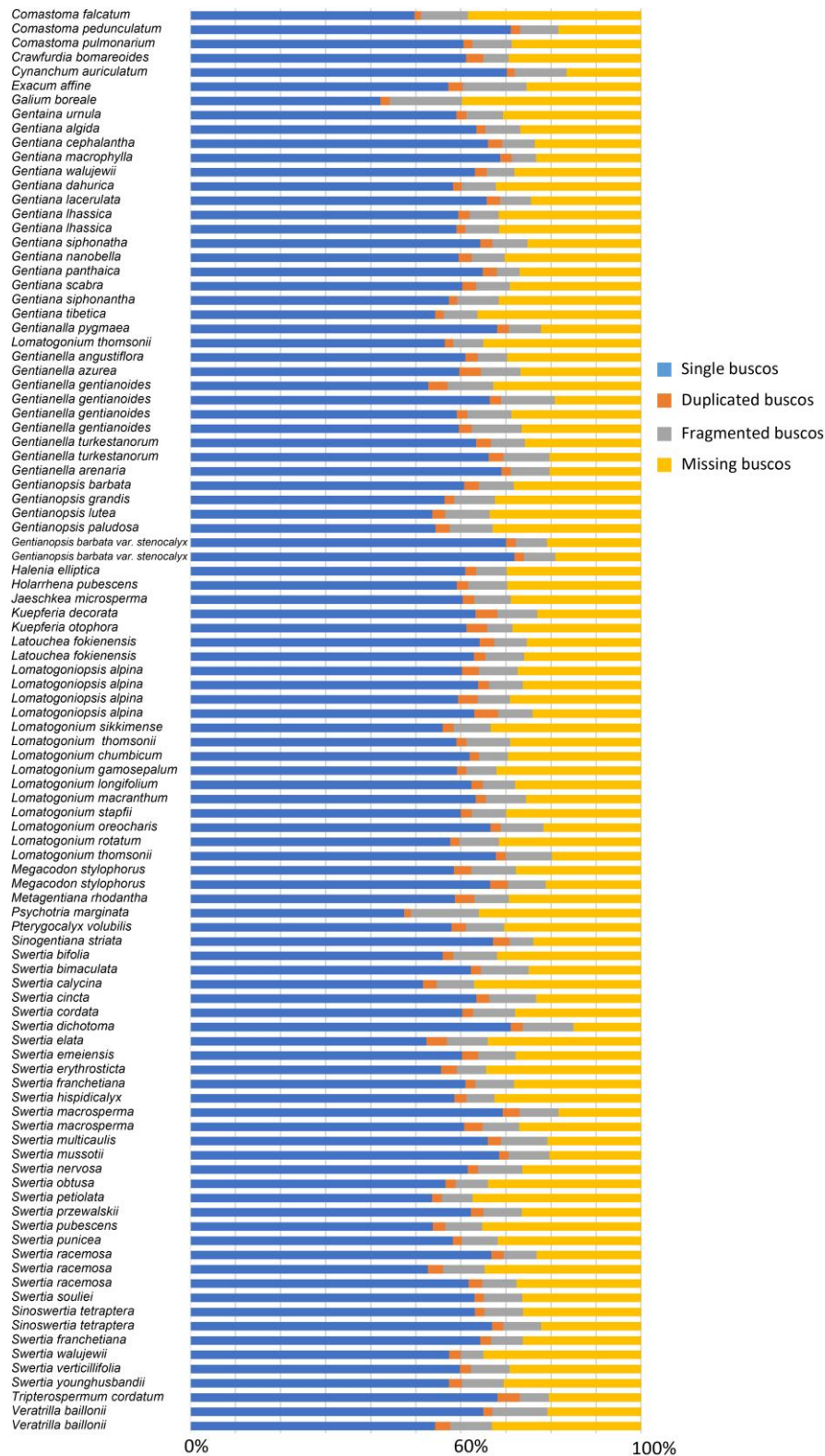


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

**Phylotranscriptomics of Swertiinae (Gentianaceae)
reveals that key floral traits are not phylogenetically
correlated**

Chunlin Chen^{1†}, Brad R. Ruhfel^{2†}, Jialiang Li¹, Zefu Wang¹, Lushui Zhang¹,
Lei Zhang¹, Xingxing Mao¹, Ji Wang¹, Dashan He¹, Yue Luo¹, Qianjun Hu¹,
Yuanwen Duan⁴, Xiaoting Xu^{1*}, Zhenxiang Xi^{1*}, Jianquan Liu^{1*}

SUPPORTING INFORMATION



19

20 **Figure S1.** BUSCO assessment of the transcriptome assemblies

21 For each sample, the bars from left to right are percentages of single
 22 BUSCOs, duplicated BUSCOs, fragmented BUSCOs and missing BUSCOs
 23 respectively.

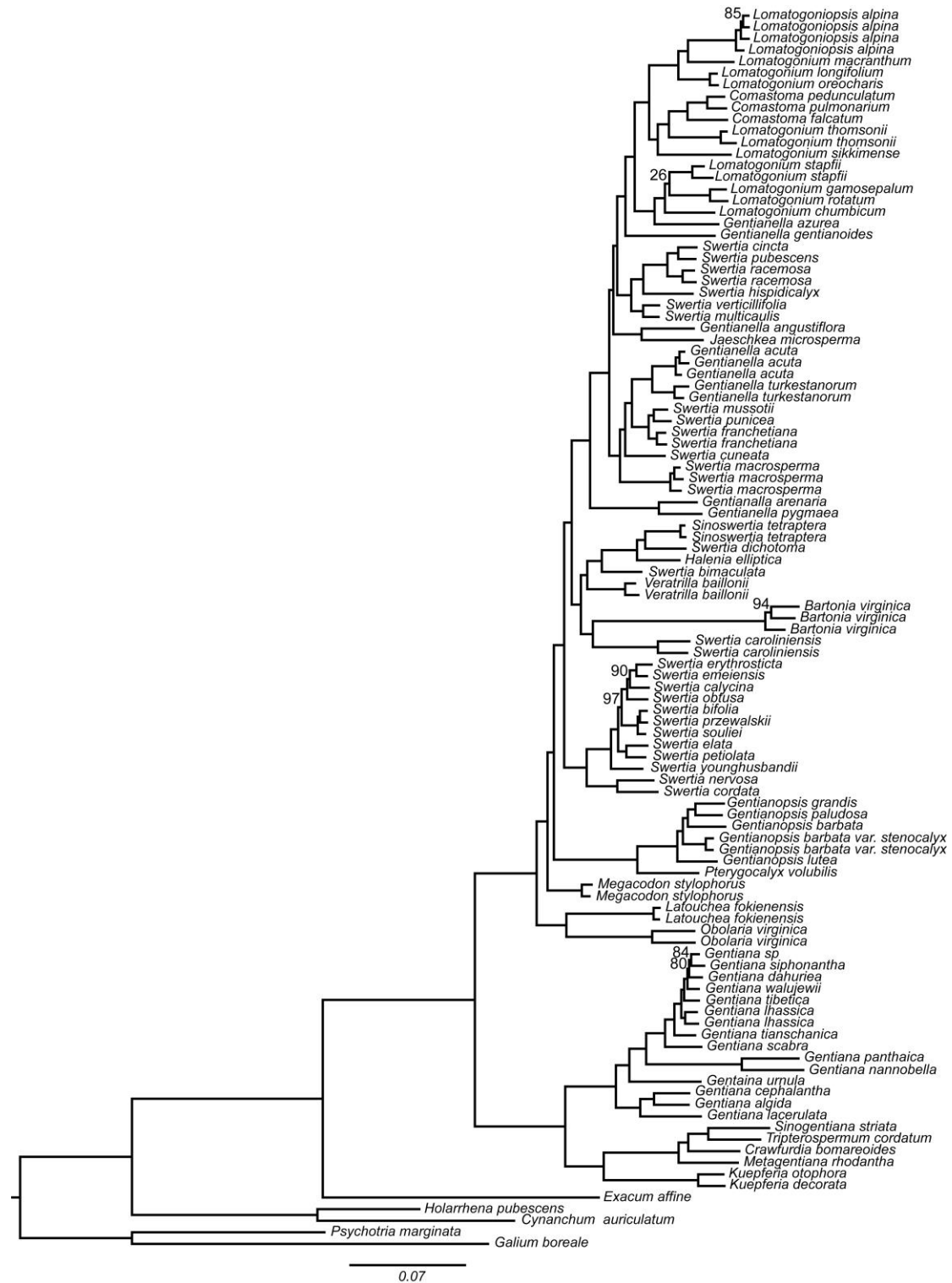


Figure S2. Species tree inferred using RAxML based on the concatenated matrix of 1280 single copy nuclear genes identified by MO pipeline. Bootstrap percentages (BP) are indicated above the branches. Only BPs less than 100 were shown.



Figure S3. Species tree inferred using Astral based on 2045 single copy nuclear genes identified by *RT* pipeline. Quartet support values are indicated above the branches. Only Quartet support values less than 1 were shown.



35

36 **Figure S4.** Species tree inferred using RAXML based on the concatenated37 matrix of 2045 single copy nuclear genes identified by *RT* pipeline

38 Bootstrap percentages (BP) are indicated above the branches. Only BPs less

39 than 100 were shown.

