***New Phytologist* Supporting Information**

Article title:

Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies

Authors:
Erik J.M. Koenen, Dario I. Ojeda, Royce Steeves, Jérémy Migliore, Freek T. Bakker, Jan J. Wieringa, Catherine Kidner, Olivier J. Hardy, R. Toby Pennington, Anne Bruneau and Colin E. Hughes

Article acceptance date: 14 September 2019.

The following Supporting Information is available for this article:

Table S1. Accession information for taxa included in the chloroplast alignment.

Table S2. Accession information for taxa included in the nuclear genomic and transcriptomic data set.

Table S3. Counts of bipartitions representing nodes A-H (Fig. 3) and conflicting bipartitions representing other subfamily relationships among 3,473 gene trees.

Figure S1. ML topology as inferred by RAxML from amino acid alignment of chloroplast genes under the LG4X model. Numbers on nodes indicate bootstrap percentages estimated from 1000 replicates.

Figure S2. Bayesian majority-rule consensus tree inferred with Phylobayes from amino acid alignment of chloroplast genes under the CATGTR model. Numbers on nodes indicate posterior probabilities (pp) from 9000 post-burn-in MCMC cycles.

Figure S3. ML topology as inferred by RAxML from nucleotide alignment of chloroplast genes under the GTR + G model. Numbers on nodes indicate bootstrap percentages estimated from 1000 replicates.

Figure S4. Bayesian majority-rule consensus tree inferred with Phylobayes from nucleotide alignment of chloroplast genes under the CATGTR model. Numbers on nodes indicate the posterior probabilities (pp) from 9000 post-burn-in MCMC cycles.

Figure S5. ML topology as inferred by RAxML from a concatenated alignment of 1,103 nuclear genes, under the LG4X model. Numbers on nodes indicate Internode Certainty All (ICA) values, as estimated from gene trees of the same 1,103 genes.

Figure S6. Bayesian gene jackknifing majority-rule consensus tree inferred with Phylobayes from a concatenated alignment of 1,103 nuclear genes. Numbers on nodes indicate posterior probabilities (pp), averaged over 500 posterior trees each, for 25 replicates (12,500 posterior trees in total).

Figure S7. Phylogeny estimated under the multi-species coalescent with ASTRAL. Support values indicated represent local posterior probability (blue rectangles) and quartet support (yellow rectangles).

Supporting Information Notes S1. Python script for counting bipartitions.

Dataset S1. The concatenated nucleotide alignment of 72 protein-coding genes from the chloroplast genome in NEXUS format.

Dataset S2. The concatenated amino acid alignment of 72 protein-coding genes from the chloroplast genome in NEXUS format.

Dataset S3. ZIP file containing amino acid alignments of 1,103 1-to-1 orthologs in NEXUS format.

Dataset S4. ZIP file containing amino acid alignments of 7,621 RT homologs in NEXUS format.

Dataset S5. ZIP file containing 1,103 gene trees estimated from 1-to-1 orthologs in newick format, with bootstrap values and branch lengths.

Dataset S6. ZIP file containing 7,621 gene trees estimated from RT homologs in newick format, with bootstrap values and branch lengths.

Table S1. Accession information for the taxa included in the chloroplast alignment.

Taxon	Herbarium voucher	Genbank accession number	Comments
Abarema jupunba	M.F. Simon 1600 (CEN)	MN579657, MN579692, MN579725, MN579759, MN579793, MN579859, MN579894, MN579914, MN590534, MN590568, MN590598, MN590632, MN590666, MN590697-MN590744	**Newly sequenced**
Acacia koa		TSA: GBYE00000000	Transcriptome
Acacia ligulata		LN555649.2	
Acrocarpus fraxinifolius			Transcriptome, available at https://ics.hutton.ac.uk/tropiTree/
Adenanthera pavonina	Ambriansyah & Arifin AA295 (K)	KX302290, MN579658, MN579693, MN579726, MN579760, MN579794, MN579827, MN579860, MN579915, MN590535, MN590569, MN590599, MN590633, MN590667, MN590745-MN590800	**Newly sequenced**
Afzelia africana	S.L.A. Donkpegan 27 (BRLU)	KX673213	
Afzelia bipindensis	S.L.A. Donkpegan 626 (BRLU)	MN540450	**Newly sequenced**
Ajuga reptans		KF709391	
Albizia adianthifolia	J.J. Wieringa 6278 (WAG)	KX302294, MN579659, MN579694, MN579727, MN579761, MN579795,	**Newly sequenced**

		MN579828, MN579861, MN579916, MN590536, MN590570, MN590600, MN590634, MN590668, MN590801-MN590857	
Albizia julibrissin	E. Koenen 601 (Z)	TSA: GHWM00000000	**Newly sequenced**; Transcriptome
Anthonotha fragrans		Bioproject: PRJNA472454	**Newly sequenced**; Transcriptome
Apios americana		KF856618	Different accession from the nuclear data set
Arabidopsis thaliana		AP000423	
Arachis hypogaea		KJ468094	
Arachis ipaensis		TSA: GBIW00000000	Transcriptome
Aralia undulata	R. Li 551 (KUN)	KC456163	
Archidendron lucidum	Wang & Lin 2534 (L)	KX302315, MN579660, MN579695, MN579728, MN579762, MN579796, MN579829, MN579862, MN579917, MN590537, MN590571, MN590601, MN590635, MN590669, MN590858-MN590911	**Newly sequenced**
Astragalus membranaceus			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Astragalus propinquus			Transcriptome, OneKP: MYMP, available at http://www.onekp.com/public_data.html
Azadirachta indica		KF986530	
Bauhinia tomentosa			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Bituminaria bituminosa			Transcriptome, OneKP: TVSH, available at http://www.onekp.com/public_data.html

Bulnesia arborea	M.J. Moore 334 (FLAS)	EU002159, EU002172, EU002205, EU002275, EU002299, EU002388, EU002478, GQ998005-GQ998073, HQ664597	
Buxus microphylla		EF380351	
Calliandra hygrophila	L.P. Queiroz 15542 (HUEFS)	KX302321, MN579661, MN579729, MN579763, MN579797, MN579830, MN579863, MN579918, MN590538, MN590602, MN590636, MN590670, MN590912-MN590963	**Newly sequenced**
Carica papaya		EU431223	
Ceratonia siliqua		KJ468096	
Cercis canadensis		KF856619	Different accession from the nuclear data set
Chamaecrista fasciculata			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Chidlowia sanguinea	J.J. Wieringa 4338 (WAG)	MN579662, MN579730, MN579764, MN579798, MN579831, MN579864, MN579895, MN579919, MN590539, MN590572, MN590603, MN590637, MN590671, MN590964-MN591018	**Newly sequenced**
Chrysobalanus icaco		KJ414480	
Cicer arietinum		EU835853	
Citrus sinensis		DQ864733	
Cladrastis lutea			Transcriptome, available at

			http://dx.doi.org/10.5061/dryad.ff1tq
Codariocalyx motorius			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Coffea arabica		EF044213	
Cojoba arborea	M.F. Simon 1545 (CEN)	KX302324, MN579663, MN579696, MN579731, MN579765, MN579799, MN579865, MN579920, MN590540, MN590573, MN590604, MN590638, MN590672, MN591019-MN591074	**Newly sequenced**
Colvillea racemosa	Kew living collection 1993-224 (K)	MN551236	**Newly sequenced**
Copaifera officinalis			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Cucumis sativus		AJ970307	
Daucus carota		DQ898156	Hoi – hoe
Desmanthus illinoensis			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Dialium guineense	T. van Andel 4184 (WAG)	KX302327, MN579664, MN579697, MN579732, MN579766, MN579800, MN579832, MN579866, MN579921, MN590541, MN590574, MN590605, MN590639, MN590673, MN591075-MN591130	**Newly sequenced**
Dichrostachys cinerea	O. Maurin 256 (JRAU)	KX302328, MN579665, MN579698, MN579733, MN579767, MN579801, MN579833, MN579867,	**Newly sequenced**

