

Supplementary Tables and Figures

- Table S1. Samples used in this study.
Table S2. Correlation tests for all pairs of gene properties.
Table S3. Distribution of loci on autosomes and Z chromosome.
Table S4. Gene concordance and discordance factors.
Table S5. Summary statistics of datasets.
Table S6. Summary of tests for stationarity, homogeneity and recombination.
Table S7. Robinson-Foulds distances.
Table S8. ABBA-BABA tests for UCEs and introns.

- Figure S1. SVDquartets tree.
Figure S2. All loci concatenated tree, under traditional or expanded models.
Figure S3. Monophyletic loci of UCEs and introns concatenated tree.
Figure S4. UCEs concatenated tree.
Figure S5. UCE monophyletic loci concatenated tree.
Figure S6. Introns concatenated tree.
Figure S7. Introns monophyletic loci concatenated tree.
Figure S8. ASTRAL species tree using all gene trees under traditional or expanded models.
Figure S9. ASTRAL species tree using UCE and intron monophyletic loci.
Figure S10. ASTRAL species tree using all UCEs.
Figure S11. ASTRAL species tree using UCE monophyletic loci.
Figure S12. ASTRAL species tree using all UCEs with non-monophyletic loci constrained.
Figure S13. ASTRAL species tree using all UCEs with all loci constrained for two clades.
Figure S14. ASTRAL species tree using all introns.
Figure S15. ASTRAL species tree using intron monophyletic loci.
Figure S16. ASTRAL species tree using all introns with non-monophyletic loci constrained.
Figure S17. ASTRAL species tree using all introns with all loci constrained for two clades.
Figure S18. Concordance and discordance factors.
Figure S19. Bayesian MCMC phylogenetic networks estimated by PhyloNet when mr = 1.
Figure S20. Bayesian MCMC phylogenetic networks estimated by PhyloNet when mr = 2.
Figure S21. Bayesian MCMC phylogenetic networks estimated by PhyloNet when mr = 3.

Table S1. Samples used in this study and their GenBank accessions.

Sequence Reads	Voucher	Sampling location	GenBank Accession
<i>Antilophia galeata</i>	LSUMNS 13882	Santa Cruz, Bolivia	SRS8892407
<i>Chiroxiphia boliviana</i>	LSUMNS 38925	Cochabamba, Bolivia	SRS8892414
<i>Chiroxiphia caudata</i>	LSUMNS 25852	Caazapa, Paraguay	SRS8892415
<i>Chiroxiphia pareola</i>	LSUMNS 5470	San Martin, Peru	SRS8892416
<i>Chloropipo unicolor</i>	LSUMNS 88978	Ucayali, Peru	SRS8892417
<i>Cryptopipo holochlora</i>	LSUMNS 88969	Ucayali, Peru	SRS8892419
<i>Lepidothrix coronata exquisita</i>	LSUMNS 74001	Junin, Peru	SRS8892427
<i>Lepidothrix coronata minuscula</i>	LSUMNS 28731	Panama, Panama	SRS8892428
<i>Lepidothrix isidorei</i>	LSUMNS 5602	San Martin, Peru	SRS8892429
<i>Lepidothrix nattereri</i>	LSUMNS 77610	Amazonas, Brazil	SRS8892430
<i>Lepidothrix serena</i>	LSUMNS 65823	Sipaliwini, Suriname	SRS8892431
<i>Machaeropterus pyrocephalus</i>	LSUMNS 76084	Ucayali, Peru	SRS8892433
<i>Masius chrysopterus</i>	LSUMNS 33760	Cajamarca, Peru	SRS8892434
<i>Xenopipo atronitens</i>	LSUMNS 42648	Loreto, Peru	SRS8892410
Genome Assembly			
<i>Chiroxiphia lanceolata</i>	MVZ:Bird:191951	Chiriqui, Panama	GCA_009829145.1
<i>Corapipo altera</i>	x	Costa Rica	GCA_003945725.1
<i>Lepidothrix coronata coronata</i>	LSUMZ Birds 110521	Loreto, Peru	GCA_001604755.1
<i>Manacus vitellinus</i>	x	Gamboa, Panama	GCA_001715985.3
<i>Neopelma chrysocephalum</i>	YPM:ORN:139661	Para, Suriname	GCA_003984885.2
<i>Pipra filicauda</i>	LSUMZ Birds 115617	Loreto, Peru	GCA_003945595.1
<i>Empidonax traillii</i> (outgroup)	x	Arizona, USA	GCA_003031625.1

Table S2. R values from Spearman's correlation tests for all pairs of standardized gene properties.

All loci	Parsimony informative sites	Proportion of informative sites	GC content	Interquartile GC variation	Missing data	Stationarity	Homogeneity
Proportion of informative sites	0.266						
GC content	0.148	0.248					
Interquartile GC variation	-0.553	-0.170	-0.147				
Missing data	0.155	-0.201	-0.604	0.048			
Stationarity	0.146	0.074	0.141	-0.049	-0.029		
Homogeneity	0.031	0.002	0.011	-0.019	0.007	-0.005	
Recombinant triplets	0.516	0.190	0.134	-0.280	0.100	0.108	0.032
UCEs	Parsimony informative sites	Proportion of informative sites	GC content	Interquartile GC variation	Missing data	Stationarity	Homogeneity
Proportion of informative sites	0.947						
GC content	0.270	0.270					
Interquartile GC variation	-0.380	-0.364	-0.162				
Missing data	0.118	0.095	-0.511	0.078			
Stationarity	0.155	0.151	0.139	0.003	0.036		
Homogeneity	0.025	0.028	0.024	-0.023	0.002	-0.003	
Recombinant triplets	0.429	0.413	0.171	-0.140	0.087	0.134	0.024
introns	Parsimony informative sites	Proportion of informative sites	GC content	Interquartile GC variation	Missing data	Stationarity	Homogeneity
Proportion of informative sites	0.267						
GC content	0.119	0.232					
Interquartile GC variation	-0.630	-0.158	-0.139				
Missing data	0.092	-0.116	-0.692	0.058			
Stationarity	0.149	0.013	0.140	-0.081	-0.057		
Homogeneity	0.023	0.014	0.002	-0.014	-0.003	-0.006	
Recombinant triplets	0.548	0.187	0.118	-0.365	0.053	0.098	0.032

Table S3. Distribution of loci on autosomes and Z chromosome. Each locus was blast against the *Chiropipha lanceolata* genome database to determine its location in the genome.

location	UCEs	UCE monophyletic	UCE non- monophyletic	introns	Intron monophyletic	Intron non- monophyletic
chr1	761	616	145	580	420	160
chr2	345	266	79	544	406	138
chr3	539	429	110	668	470	198
chr4	187	150	37	428	302	126
chr5	247	197	50	503	365	138
chr6	415	340	75	550	395	155
chr7	275	217	58	255	184	71
chr8	296	235	61	219	167	52
chr9	226	190	36	257	180	77
chr10	107	82	25	282	197	85
chr11	118	101	17	194	140	54
chr12	59	51	8	235	154	81
chr13	206	156	50	197	144	53
chr14	37	31	6	138	95	43
chr15	92	78	14	178	119	59
chr16	31	29	2	268	166	102
chr17	80	58	22	160	102	58
chr18	29	23	6	219	140	79
chr19	38	24	14	158	100	58
chr20	39	30	9	267	181	86
chr21	40	35	5	140	101	39
chr22	25	22	3	89	58	31
chr23	24	22	2	28	18	10
chr24	17	14	3	64	52	12
chr25	21	18	3	38	22	16
chr26	3	3	0	25	18	7
chr27	2	2	0	59	33	26
chr28	21	16	5	23	16	7
chr29	0	0	0	0	0	0
chr30	4	2	2	4	3	1
chr31	0	0	0	1	0	1
chrZ	315	268	47	118	85	33
uncertain	7	6	1	6	5	1

Table S4. Gene concordance and discordance factors. gCF shows gene concordance factor, gDFP shows gene discordance factor due to polyphyly, and gDF1 and gDF2 are the percentage of gene trees concordant with the two alternative topologies around the branch. gCF, gDF1, gDF2 and gDFP should add up to 100 for a specific branch. ID corresponds to the node number on tree in Figure S18.

Gene trees under traditional models

ID	gCF_UCE	gDF1_UCE	gDF2_UCE	gDFP_UCE	gCF_intron	gDF1_intron	gDF2_intron	gDFP_intron
24	29.51	6.62	5.99	57.88	26.51	5.29	4.68	63.51
25	48.89	5.97	5.49	39.64	39.72	10.76	3.8	45.71
26	67.35	3.71	3.52	25.42	56.81	4.02	4.03	35.14
27	29.98	6.86	6.56	56.6	27.77	5.09	5.18	61.96
28	20.3	12.92	11.42	55.36	17.22	11.94	11.23	59.61
29	36.89	2.63	3.91	56.58	29.89	2.39	4.38	63.34
30	29.02	16.32	18.32	36.35	24.65	16.98	16.18	42.18
31	65.11	1.76	1	32.13	51.94	2.02	1.52	44.53
32	22.9	14.79	21.36	40.95	24.28	16.07	13.31	46.34
33	24.79	9.57	16.91	48.72	21.71	8.15	17.38	52.76
34	78.47	2.36	1.61	17.56	75.49	2.44	2.46	19.61
35	78.55	0.39	0.63	20.43	67.48	1.55	0.71	30.25
36	84.23	2.73	3.39	9.65	78.31	3.5	3.43	14.77
37	81.13	0.15	0.41	18.3	70.17	0.26	0.59	28.98
38	15.59	15.59	14.26	54.56	16.19	15.07	13.01	55.74
39	11.77	9.1	13.66	65.48	11.59	8.08	10.81	69.52
40	34.17	9.57	8.15	48.12	27.92	11.69	9.71	50.67
41	33.55	27.97	7.25	31.24	27.7	20.4	14.73	37.17

Gene trees under expanded models

ID	gCF_UCE	gDF1_UCE	gDF2_UCE	gDFP_UCE	gCF_intron	gDF1_intron	gDF2_intron	gDFP_intron
24	28.9	6.6	5.92	58.58	26.5	5.31	4.5	63.7
25	48.7	5.86	5.32	40.12	39.74	10.91	3.89	45.47
26	66.37	3.69	3.58	26.36	56.58	4.1	3.93	35.39
27	29.42	6.77	6.25	57.56	27.88	5.08	5.28	61.77
28	20.34	12.87	11.2	55.58	17.38	11.93	11.29	59.41
29	36.24	2.71	3.8	57.25	30.18	2.52	4.38	62.92
30	28.49	16.25	18.18	37.07	24.6	16.65	16.29	42.47
31	64.55	1.72	0.96	32.78	51.91	2.09	1.51	44.5
32	22.86	15.15	21.34	40.64	23.94	15.97	13.3	46.79
33	25.53	9.47	16.46	48.55	21.01	7.97	17.3	53.73
34	76.86	2.91	1.81	18.42	76.31	2.11	2.43	19.15

35	78.27	0.33	0.54	20.86	67.56	1.64	0.73	30.08
36	83.25	3.15	3.6	10	78.5	3.48	3.44	14.58
37	80.57	0.15	0.43	18.84	70.36	0.25	0.59	28.8
38	15.87	15.33	14.33	54.47	15.91	15.37	13.14	55.58
39	11.59	9.03	13.5	65.87	11.87	8.17	10.68	69.29
40	33.73	9.48	8.37	48.42	28.81	11.72	9.47	50
41	32.44	29.44	7.14	30.99	28.64	19.73	14.58	37.05

Table S5. Summary statistics of locus characteristics. Note that here we only showed the monophyletic and non-monophyletic loci that were grouped based on the topologies of gene trees estimated under traditional models, since the general patterns are the same to those of gene trees under expanded models.

Dataset		Loci	Length	Average GC%	Interquartile GC% variation	Parsimony informative sites	Proportion of informative sites	Missing characters
UCEs	ALL	4606	2277 ± 359	39.9% ± 5.4%	5.8% ± 2.6%	111 ± 52	4.9% ± 2.2%	14.7% ± 5.7%
	monophyletic	3711	2284 ± 382	40.1% ± 5.5%	5.6% ± 2.2%	109 ± 41	4.8% ± 1.7%	14.3% ± 5.4%
	non-monophyletic	895	2250 ± 240	39.4% ± 5.3%	6.7% ± 3.7%	119 ± 83	5.2% ± 3.6%	16.3% ± 6.5%
introns	ALL	6895	1664 ± 1665	40.4% ± 5.4%	6.2% ± 3.0%	108 ± 106	6.7% ± 1.8%	10.8% ± 6.0%
	monophyletic	4838	1938 ± 1860	40.6% ± 5.1%	5.8% ± 2.8%	123 ± 116	6.5% ± 1.5%	10.3% ± 5.6%
	500-999	1637	752 ± 142	41.2% ± 6.0%	7.6% ± 3.1%	51 ± 16	6.9% ± 1.7%	9.2% ± 5.7%
	1000-2999	2418	1694 ± 541	40.3% ± 4.8%	5.4% ± 2.1%	108 ± 40	6.4% ± 1.4%	10.7% ± 5.6%
	>=3000	783	5172 ± 2618	40.7% ± 4.0%	3.6% ± 1.4%	319 ± 165	6.2% ± 1.3%	11.5% ± 5.1%
	non-monophyletic	2057	1019 ± 754	39.7% ± 5.8%	7.1% ± 3.2%	73 ± 66	7.0% ± 2.3%	11.9% ± 6.7%
	500-999	1397	710 ± 132	39.7% ± 6.0%	7.9% ± 3.2%	49 ± 19	6.9% ± 2.3%	11.2% ± 6.4%
	1000-2999	611	1446 ± 428	39.6% ± 5.4%	5.6% ± 2.3%	103 ± 49	7.1% ± 2.4%	13.0% ± 7.0%
	>=3000	49	4518 ± 1973	41.2% ± 4.2%	3.4% ± 1.4%	365 ± 180	8.2% ± 2.8%	16.2% ± 7.9%

Table S6. Summary of tests for stationarity, homogeneity and recombination. For the first two tests, *loci* and *proportion* indicate the number of loci that have significant sequence pairs by each test (i.e., the assumption is rejected) and its proportion relative to the entire dataset; *mean* (and associated standard deviation) indicates the average number of significant sequence pairs by each test with zeros removed. For recombination test, *loci* indicates the number of loci that contain recombinant triplets; *mean* indicates the average number of recombinant triplets across the dataset with zeros removed.

Dataset		Loci	Stationarity rejected				Homogeneity rejected				Recombinant triplets			
			loci	proportion	mean	STD	loci	proportion	mean	STD	loci	proportion	mean	STD
UCEs	ALL	4606	705	15.31%	27.43	16.86	229	4.97%	13.38	8.34	2327	50.52%	77.07	202.28
	monophyletic	3711	536	14.44%	24.15	11.36	181	4.88%	13.60	8.55	1786	48.13%	39.19	96.96
	non-monophyletic	895	169	18.88%	37.86	25.23	48	5.36%	12.56	7.51	541	60.45%	202.13	353.20
introns	ALL	6895	1110	16.10%	24.37	12.27	266	3.86%	12.1	7.5	2613	37.90%	34.84	112.97
	monophyletic	4838	838	17.32%	25.11	12.14	185	3.82%	12.3	8.2	1960	40.51%	21.15	52.96
	non-monophyletic	2057	272	13.22%	22.07	12.42	81	3.94%	11.7	5.7	653	31.75%	75.94	201.12

Table S7. Average Robinson-Foulds distances from each gene tree set to the reference species trees with T3/Ta and T1/Tb, with standard deviation.