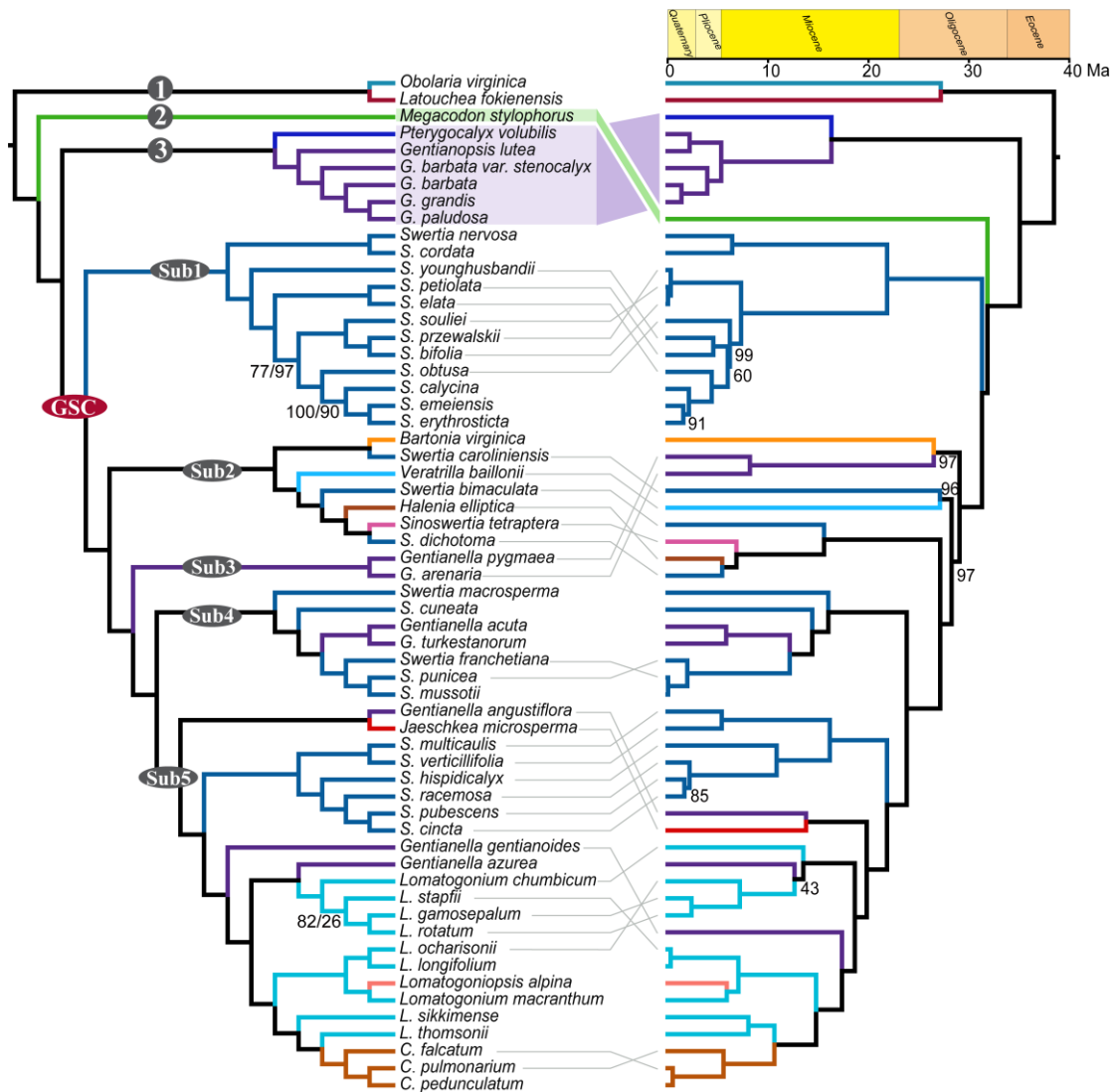


Figure S5. Phylogeny of the subtribe Swertiinae based on 1280 single copy nuclear genes (left, cladogram) and plastome data (right, chronogram)

The GSC clade denotes the *Gentinella*-*Swertia* complex clade. Different branch colors represent the different genera. Bootstrap percentages (BPs) are indicated below the branch, and only values less than 100 are shown. The grey lines and parallelogram between the two trees indicate the corresponding placement of taxa in the plastome phylogeny. For nuclear genes, BPs from concatenation and coalescent-based analyses are shown before and after the slash, respectively.

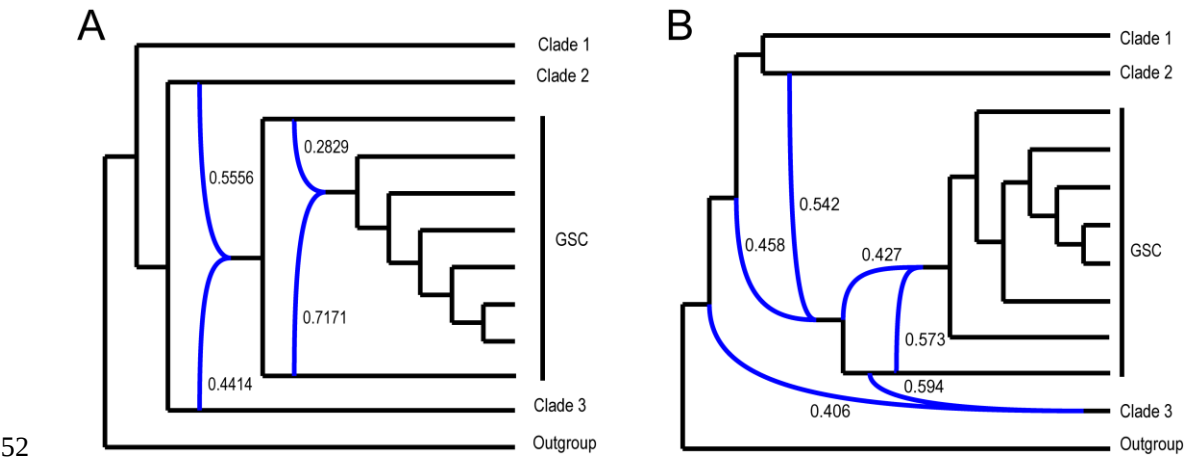
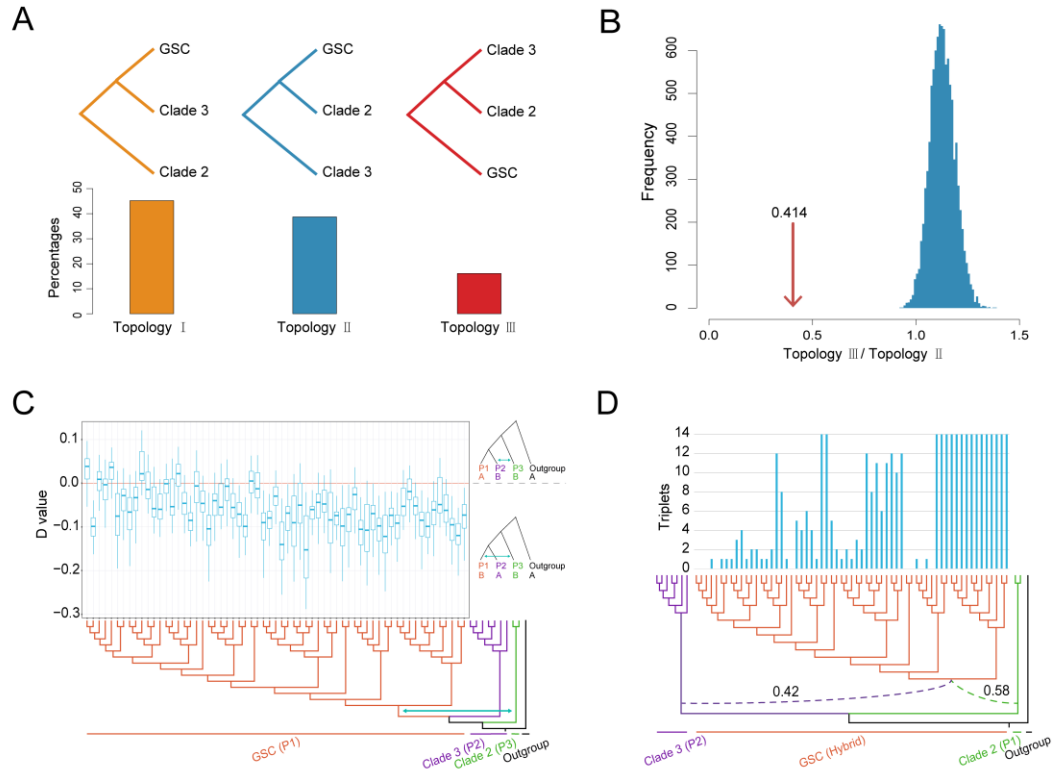


Figure S6. Networks (inferred by PhyloNet) with the highest respective log probability when two **(A)** and three **(B)** reticulations were present

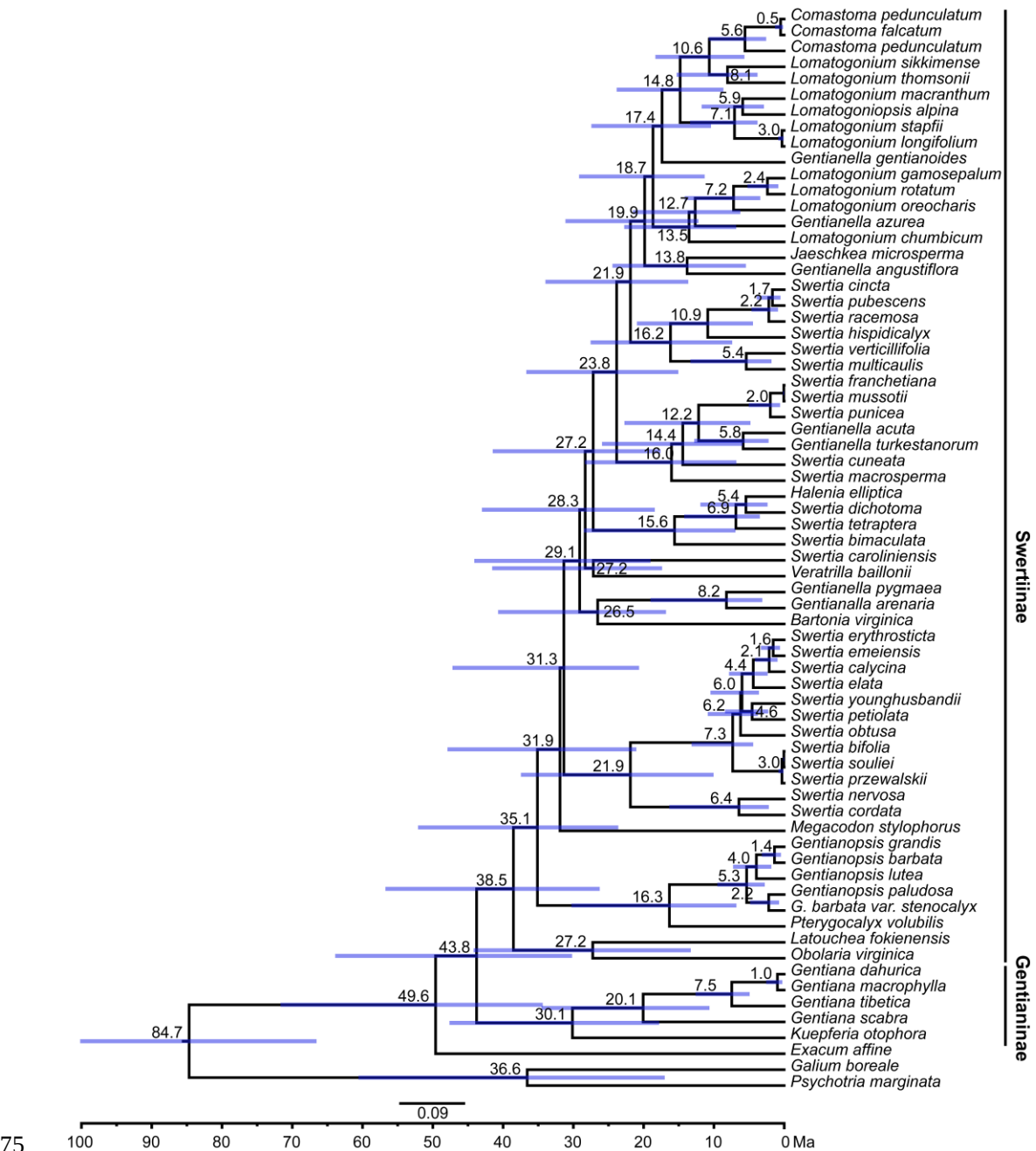
Blue lines denote the inferred parental lineage of the *Gentianella-Swertia* complex (GSC) clade and other hybrid lineage in the GSC clade, values to the right of each line represent the inheritance probability.