		MN579922, MN590542, MN590575, MN590606, MN590640, MN590674, MN591131-MN591185	
Dimorphandra macrostachya	J.R. Iganҫi 877 (RB)	KY046043, MN579666, MN579699, MN579734, MN579768, MN579802, MN579834, MN579868, MN579923, MN590543, MN590576, MN590607, MN590641, MN590675, MN591186-MN591241	**Newly sequenced**
Diptychandra aurantiaca	J.R.I. Wood 26513 (K)	MN579667, MN579700, MN579735, MN579769, MN579803, MN579835, MN579869, MN579896, MN579924, MN590544, MN590577, MN590608, MN590642, MN590676, MN591242-MN591298	**Newly sequenced**
Distemonanthus benthamianus	G. Dauby 728 (BRLU)	MN604403	**Newly sequenced**
Duparquetia orchidacea	J.J. Wieringa 7805 (L)	MN579668, MN579701, MN579736, MN579770, MN579804, MN579836, MN579870, MN579897, MN579925, MN590545, MN590609, MN590643, MN591299-MN591355	**Newly sequenced**
Entada abyssinica	MSB 0133199 (K)	TSA: GHWN00000000	**Newly sequenced**; Transcriptome
Entada rheedei	E. Koenen 496 (Z)	MN579669, MN579702, MN579737, MN579771, MN579805, MN579837,	**Newly sequenced**

		MN579871, MN579898, MN579926, MN590546, MN590578, MN590610, MN590644, MN590677, MN591356-MN591412	
Erythrophleum ivorense	J.J. Wieringa 5487 (WAG)	MN579670, MN579703, MN579738, MN579772, MN579806, MN579838, MN579872, MN579899, MN579927, MN590547, MN590579, MN590611, MN590645, MN590678, MN591413-MN591468	**Newly sequenced**
Erythrostemon gilliesii	R. Steeves 852 (MT)	MN551237	**Newly sequenced**
Eucalyptus grandis		HM347959	
Euonymus americanus	W. Judd 8071 (FLAS)	EU002160, EU002170, EU002193, EU002277, EU002321, EU002409, EU002500, GQ998147-GQ998219, HQ664608	
Fagopyrum esculentum		EU254477	
Faidherbia albida	O. Maurin 3495 (JRAU)	MN579671, MN579704, MN579739, MN579773, MN579807, MN579839, MN579873, MN579900, MN579928, MN590548, MN590580, MN590612, MN590646, MN590679, MN591469-MN591522	**Newly sequenced**
Garcinia mangostana		HQ331601, HQ331906, HQ332057, HQ848709, JX661816, JX661859,	

JX661902, JX661944,
JX661980, JX662020,
JX662065, JX662109,
JX662151, JX662196,
JX662237, JX662279,
JX662320, JX662359,
JX662399, JX662434,
JX662467, JX662502,
JX662543, JX662580,
JX662622, JX662666,
JX662710, JX662752,
JX662799, JX662841,
JX662880, JX662914,
JX662955, JX662996,
JX663032, JX663071,
JX663104, JX663149,
JX663196, JX663237,
JX663280, JX663322,
JX663365, JX663410,
JX663583, JX663630,
JX663677, JX663721,
JX663763, JX663804,
JX663841, JX663874,
JX663915, JX663962,
JX664006, JX664049,
JX664091, JX664127,
JX664165, JX664209,
JX664252, JX664297,
JX664341, JX664385,
JX664458, JX664495,
JX664535, JX664580,
JX664623, JX664659,
JX664694, JX664726,
JX664771, JX664812,
JX664852, JX664895,
JX664939, JX665004,

		KF783277, U92876, U92877, U92878	
Gleditsia sinensis			Transcriptome, OneKP: VHZV, available at http://www.onekp.com/public_data.html
Gleditsia triacanthos			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Gliricidia sepium			Transcriptome, available at https://ics.hutton.ac.uk/tropiTree/
Glycine canescens		KC893635	
Glycine max		DQ317523	
Glycyrrhiza glabra		KF201590	
Glycyrrhiza lepidota			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Gompholobium polymorphum			Transcriptome, OneKP: VLNB, available at http://www.onekp.com/public_data.html
Gossypium hirsutum		DQ345959	
Guibourtia ehie	F. Tosso 272 (BRLU)	MN540451	**Newly sequenced**
Guibourtia tessmannii		MN540452	**Newly sequenced**
Guilfoylia monostylis	P.I. Forster 28103 (Z)	MN579672, MN579705, MN579740, MN579774, MN579808, MN579840, MN579874, MN579901, MN579929, MN590549, MN590581, MN590613, MN590647, MN590680, MN591523-MN591579	**Newly sequenced**
Gymnocladus dioicus			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Haematoxylum brasiletto		KJ468097	
Helianthus annuus		DQ383815	

Hymenostegia brachyura	Zenker 4481 (WAG)	MN579673, MN579706, MN579741, MN579809, MN579841, MN579875, MN579902, MN579930, MN590550, MN590614, MN590648, MN591580-MN591626	**Newly sequenced**
Hymenostegia felicis	Jacques-Félix 5129 (WAG)	MN579674, MN579707, MN579742, MN579775, MN579876, MN579931, MN590551, MN590615, MN590649, MN591627-MN591667	**Newly sequenced**
Indigofera tinctoria		KJ468098	
Inga leiocalycina	T.D. Pennington 13822 (K)	KT428296	
Inga spectabilis	T.D. Pennington 15061 (K)	MN579675, MN579708, MN579743, MN579776, MN579810, MN579842, MN579877, MN579903, MN579932, MN590552, MN590582, MN590616, MN590650, MN590681, MN591668-MN591725	**Newly sequenced**
Intsia bijuga		KX673214	
Lathyrus graminifolius		KJ806193	
Lathyrus sativus		HM029371	Different accession from the nuclear data set
Lens culinaris		KF186232	Different accession from the nuclear data set
Leucaena trichandra		KT428297	
Libidibia coriaria		KJ468095	
Lotus japonicus		AP002983	
Lupinus luteus		KC695666	

Lupinus polyphyllus			Transcriptome, OneKP: CMFF, available at http://www.onekp.com/public_data.html
Macadamia integrifolia		KF862711	
Manihot esculenta		EU117376	
Medicago hybrida		KJ850240	
Medicago truncatula		AC093544	
Microlobius foetidus	C.E. Hughes 1219 (FHO)	TSA: GHWO00000000	**Newly sequenced**; Transcriptome
Millettia pinnata		JN673818	
Mimosa tenuiflora	L.P. Queiroz 15498 (HUEFS)	MN579676, MN579709, MN579744, MN579777, MN579811, MN579843, MN579878, MN579904, MN579933, MN590553, MN590583, MN590617, MN590651, MN590682, MN591726-MN591782	**Newly sequenced**
Morus indica		DQ226511	
Nelumbo nucifera		JQ336993	Different accession from the nuclear data set
Nerium oleander	W. Judd 8076 (FLAS)	KJ953907	
Newtonia hildebrandtii	O. Maurin 2457 (JRAU)	HM386340, MN579677, MN579710, MN579745, MN579778, MN579812, MN579844, MN579879, MN579934, MN590554, MN590584, MN590618, MN590652, MN590683, MN591783-MN591840	**Newly sequenced**
Oenothera biennis		EU262889	
Olea europaea		GU228899	
Oxalis latifolia	M.J. Moore 316 (FLAS)	EU002165, EU002186,	