Gene tree set	RF to T3/Ta	SD	RF to T1/Tb	SD
All UCEs	18.63	5.95	18.70	5.68
Monophyletic UCEs	17.04	4.63	17.18	4.39
Non-monophyletic UCEs	25.21	6.29	24.98	6.09
All introns	20.01	5.78	20.18	5.58
Monophyletic introns	18.06	4.91	18.27	4.67
Non-monophyletic introns	24.59	5.06	24.67	4.92

Table S8. ABBA-BABA tests for all possible combinations of *C. pareola*, *C. lanceolata*, *A. galeata*, *C. boliviiana* and *C. caudata* with the relationship of ((P1,P2),P3) using *Masius chrysopterus* as the outgroup. The underlying topology for the tests is indicated (T1 for three triplets, T3 for three triplets, and it is compatible with either T1 or T3 for seven triplets). N is the number of tests with significant D statistics (p-value < 0.05) out of a total of 100 random SNP selections from all loci, D stat. is the average D statistic across N tests, and SD is the standard deviation. Positive D statistics indicate excess shared ancestry for P2 and P3 and negative D statistics indicate excess shared ancestry between P1 and P3.

Underlying topology	Taxon set			UCEs			Introns		
	P1	P2	P3	N	D stat.	SD	N	D stat.	SD
T1	<i>C. boliviiana</i>	<i>A. galeata</i>	<i>C. caudata</i>	3	0.191	0.055	5	0.121	0.007
	<i>C. boliviiana</i>	<i>C. lanceolata</i>	<i>C. caudata</i>	9	0.183	0.029	3	-0.132	0.032
	<i>C. boliviiana</i>	<i>C. pareola</i>	<i>C. caudata</i>	4	0.001	0.189	9	0.113	0.093
T3	<i>C. caudata</i>	<i>C. boliviiana</i>	<i>A. galeata</i>	26	0.179	0.032	34	0.144	0.031
	<i>C. caudata</i>	<i>C. boliviiana</i>	<i>C. lanceolata</i>	39	0.191	0.044	59	0.156	0.034
	<i>C. caudata</i>	<i>C. boliviiana</i>	<i>C. pareola</i>	27	0.176	0.076	53	0.143	0.031
T1 or T3	<i>A. galeata</i>	<i>C. lanceolata</i>	<i>C. boliviiana</i>	7	0.182	0.022	5	-0.105	0.127
	<i>C. lanceolata</i>	<i>A. galeata</i>	<i>C. caudata</i>	6	-0.006	0.200	10	0.152	0.026
	<i>C. pareola</i>	<i>A. galeata</i>	<i>C. boliviiana</i>	4	-0.095	0.182	13	-0.149	0.030
	<i>C. pareola</i>	<i>A. galeata</i>	<i>C. caudata</i>	3	0.201	0.037	17	-0.141	0.077
	<i>C. pareola</i>	<i>C. lanceolata</i>	<i>C. boliviiana</i>	15	0.194	0.030	19	-0.159	0.034
	<i>C. pareola</i>	<i>C. lanceolata</i>	<i>C. caudata</i>	4	0.097	0.187	37	-0.160	0.028
	<i>C. pareola</i>	<i>C. lanceolata</i>	<i>A. galeata</i>	14	0.206	0.027	5	-0.025	0.149

Figure S1. SVDquartets species tree, which shows consistent topologies across all SVDquartets analyses (T1 & Tb). Values at nodes are bootstrap support for UCEs, UCE monophyletic loci, introns, and intron monophyletic loci respectively. Full supports were not labeled on node. Note that within *Lepidothrix coronata*, SVDquartets trees give an unexpected topology, based on known relationships among subspecies and geography.

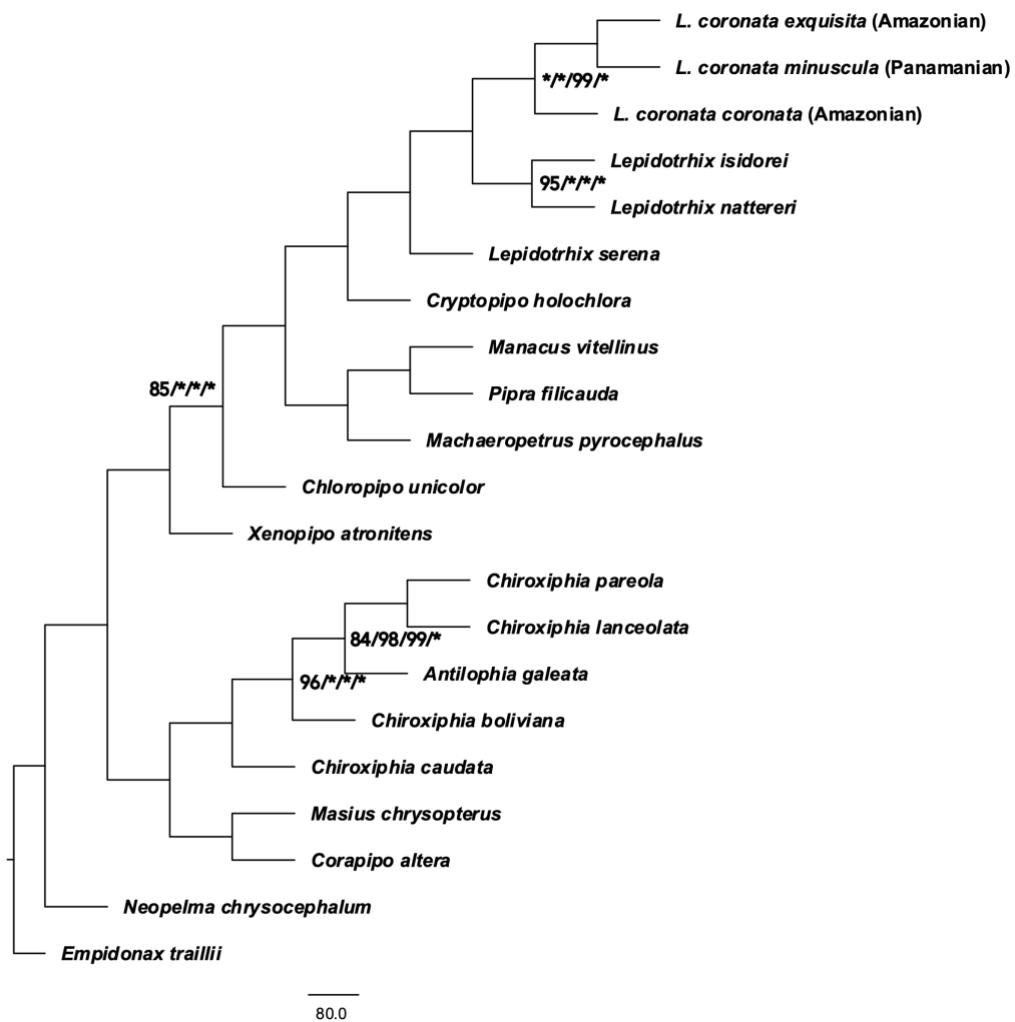
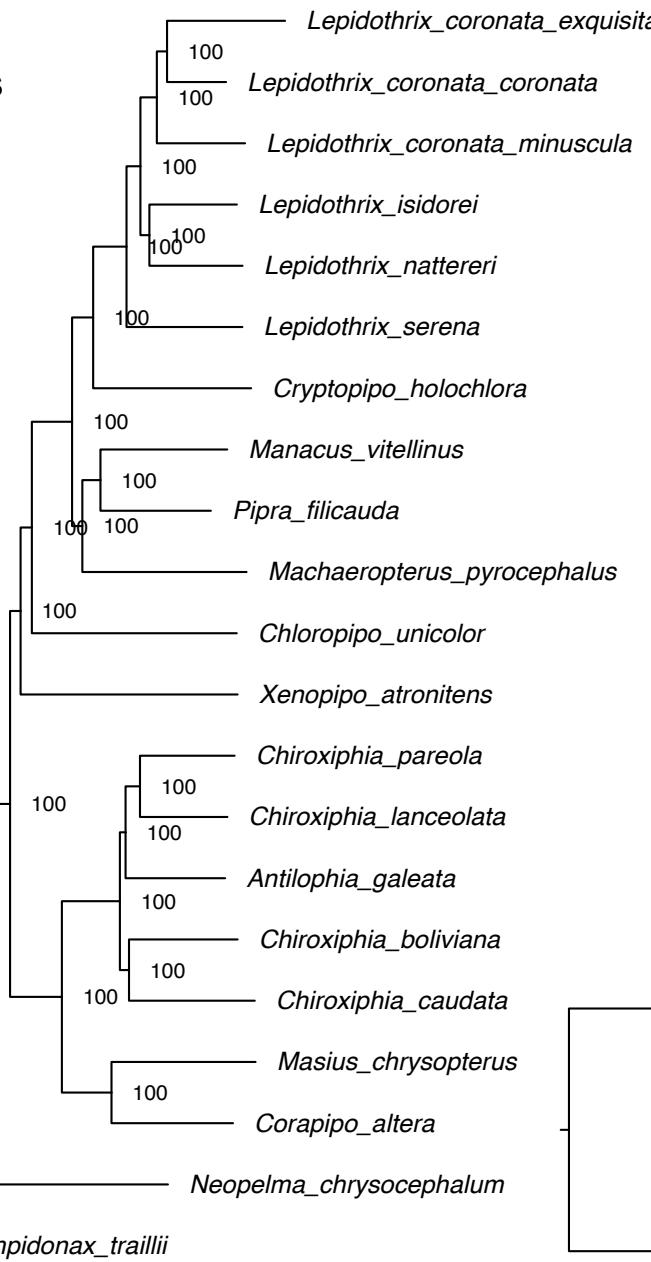
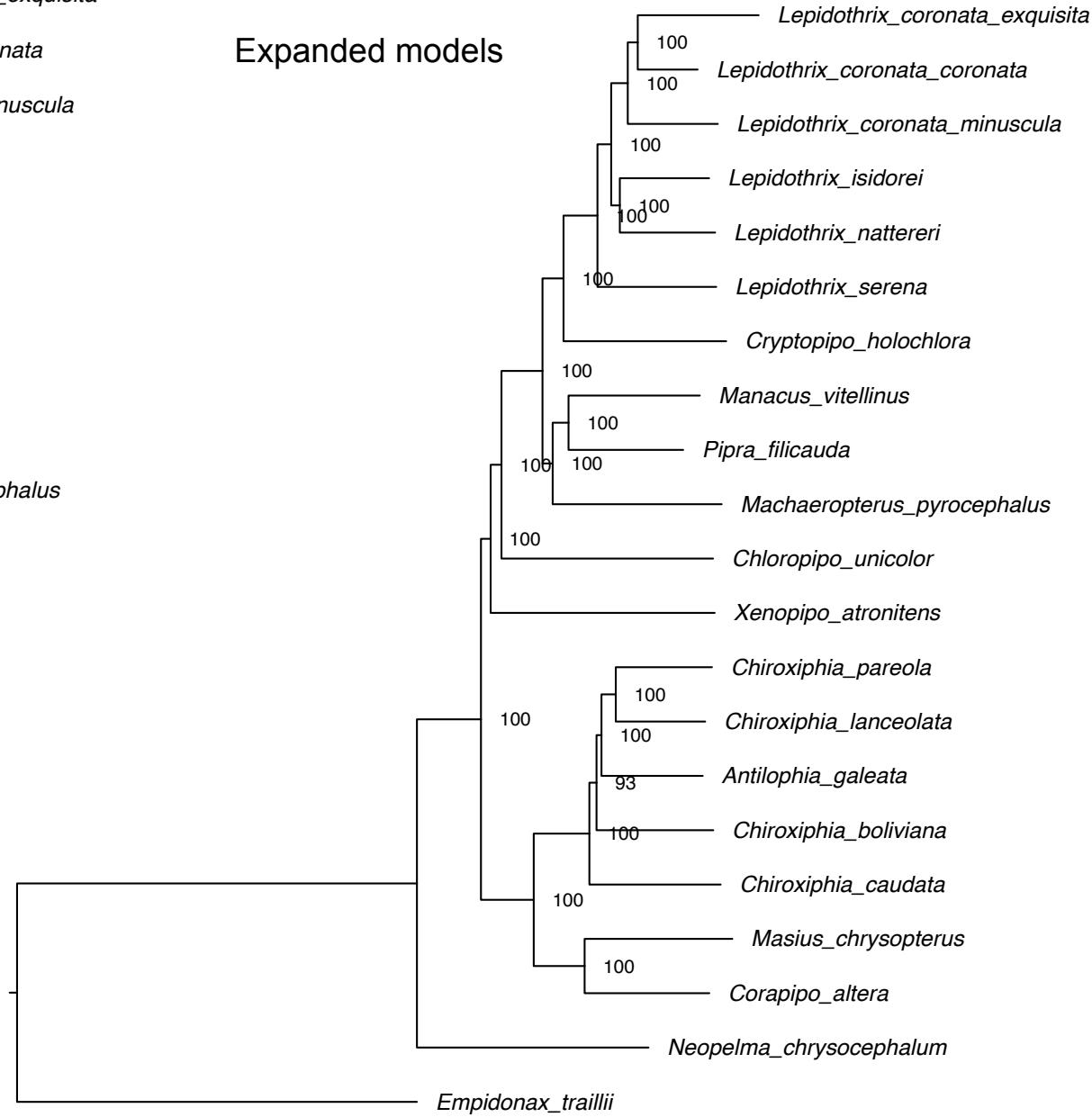


Figure S2. Concatenated trees estimated under two different models sets in IQ-TREE using all UCEs and introns combined.

Traditional models



Expanded models



0.008

0.008

Figure S3. Concatenated trees estimated under two different models sets in IQ-TREE using monophyletic loci of UCEs and introns combined. Removal of the non-monophyletic loci had little impact on the relevant branch length compared to the trees in Figure S2.

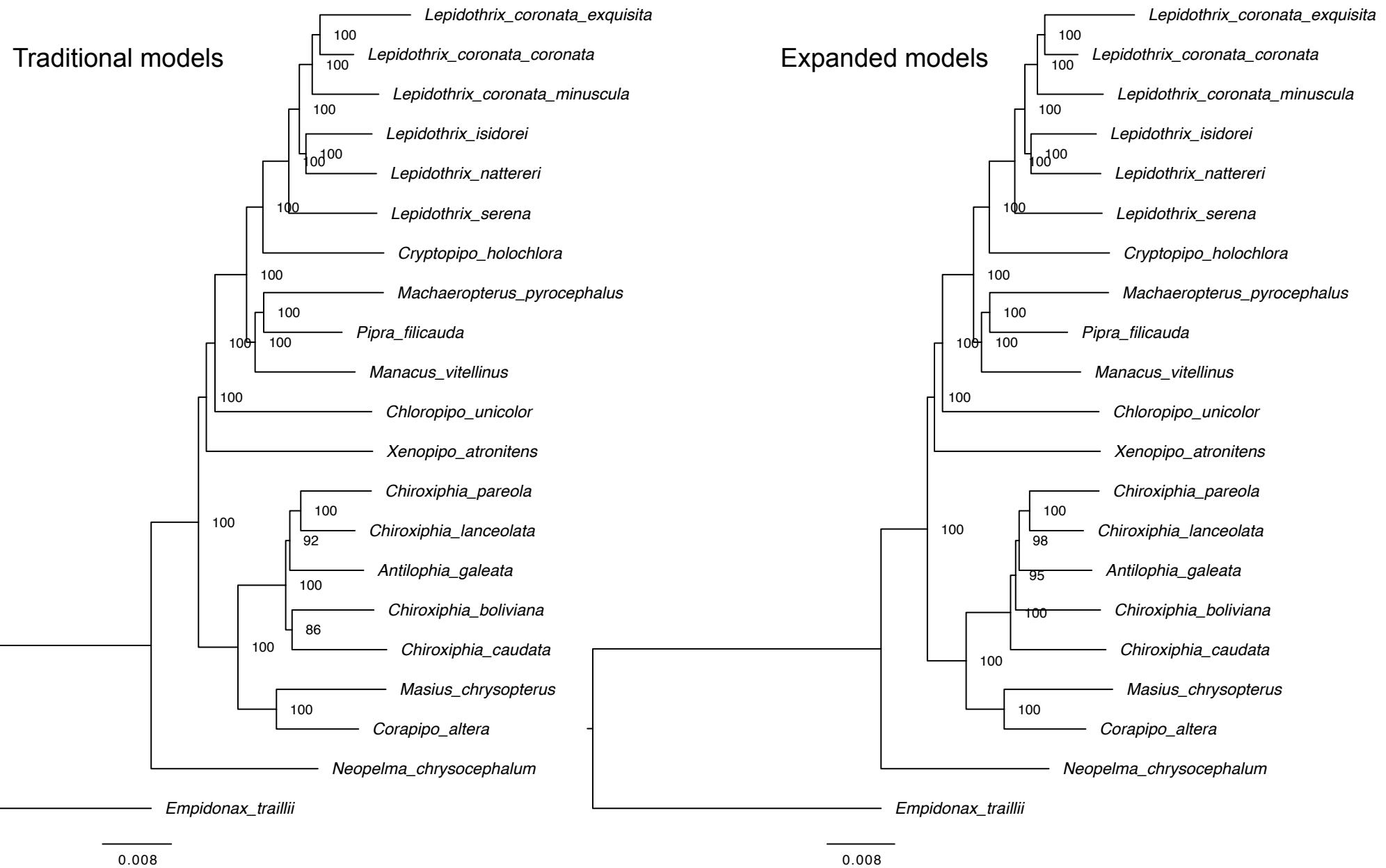


Figure S4. Concatenated trees estimated under two different models sets in IQ-TREE using all UCEs.