60 **Figure S7. Hybridization test analysis**

61 **(A)** Proportion of gene trees supporting different relationships between
 62 *Gentianella-Swertia* complex (GSC) clade, clade 2 and clade 3. **(B)**
 63 Simulations under solely ILS scenario. The red arrow indicates the observed
 64 ratio (Topology III/Topology II) from 1025 ortholog groups (see in A). The
 65 histogram are the ratios obtained in 10,000 simulated datasets. Values on the
 66 x-axis denote the ratio of the number of gene trees support topology III and
 67 topology II **(C)** Results of D statistics from the 882 triplets with boxplot of D
 68 value for each taxon from GSC on the top left, illustration of gene flow with the
 69 positive and negative D value on the top right, and schematic genealogy of
 70 the taxa used for D statistics. **(D)** Results of HyDe analysis from the 882
 71 triplets. Distributions of the number triplets (out of 14 triplets for each taxon)
 72 tested significant for each species are on the top. The schematic genealogy of
 73 the taxa used in the HyDe analysis was on the bottom. Values on the dotted

74 curve denote the inferred ancestry from the PhyloNet analysis.



75
76 **Figure S8.** The chronogram obtained using MCMCTree based on the
77 concatenated plastome gene matrix
78 Values next to each node indicate it's ages. Blue bars indicate the 95%
79 highest posterior density (HPD) credibility intervals for node ages.

80

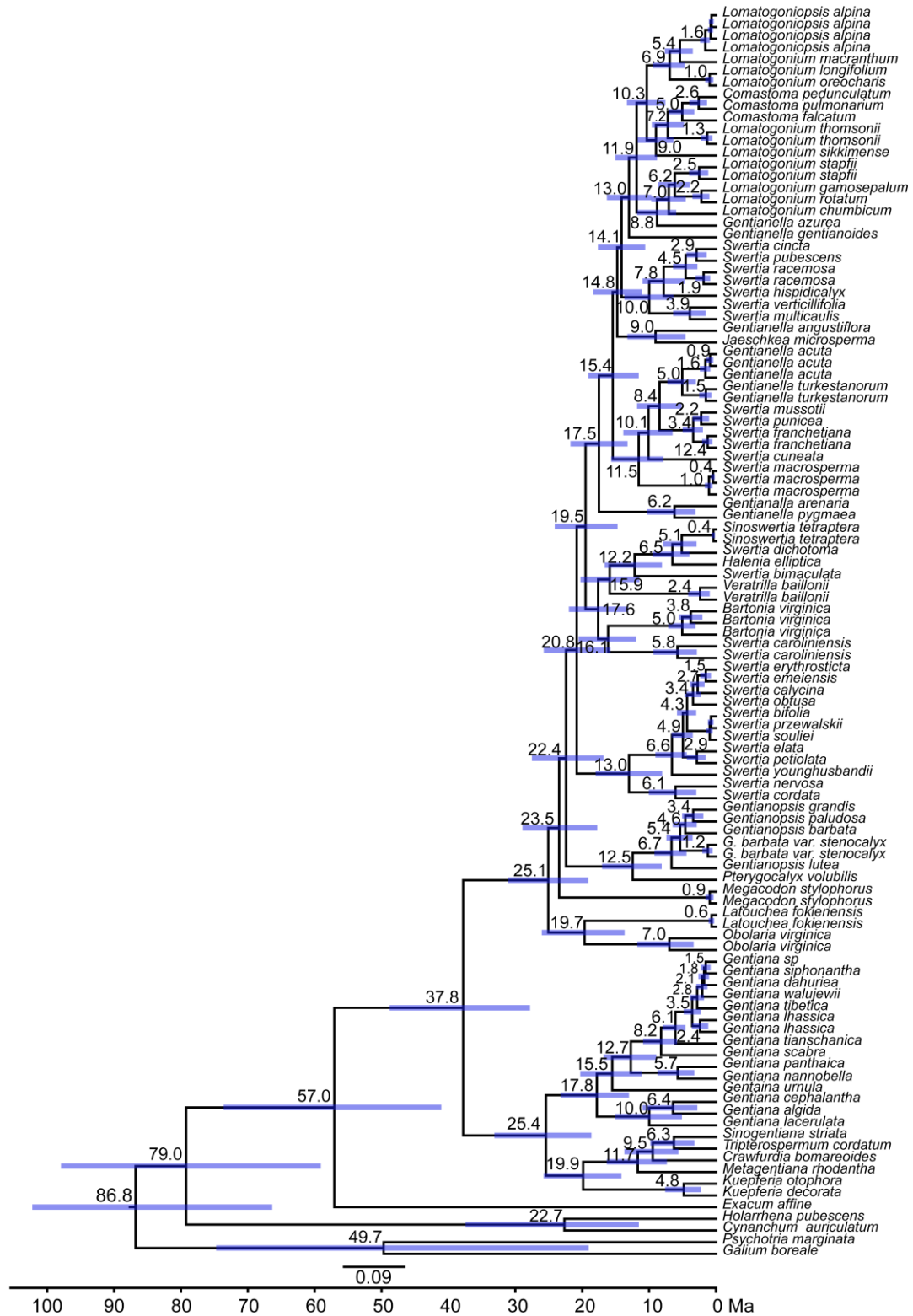
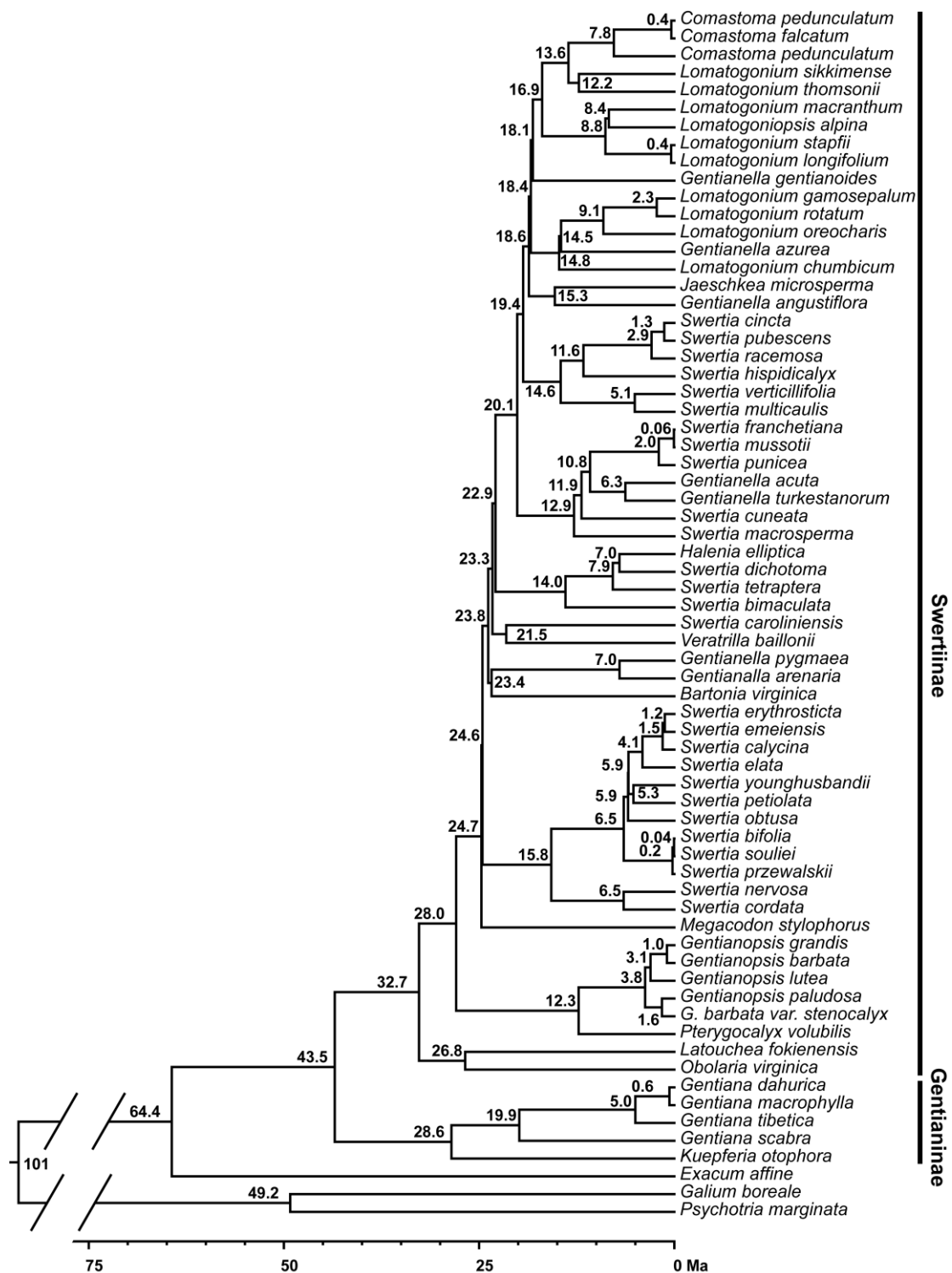


Figure S9. The chronogram obtained using MCMCTree based on the concatenated nuclear gene matrix

Values next to each node indicate it's ages. Blue bars indicate the 95% highest posterior density (HPD) credibility intervals for node



87
88 **Figure S10.** The chronogram obtained using TreePL based on the branch
89 length estimated by RAxML from concatenated plastome gene matrix
90 Values next to each node indicate.