		EU002248, EU002282, EU002350, EU002438, EU002528, GQ998511-GQ998580, HQ664602, KF783277, U92876, U92877, U92878	
Pachyelasma tessmannii	J.J. Wieringa 5229 (WAG)	MN579678, MN579711, MN579746, MN579779, MN579813, MN579845, MN579880, MN579905, MN579935, MN590555, MN590585, MN590619, MN590653, MN590684, MN591841-MN591898	**Newly sequenced**
Pachyrhizus erosus		KJ468100	
Paeonia obovata		KJ206533	
Parkia panurensis	J.R. Iganҫi 842 (RB)	KY046211, MN579679, MN579712, MN579747, MN579780, MN579814, MN579846, MN579881, MN579936, MN590556, MN590586, MN590620, MN590654, MN590685, MN591899-MN591952	**Newly sequenced**
Pelargonium alternans		KF240617	
Peltophorum africanum	Koenen 601 (Z)	KX302342, MN579680, MN579713, MN579748, MN579781, MN579815, MN579847, MN579882, MN579937, MN590557, MN590587, MN590621, MN590655, MN590686,	**Newly sequenced**

		MN591953-MN592008	
Pentaclethra macrophylla	Galeuchet & Balthazar 10 (Z)	KX302343, MN579681, MN579714, MN579749, MN579782, MN579816, MN579848, MN579883, MN579938, MN590558, MN590588, MN590622, MN590656, MN590687, MN592009-MN592062	**Newly sequenced**
Phaseolus vulgaris		DQ886273	
Piptadeniastrum africanum	E. Koenen 152 (WAG)	MN579682, MN579715, MN579750, MN579783, MN579817, MN579849, MN579884, MN579906, MN579939, MN590559, MN590589, MN590623, MN590657, MN590688, MN592063-MN592117	**Newly sequenced**
Pisum sativum		HM029370	Different accession from the nuclear data set
Pithecellobium dulce	B. Marazzi BM309 (ARIZ)	MN579683, MN579716, MN579751, MN579784, MN579818, MN579850, MN579885, MN579907, MN579940, MN590560, MN590590, MN590624, MN590658, MN590689, MN592118-MN592174	**Newly sequenced**
Poeppigia procera	Hernández 558 (Z)	MN579684, MN579717, MN579752, MN579785, MN579819, MN579851, MN579886, MN579908, MN579941, MN590561, MN590591, MN590625,	**Newly sequenced**

		MN590659, MN590690, MN592175-MN592230	
Polygala lutea		EF489041	Different accession from the nuclear data set
Populus trichocarpa		KF753634	
Primula poissonii		KF753634	
Prioria balsamifera		Bioproject: PRJNA472454	**Newly sequenced**; Transcriptome
Prosopis alba		TSA: GAOO00000000	Transcriptome
Prosopis glandulosa		KJ468101	
Prunus persica		HQ336405	
Quercus rubra		JX970937	
Quillaja saponaria			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Ranunculus macranthus		DQ359689	
Robinia pseudoacacia		KJ468102	
Samanea saman	C.E. Hughes 421 (FHO)	MN579685, MN579718, MN579753, MN579786, MN579820, MN579852, MN579887, MN579909, MN579942, MN590562, MN590592, MN590626, MN590660, MN590691, MN592231-MN592288	**Newly sequenced**
Sapindus mukorossi		KM454982	
Saraca indica	Kew living collection 2011-1421 (K)	MN551238	**Newly sequenced**
Schotia brachypetala	R. Steeves 846 (MT)	MN551239	**Newly sequenced**
Sedum sarmentosum		JX427551	
Senegalia ataxacantha	C. Jongkind 10603 (WAG)	KX302351, MN579686,	**Newly sequenced**

		MN579719, MN579754, MN579787, MN579821, MN579853, MN579888, MN579943, MN590563, MN590593, MN590627, MN590661, MN590692, MN592289-MN592346	
Senna hebecarpa			Transcriptome, available at http://dx.doi.org/10.5061/dryad.ff1tq
Senna siamea			Transcriptome, available at https://ics.hutton.ac.uk/tropiTree/
Sesbania macrantha			Transcriptome, available at https://ics.hutton.ac.uk/tropiTree/
Sesbania sesban			Transcriptome, available at https://ics.hutton.ac.uk/tropiTree/
Silene latifolia		JF715055	
Solanum lycopersicum		KP331414	
Styphnolobium japonicum			Transcriptome, available at https://www.hindawi.com/journals/bmri/2014/750961/sup/
Swartzia emarginata	M.P. Morim 576 (RB)	MN579687, MN579720, MN579755, MN579788, MN579822, MN579854, MN579889, MN579910, MN579944, MN590564, MN590594, MN590628, MN590662, MN590693, MN592347-MN592401	**Newly sequenced**
Tachigali odoratissima	M.P. Morim 562 (RB)	KY045941, MN579688, MN579721, MN579756, MN579789, MN579823, MN579855, MN579890,	**Newly sequenced**

		MN579945, MN590565, MN590595, MN590629, MN590663, MN590694, MN592402-MN592458	
Tamarindus indica		KJ468103	
Theobroma cacao		HQ244500	
Tipuana tipu			Transcriptome, available at https://ics.hutton.ac.uk/tropiTree/
Trachelium caeruleum		EU090187	
Trifolium aureum		KC894708	
Trifolium repens		KC894706	
Trochodendron aralioides		KC608753	
Vaccinium macrocarpon		JQ757046	
Vachellia tortilis	E. Koenen 603 (Z)	MN579689, MN579722, MN579757, MN579790, MN579824, MN579856, MN579891, MN579911, MN579946, MN590566, MN590596, MN590630, MN590664, MN590695, MN592459-MN592510	**Newly sequenced**
Vicia faba		KF042344	Different accession from the nuclear data set
Vicia sativa		KJ850242	
Vigna radiata		GQ893027	
Vigna unguiculata		JQ755301	
Vitis vinifera		DQ424856	
Wisteria floribunda			Transcriptome, OneKP: RMWJ, available at http://www.onekp.com/public_data.html
Xanthocercis zambesiaca			Transcriptome, available at

			http://dx.doi.org/10.5061/dryad.ff1tq
Xanthophyllum eurhynchum	P. Herendeen H.416 (F)	MN579690, MN579723, MN579758, MN579791, MN579825, MN579857, MN579892, MN579912, MN579947, MN590567, MN590597, MN590631, MN590665, MN590696, MN592511-MN592563	**Newly sequenced**
Xylia hoffmannii	E. Koenen 402 (Z)	MN579691, MN579724, MN579792, MN579826, MN579858, MN579893, MN579913, MN592564-MN592619	**Newly sequenced**
Zenia insignis	Averyanov et al. 5748 (MT)	MN551240	**Newly sequenced**; Different accession from the nuclear data set

Table S2. Accession information for the taxa included in the nuclear genomic and transcriptomic data set.

Taxon	Source	Citation
Acacia koa	Genbank BioProject: PRJNA268386	Ishihara et al.
Acrocarpus fraxinifolius	TropiTree: https://ics.hutton.ac.uk/tropiTree/	Russel et al. 2014
Afzelia bella	Genbank BioProject: PRJNA472454	**Newly sequenced**
Albizia julibrissin	Genbank BioProject: PRJNA574148	**Newly sequenced**
Alnus serrulata	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Amaranthus hypochondriacus	Phytozome v11: v1.0	Clouse et al. 2016
Anthonotha fragrans	Genbank BioProject: PRJNA472454	**Newly sequenced**
Apios americana	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Aquilegia coerulea	Phytozome v11: v3.1	Filiault et al. 2018
Arabidopsis thaliana	Phytozome v11: TAIR10	Lamesch et al. 2012
Arachis ipaensis	Peanutbase.org: K30076.a1.M1	Bertioli et al. 2016
Astragalus membranaceus	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Bauhinia tomentosa	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Bituminaria bituminosa	OneKP: TVSH	Wicket et al. 2014
Cajanus cajan	http://gigadb.org/dataset/100028	Varshney et al. 2012
Cannabis sativa	Genbank BioProject: PRJNA74271	van Bakel et al. 2011
Carica papaya	Phytozome v11: ASGPBv4.0	Ming et al. 2008
Castanea mollissima	https://www.hardwoodgenomics.org/Genome-assembly/1962958	*Not available*
Cercis canadensis	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Chamaecrista fasciculata	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Cicer arietinum	http://gigadb.org/dataset/100076	Varshney et al. 2013
Citrus sinensis	Phytozome v11: v1.1	Wu et al. 2014