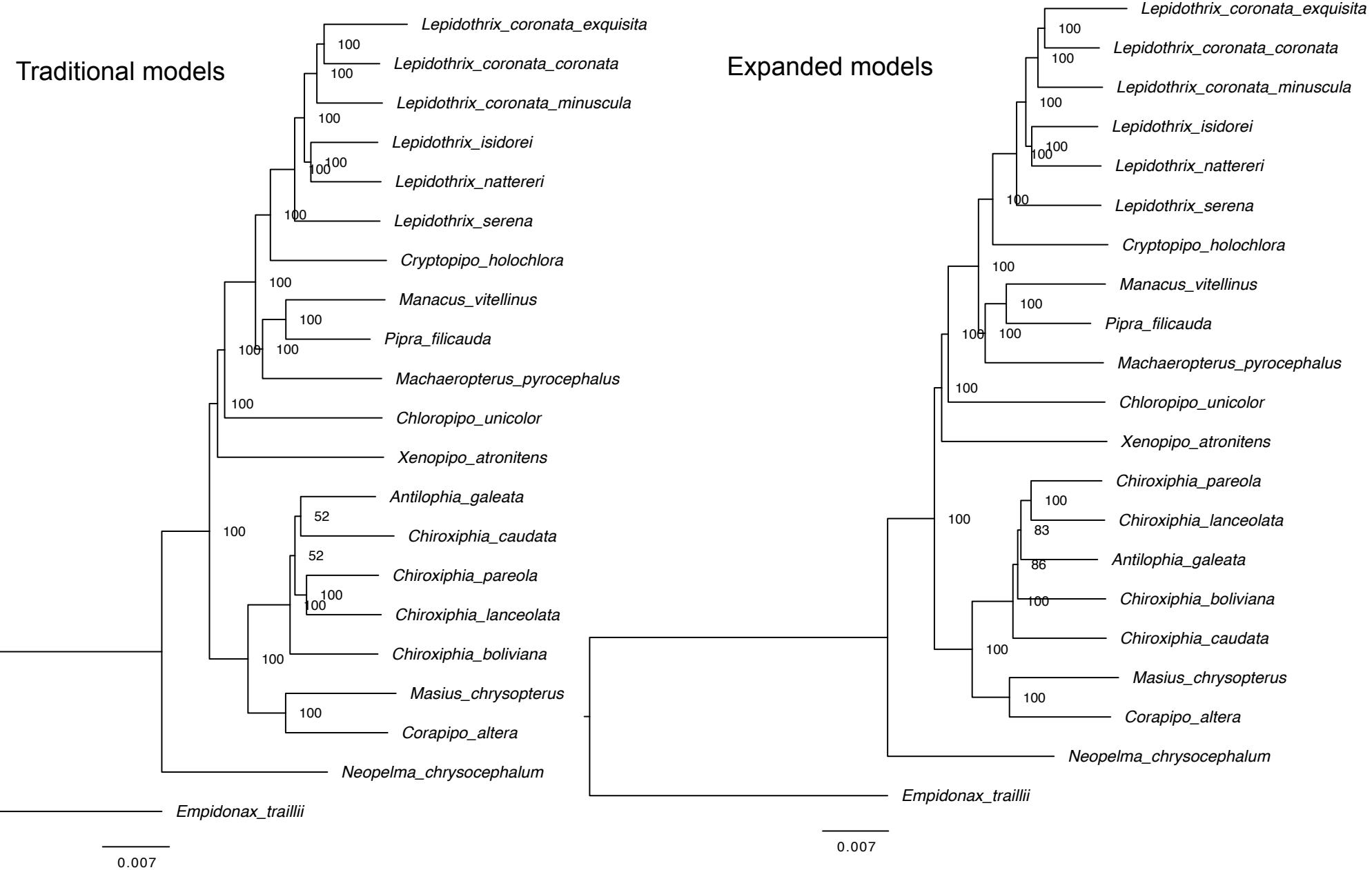


Figure S5. Concatenated trees estimated under two different models sets in IQ-TREE using UCE monophyletic loci. Removal of the non-monophyletic loci had little impact on the relevant branch length compared to the trees in Figure S4.

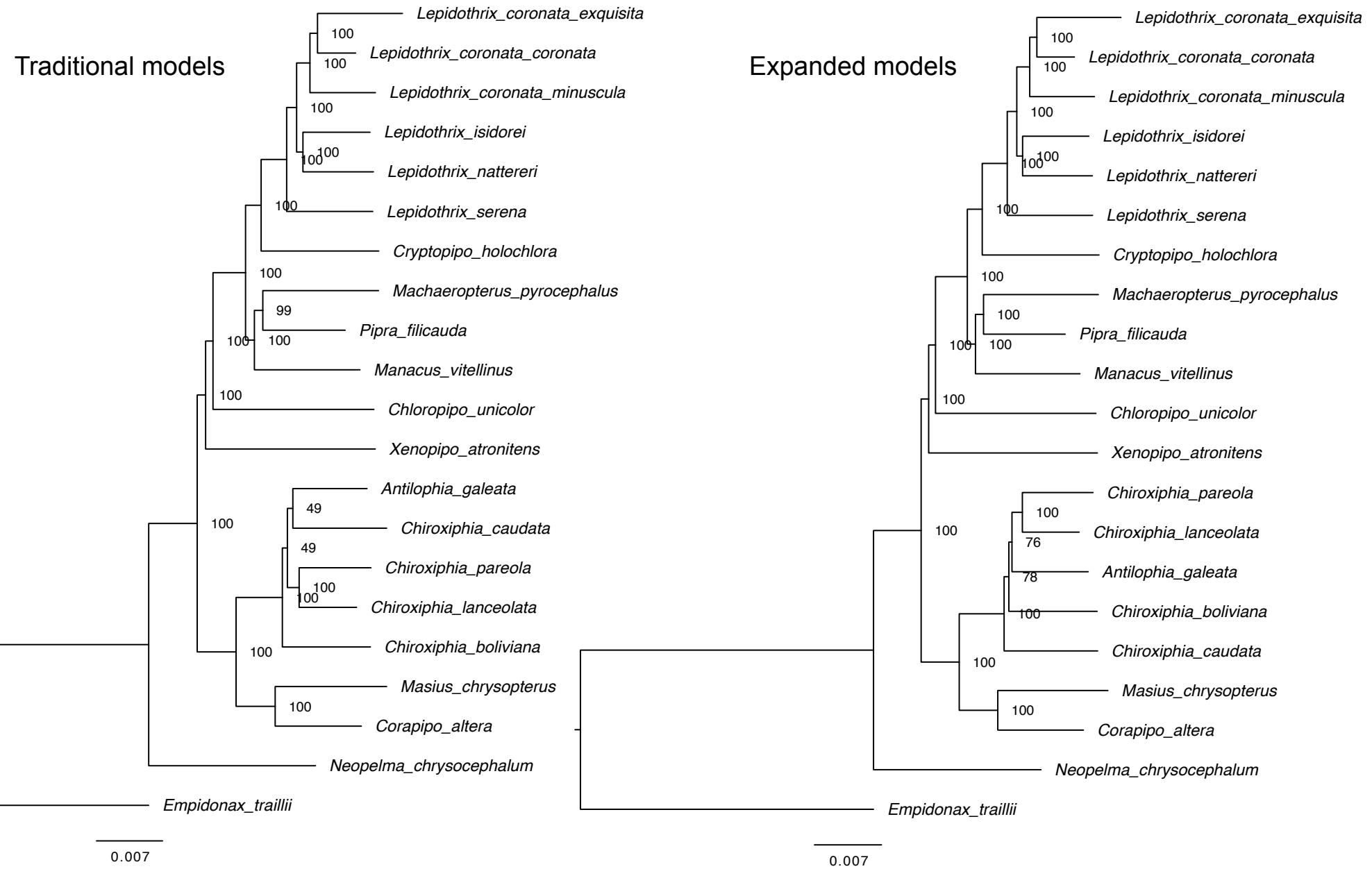


Figure S6. Concatenated trees estimated under two different models sets in IQ-TREE using all introns.

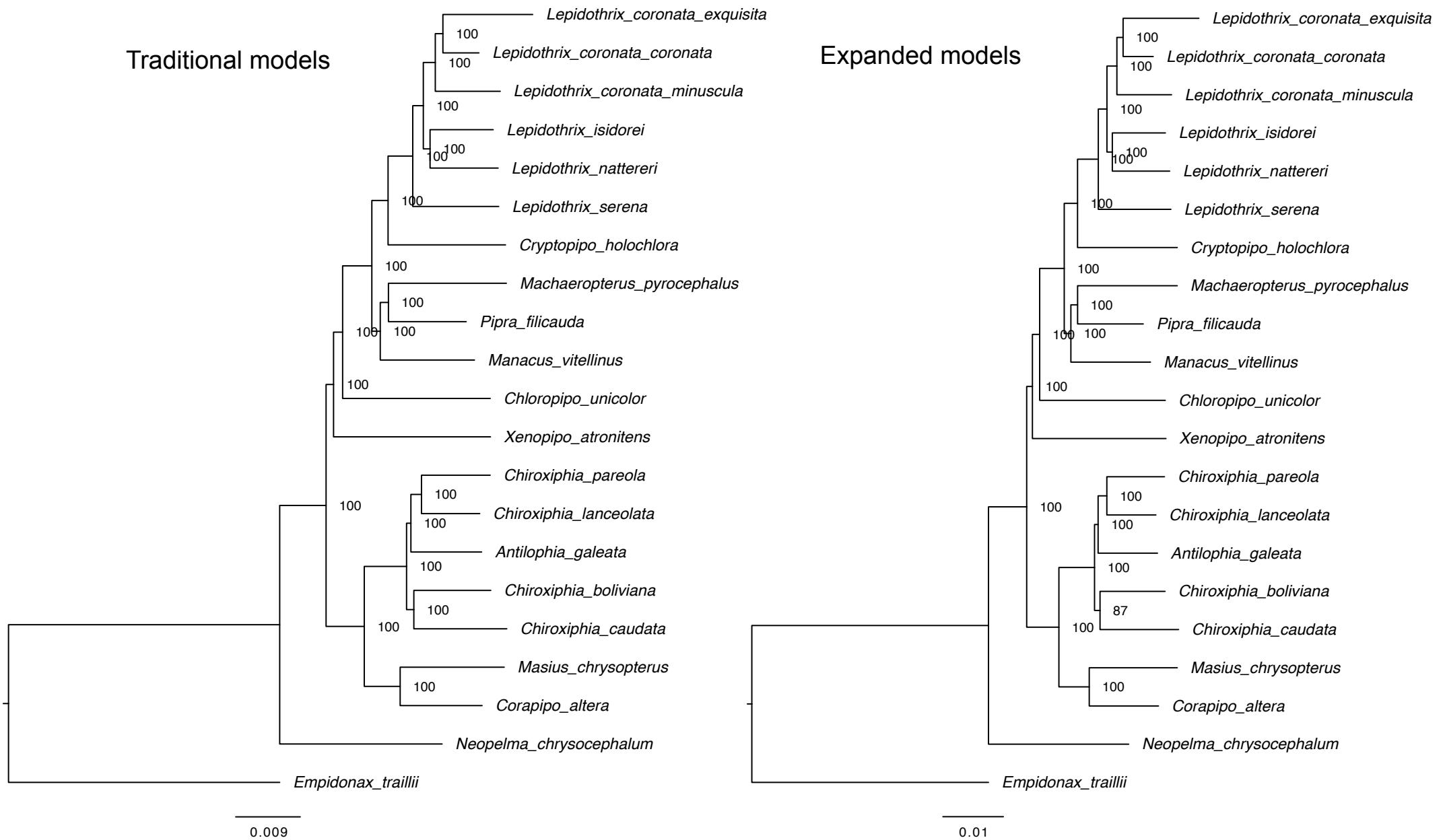


Figure S7. Concatenated trees estimated under two different models sets in IQ-TREE using intron monophyletic loci. Removal of the non-monophyletic loci had little impact on the relevant branch length compared to the trees in Figure S6.