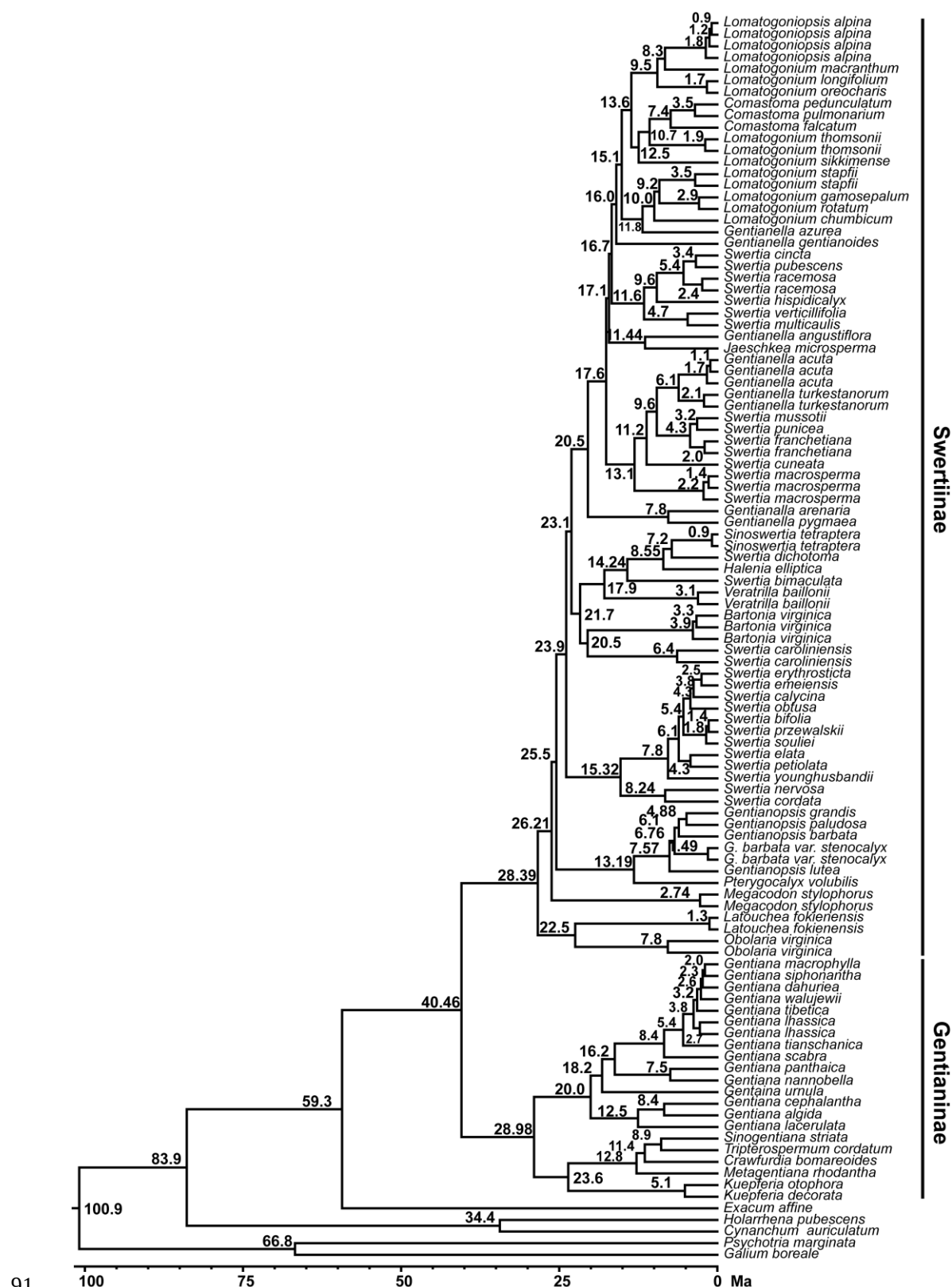


Fig. S11. The chronogram obtained using TreePL based on the branch length estimated by RAxML from concatenated nuclear gene matrix. Values next to each node indicate its ages.

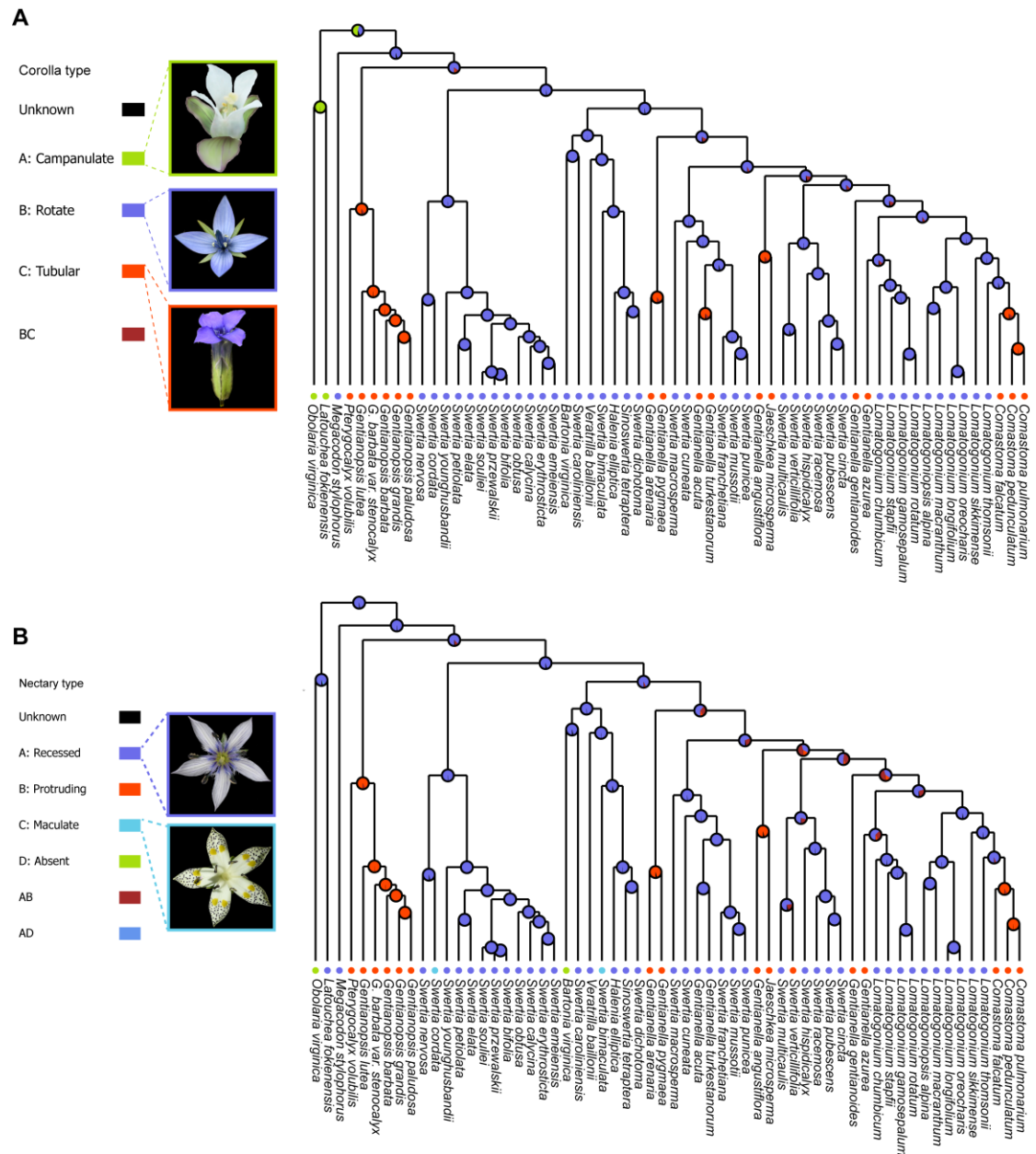


Fig S12. Ancestral-state reconstructions for the two key floral traits of subtribe Swertiinae for generic circumscriptions based on the nuclear phylogeny

(A) Corolla type. The three photos from top to bottom are *Bartonia virginica*, *Gentianopsis paludosa* and *Lomatogonium brachyantherum* with campanulate, rotate and tubular corollas, respectively. **(B)** Nectary type. The two photos from top to bottom are *Swertia franchetiana* and *S. bimaculata* with recessed and maculate nectaries, respectively.