Cladrastis lutea	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Codariocalyx motorius	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Copaifera officinalis	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Cucumis sativus	Phytozome v11: v1.0	*Not available*
Desmanthus illinoensis	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Elaeocarpus photiniifolia	Genbank BioProject: PRJDA67329	Sugai et al. 2012
Entada abyssinica	Genbank BioProject: PRJNA574148	**Newly sequenced**
Eucalyptus grandis	Phytozome v11: v2.0	Bartholomé et al. 2015
Fragaria vesca	Phytozome v11: v1.1	Shulaev et al. 2011
Gleditsia triacanthos	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Glycine max	Phytozome v11: Wm82.a2.v1	Schmutz et al. 2010
Glycyrrhiza lepidota	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Gossypium raimondii	Phytozome v11: v2.1	Paterson et al. 2012
Gymnocladus dioicus	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Inga spectabilis	https://doi.org/10.5061/dryad.r9c12	Nicholls et al. 2015
Juglans regia	https://www.hardwoodgenomics.org/Genome-assembly/2209485	Martínez-García et al. 2016
Lactuca sativa	Genbank BioProject: PRJNA65477	*Not available*
Lathyrus sativus	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Lens culinaris	Genbank BioProject: PRJNA65667	Kaur et al. 2011
Linum usitatissimum	Phytozome v11: v1.0	Wang et al. 2012
Lotus japonicus	http://www.plantgdb.org/LjGDB/	Sato et al. 2008
Lupinus angustifolius	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Lupinus polyphyllus	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Manihot esculenta	Phytozome v11: v6.1	Bredeson et al. 2016
Medicago truncatula	Phytozome v11: Mt4.0v1	Young et al. 2011

Microlobius foetidus	Genbank BioProject: PRJNA574148	**Newly sequenced**
Mimulus guttatus	Phytozome v11: v2.0	Hellsten et al. 2013
Morus notabilis	Genbank BioProject: PRJNA202089 (assembly version ASM41409v2)	He et al. 2013
Nelumbo nucifera	Genbank BioProject: PRJNA264089 (assembly version 1.1)	Ming et al. 2013
Paeonia lactiflora	Genbank BioProject: PRJNA245064	Zhang et al. 2015
Panax ginseng	Genbank BioProject: PRJNA173906	Li et al. 2013
Papaver somniferum	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Phaseolus vulgaris	Phytozome v11: v1.0	Schmutz et al. 2014
Pisum sativum	Genbank BioProject: PRJNA211622	Duarte et al. 2014
Polygala lutea	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Populus trichocarpa	Phytozome v11: v3.0	Tuskan et al. 2006
Primula veris	https://doi.org/10.5061/dryad.2s200	Nowak et al. 2015
Prioria balsamifera	Genbank BioProject: PRJNA472454	**Newly sequenced**
Prosopis alba	Genbank BioProject: PRJNA218545	Torales et al. 2013
Prunus persica	Phytozome v11: v2.1	International Peach Genome Initiative et al., 2013
Punica granatum	Genbank BioProject: PRJNA231033	Ophir et al. 2014
Quillaja saponaria	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Salix purpurea	Phytozome v11: v1.0	Zhou et al. 2018
Senna hebecarpa	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Solanum tuberosum	Phytozome v11: v3.4	Sharma et al. 2013
Styphnolobium japonicum	https://www.hindawi.com/journals/bmri/2014/750961/sup/	Zhu et al. 2014
Theobroma cacao	Phytozome v11: v1.1	Motamayor et al. 2013
Trifolium pratense	Genbank BioProject: PRJNA219226	Yates et al. 2014
Tripterygium wilfordii	Genbank BioProject: PRJNA218574	*Not available*

Vicia faba	Genbank BioProject: PRJNA81211	Kaur et al. 2012
Vigna radiata	ftp://plantgenomics.snu.ac.kr/mungbean_data/	Kang et al. 2014
Vitis vinifera	Phytozome v11: Genoscope.12X	Jaillon et al. 2007
Xanthocercis zambesiaca	Dryad: http://dx.doi.org/10.5061/dryad.ff1tq	Cannon et al. 2015
Zenia insignis	Genbank BioProject: PRJNA285444	*Not available*

Phytozome is available at https://phytozome.jgi.doe.gov
OneKP data is available at http://www.onekp.com/public_data.html

References

Bertioli DJ, Cannon SB, Froenicke L, Huang G, Farmer AD, Cannon EKS, et al. The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nat Genet. 2016 Apr;48(4):438–446.

Bredeson JV, Lyons JB, Prochnik SE, Wu GA, Ha CM, Edsinger-Gonzales E, et al. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. Nat Biotechnol. 2016 May;34(5):562–570.

Clouse JW, Adhikary D, Page JT, Ramaraj T, Deyholos MK, Udall JA, et al. The Amaranth Genome: Genome, Transcriptome, and Physical Map Assembly. Plant Genome. 2016 Mar;9:1.

Duarte J, Rivière N, Baranger A, Aubert G igoire, Burstin J, Cornet L, et al. Transcriptome sequencing for high throughput SNP development and genetic mapping in Pea. BMC Genomics. 2014 Feb;15:126.

Filiault DL, Ballerini ES, Mandáková T, Aköz G, Derieg NJ, Schmutz J, et al. The Aquilegia genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. eLife. 2018 Oct;

He N, Zhang C, Qi X, Zhao S, Tao Y, Yang G, et al. Draft genome sequence of the mulberry tree Morus notabilis. Nat Commun. 2013;4:2445.

Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, Wessler SR, et al. Fine-scale variation in meiotic recombination in Mimulus inferred from population shotgun sequencing. PNAS. 2013 Nov;110(48):19478–19482.

Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007 Sep;449(7161):463–467.

Kaur S, Cogan NOI, Pembleton LW, Shinozuka M, Savin KW, Materne M, et al. Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. BMC Genomics. 2011 May;12:265.

Kaur S, Pembleton LW, Cogan NOI, Savin KW, Leonforte T, Paull J, et al. Transcriptome sequencing of field pea and faba bean for discovery and validation of SSR genetic markers. BMC Genomics. 2012 Mar;13:104.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 2012 Jan;40(Database):issue.

Li C, Zhu Y, Guo X, Sun C, Luo H, Song J, et al. Transcriptome analysis reveals ginsenosides biosynthetic genes, microRNAs and simple sequence repeats in Panax ginseng C. A. Meyer. BMC Genomics. 2013 Apr;14:245.

Martínez-García PJ, Crepeau MW, Puiu D, Gonzalez-Ibeas D, Whalen J, Stevens KA, et al. The walnut (Juglans regia) genome sequence reveals diversity in genes coding for the biosynthesis of non-structural polyphenols. Plant J. 2016 Sep;87(5):507–532.

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, et al. The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature. 2008 Apr;452(7190):991–996.

Ming R, VanBuren R, Liu Y, Yang M, Han Y, Li L-T, et al. Genome of the long-living sacred lotus (Nelumbo nucifera Gaertn.). Genome Biol. 2013 May;14(5):R41.

Moore MJ, Hassan N, Gitzendanner MA, Bruenn RA, Croley M, Vandeventer A, et al. Phylogenetic Analysis of the Plastid Inverted Repeat for 244 Species: Insights into Deeper-Level Angiosperm Relationships from a Long, Slowly Evolving Sequence Region. International Journal of Plant Sciences. 2011 May;172(4):541–558.

Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. PNAS. 2010 Mar;107(10):4623–4628.

Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D, Cornejo O, et al. The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. Genome Biol. 2013 Jun;14(6):r53.

Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, et al. The genome of Eucalyptus grandis. Nature. 2014 Jun;510(7505):356–362.

Nicholls JA, Pennington RT, Koenen EJ, Hughes CE, Hearn J, Bunnefeld L, et al. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus Inga (Leguminosae: Mimosoideae). Front Plant Sci. 2015 Sep;6.

Nowak MD, Russo G, Schlapbach R, Huu CN, Lenhard M, Conti E. The draft genome of Primula veris yields insights into the molecular basis of heterostyly. Genome Biol. 2015 Jan;16:12.

Ophir R, Sherman A, Rubinstein M, Eshed R, Sharabi Schwager M, Harel-Beja R, et al. Single-nucleotide polymorphism markers from de-novo assembly of the pomegranate transcriptome reveal germplasm genetic diversity. PLoS One. 2014 Feb;9(2):e88998.

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, et al. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature. 2012 Dec;492(7429):423–427.

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, et al. Genome sequence of the palaeopolyploid soybean. Nature. 2010 Jan;463(7278):178–183.

Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet. 2014 Jun;46(7):707.

Sharma SK, Bolser D, de Boer J, Sønderkær M, Amoros W, Carboni MF, et al. Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. G3 (Bethesda). 2013 Nov;3(11):2031–2047.

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, et al. The genome of woodland strawberry (Fragaria vesca). Nat Genet. 2011 Feb;43(2):109–116.

Sugai K, Setsuko S, Uchiyama K, Murakami N, Kato H, Yoshimaru H. Development of EST-SSR markers for Elaeocarpus photiniifolia (Elaeocarpaceae), an endemic taxon of the Bonin Islands. Am J Bot. 2012 Feb;99(2):84–87.

Torales SL, Rivarola M iximo, Pomponio M ia F, Gonzalez S, Acuña CV, Fernández P, et al. De novo assembly and characterization of leaf transcriptome for the development of functional molecular markers of the extremophile multipurpose tree species Prosopis alba. BMC Genomics. 2013 Oct;14:705.

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science. 2006 Sep;313(5793):1596–1604.

van Bakel H, Stout JM, Cote AG, Tallon CM, Sharpe AG, Hughes TR, et al. The draft genome and transcriptome of Cannabis sativa. Genome Biol. 2011 Oct;12(10):R102.

Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, et al. Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat Biotechnol. 2013 Mar;31(3):240–246.

Wang H, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, et al. Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc Natl Acad Sci USA. 2009 Mar;106(10):3853–3858.

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, et al. The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J. 2012 Nov;72(3):461–473.

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc Natl Acad Sci USA. 2014 Nov;111(45):E4859–E4868.

Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat Biotechnol. 2014 Jul;32(7):656–662.

Xi Z, Ruhfel BR, Schaefer H, Amorim A i M, Sugumaran M, Wurdack KJ, et al. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. PNAS. 2012 Oct;109(43):17519–17524.

Yates SA, Swain MT, Hegarty MJ, Chernukin I, Lowe M, Allison GG, et al. De novo assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC Genomics. 2014 Jun;15:453.

Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, et al. The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature. 2011 Nov;480(7378):520–524.

Zhang J, Wu Y, Li D, Wang G, Li X, Xia Y. Transcriptomic analysis of the underground renewal buds during dormancy transition and release in “Hangbaishao” peony (Paeonia lactiflora). PLoS One. 2015 Mar;10(3):e0119118.

Zhou R, Macaya-Sanz D, Rodgers-Melnick E, Carlson CH, Gouker FE, Evans LM, et al. Characterization of a large sex determination region in Salix purpurea L. (Salicaceae). Mol Genet Genomics. 2018 Jul;1–16.

Zhu L, Zhang Y, Guo W, Xu X-J, Wang Q. De Novo Assembly and Characterization of Sophora japonica Transcriptome Using RNA-seq. Biomed Res Int. 2014 Jan;2014.

Table S3. Counts and percentages of bipartitions representing nodes A-H and conflicting bipartitions representing other subfamily relationships among 3,473 gene trees.

Clade	**ML**		**>50% bootstrap support**		**>80% bootstrap support**	
	count	**percentage**	**count**	**percentage**	**count**	**percentage**
bipartitions of best supported topology						
Leguminosae (node A)	2669	76.85%	2254	64.90%	1660	47.80%
Cerc + Detar (node B)	744	21.42%	325	9.36%	48	1.38%
Cercidoideae (node C)	1815	52.26%	1705	49.09%	1394	40.14%
Detarioideae (node D)	3041	87.56%	2918	84.02%	2585	74.43%
Pap + Caes + Dial (node E)	794	22.86%	360	10.37%	91	2.62%
Pap + Caes (node F)	599	17.25%	231	6.65%	42	1.21%
Caesalpinioideae (node G)	2114	60.87%	1712	49.29%	1151	33.14%
Papilionoideae (node H)	2456	70.72%	1957	56.35%	1248	35.93%
conflicting bipartitions						
Pap + Caes + Dial + Detar	625	18.00%	258	7.43%	34	0.98%
Pap + Caes + Dial + Cerc	546	15.72%	194	5.59%	20	0.58%
Pap + Dial	446	12.84%	133	3.83%	20	0.58%
Caes + Dial	448	12.90%	132	3.80%	16	0.46%
Dial + Cerc	295	8.49%	93	2.68%	7	0.20%
Dial + Detar	307	8.84%	96	2.76%	4	0.12%
Caes + Dial + Cerc + Detar	247	7.11%	47	1.35%	4	0.12%
Pap + Dial + Cerc + Detar	196	5.64%	29	0.84%	4	0.12%
Caes + Detar	200	5.76%	44	1.27%	3	0.09%
Pap + Caes + Cerc + Detar	234	6.74%	46	1.32%	2	0.06%
Pap + Detar	189	5.44%	41	1.18%	2	0.06%
Caes + Cerc	163	4.69%	37	1.07%	2	0.06%
Pap + Cerc	173	4.98%	30	0.86%	2	0.06%
Pap + Caes + Detar	153	4.41%	27	0.78%	1	0.03%
Dial + Cerc + Detar	202	5.82%	21	0.60%	1	0.03%
Pap + Dial + Detar	122	3.51%	12	0.35%	1	0.03%
Caes + Dial + Cerc	121	3.48%	16	0.46%	0	0.00%
Pap + Caes + Cerc	132	3.80%	15	0.43%	0	0.00%
Caes + Cerc + Detar	127	3.66%	14	0.40%	0	0.00%
Pap + Dial + Cerc	110	3.17%	12	0.35%	0	0.00%
Caes + Dial + Detar	134	3.86%	11	0.32%	0	0.00%
Pap + Cerc + Detar	115	3.31%	9	0.26%	0	0.00%