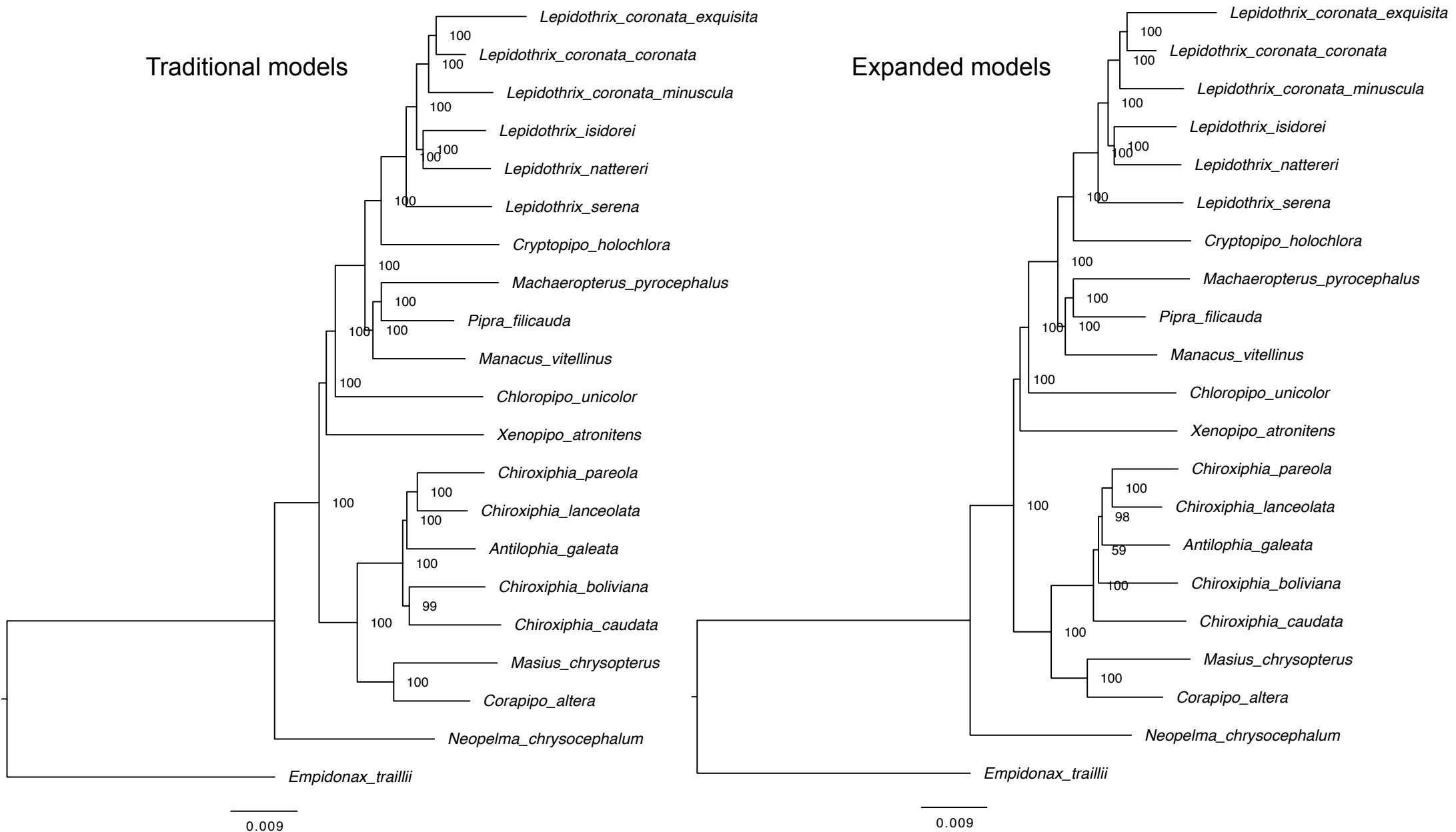


Figure S8. ASTRAL species tree using gene trees of all loci esitmated under traditional or expanded models.

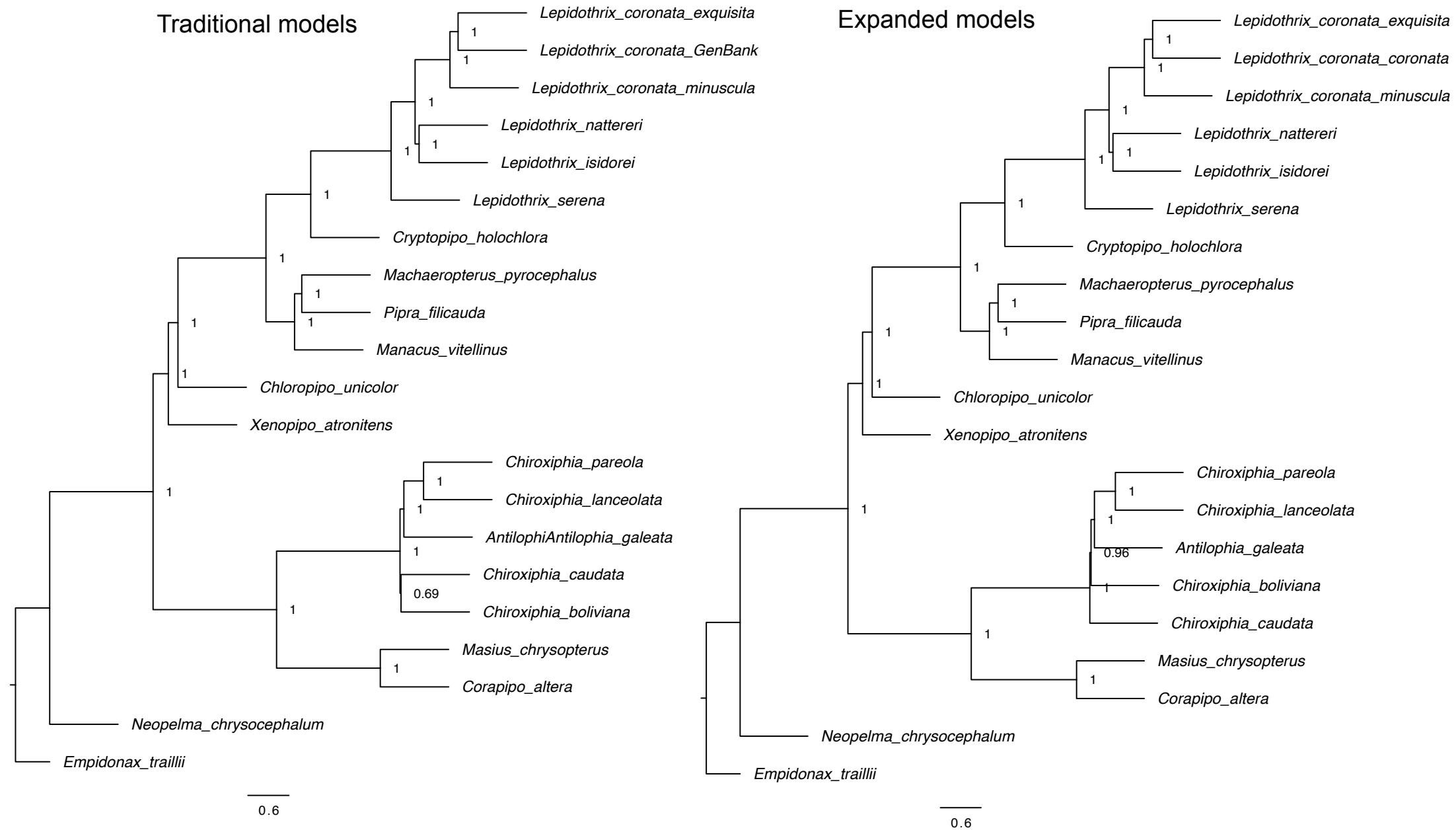
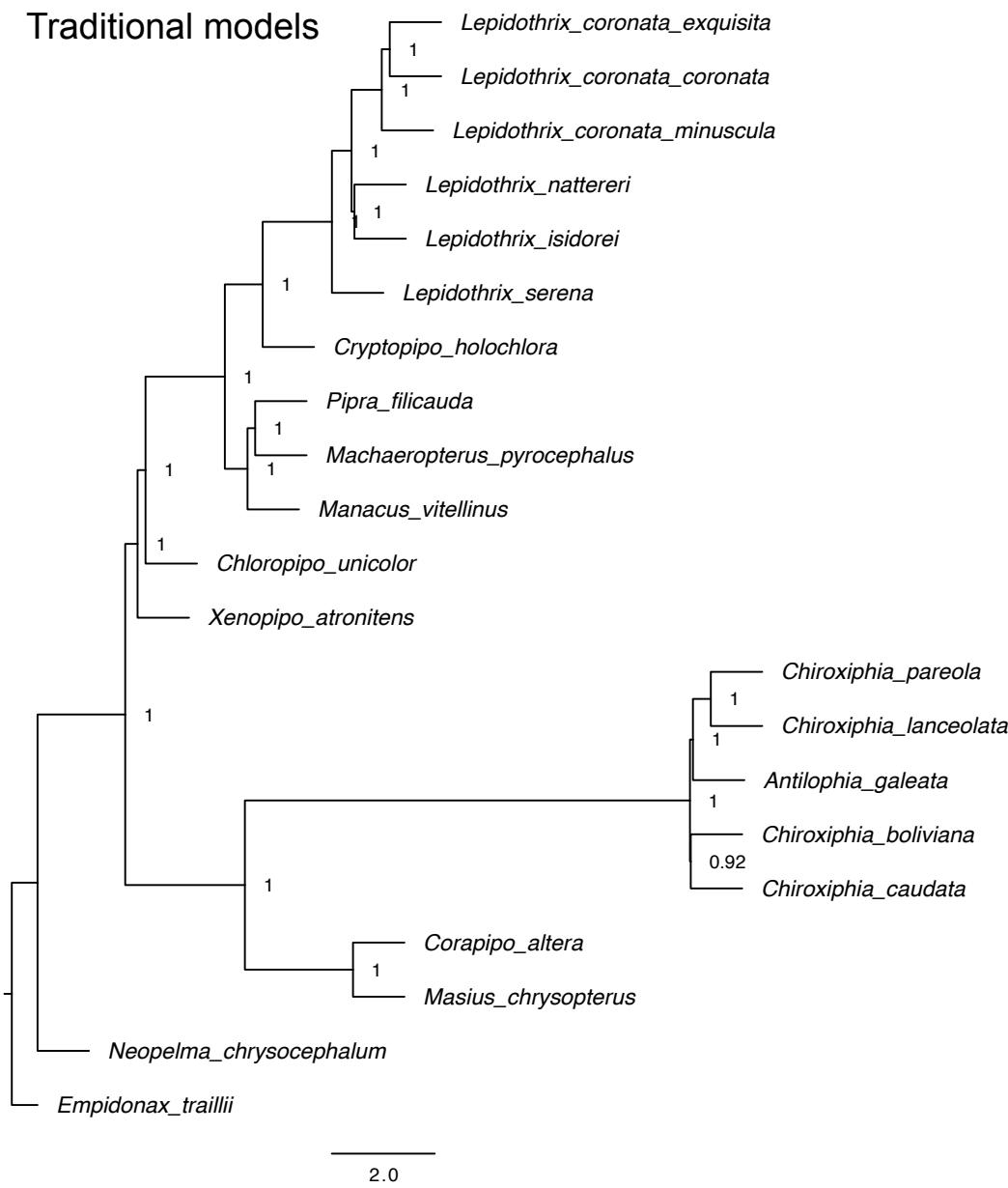


Figure S9. ASTRAL species tree using UCE and intron monophyletic loci.

Traditional models



Expanded models

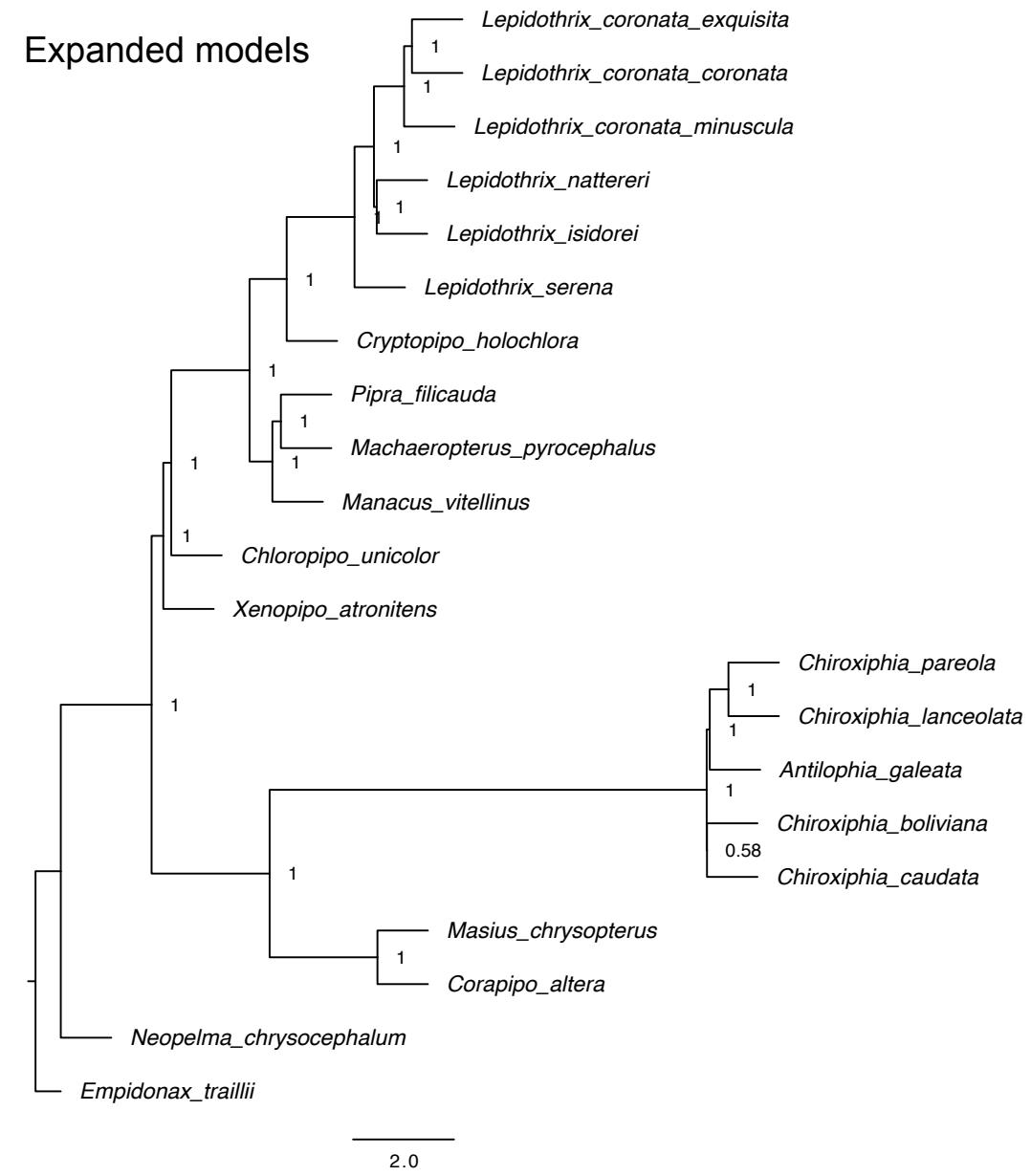
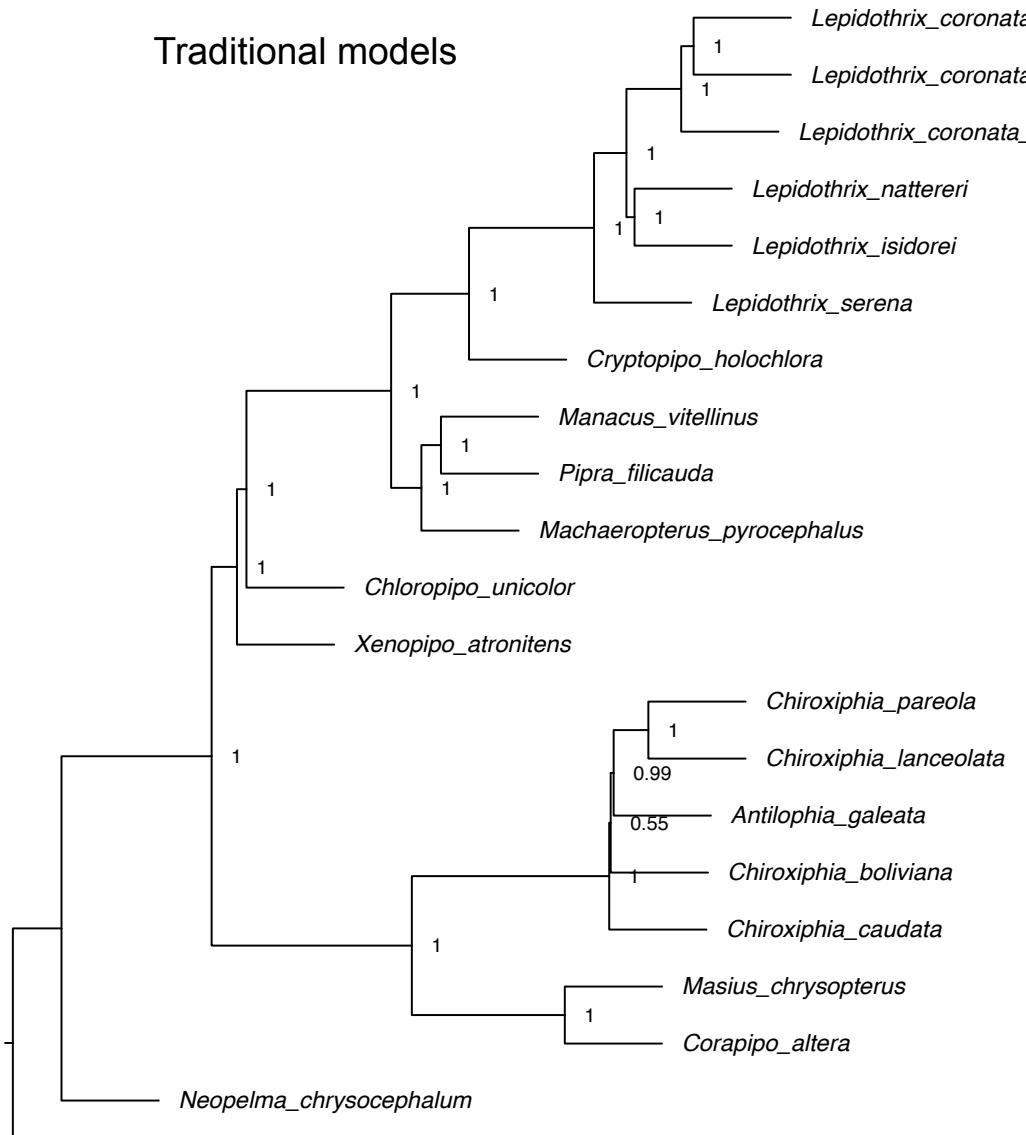
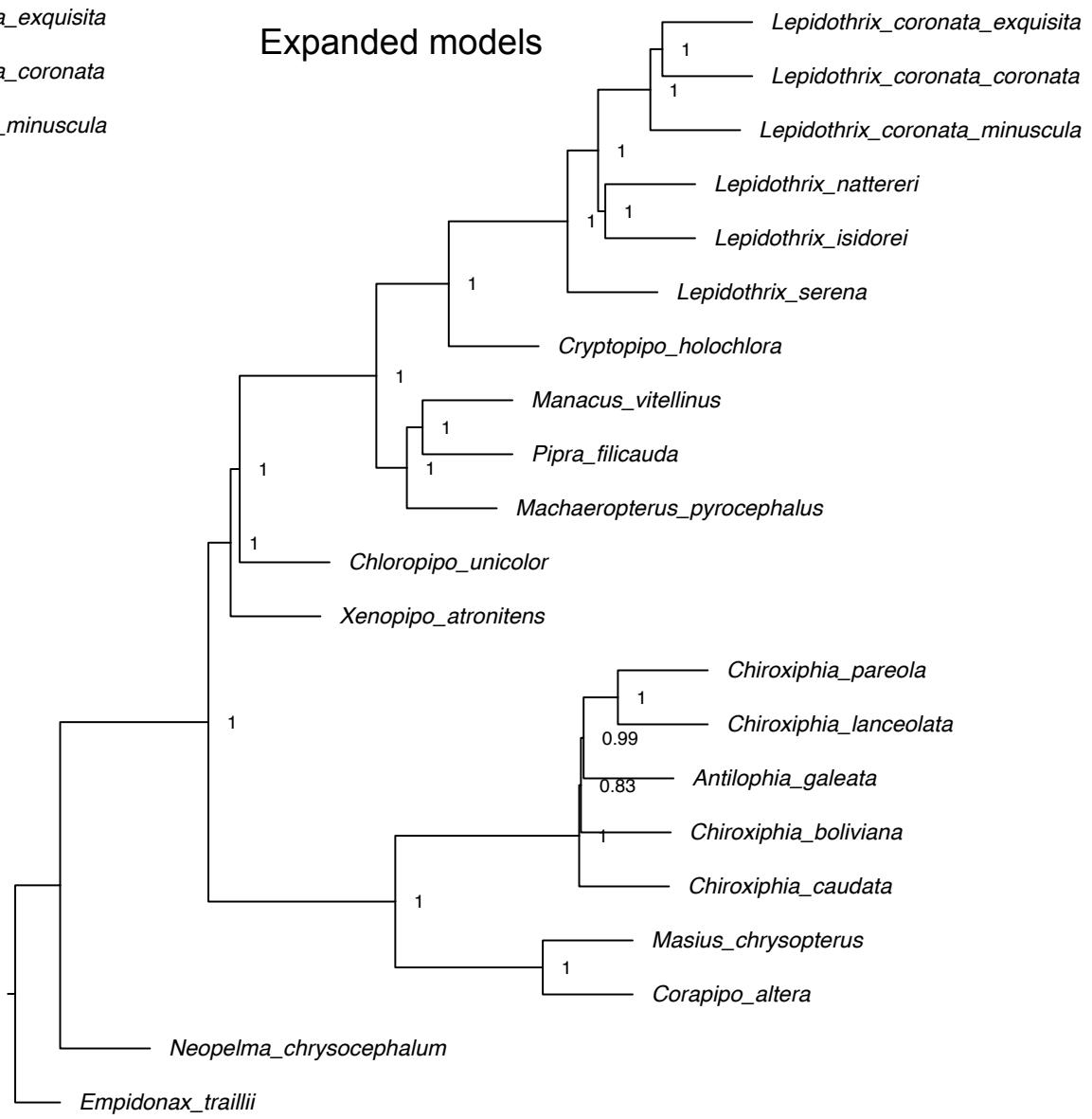


