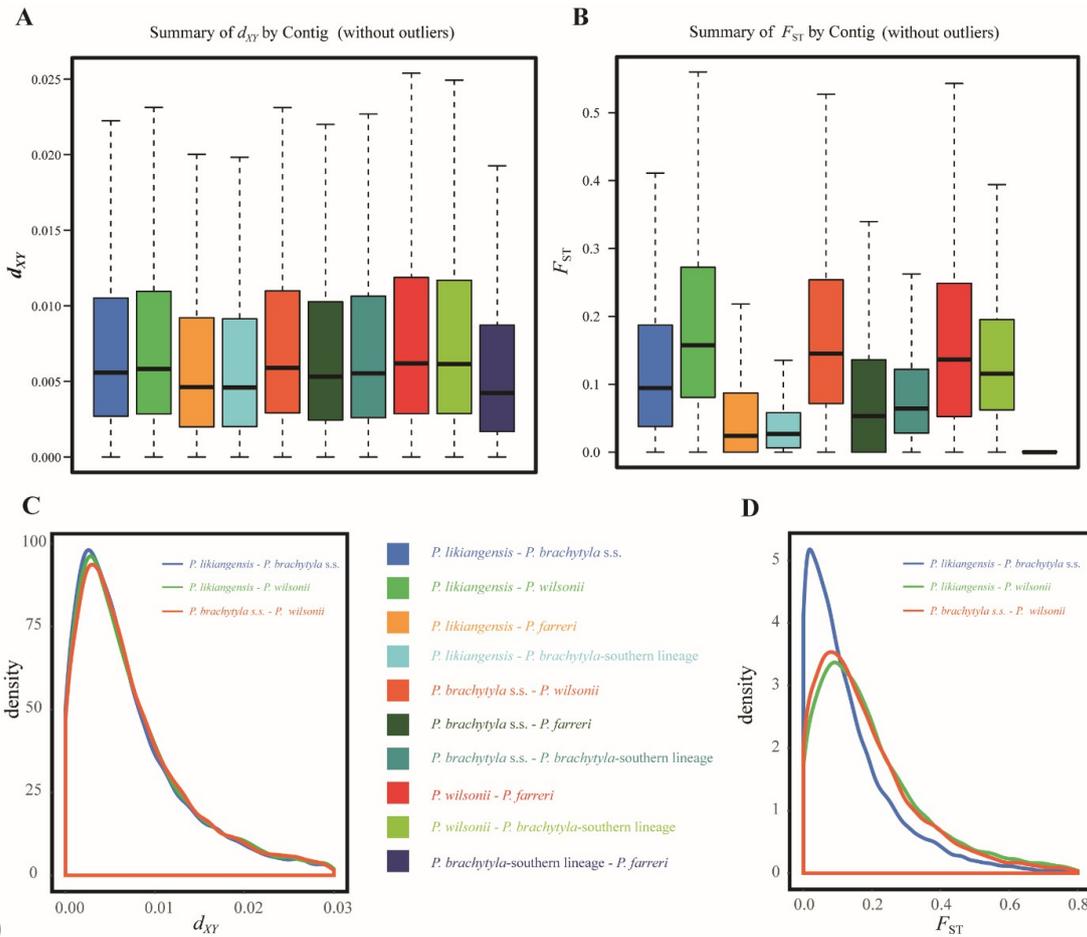
910

911 Figure 2. (A, B) Comparison of d_{XY} and F_{ST} among 10 population pairs evaluated for the nuclear **912**SNP data set. (C, D) Distribution of genetic differentiation (d_{XY} and F_{ST}) among 3 population pairs **913**based on the nuclear SNP data set.