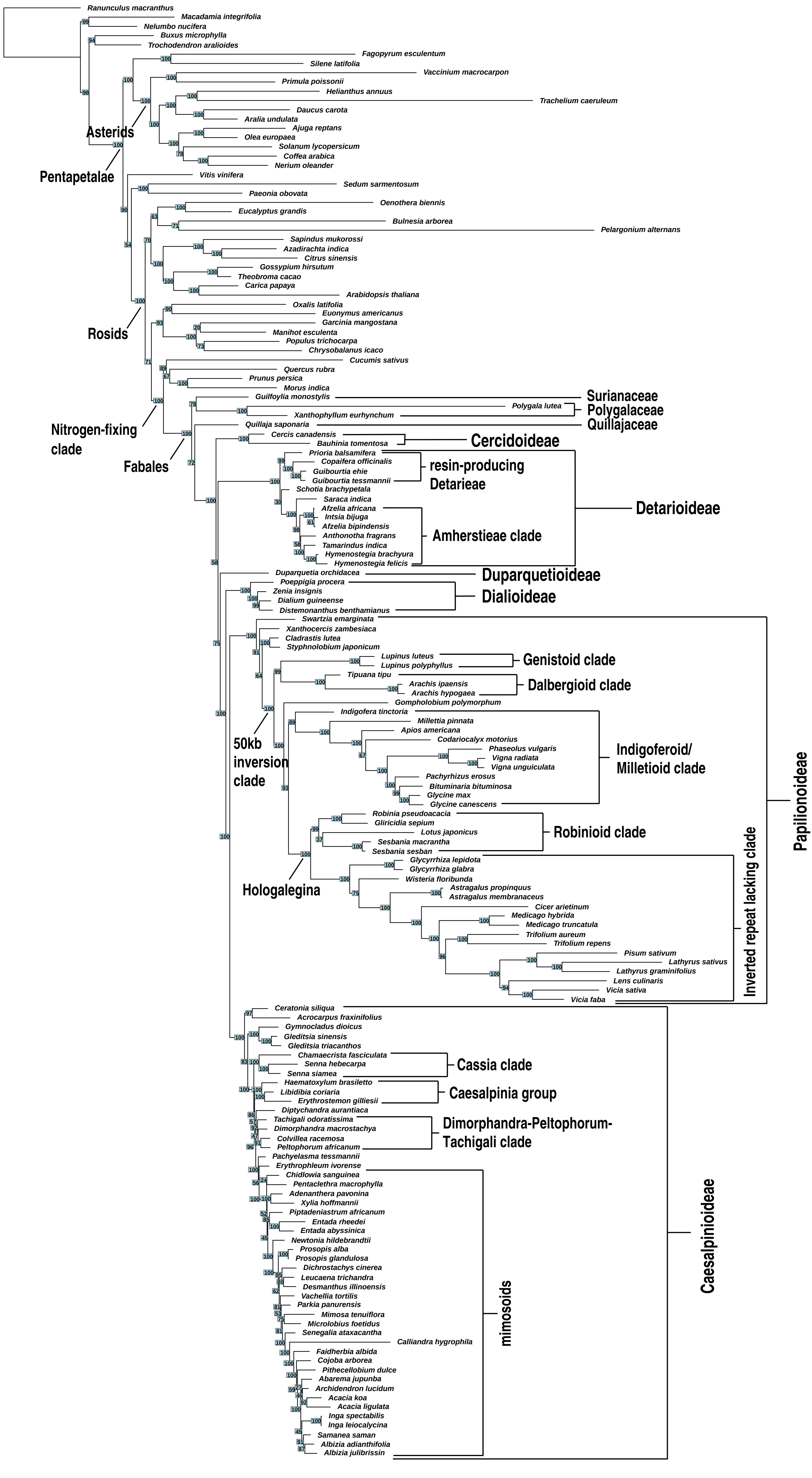

Figure S1. ML topology as inferred by RAxML from amino acid alignment of chloroplast genes under the LG4X model. Numbers on nodes indicate bootstrap percentages estimated from 1000 replicates.

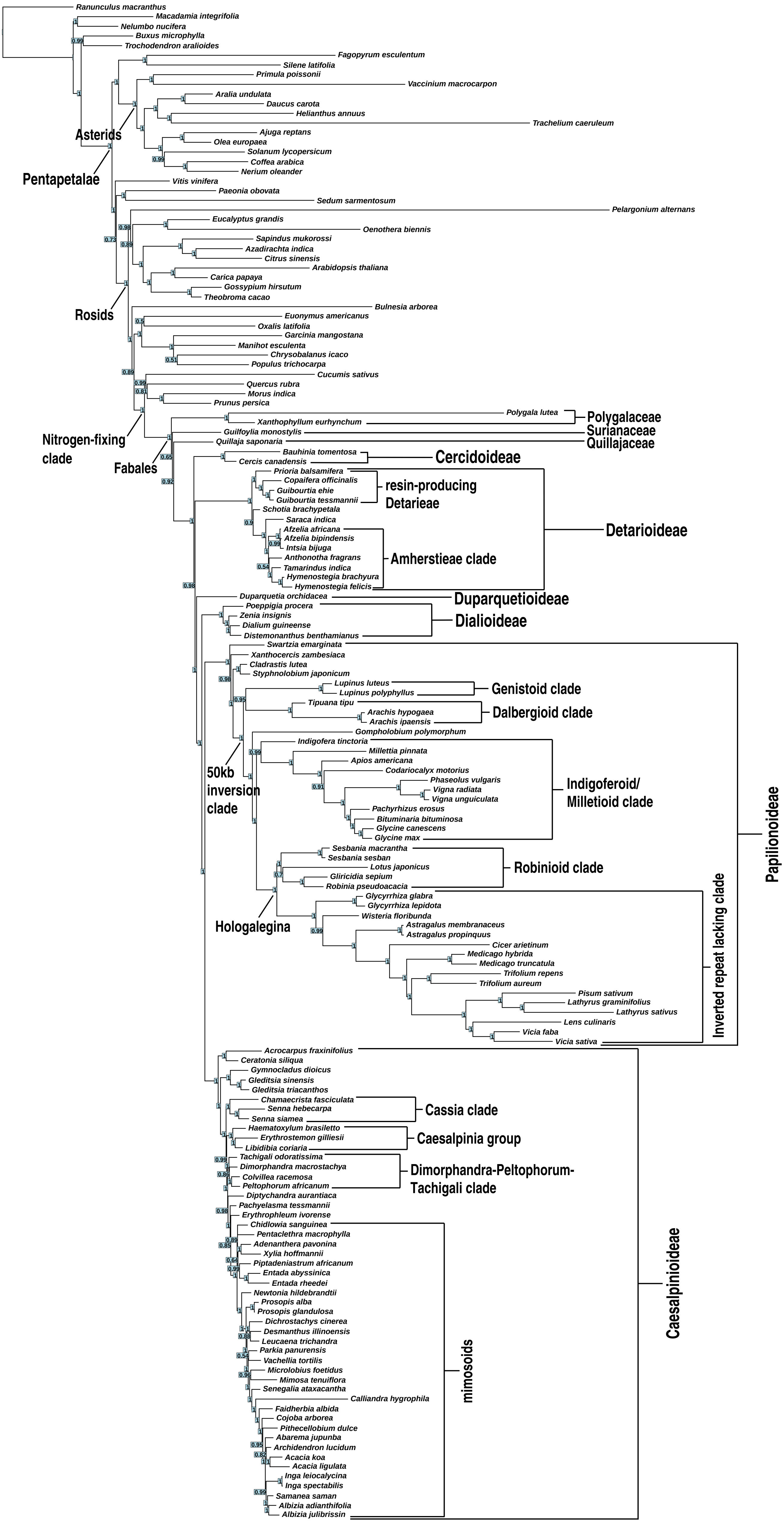

Figure S2. Bayesian majority-rule consensus tree inferred with Phylobayes from amino acid alignment of chloroplast genes under the CATGTR model. Numbers on nodes indicate posterior probabilities (pp) from 9000 post-burn-in MCMC cycles.