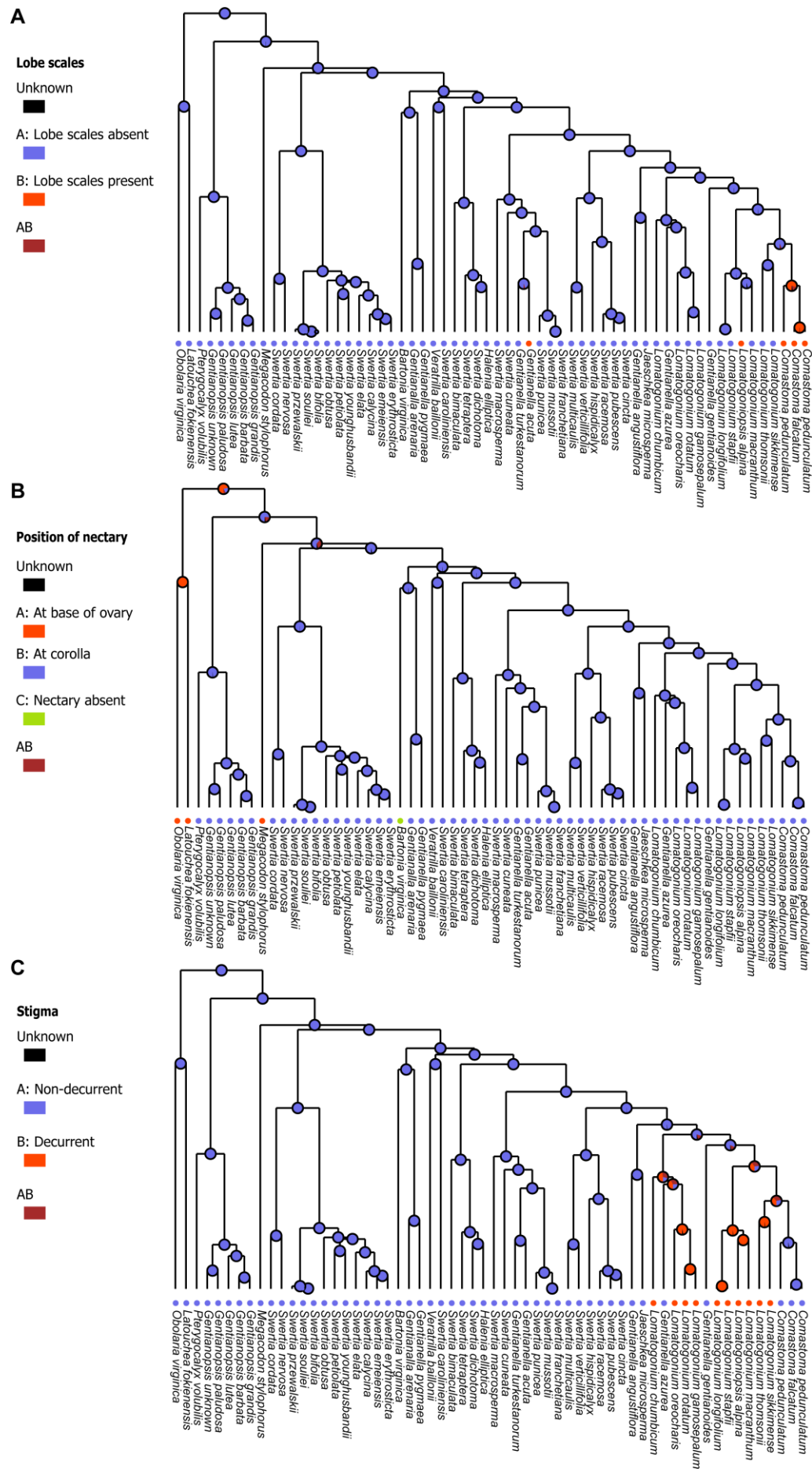


Fig. S13. Ancestral-state reconstruction for the three key floral traits of subtribe Swertiinae on the plastome phylogeny
(A) Lobe scales, absent versus present. **(B)** Position of nectary, at base of ovary, at corolla and nectary absent. **(C)** Stigma type, decurrent versus non-decurrent.

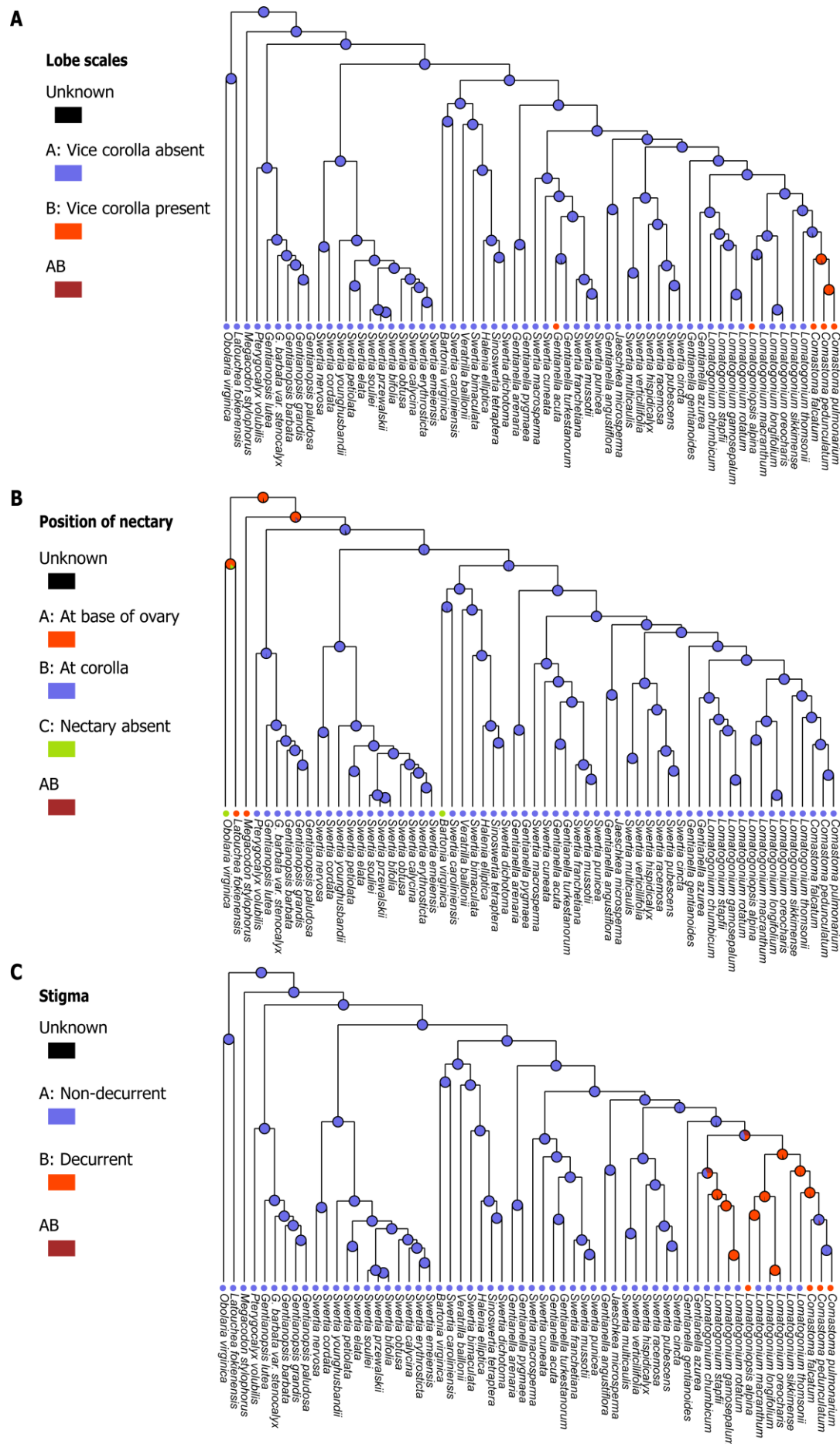


Figure S14. Ancestral-state reconstruction for the three key floral traits of subtribe Swertiinae on the nuclear phylogeny
(A) Lobe scales, absent versus present. **(B)** Position of nectary, at base of ovary, at corolla and nectary absent. **(C)** Stigma type, decurrent versus non-decurrent.

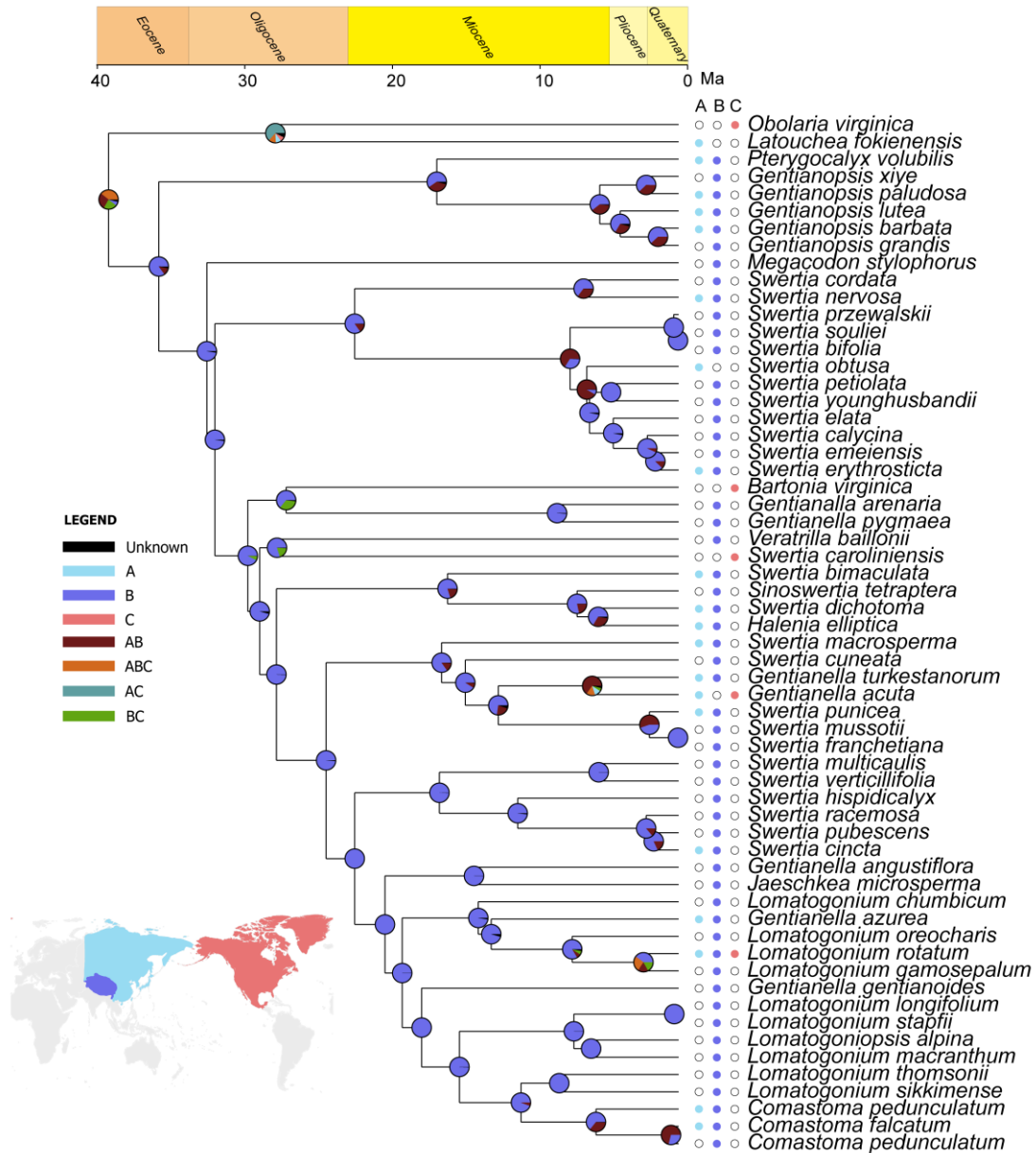


Figure S15. Ancestral distribution reconstructions based on the plastome phylogeny from DEC model inferred by Biogeobears implemented in RASP

The three defined biogeographic regions are A, other parts of Asia; B, Qinghai-Tibet Plateau and C, North America.