Figure S10. ASTRAL species tree using all UCEs.

Traditional models



Expanded models

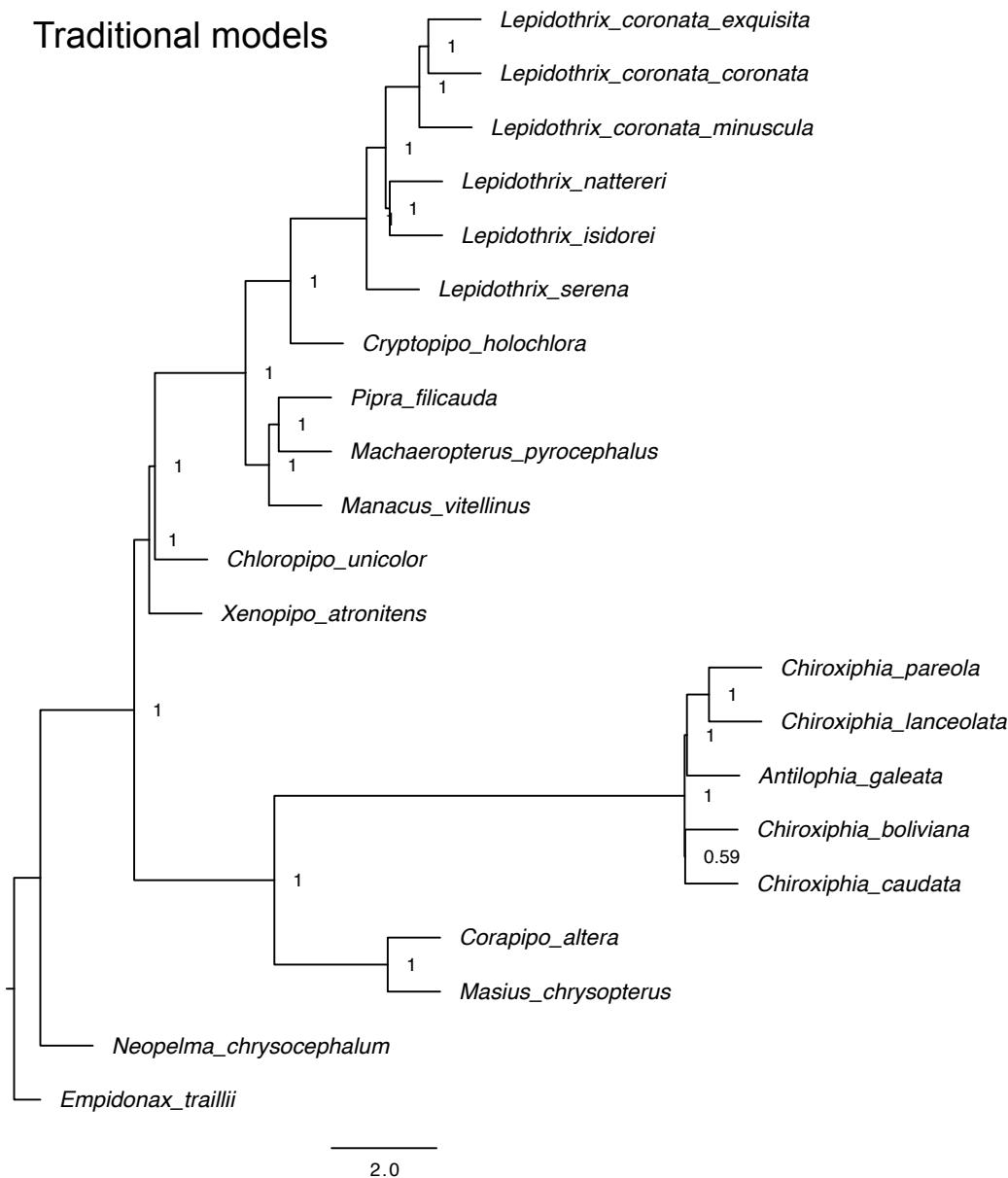


0.6

0.6

Figure S11. ASTRAL species tree using UCE monophyletic loci.

Traditional models



Expanded models

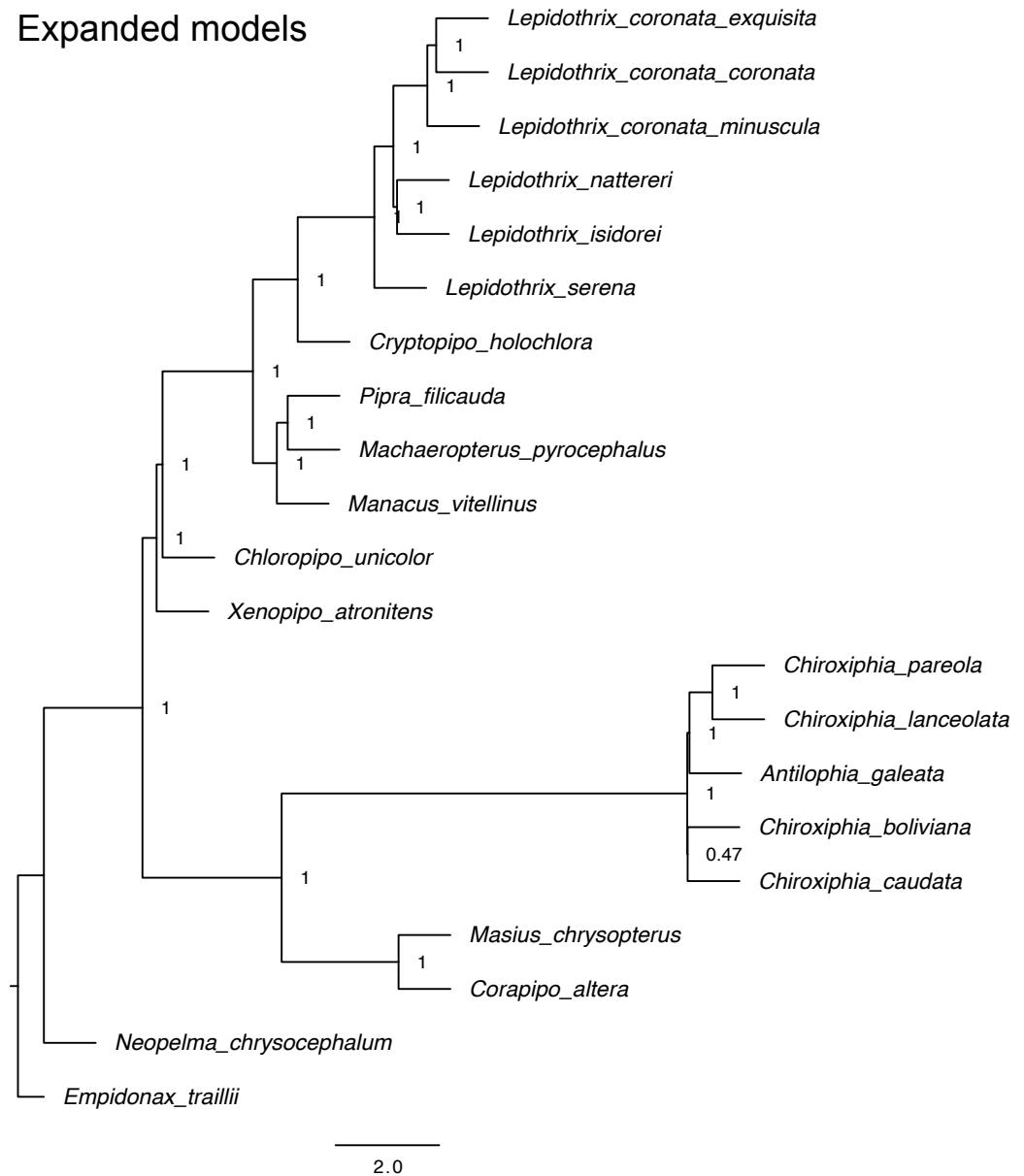


Figure S12. ASTRAL species tree using all UCEs with non-monophyletic loci constrained to be monophyletic for Antilophia+Chiroxiphia (constr1).

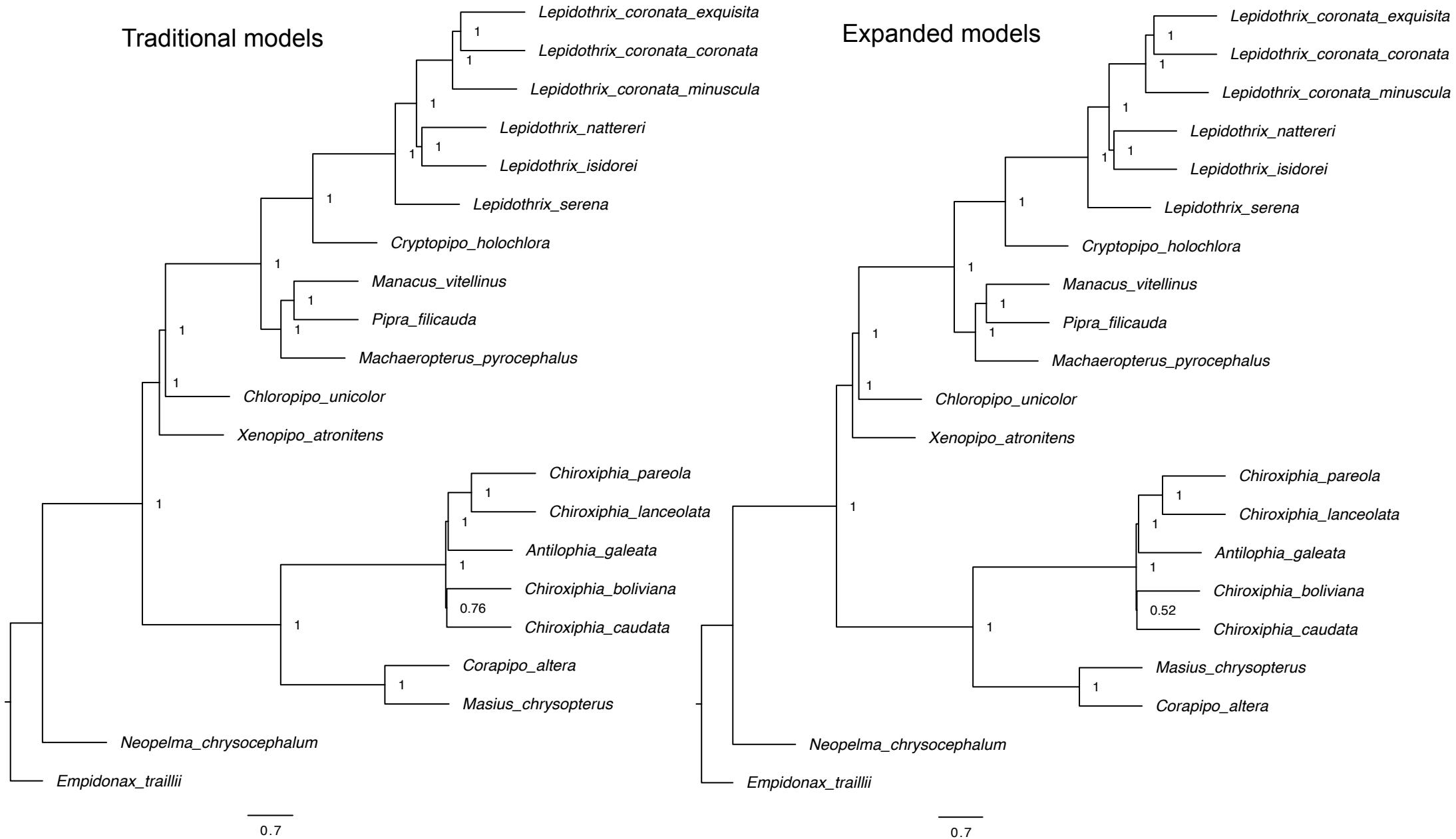


Figure S13. ASTRAL species trees using all UCEs with all loci constrained to be monophyletic for Antilophia+Chiroxiphia and Manacus+Pipra+Machaeropterus (constr2).