914

915
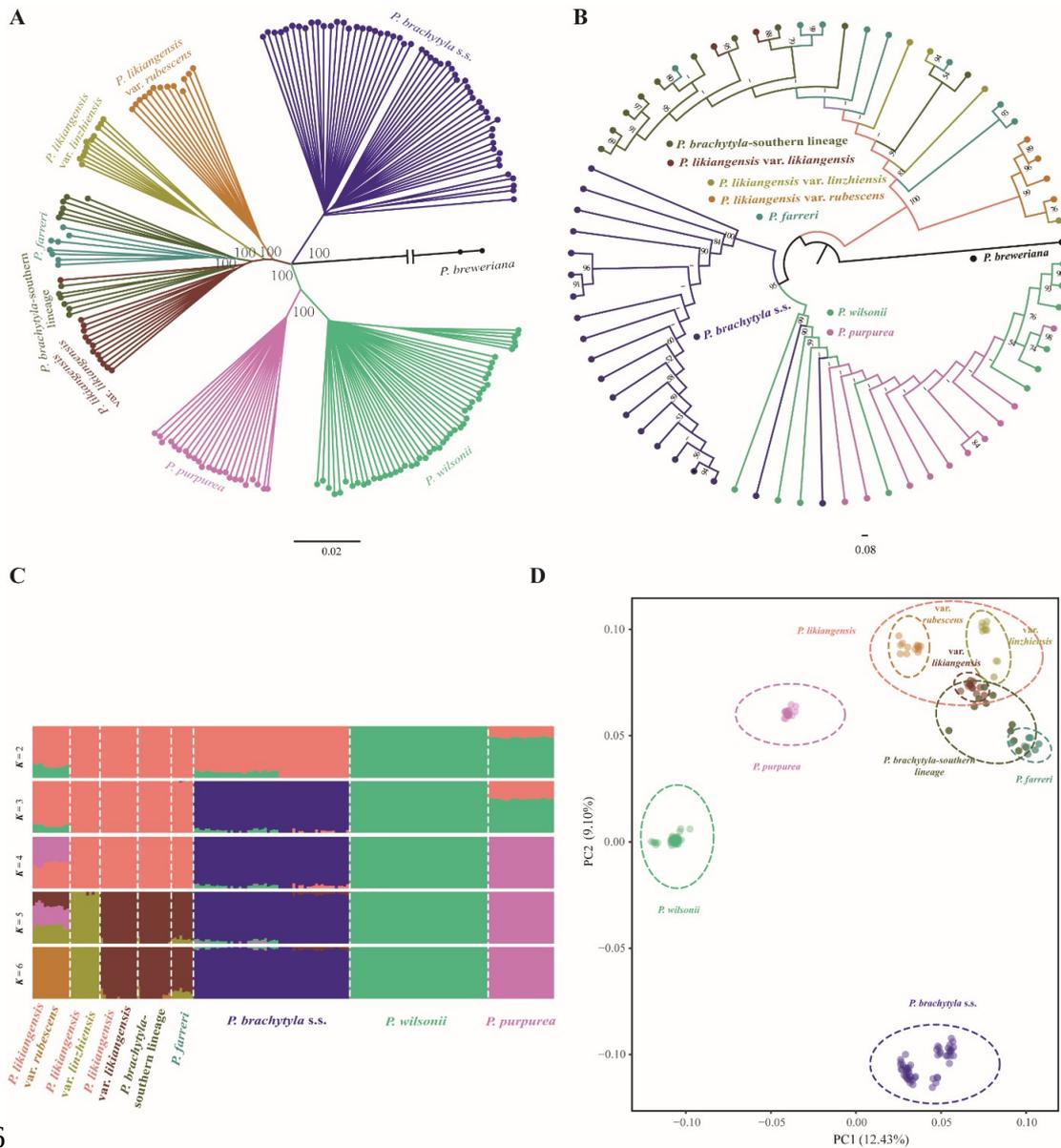
916

917 Figure 3. (A) Neighbor Joining tree based on all 186 transcriptome sequences with *P. breweriana*
918 as the outgroup. (B) Maximum Likelihood (ML) tree based on 69 transcriptome sequences with *P.*
919 *breweriana* as the outgroup. (C) Bar plots indicative of assignment probabilities from
920 ADMIXTURE analysis of 184 transcriptomes from K =2 to K=6. (D) Principal component
921 analysis (PCA) plots showing the first two principal components.
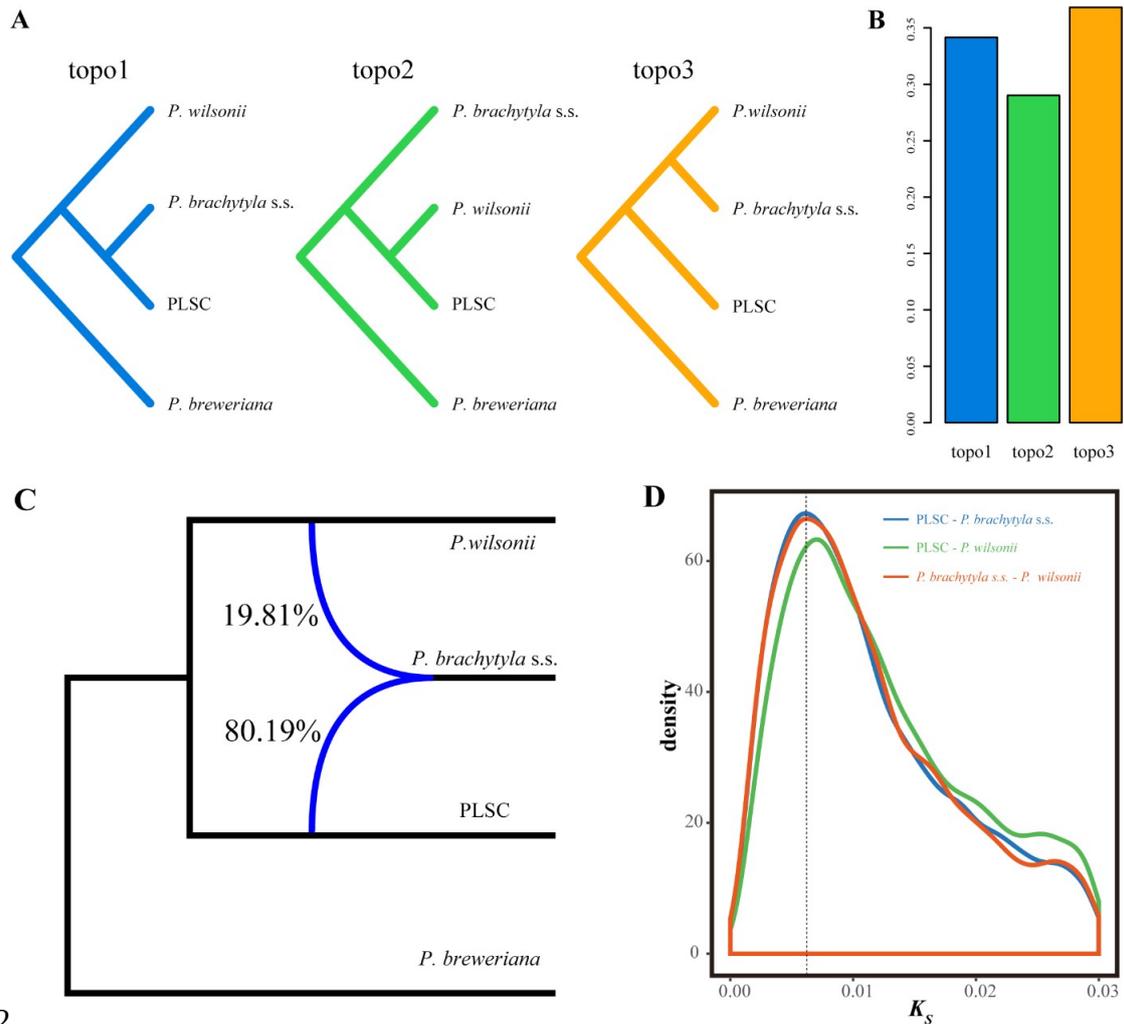
922

923 Figure 4. (A) Relationships between *P. likiangensis* species complex (PLSC), *P. brachytyla* s.s.
924 and *P. wilsonii* (using *P. breweriana* as outgroup) according to ML analysis of 3,305 orthologous
925 gene sequences identified by OrthoMCL as having a ratio of 1:1:1:1. (B) The proportion of each
926 topology is based on the ortholog groups. (C) ML-bootstrap network for 3,305 orthologous gene
927 trees generated by PhyloNet after runs allowing 0, 1 or 2 reticulations. Reticulations are shown in
928 blue with inheritance probabilities. Note: Only trees with branch bootstrap values >70% were
929 analysed. (D) K_S age distributions for one-to-one orthologs between PLSC and *P. brachytyla* s.s.
930 and *P. wilsonii*.