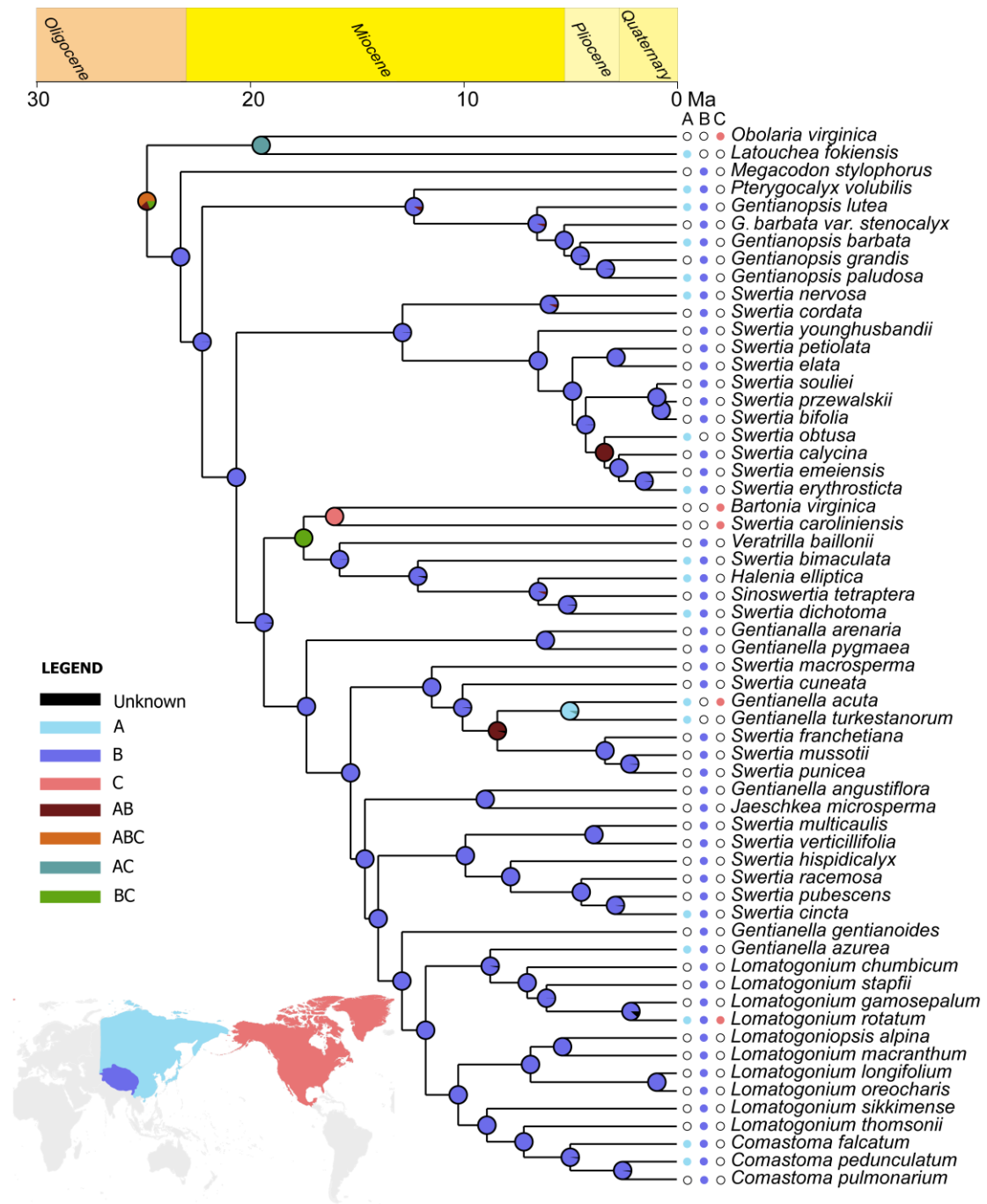


Figure S16. Ancestral distribution reconstructions on the nuclear phylogeny from DIVALIKE model inferred by Biogeobears implemented in RASP

The three defined biogeographic regions are A, other parts of Asia; B, Qinghai-Tibet Plateau and C, North America.

Table S1. Voucher information/accession ID, sequencing platforms and transcripts count of each taxon included

Family	Taxon name	Voucher ID/ accession ID	Sequencing platform/source	Library source	Assembler	CongithN50	Transcript count
Gentianaceae	<i>Bartonia virginica</i>	Bartonia.146.1A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Gentianaceae	<i>Bartonia virginica</i>	Bartonia.146.2A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Gentianaceae	<i>Bartonia virginica</i>	Bartonia.146.3A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Gentianaceae	<i>Comastoma falcatum</i>	QTP041	Illumina Hiseq X Ten	Transcriptomic	Trinity	1471	33835
Gentianaceae	<i>Comastoma pedunculatum</i>	liu19-29A	BGI DNBSEQ™	Transcriptomic	Trinity	1969	24524
Gentianaceae	<i>Comastoma pulmonarium</i>	QTP014	Illumina Hiseq X Ten	Transcriptomic	Trinity	1835	29138
Gentianaceae	<i>Crawfurdia bomareoides</i>	RC47	Illumina Novaseq 6000	Transcriptomic	Trinity	1611	37132
Gentianaceae	<i>Cynanchum auriculatum</i>	SRX6644067	NCBI SRA	Transcriptomic	Trinity	1961	37447
Apocynaceae	<i>Exacum affine</i>	ERR2040542	NCBI SRA	Transcriptomic	Trinity	1403	32633
Rubiaceae	<i>Galium boreale</i>	ERR2040548	NCBI SRA	Transcriptomic	Trinity	1233	23170
Gentianaceae	<i>Gentiana urnula</i>	SRX6591655	NCBI SRA	Transcriptomic	Trinity	1813	30163
Gentianaceae	<i>Gentiana algida</i>	SRX6490704	NCBI SRA	Transcriptomic	Trinity	1809	27415
Gentianaceae	<i>Gentiana cephalantha</i>	SRX6591658	NCBI SRA	Transcriptomic	Trinity	1669	44078
Gentianaceae	<i>Gentiana dahurica</i>	SRX6591649	NCBI SRA	Transcriptomic	Trinity	1832	26491
Gentianaceae	<i>Gentiana lacerulata</i>	QTPXJ_19_056	Illumina Novaseq 6000	Transcriptomic	Trinity	1870	29657
Gentianaceae	<i>Gentiana lhassica</i>	SRX6610690	NCBI SRA	Transcriptomic	Trinity	1715	27860
Gentianaceae	<i>Gentiana lhassica</i>	QTP085	Illumina Hiseq X Ten	Transcriptomic	Trinity	1817	31165
Gentianaceae	<i>Gentiana macrophylla</i>	liu19-28B	BGI DNBSEQ™	Transcriptomic	Trinity	1650	33346
Gentianaceae	<i>Gentiana nanobella</i>	SRX6610697	NCBI SRA	Transcriptomic	Trinity	1614	37395
Gentianaceae	<i>Gentiana panthaica</i>	ZMC321	Illumina Novaseq 6000	Transcriptomic	Trinity	1653	41010
Gentianacea	<i>Gentiana scabra</i>	SXCZ02	Illumina Hiseq X Ten	Transcriptomic	Trinity	1575	34841

e							
Gentianaceae	<i>Gentiana siphonantha</i>	QTP059	Illumina Hiseq X Ten	Transcriptomic	Trinity	1697	29990
Gentianaceae	<i>Gentiana tianschanica</i>	QTPXJ_19_117	Illumina Novaseq 6000	Transcriptomic	Trinity	1723	35498
Gentianaceae	<i>Gentiana tibetica</i>	QTPXJ-19-046	Illumina Novaseq 6000	Transcriptomic	Trinity	1713	25396
Gentianaceae	<i>Gentiana walujewii</i>	QTPXJ_19_125	Illumina Novaseq 6000	Transcriptomic	Trinity	1892	34038
Gentianaceae	<i>Gentianella acuta</i>	20_2_2_1	BGI DNBSEQ™	Transcriptomic	Trinity	2058	42386
Gentianaceae	<i>Gentianella acuta</i>	CCL1	Illumina Novaseq 6000	Transcriptomic	Trinity	1733	34770
Gentianaceae	<i>Gentianella acuta</i>	CCL3	Illumina Novaseq 6000	Transcriptomic	Trinity	1723	39188
Gentianaceae	<i>Gentianella angustiflora</i>	QTPXJ_19_082	Illumina Novaseq 6000	Transcriptomic	Trinity	1953	24355
Gentianaceae	<i>Gentianella arenaria</i>	liu19-19A	BGI DNBSEQ™	Transcriptomic	Trinity	2015	25133
Gentianaceae	<i>Gentianella azurea</i>	QTP038	Illumina Hiseq X Ten	Transcriptomic	Trinity	1601	39552
Gentianaceae	<i>Gentianella gentianoides</i>	RC57	Illumina Novaseq 6000	Transcriptomic	Trinity	1694	36010
Gentianaceae	<i>Gentianella pygmaea</i>	liu19-21A	BGI DNBSEQ™	Transcriptomic	Trinity	1939	24753
Gentianaceae	<i>Gentianella turkestanorum</i>	QTPXJ_19_112	Illumina Novaseq 6000	Transcriptomic	Trinity	1993	30214
Gentianaceae	<i>Gentianella turkestanorum</i>	QTPXJ_19_114	Illumina Novaseq 6000	Transcriptomic	Trinity	1809	34599
Gentianaceae	<i>Gentianopsis barbata</i>	RC58	Illumina Novaseq 6000	Transcriptomic	Trinity	1703	35687
Gentianaceae	<i>Gentianopsis grandis</i>	RC14	Illumina Novaseq 6000	Transcriptomic	Trinity	1403	43877
Gentianaceae	<i>Gentianopsis lutea</i>	ZMC296	Illumina Novaseq 6000	Transcriptomic	Trinity	1686	31266
Gentianaceae	<i>Gentianopsis paludosa</i>	XL0819-3	Illumina Novaseq 6000	Transcriptomic	Trinity	1663	32584
Gentianaceae	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	liu19-48A	BGI DNBSEQ™	Transcriptomic	Trinity	1984	28504
Gentianaceae	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	liu19-68A	BGI DNBSEQ™	Transcriptomic	Trinity	1986	28488
Gentianaceae	<i>Halenia elliptica</i>	QTP032	Illumina Hiseq X Ten	Transcriptomic	Trinity	1889	20548
Apocynaceae	<i>Holarrhena pubescens</i>	ERR2040540	NCBI SRA	Transcriptomic	Trinity	1713	20817