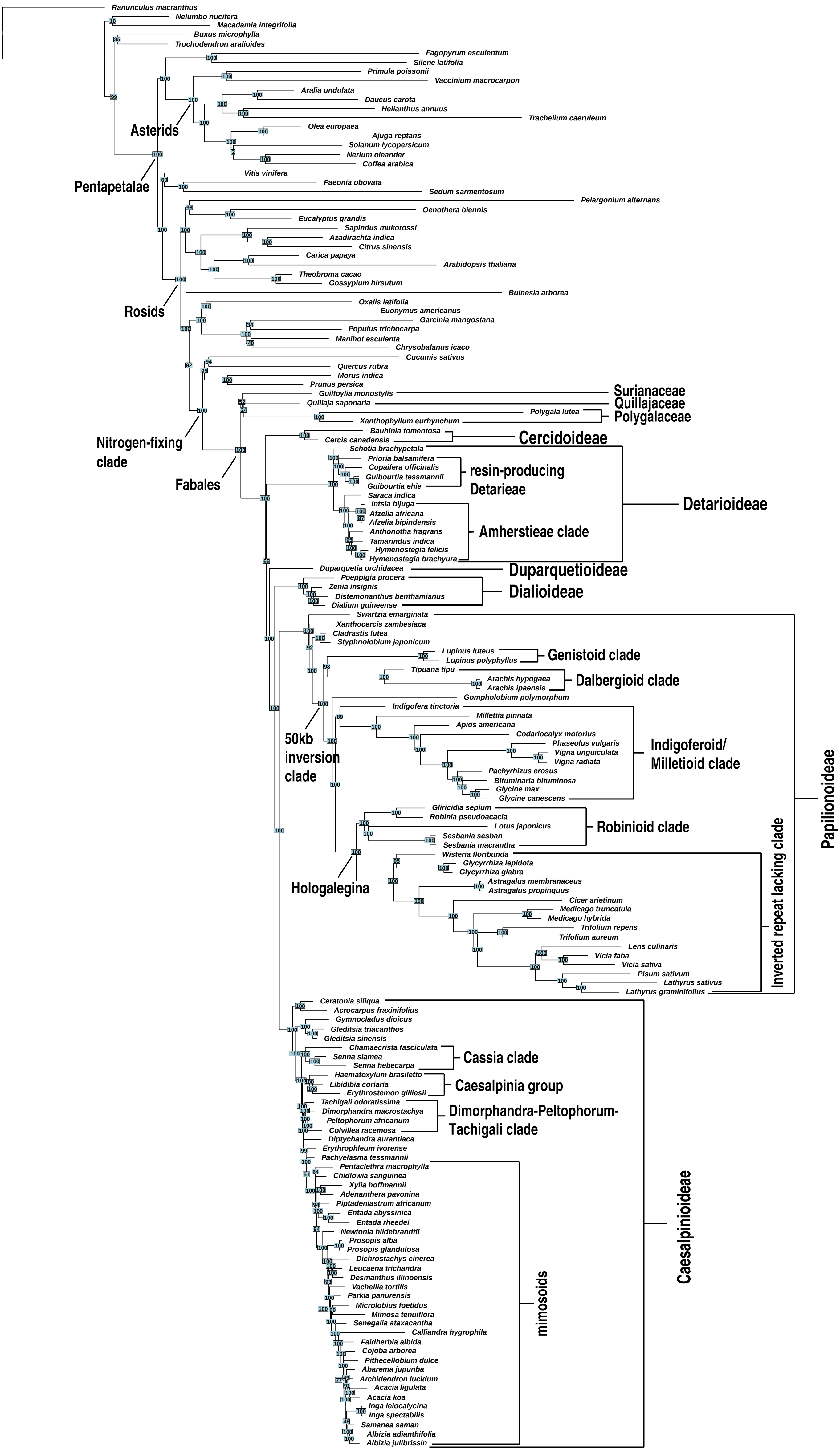

Figure S3. ML topology as inferred by RAxML from nucleotide alignment of chloroplast genes under the GTR + G model. Numbers on nodes indicate bootstrap percentages estimated from 1000 replicates.

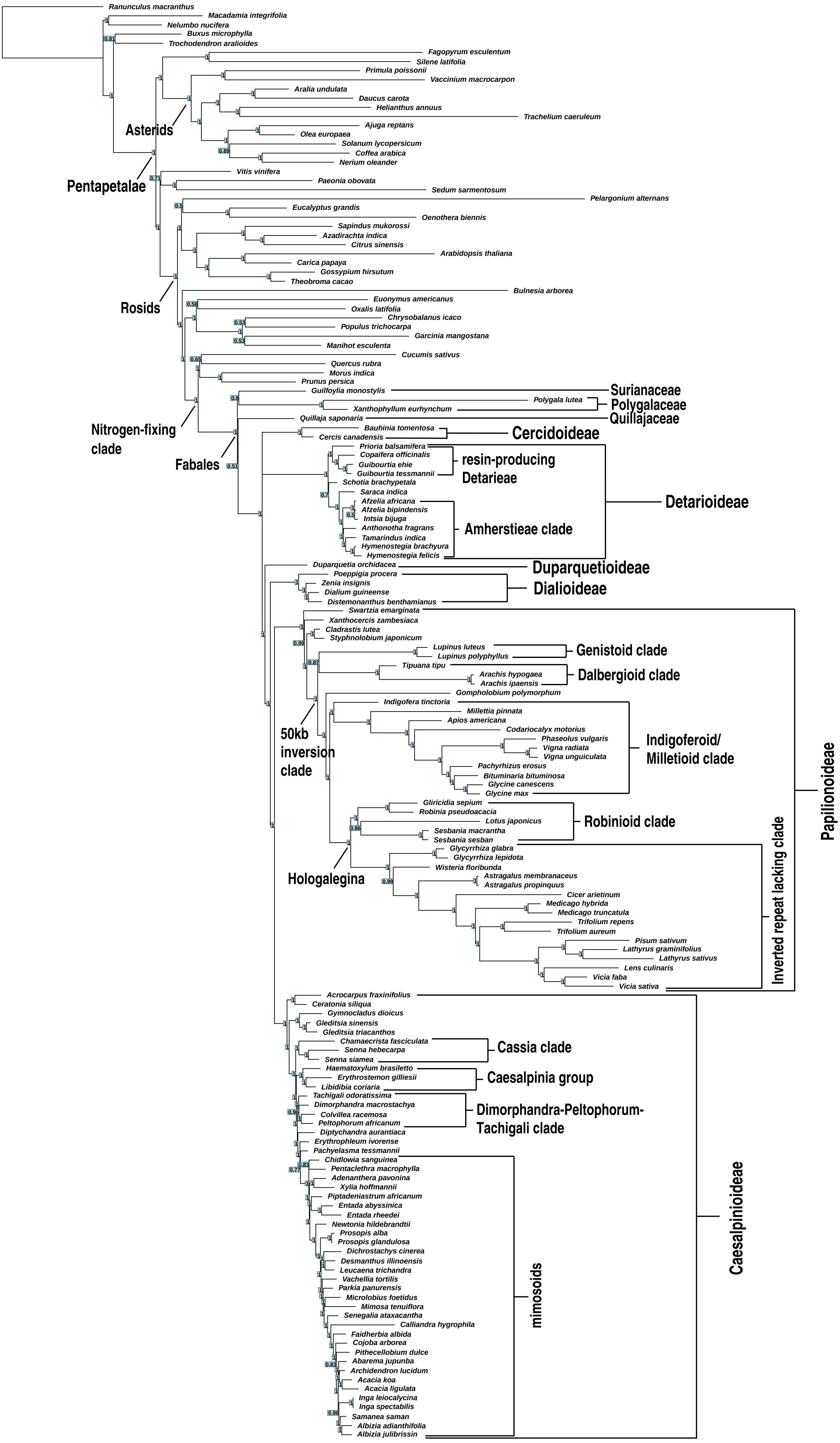

Figure S4. Bayesian majority-rule consensus tree inferred with Phylobayes from nucleotide alignment of chloroplast genes under the CATGTR model. Numbers on nodes indicate the posterior probabilities (pp) from 9000 post-burn-in MCMC cycles.