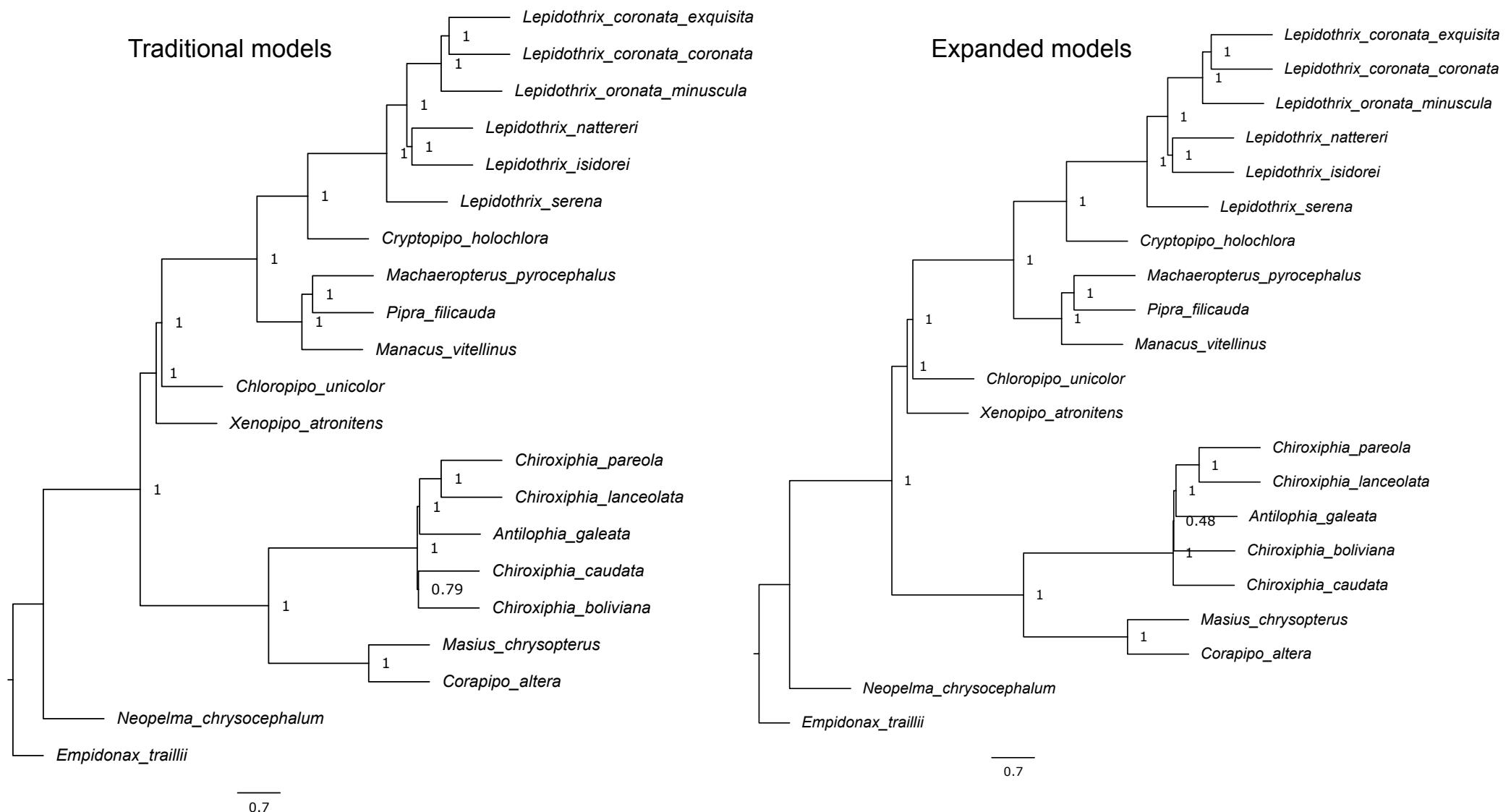


Figure S14. ASTRAL species trees using all introns.

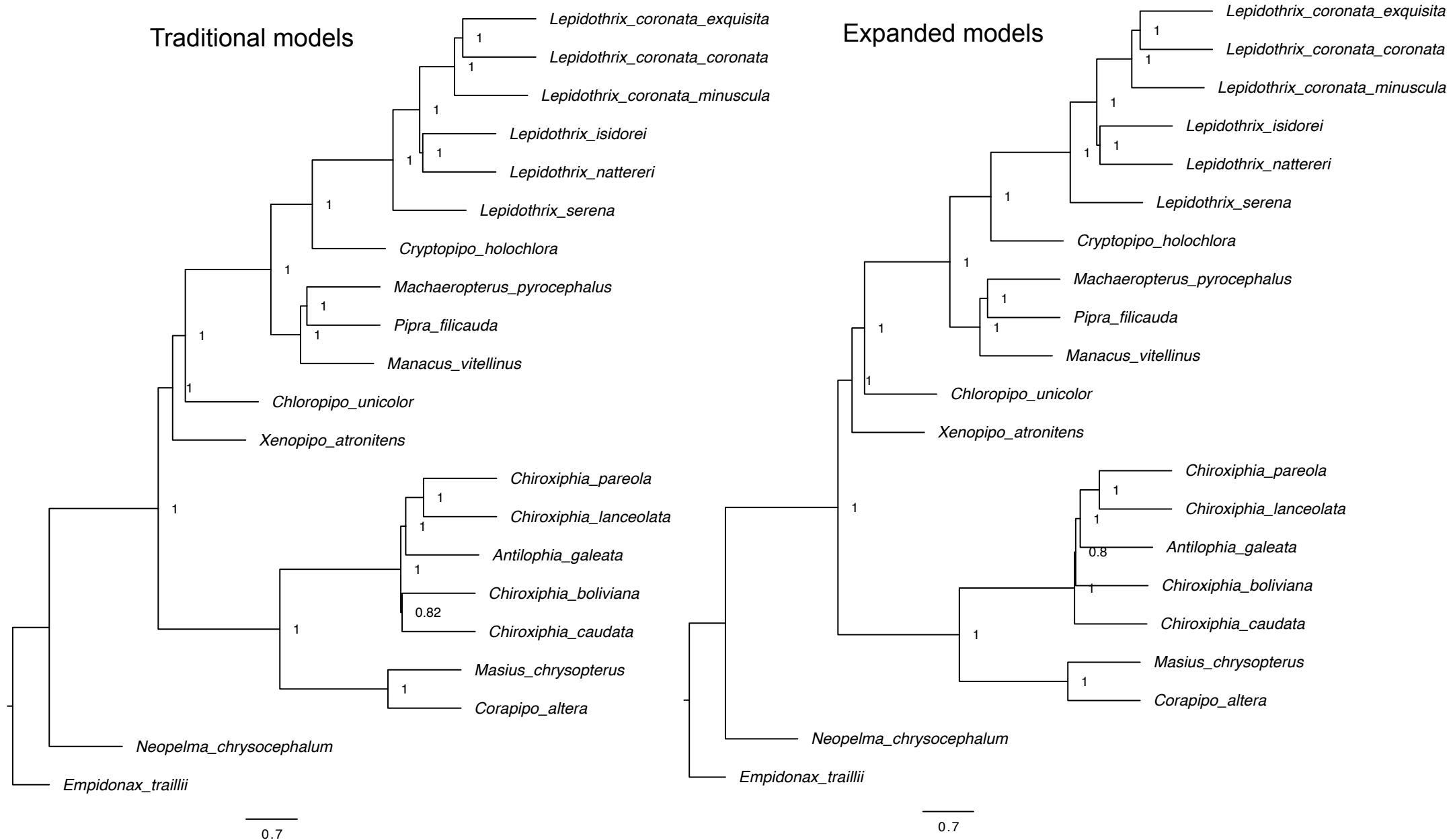
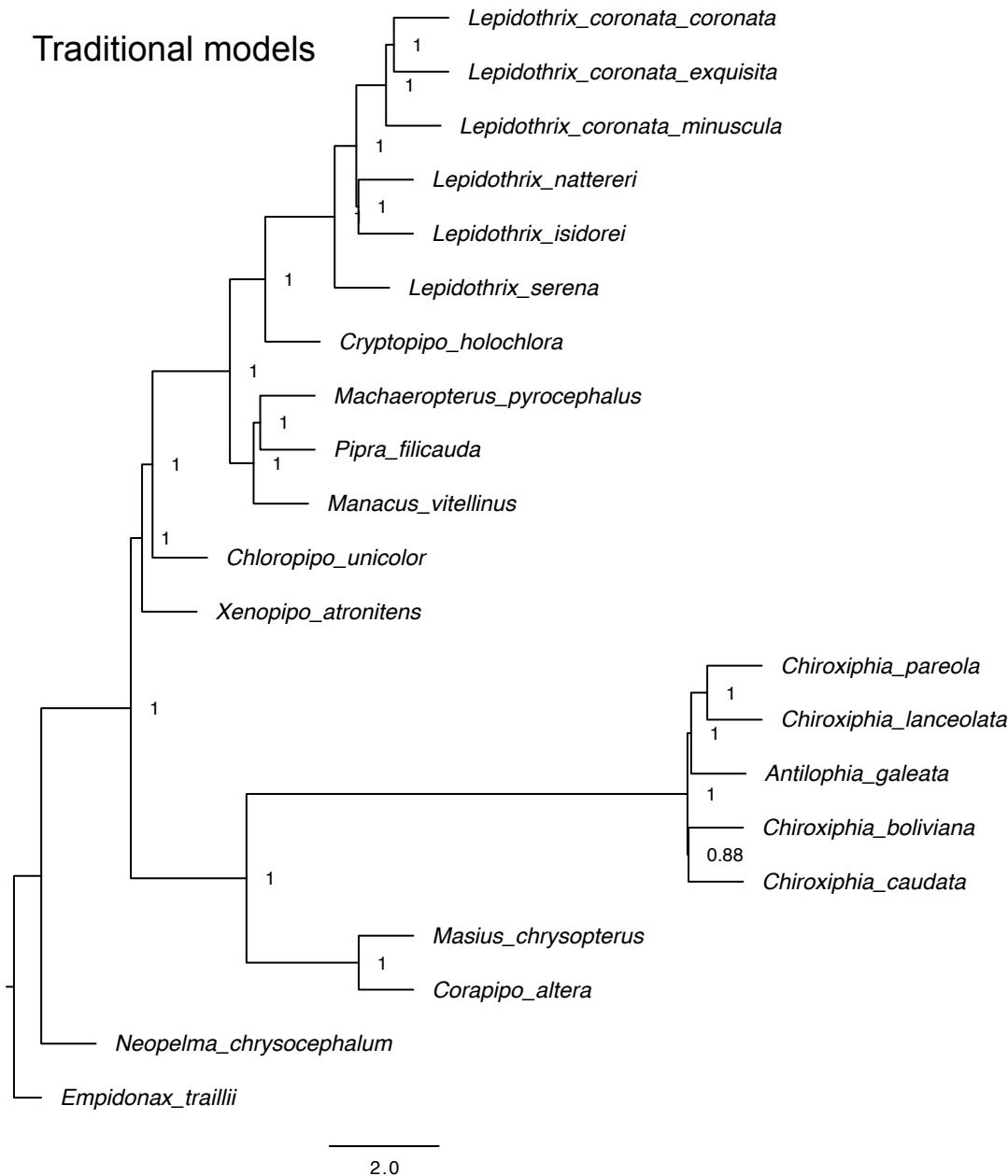


Figure S15. ASTRAL species trees using intron monophyletic loci.

Traditional models



Expanded models

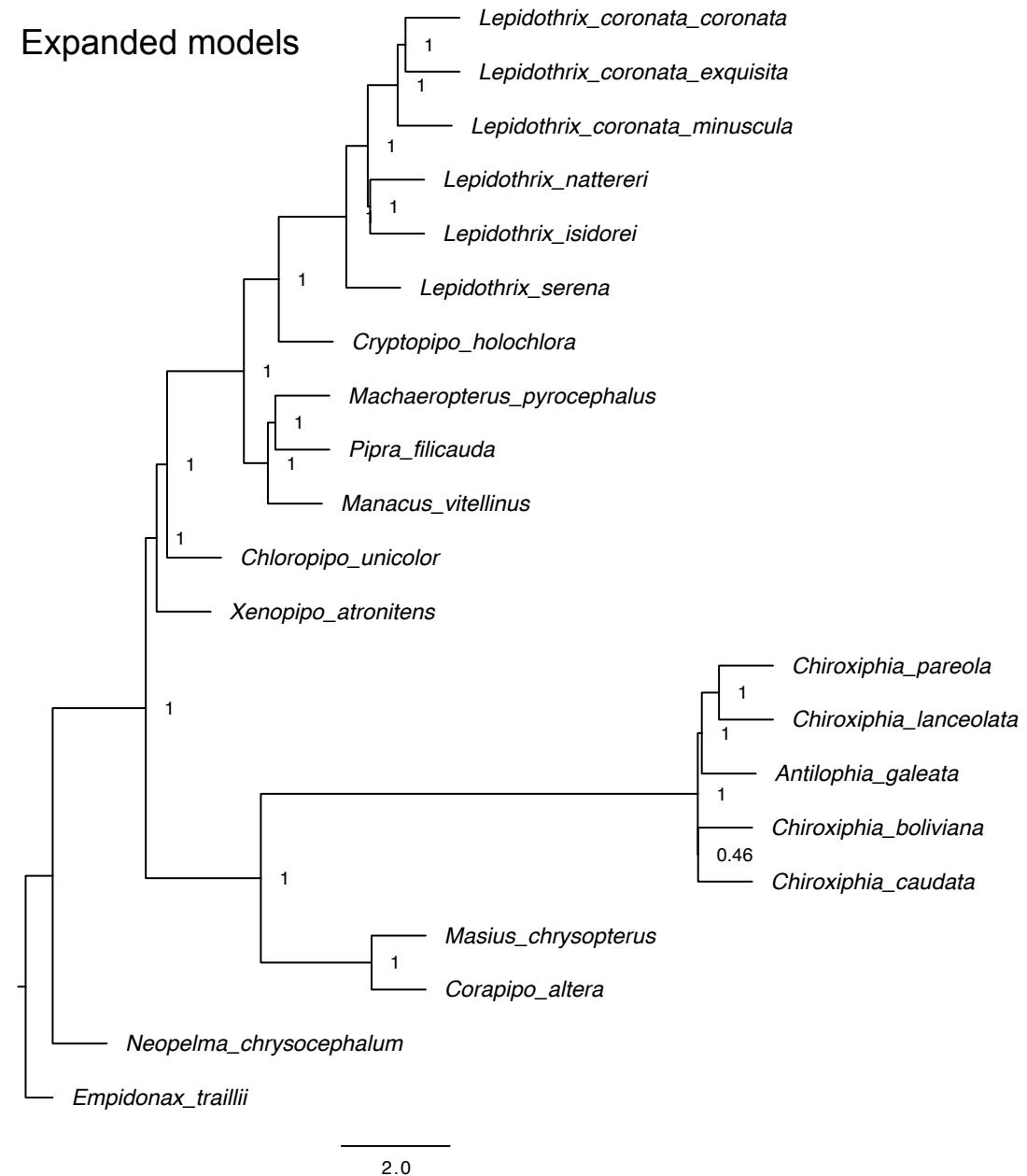


Figure S16. ASTRAL species trees using all introns with non-monophyletic loci constrained to be monophyletic for Antilophia+Chiroxiphia (constr1).