Gentianaceae	<i>Jaeschkea microsperma</i>	QTPXJ_19_093	Illumina Novaseq 6000	Transcriptomic	Trinity	1829	33166
Gentianaceae	<i>Kuepferia decorata</i>	ZMC318	Illumina Novaseq 6000	Transcriptomic	Trinity	1704	46026
Gentianaceae	<i>Kuepferia otophora</i>	ZMC271	Illumina Novaseq 6000	Transcriptomic	Trinity	1628	39313
Gentianaceae	<i>Latouchea fokienensis</i>	MXXC16	Illumina Hiseq X Ten	Transcriptomic	Trinity	1786	36679
Gentianaceae	<i>Latouchea fokienensis</i>	MXXC50	Illumina Hiseq X Ten	Transcriptomic	Trinity	1783	37620
Gentianaceae	<i>Lomatogoniopsis alpina</i>	CCL1801_1	Illumina Novaseq 6000	Transcriptomic	Trinity	1667	52572
Gentianaceae	<i>Lomatogoniopsis alpina</i>	CCL1801_2	Illumina Novaseq 6000	Transcriptomic	Trinity	1692	33029
Gentianaceae	<i>Lomatogoniopsis alpina</i>	CCL1801_3	Illumina Novaseq 6000	Transcriptomic	Trinity	1675	41528
Gentianaceae	<i>Lomatogoniopsis alpina</i>	QTPXJ-19-001	Illumina Novaseq 6000	Transcriptomic	Trinity	1954	24350
Gentianaceae	<i>Lomatogoniopsis ovatifolia</i>	QTPXJ_19_067	Illumina Novaseq 6000	Transcriptomic	Trinity	2036	23171
Gentianaceae	<i>Lomatogonium thomsonii</i>	QTPXJ_19_106	Illumina Novaseq 6000	Transcriptomic	Trinity	2020	29528
Gentianaceae	<i>Lomatogonium chumbicum</i>	QTPXJ_19_084	Illumina Novaseq 6000	Transcriptomic	Trinity	2015	19382
Gentianaceae	<i>Lomatogonium gamosepalum</i>	QTPXI-19-025	Illumina Novaseq 6000	Transcriptomic	Trinity	1914	24729
Gentianaceae	<i>Lomatogonium longifolium</i>	ZMC230	Illumina Novaseq 6000	Transcriptomic	Trinity	1752	33449
Gentianaceae	<i>Lomatogonium macranthum</i>	QTP086	Illumina Hiseq X Ten	Transcriptomic	Trinity	1733	29277
Gentianaceae	<i>Lomatogonium stapfii</i>	QTPXJ-19-033	Illumina Novaseq 6000	Transcriptomic	Trinity	1881	23318
Gentianaceae	<i>Lomatogonium stapfii</i>	RC24	Illumina Novaseq 6000	Transcriptomic	Trinity	2012	37477
Gentianaceae	<i>Lomatogonium rotatum</i>	QTP028	Illumina Hiseq X Ten	Transcriptomic	Trinity	1837	26482
Gentianaceae	<i>Lomatogonium thomsonii</i>	QTP070	Illumina Hiseq X Ten	Transcriptomic	Trinity	2070	20097
Gentianaceae	<i>Lomatogonium oreocharis</i>	RC23	Illumina Novaseq 6000	Transcriptomic	Trinity	1730	46520
Gentianaceae	<i>Megacodon stylophorus</i>	SRR11743765	NCBI SRA	Transcriptomic	Trinity	2000	36533
Gentianaceae	<i>Megacodon stylophorus</i>	ZMC270	Illumina Novaseq 6000	Transcriptomic	Trinity	1798	38440
Gentianaceae	<i>Metagentiana rhodantha</i>	ZMC136	Illumina Novaseq 6000	Transcriptomic	Trinity	1718	34127
Gentianaceae	<i>Obolaria virginica</i>	Obolaria1A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Gentianaceae	<i>Obolaria virginica</i>	Obolaria2A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Rubiaceae	<i>Psychotria marginata</i>	ERR2040549	NCBI SRA	Transcriptomic	Trinity	1315	28034

Gentianaceae	<i>Pterygocalyx volubilis</i>	ZMC132	Illumina Novaseq 6000	Transcriptomic	Trinity	1571	40773
Gentianaceae	<i>Sinogentiana striata</i>	QTP007	Illumina Hiseq X Ten	Transcriptomic	Trinity	1739	32757
Gentianaceae	<i>Sinoswertia tetraptera</i>	liu19-10A	BGI DNBSEQ™	Transcriptomic	Trinity	1986	22221
Gentianaceae	<i>Sinoswertia tetraptera</i>	QTPXJ-19-003	Illumina Novaseq 6000	Transcriptomic	Trinity	1999	24017
Gentianaceae	<i>Swertia bifolia</i>	QTP120	Illumina Hiseq X Ten	Transcriptomic	Trinity	1786	26542
Gentianaceae	<i>Swertia bimaculata</i>	RC37	Illumina Novaseq 6000	Transcriptomic	Trinity	1822	38657
Gentianaceae	<i>Swertia calycina</i>	ZMC178	Illumina Novaseq 6000	Transcriptomic	Trinity	1575	36668
Gentianaceae	<i>Swertia caroliniensis</i>	Frasera.147A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Gentianaceae	<i>Swertia caroliniensis</i>	Frasera.148A	BGI DNBSEQ™	Whole genomic	ATRAM	—	—
Gentianaceae	<i>Swertia cincta</i>	RC42	Illumina Novaseq 6000	Transcriptomic	Trinity	1655	44816
Gentianaceae	<i>Swertia cordata</i>	RC56	Illumina Novaseq 6000	Transcriptomic	Trinity	1684	31697
Gentianaceae	<i>Swertia dichotoma</i>	20_1_2	BGI DNBSEQ™	Transcriptomic	Trinity	1901	45524
Gentianaceae	<i>Swertia elata</i>	ZMC123	Illumina Novaseq 6000	Transcriptomic	Trinity	1658	38359
Gentianaceae	<i>Swertia emeiensis</i>	XL0819-2	Illumina Novaseq 6000	Transcriptomic	Trinity	1571	43262
Gentianaceae	<i>Swertia erythrosticta</i>	XL0819-1	Illumina Novaseq 6000	Transcriptomic	Trinity	1556	38494
Gentianaceae	<i>Swertia franchetiana</i>	QTP036	Illumina Hiseq X Ten	Transcriptomic	Trinity	1805	28275
Gentianaceae	<i>Swertia franchetiana</i>	QTPXJ-19-009	Illumina Novaseq 6000	Transcriptomic	Trinity	1945	27099
Gentianaceae	<i>Swertia hispidicalyx</i>	QTP080	Illumina Hiseq X Ten	Transcriptomic	Trinity	2154	19898
Gentianaceae	<i>Swertia macrosperma</i>	ZMC204	Illumina Novaseq 6000	Transcriptomic	Trinity	1799	49104
Gentianaceae	<i>Swertia macrosperma</i>	ZMC231	Illumina Novaseq 6000	Transcriptomic	Trinity	1836	33675
Gentianaceae	<i>Swertia multicaulis</i>	QTPXJ-19-071	Illumina Novaseq 6000	Transcriptomic	Trinity	2027	33964
Gentianaceae	<i>Swertia mussotii</i>	liu19-47A	BGI DNBSEQ™	Transcriptomic	Trinity	2166	23431
Gentianaceae	<i>Swertia nervosa</i>	RC36	Illumina Novaseq 6000	Transcriptomic	Trinity	1826	36227
Gentianaceae	<i>Swertia obtusa</i>	QTPXJ_19_126	Illumina Novaseq 6000	Transcriptomic	Trinity	1863	22917
Gentianaceae	<i>Swertia petiolata</i>	QTPXJ-19-044	Illumina Novaseq 6000	Transcriptomic	Trinity	1892	20432

Gentianaceae	<i>Swertia przewalskii</i>	QTP042	Illumina Hiseq X Ten	Transcriptomic	Trinity	1769	31160
Gentianaceae	<i>Swertia pubescens</i>	ZMC180	Illumina Novaseq 6000	Transcriptomic	Trinity	1664	32214
Gentianaceae	<i>Swertia punicea</i>	ZMC214	Illumina Novaseq 6000	Transcriptomic	Trinity	1821	32683
Gentianaceae	<i>Swertia racemosa</i>	QTPXJ_19_057	Illumina Novaseq 6000	Transcriptomic	Trinity	2130	21487
Gentianaceae	<i>Swertia racemosa</i>	QTPXJ_19_104	Illumina Novaseq 6000	Transcriptomic	Trinity	1920	26245
Gentianaceae	<i>Swertia racemosa</i>	ZMC202	Illumina Novaseq 6000	Transcriptomic	Trinity	1740	31520
Gentianaceae	<i>Swertia souliei</i>	QTP009	Illumina Hiseq X Ten	Transcriptomic	Trinity	1692	31655
Gentianaceae	<i>Swertia unknown</i>	QTPXJ_19_086	Illumina Novaseq 6000	Transcriptomic	Trinity	1953	17988
Gentianaceae	<i>Swertia verticillifolia</i>	QTPXJ_19_54	Illumina Novaseq 6000	Transcriptomic	Trinity	1948	21492
Gentianaceae	<i>Swertia younghusbandii</i>	QTPXJ_19_094	Illumina Novaseq 6000	Transcriptomic	Trinity	1676	35100
Gentianaceae	<i>Tripterospermum cordatum</i>	XL0819-9	Illumina Novaseq 6000	Transcriptomic	Trinity	1743	39355
Gentianaceae	<i>Veratilla baillonii</i>	SRR2032358	NCBI SRA	Transcriptomic	Trinity	1867	35085
Gentianaceae	<i>Veratilla baillonii</i>	ZMC223	Illumina Novaseq 6000	Transcriptomic	Trinity	1779	33405

128 Note: the taxon with whole genomic sequencing data were only used for orthologous gene assembly, and only contain the orthologous genes identified by MO pipeline, thus
129 were not assessed regarding the contig N50 and transcripts count.