931

932
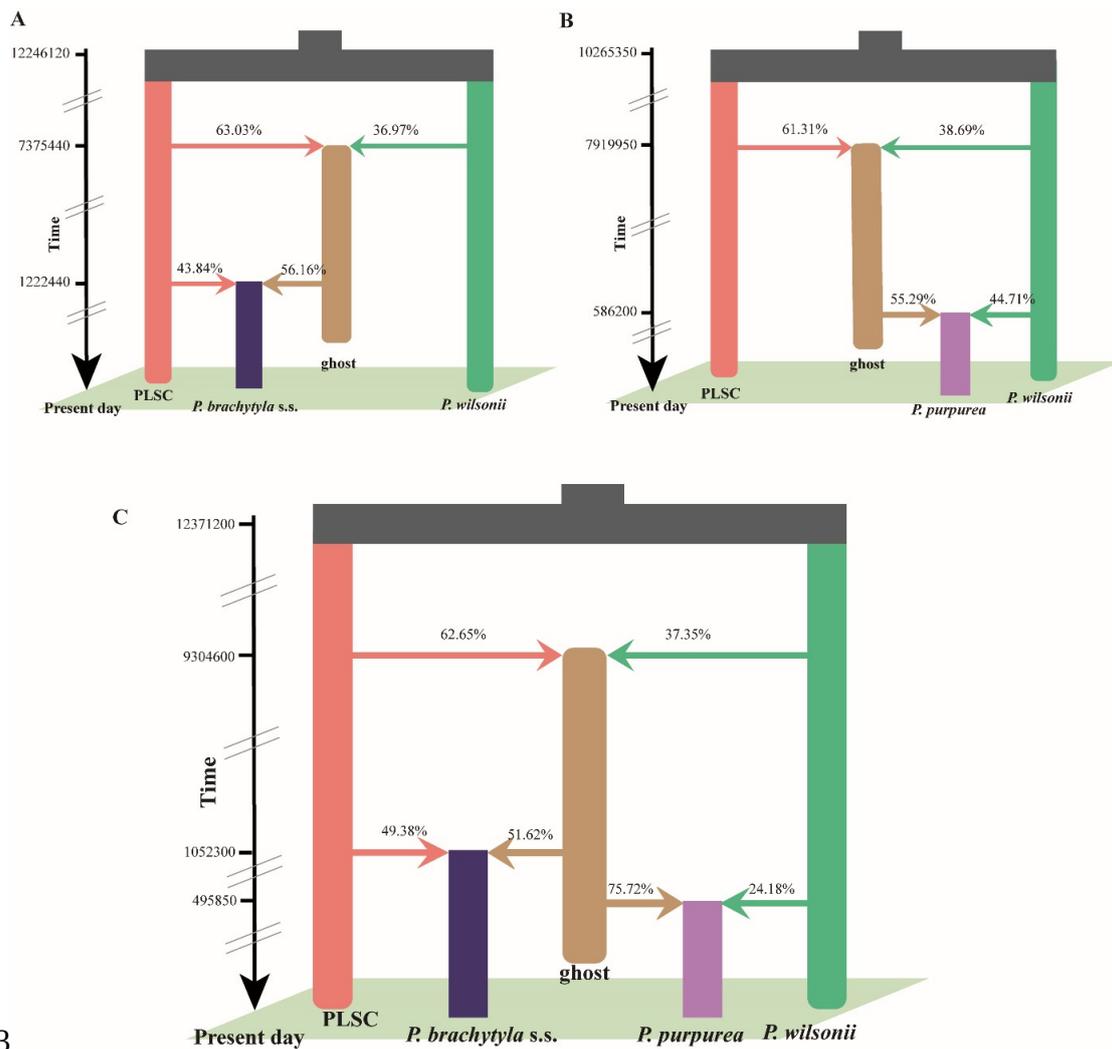
933

934Figure 5. (A) Simplified graphical summary of the best-fitting demographic model inferred by
935fastsimcoal2 for the reduced PLSC, *P. brachytyla* s.s., and *P. wilsonii*. (B) Simplified graphical
936summary of the best-fitting demographic model inferred by fastsimcoal2 for the reduced PLSC, *P.*
937*purpurea*, and *P. wilsonii*. (C) Simplified graphical summary of the best-fitting demographic
938model inferred by fastsimcoal2 for the reduced PLSC, *P. brachytyla* s.s., *P. purpurea* and *P.*
939*wilsonii*. The percentages indicate nuclear genomic compositions contributed from parents to the
940hybrid offspring. The right-hand axis indicates the timescale in years before the present.

941

942

Tables

Table 1 Locations of *Picea* individuals sampled for this study

Species	Latitude	Longitude	Altitude(m)	Collection site	Source	n
P. breweriana						**1**
brew	35.93306	104.15008	1774	Common garden, Gansu	Ru et al.,, 2018	1
P. likiangensis						**30**
MSZ-04	29.99958	100.87139	4178.94	Jianziwan mountain, Sichuan	Ru et al.,, 2018	1
MSZ-05	30.28694	99.519222	4252	Heni, Sichuan	Ru et al.,, 2018	3