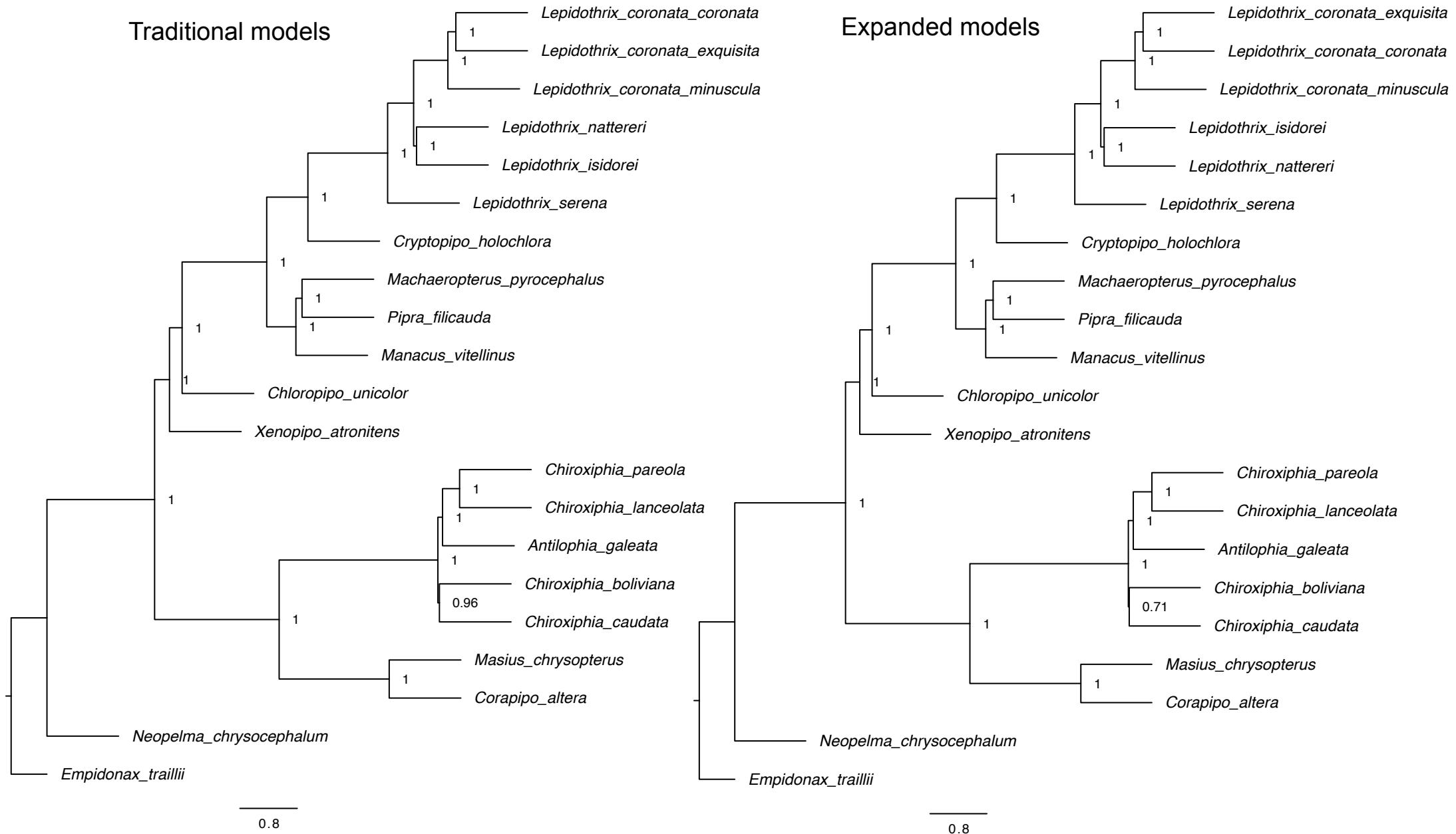


Figure S17. ASTRAL species trees using all introns with all loci constrained to be monophyletic for Antilophia+Chiroxiphia and Manacus+Pipra+Machaeropterus (constr2).

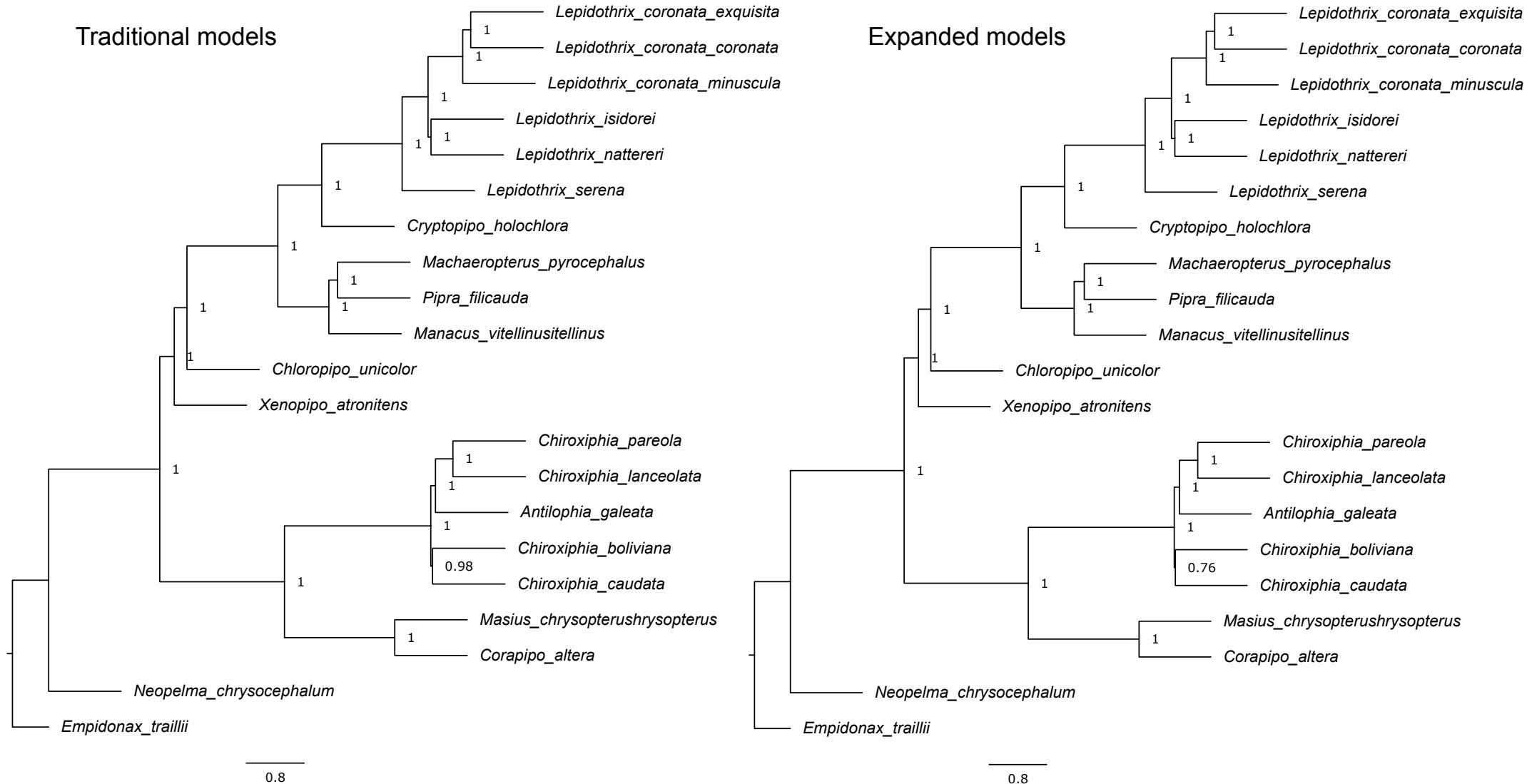


Figure S18. Here shows the gene concordance and discordance factors for the gene trees estimated under traditional models. gCF, gene concordance factor, is the percentage of decisive gene trees that contain a specific branch. gDFP is the gene discordance factor due to polyphyly, i.e., the percentage of gene trees that do not support monophyly for the clade that a specific branch connects. Node numbers were used to label a branch, for example, node 37 represents the branch from node 35 to 37. Data that were used to plot the bar charts can be found in Supplementary Table S3.

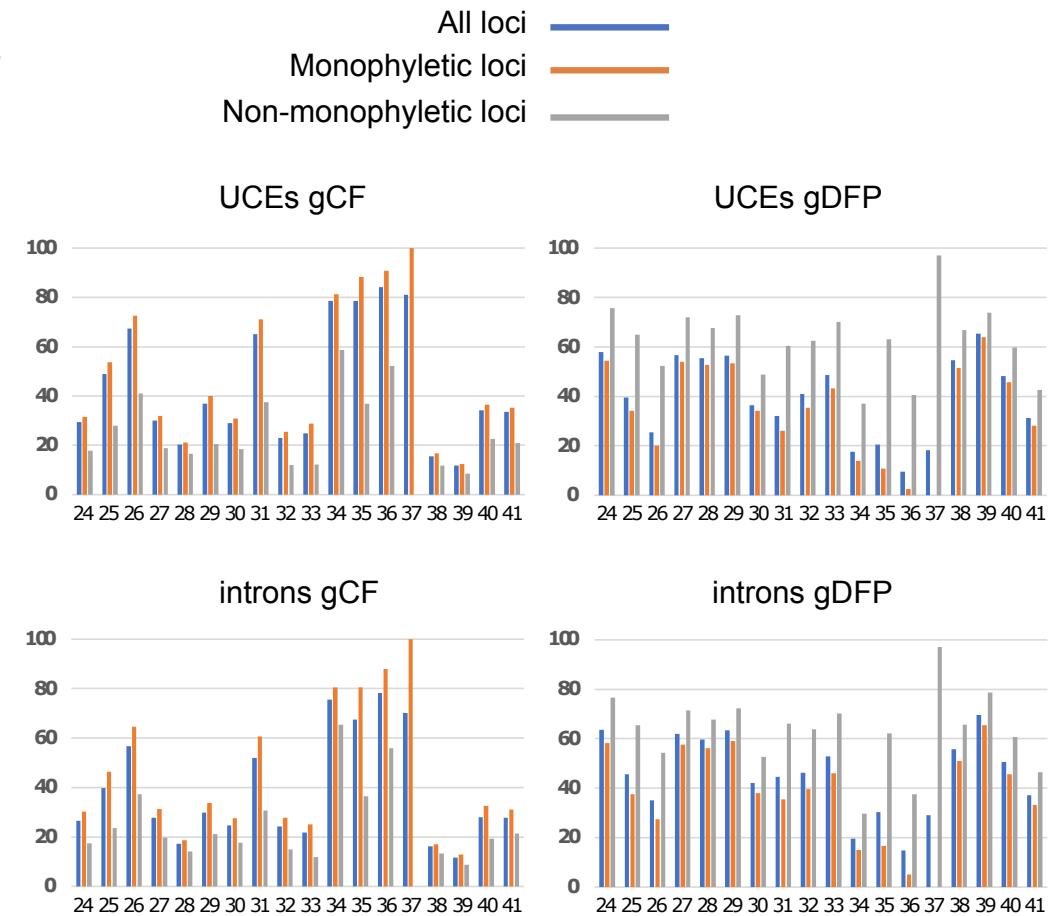
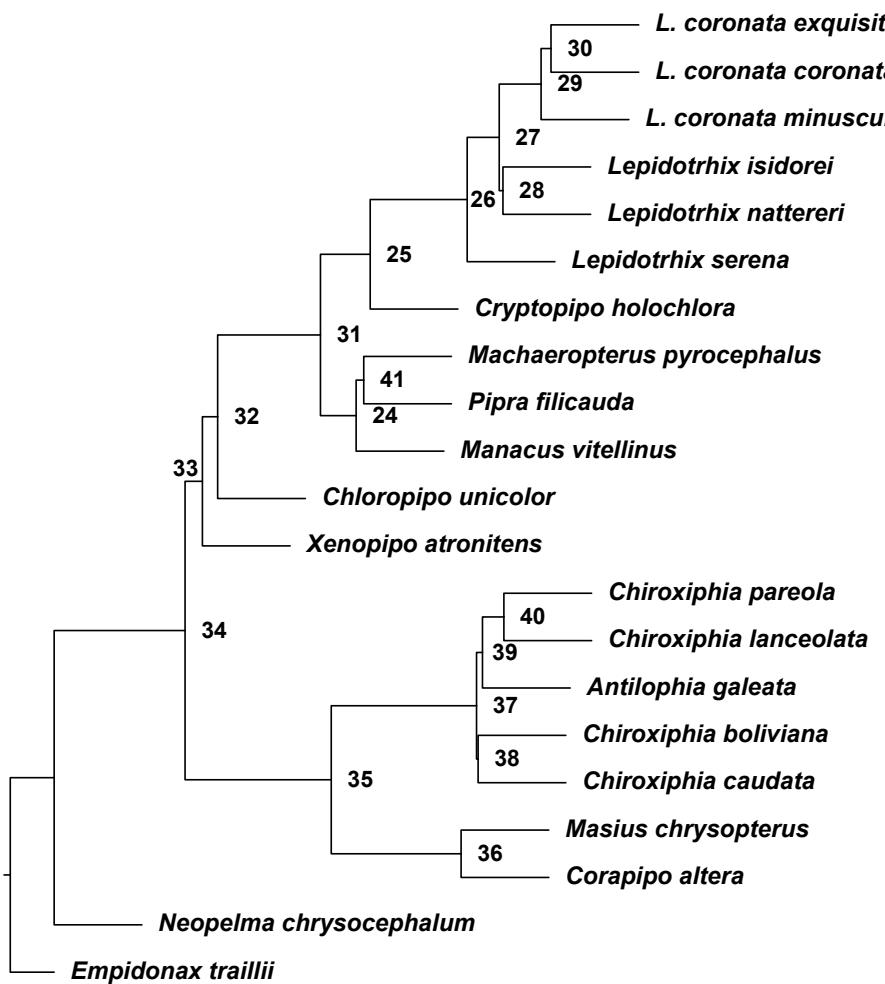


Figure S19. The top three most probable Bayesian MCMC phylogenetic networks within the 95% credible set for each dataset when the maximum number of reticulation events was set to be one. The pp values indicate posterior probabilities (rounded to two decimal places). The blue and red lines indicate possible reticulation event.