Table S2. NCBI accession ID of the chloroplast genome used in the plastome phylogeny construction

NCBI accession	species name
MN627282	<i>Comastoma pedunculatum</i>
MT228728.1	<i>Lomatogoniopsis alpina</i>
MZ261905.1	<i>Swertia pubescens</i>
NC_044474.1	<i>Swertia hispidicalyx</i>
MF795137.1	<i>Swertia verticillifolia</i>
NC_050660.1	<i>Swertia multicaulis</i>
NC_031155.1	<i>Swertia mussotii</i>
MZ261901.1	<i>Swertia franchetiana</i>
MT228726.1	<i>Halenia elliptica</i>
NC_054249.1	<i>Sinoswertia tetraptera</i>
MW344296.1	<i>Swertia bimaculata</i>
MW292558.1	<i>Swertia erythrosticta</i>
MZ261897.1	<i>Swertia bifolia</i>
MZ261904.1	<i>Swertia przewalskii</i>
NC_054359.1	<i>Swertia cordata</i>
MW406478.1	<i>Swertia nervosa</i>
MT921831.1	<i>Gentianopsis paludosa</i>
MK993635.1	<i>Pterygocalyx volubilis</i>
NC_035719.1	<i>Gentiana macrophylla</i>
NC_039572.1	<i>Gentiana dahurica</i>
NC_030319.1	<i>Gentiana tibetica</i>
NC_053842.1	<i>Gentiana scabra</i>
NC_051950.1	<i>Kuepferia otophora</i>

132 Table S3. Taxa selected for the 12-taxon dataset

clade/subclade	taxon name
Outgroup	<i>Kueperia decorata</i>
Clade1	<i>Latouchea fokienensis</i>
Clade2	<i>Megacodon stylophorus</i>
Clade3	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>
Subclade1	<i>Swertia souliei</i>
Subclade2	<i>Veratrilla baillonii</i>
Subclade3	<i>Gentianella pygmaea</i>
Subclade4	<i>Gentianella acuta</i>
Subclade5	<i>Jaeschkea microsperma</i>
Subclade5	<i>Swertia multicaulis</i>
Subclade5	<i>Gentianella gentianoides</i>
Subclade5	<i>Comastoma pedunculatum</i>

Table S4. Results of ABBA-BABA analysis based on the 8 triplets from the 12-taxon dataset using evobiR

P1	P2	P3	P4/out group	Gene count	D raw	Z score	P-value (that D=0)	SD of D statistic	gene flow
<i>Comastoma pedunculatum</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3088	-0.11305	69.69261	0	0.001622102	Clade 1 ⇌ Clade 10
<i>Gentianella gentianoides</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3429	-0.15862	93.17538	0	0.00170236	Clade 2 ⇌ Clade 10
<i>Swertia multicaulis</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3183	-0.12681	91.67074	0	0.001383274	Clade 3 ⇌ Clade 10
<i>Jaeschkea microsperma</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	2902	-0.09065	54.31	0	0.001669039	Clade 4 ⇌ Clade 10
<i>Gentianella acuta</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3441	-0.15098	122.7184	0	0.001230295	Clade 5 ⇌ Clade 10
<i>Gentianella pygmaea</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3402	-0.04742	34.42299	0	0.001377525	Clade 6 ⇌ Clade 10
<i>Veratrilla baillonii</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3503	-0.08934	53.15193	0	0.001680764	Clade 7 ⇌ Clade 10
<i>Swertia souliei</i>	<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Megacodon stylophorus</i>	<i>Kuepferia decorata</i>	3345	-0.12676	92.14727	0	0.001375577	Clade 8 ⇌ Clade 10

NOTE: P1: taxa from *Gentianella*-*Swertia* complex (GSC) clade; P2 taxa from clade 3; P3 taxa from clade 2; P4: outgroup from subtribe *Gentianinae*

136

Table S6. Results of HyDe analysis at the individual level based on the 12-taxon dataset in testing the hybrid origin of Gentianella-Swertia complex (GSC)

137

clade

P1	Hybrid	P2	Z-score	Pvalue	Gamma	AAAA	AAAB	AABA	AABB	AABC	ABAA	ABAB	ABAC	ABBA	BAAA	ABBC	CABC	BACA	BCAA	ABCD
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Comastoma pedunculatum</i>	<i>Megacodon stylophorus</i>	21.09283	0	0.552552	86232	1912	10952	7278	4012	32016	4911	1300	7834	15821	944	5534	2812	1237	402
						8	6	3												
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Swertia multicaulis</i>	<i>Megacodon stylophorus</i>	18.70829	0	0.451932	95208	2192	55659	7055	2500	37254	4615	1375	6627	18239	976	3394	1670	1460	273
						2	2													
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Jaeschkea microsperma</i>	<i>Megacodon stylophorus</i>	11.52128	0	0.450059	95932	2364	47876	6018	2054	38633	4614	1523	5763	18191	930	2903	1527	1319	251
						7	3													
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Gentianella acuta</i>	<i>Megacodon stylophorus</i>	4.32386	7.67E-06	0.972969	97329	2693	34640	3357	1420	29621	3016	1162	15290	20681	1744	2587	1096	946	196
						4	2													
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Swertia souliei</i>	<i>Megacodon stylophorus</i>	4.185385	1.42E-05	0.941138	96284	2576	39872	4131	1672	34727	3769	1411	9557	19555	1312	2795	1247	1096	230
						9	8													
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Veratrilla baillonii</i>	<i>Megacodon stylophorus</i>	-6.62205	1	1.043359	99364	2905	14856	2209	633	32329	2684	1335	14114	21421	1687	1455	542	969	118
						6	0													
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Gentianella pygmaea</i>	<i>Megacodon stylophorus</i>	-6.95108	1	1.057682	95045	2729	23750	2595	941	31316	3139	1288	13114	19813	1637	1854	795	921	146
						8	8													
<i>Gentianopsis barbata</i> var. <i>stenocalyx</i>	<i>Gentianella gentianoides</i>	<i>Megacodon stylophorus</i>	-99999.9	1	1.303509	99765	2599	18236	5004	951	43782	4658	1776	3172	19346	783	1183	673	1445	137
						2	4													

138

139 **Table S7.** Results of HyDe analysis at the population level based on the 12-taxon dataset in testing the hybrid origin of Gentianella-Swertia complex (GSC)

140 clade

P1	Hybrid	P2	Z-score	Pvalue	Gamma	AAAA	AAAB	AABA	AABB	AABC	ABAA	ABAB	ABAC	ABBA	BAAA	ABBC	CABC	BACA	BCAA	ABCD
pop2	pop1	pop3	3.81275	6.87E-	0.96024	805088	235004	387394	461568	168697	330190	422069	142648	137618	297572	215778	303104	198294	179135	30659
			5	05	7	84	7	9			3			1	9					

Table S9. Comparison of divergence time (shown as 95% credible interval in million years ago) estimates in Gentianeae

	Plastome dataset	Nuclear dataset	Matuszak et al., 2015	Favre et al., 2016
Gentianeae stem	34.4-71.7	41.1-73.6	26.73-59.53	32.2-60.7
Gentianeae crown	30.2-63.9	27.8-48.8	22.53-50.75	29.7-52.8
Swertiinae stem	30.2-63.9	27.8-48.8	22.53-50.75	29.7-52.8
Swertiinae crown	26.3-56.7	19.1-31.1	9.5-25.5	not available
Gentianinae stem	30.2-63.9	27.8-48.8	22.53-50.75	29.7-52.8
Gentianinae crown	17.8-47.6	18.6-33.1	17.8-39.8	25.3-45.0

145

Table S10. Model test results of BioGeoBEARS analysis from the plastome dataset

Model	LnL	numparams	d	e	j	AICc	AICc_wt
DEC	-87.72	2	1.61	0.32	0	179.6	0.71
DEC+J	-87.67	3	1.61	0.27	0.0039	181.8	0.25
DIVALIKE	-90.95	2	1.65	0.12	0	186.1	0.028
DIVALIKE+J	-90.84	3	1.56	0.012	0.0064	188.1	0.01
BAYAREALIKE	-93.78	2	1.37	1.48	0	191.8	0.0017
BAYAREALIKE+J	-93.34	3	1.31	1.33	0.0067	193.1	0.0009

146

Table S11. Model test results of BioGeoBEARS analysis from the nuclear dataset

Model	LnL	numparams	d	e	j	AICc	AICc_wt
DEC	-78.04	2	1.77	1.00E-12	0	160.3	0.15
DEC+J	-77	3	1.64	1.00E-12	0.012	160.4	0.14
DIVALIKE	-76.78	2	1.89	1.00E-12	0	157.8	0.52
DIVALIKE+J	-76.7	3	1.83	1.00E-12	0.0035	159.8	0.19
BAYAREALIKE	-89.72	2	1.78	1.85	0	183.6	1.30E-06
BAYAREALIKE+J	-83.79	3	1.27	5.90E-01	0.023	174	0.0002

Other Supplementary Materials for this manuscript in a separate excel file include the following:

Table S5. Results of ABBA-BABA analysis based on the 882 triplets using evobiR

Table S8. Results of HyDe analysis at the individual level from the 882 triplets in testing the hybrid origin of the *Gentianella-Swertia* complex (GSC) clade