MSZ-06	29.72983	98.62975	4026.58	Zongla mountain, Sichuan	Ru et al.,, 2018	3
MSZ-07	29.61128	98.156944	4104.2	Rumei, Tibet	Ru et al.,, 2018	2
MSZ-08	29.68322	97.931917	4122.95	Zuogong, Tibet	Ru et al.,, 2018	2
MSZ-15	29.18475	93.978556	2988.12	Milin, Linzhi, Tibet	Ru et al.,, 2018	1
MSZ-25	29.46339	94.61775	2913.14	Milin, Linzhi, Tibet	Ru et al.,, 2018	1
MSZ-30	29.56961	94.557972	3421.91	Sejila mountain, Linzhi, Tibet	Ru et al.,, 2018	1
MSZ-31	29.67392	94.720028	3663.2	Sejila mountain, Linzhi, Tibet	Ru et al.,, 2018	3
MSZ-33	29.8905	95.523278	2698.53	Bomi, Tibet	Ru et al.,, 2018	1
MSZ-34	29.82383	95.711528	3262.82	Bomi, Tibet	Ru et al.,, 2018	2
MSZ-40	27.93083	99.616472	3511.8	Napahai, Yunnan	Ru et al.,, 2018	1
MSZ-42	27.569	100.02383	3025.85	Pudacuo, Yunnan	Ru et al.,, 2018	1
MSZ-46	27.19836	100.27886	3260.41	Daju mountain, Yunnan	Ru et al.,, 2018	1
MSZ-47	27.13161	100.23303	2947.51	Yulongxue mountain, Yunnan	Ru et al.,, 2018	3
MSZ-48	27.14214	100.2335	3197.45	Yulongxue mountain, Yunnan	Ru et al.,, 2018	2
MSZ-50	27.02508	100.20897	2845.13	Yuhu village, Yunnan	Ru et al.,, 2018	2
P. purpurea						**20**