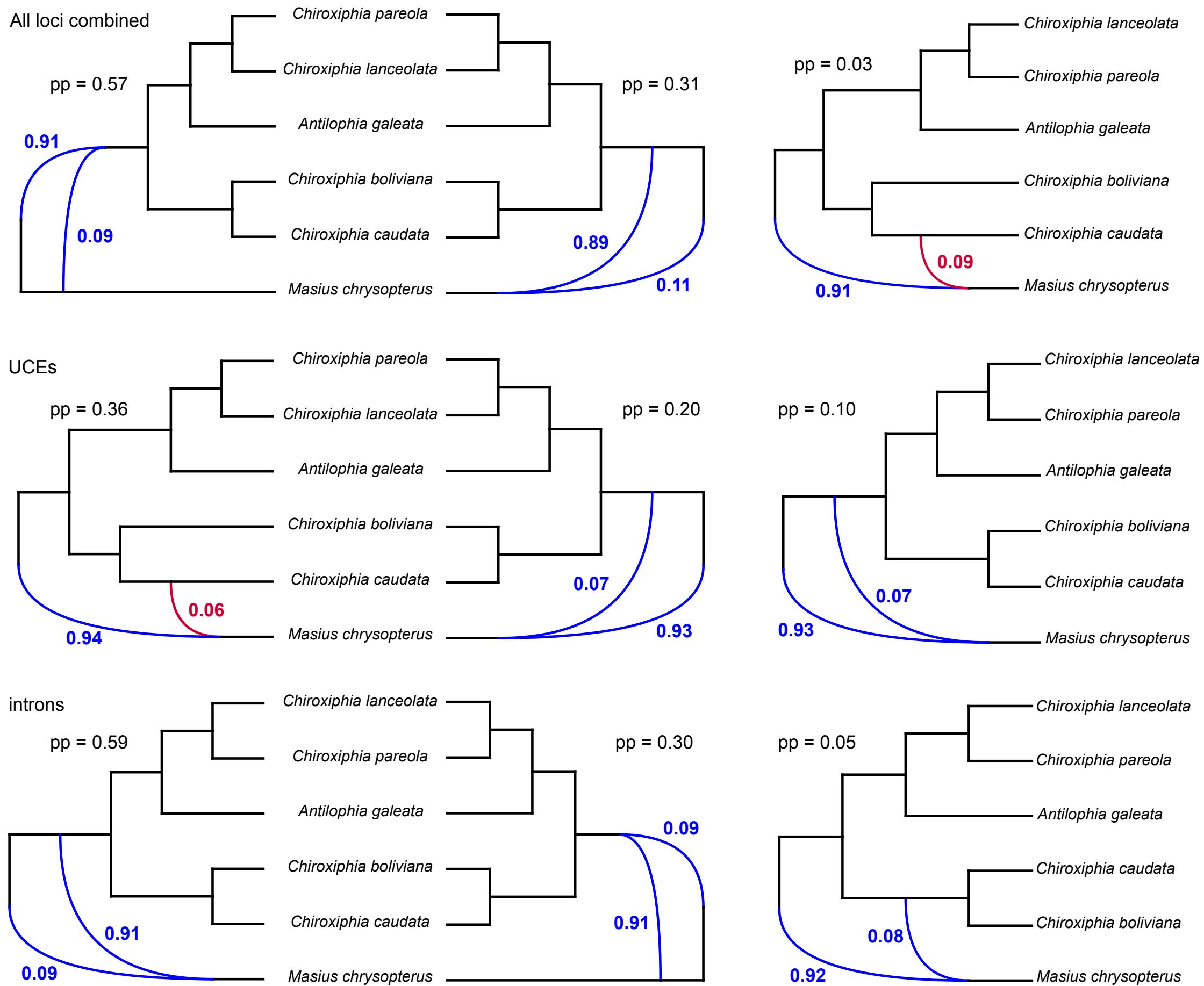


Figure S20. The top three most probable Bayesian MCMC phylogenetic networks within the 95% credible set for each dataset when the maximum number of reticulation events was set to be two. The pp values indicate posterior probabilities (rounded to two decimal places). The blue lines indicate possible reticulation event.

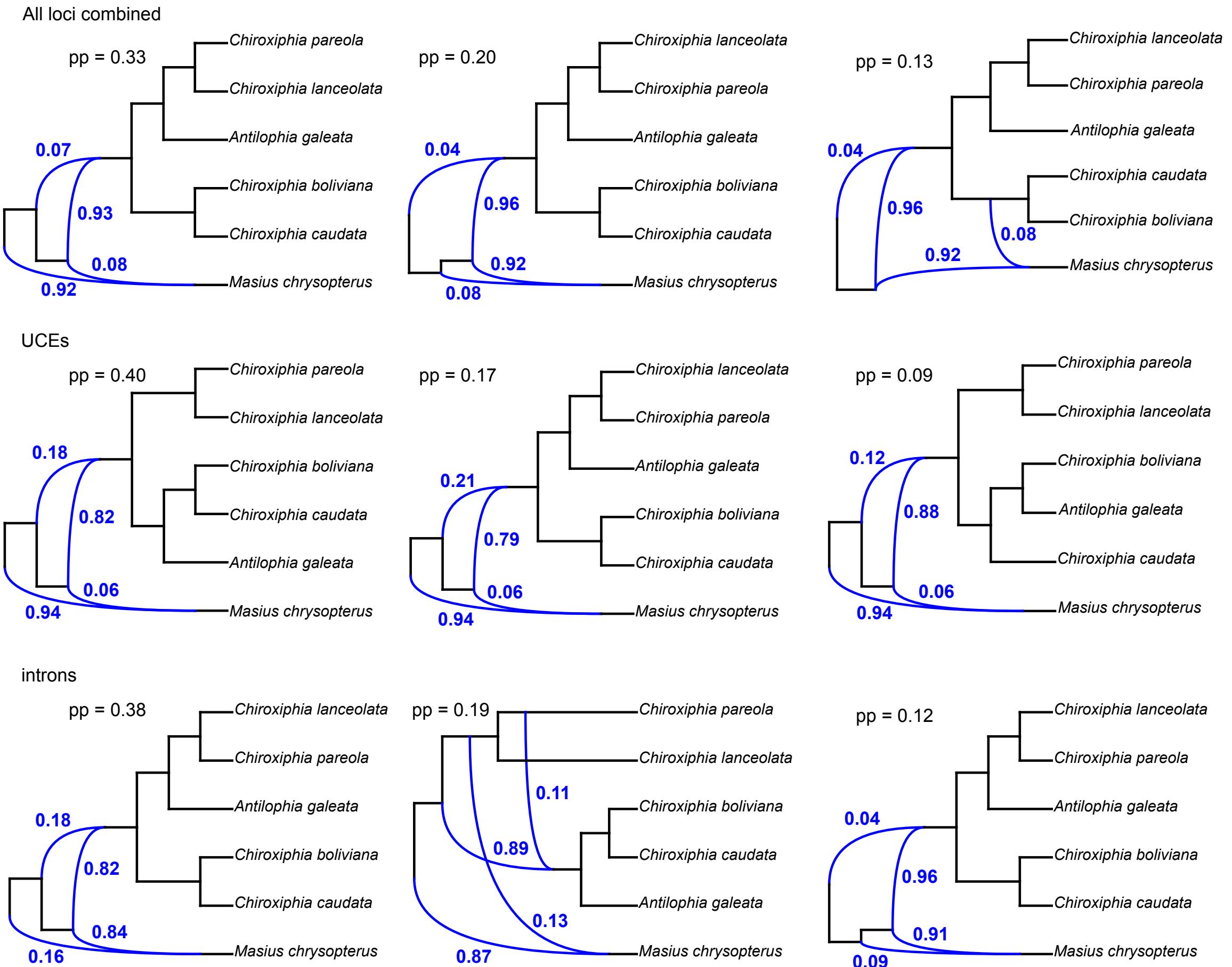
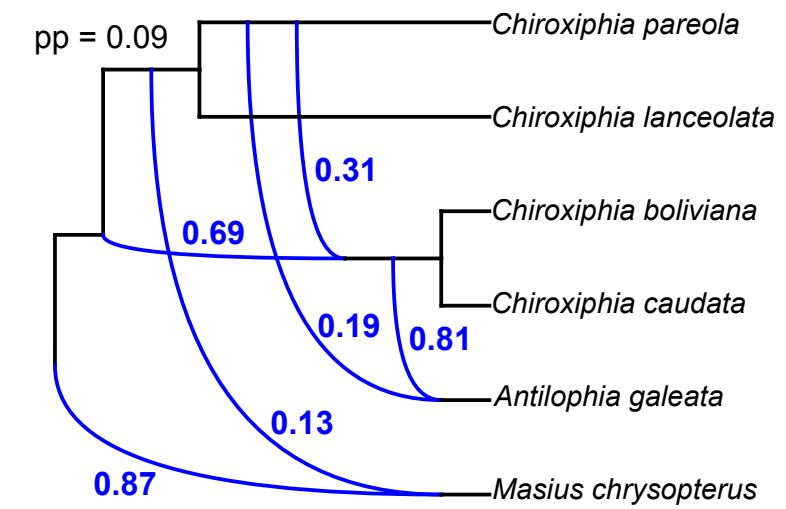
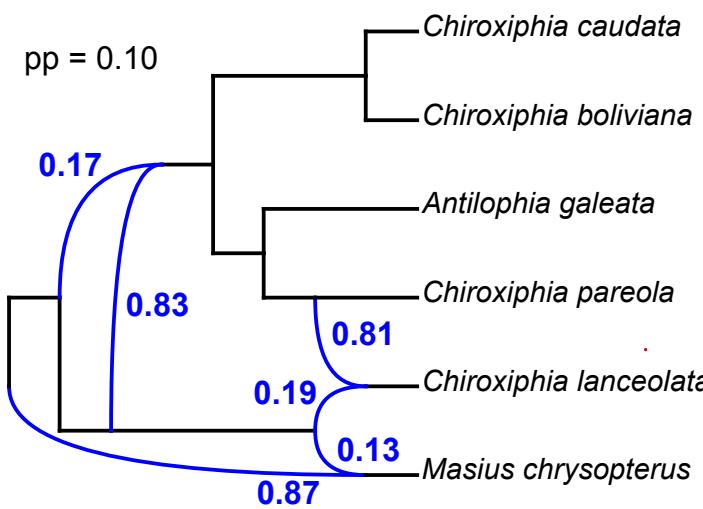
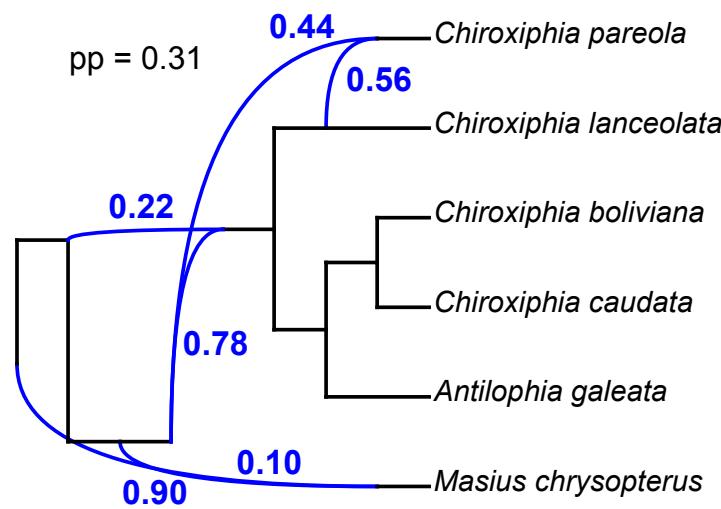
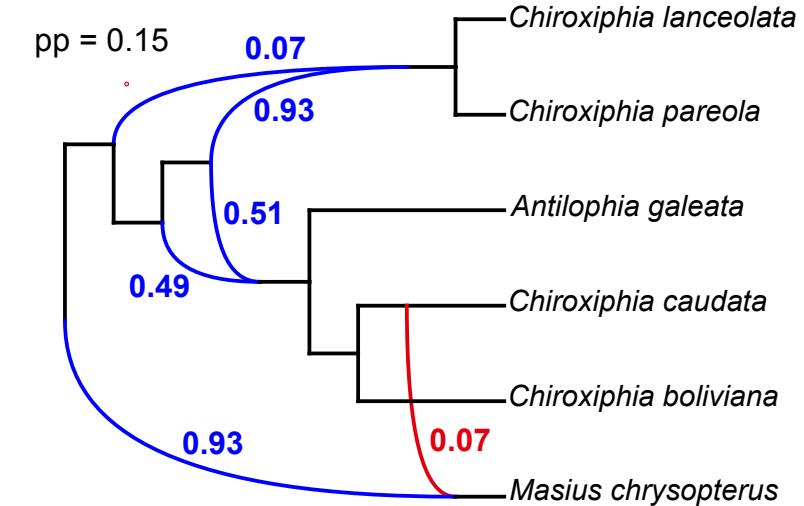
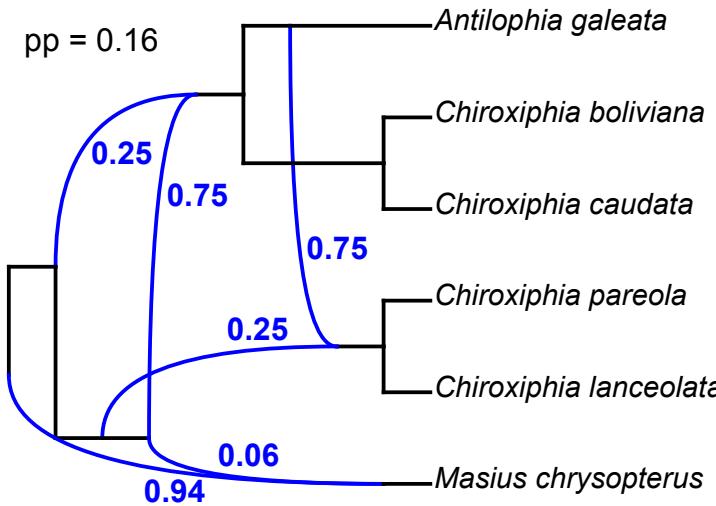
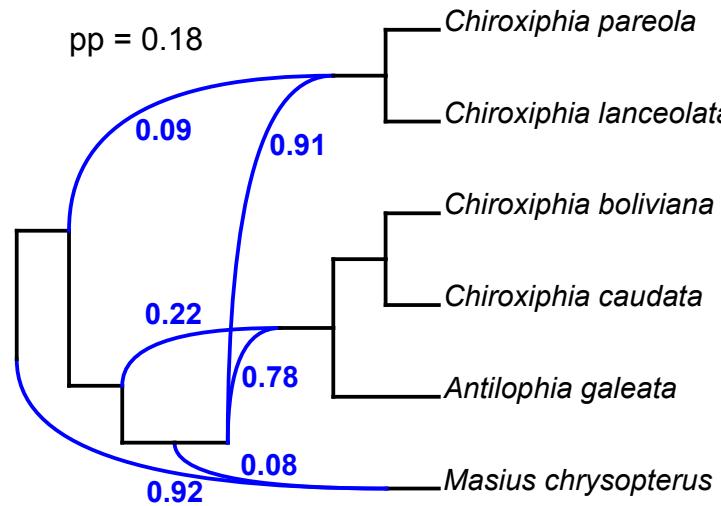


Figure S21. The top three most probable Bayesian MCMC phylogenetic networks within the 95% credible set for each dataset when the maximum number of reticulation events was set to be three. The pp values indicate posterior probabilities (rounded to two decimal places). The blue and red lines indicate possible reticulation event.

All loci combined



UCEs



introns