22-WDL-17	33.07804	102.85164	3568	Hongyuan, Sichuan	The present study	5
ZR_08	34.02296	102.73741	3526	Rierlang mountain, Ruoergai, Sichuan	Ru et al.,, 2018	3
ZR_10	34.27802	103.00059	3556	Niba,Gansu	Ru et al.,, 2018	3
ZR_11	34.45132	102.69788	3132	Duosongben mountain, Gansu	Ru et al.,, 2018	3
ZR_14	35.53111	102.24462	3085	Tongren, Qinghai	Ru et al.,, 2018	3
ZR_25	33.0443	103.72414	3497	Songpan, Sichuan	Ru et al.,, 2018	3
P. wilsonii						**46**
01-WDL-17	33.621871	104.365875	2380	Wuping, Gansu	The present study	1
02-WDL-17	33.617646	104.368253	2360	Wuping, Gansu	The present study	3
07-WDL-17	33.549964	104.336308	2142	Majiazhuang, Gansu	The present study	3
10-WDL-17	33.576466	104.37523	1745	Jue'er Mountain, Gansu	The present study	1
18_WDL-17	32.913075	104.154777	2450	Pingwu, Sichuan	The present study	6
ZR_09	34.04897	103.2207	2390	Diebu, Gansu	Ru et al.,, 2018	5
ZR_15	36.95562	102.46394	2306	Huzhu, Qinghai	Ru et al.,, 2018	6
ZR_16	35.78197	104.05484	2304	Yuzhong, Gansu	Ru et al.,, 2018	7
ZR_24	33.29427	104.47862	2389	Longnan, Gansu	Ru et al.,, 2018	7
ZR_26	34.1588	102.90655	2769	Ruoergai, Sichuan	Ru et al.,, 2018	7
P. brachytyla s.s.						**54**
06-WDL-17	33.552535	104.336832	2120	Heilingou, Gansu	The present study	6
12-WDL-17	32.92212	104.3258	2250	Lianghekou, Gansu	The present study	4
14-WDL-17	32.641667	104.819167	2290	Tangjiahe, Sichuan	The present study	6
15-WDL-17	32.603561	104.662272	1997	Huangjiawan, Sichuan	The present study	5
21-WDL-17	32.952329	104.127202	2536	Pingwu, Sichuan	The present study	5
WDL-17-CQ	31.63194	108.71	2243	Congziping, Chongqing	The present study	4
23-WDL-17	30.440295	102.558357	2323	Dongla grand canyon, Sichuan	The present study	3
26-WDL-17	30.183056	102.475278	2020	Tianquan, Sichuan	The present study	1

WDL-17-HLG	29.580833	102.0175	2780	Hailuogou, Sichuan	The present study	6
28-WDL-17	30.1765	102.47528	2010	Tianquan, Sichuan	The present study	1
W17-01	29.03865	102.98435	2062	Heizhugou, Sichuan	The present study	1
W17-02	28.90082	102.99892	2287	Kejuenapa mountain, Sichuan	The present study	3
W17-03	28.73573	103.05202	2461	Shengliping, Sichuan	The present study	6
YS-1	28.71222	103.21398	2370	Meigu, Sichuan	The present study	3
P. brachytyla-southern lineage						**10**
NPH	27.927778	99.614222	3506	Napahai, Yunnan	Ru et al.,, 2018	1
XGLL	27.8	99.65	3329	Xianggelila, Yunnan	Ru et al.,, 2018	2
XSQ	26.457167	99.313833	2925	Xinshengqiao, Yunnan	Ru et al.,, 2018	4
Zhong	27.45665	99.893133	3128	Xiaozhongdian, Yunnan	Ru et al.,, 2018	3
P. farreri						**5**
ML2019350-1	27.7871	98.51	3070	Bingzhongluo, Yunnan	The present study	1
ML2019350-2	27.7841	98.5101	3060	Bingzhongluo, Yunnan	The present study	1
ML2019351	27.8484	98.4662	2880	Bingzhongluo, Yunnan	The present study	1
ML2019382	28.16275	98.2556	2210	Bingzhongluo, Yunnan	The present study	1
ML2019408	27.7567	98.59003	2750	Bingzhongluo, Yunnan	The present study	1

945

946 **Table 2 BUSCO results for assembly completeness of four spruce transcriptomes**

Species	Classification	BUSCO results
PLSC	Spermatophytes	C: 84.4% [S: 47.1%, D: 37.3%], F: 2.7%, M:12.9%, n: 1440
P. brachytyla s.s.	Spermatophytes	C: 81.2% [S: 71.7%, D: 9.5%], F: 3.0%, M: 15.8%, n: 1440
P. wilsonii	Spermatophytes	C: 86.1% [S: 15.5%, D: 70.6%], F: 2.3%, M:11.6%, n: 1440
P. breweriana	Spermatophytes	C: 85.1% [S: 10.1%, D: 75.0%], F: 2.4%, M:12.5%, n: 1440

947

948 Table 3 Inferred demographic parameters of the best-fitting demographic model in Fig. S3

Parameters	Point estimation	95% CI Lower bound	95% CI Upper bound
$N_{e\text{-PLSC}}$	16614	11301	22342
$N_{e\text{-}P.\ wilsonii}$	32321	21375	36121
$N_{e\text{-}P.\ brachytyla\ s.s.}$	18111	13031	21685
$N_{e\text{-ghost}}$	78107	10442	172368
$m^1_{\text{PLSC->}P.\ wilsonii}$	2.48e-06	1.29e-08	9.83e-04
$m^1_{P.\ wilsonii\text{->PLSC}}$	3.37e-06	2.22e-08	9.25e-04
$m^2_{\text{PLSC->}P.\ wilsonii}$	2.63e-03	6.85e-04	4.84e-02
$m^2_{P.\ wilsonii\text{->PLSC}}$	1.36e-07	6.99e-09	2.58e-02
$m_{\text{PLSC->}P.\ brachytyla\ s.s.}$	4.84e-05	2.65e-05	9.42e-05
$m_{P.\ brachytyla\ s.s.\text{>PLSC}}$	6.09e-05	1.59e-05	1.31e-04
$m_{P.\ wilsonii\text{->}\ P.\ brachytyla\ s.s.}$	7.70e-05	3.22e-05	1.21e-04
$m_{P.\ brachytyla\ s.s.\text{->}P.\ wilsonii}$	1.27e-05	9.32e-07	2.87e-05
$m_{P.\ wilsonii\text{->PLSC}}$	2.94e-05	7.73e-06	2.31e-02
$m_{\text{PLSC->}P.\ wilsonii}$	1.42e-05	2.34e-07	6.82e-05
T_{ADM1}	1222440	1005800	5478200
T_{ADM2}	7375440	5817900	13122100
T_{DIV}	12246120	10157600	16086500

949 $N_{e\text{-PLSC}}$, $N_{e\text{-}P.\ brachytyla\ s.s.}$, $N_{e\text{-}P.\ wilsonii}$, $N_{e\text{-ghost}}$ indicate the effective population sizes of the reduced *P.* 950 *likiangensis* species complex (PLSC), *P. brachytyla* s.s., *P. wilsonii*, the ghost intermediate 951 lineage and ancestral population respectively. $m^1_{\text{PLSC->}P.\ wilsonii}$, $m^1_{P.\ wilsonii\text{->PLSC}}$, $m^2_{\text{PLSC->}P.\ wilsonii}$, $m^2_{P.}$ 952 _wilsonii->PLSC_ indicate migration per generation before and after hybridization between PLSC and *P.* 953 *wilsonii*; $m_{PLSC\text{->ghost}}$: migration per generation from PLSC to the ghost lineage and $m_{\text{ghost->}PLSC}$ 954 migration per generation from ghost lineage to PLSC. $m_{\text{PLSC->}P.\ brachytyla\ s.s.}$, $m_{P.\ brachytyla\ s.s.\text{>PLSC}}$, $m_{P.\ wilsonii\text{-}}$ 955 _> P. brachytyla s.s._ and $m_{P.\ brachytyla\ s.s.\text{->}P.\ wilsonii}$ indicate, respectively, migration per generation between *P.* 956 *brachytyla* s.s. and PLSC or *P. wilsonii* in both directions. TADM1 indicates time (years) of 957 backcrossing of the ghost lineage to *P. wilsonii* that gave rise to *P. brachytyla* s.s., while TADM2 958 indicates time (years) of formation of the ghost lineage between PLSC and *P. wilsonii*. TDIV 959 indicates the estimated divergence time (years) between PLSC and *P. wilsonii* obtained from 960 fastsimcoal2.