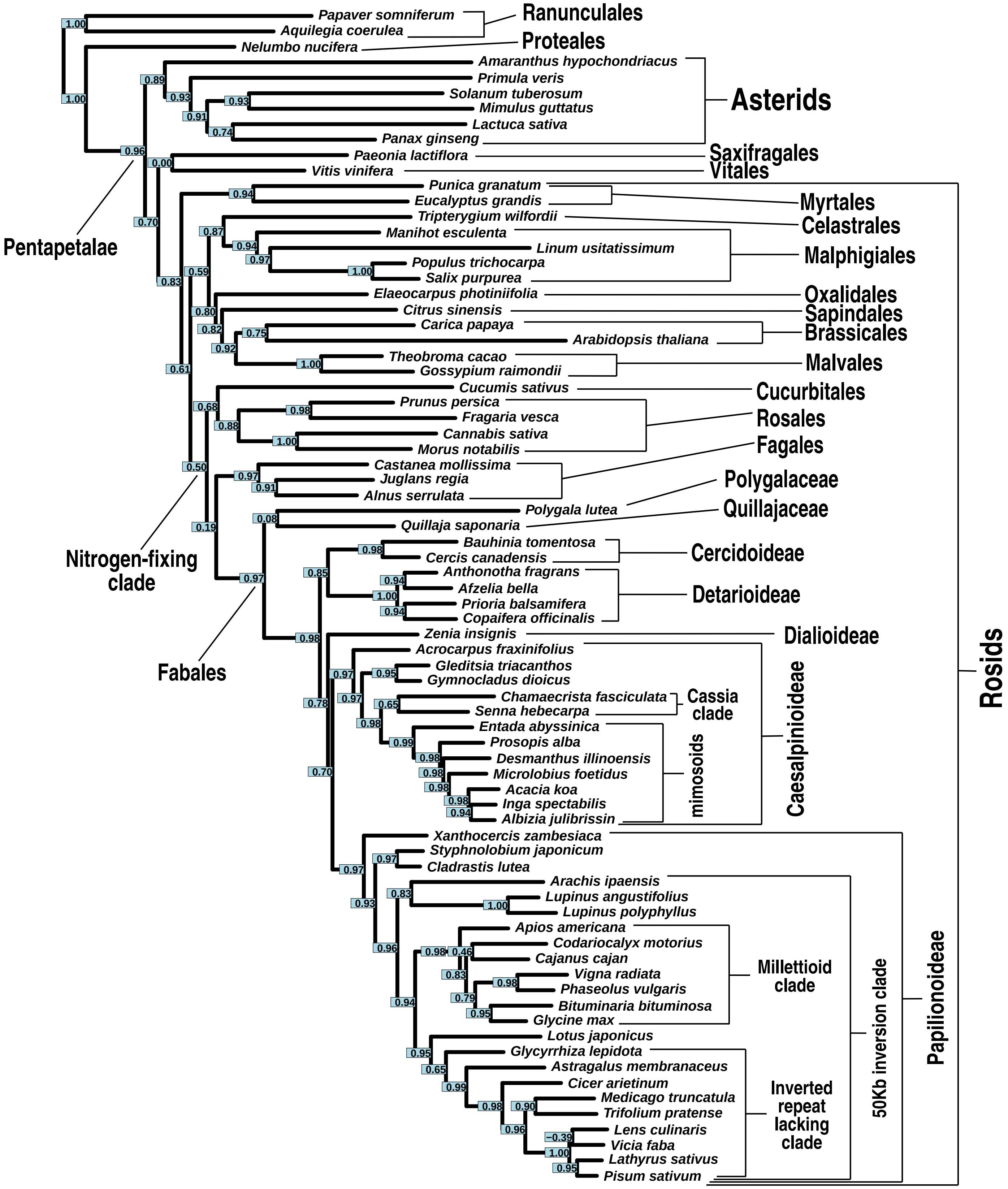

Figure S5. ML topology as inferred by RAxML from a concatenated alignment of 1,103 nuclear genes, under the LG4X model. Numbers on nodes indicate Internode Certainty All (ICA) values, as estimated from gene trees of the same 1,103 genes.

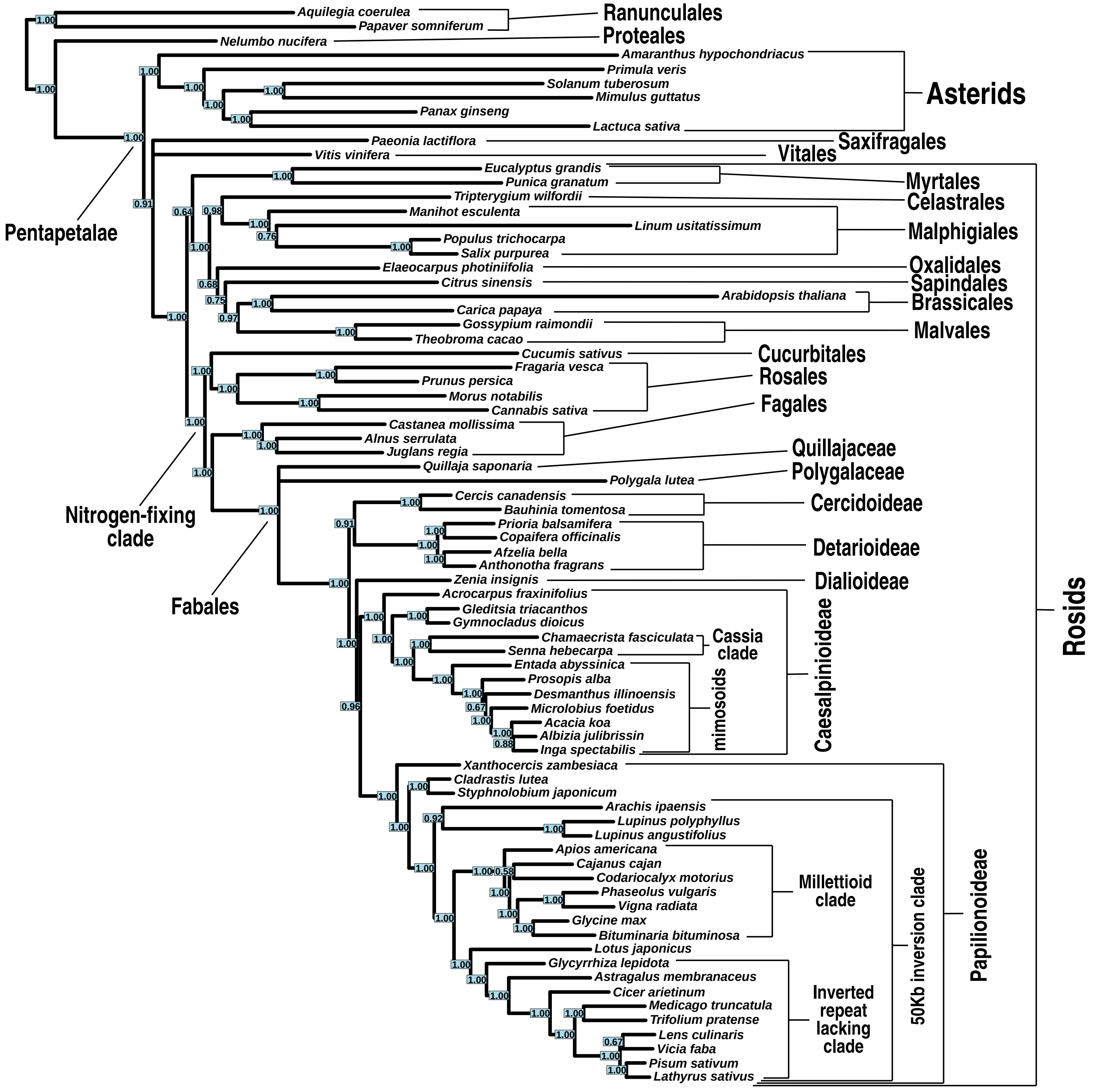

Figure S6. Bayesian gene jackknifing majority-rule consensus tree inferred with Phylobayes from a concatenated alignment of 1,103 nuclear genes. Numbers on nodes indicate posterior probabilities (pp), averaged over 500 posterior trees each, for 25 replicates (12,500 posterior trees in total).

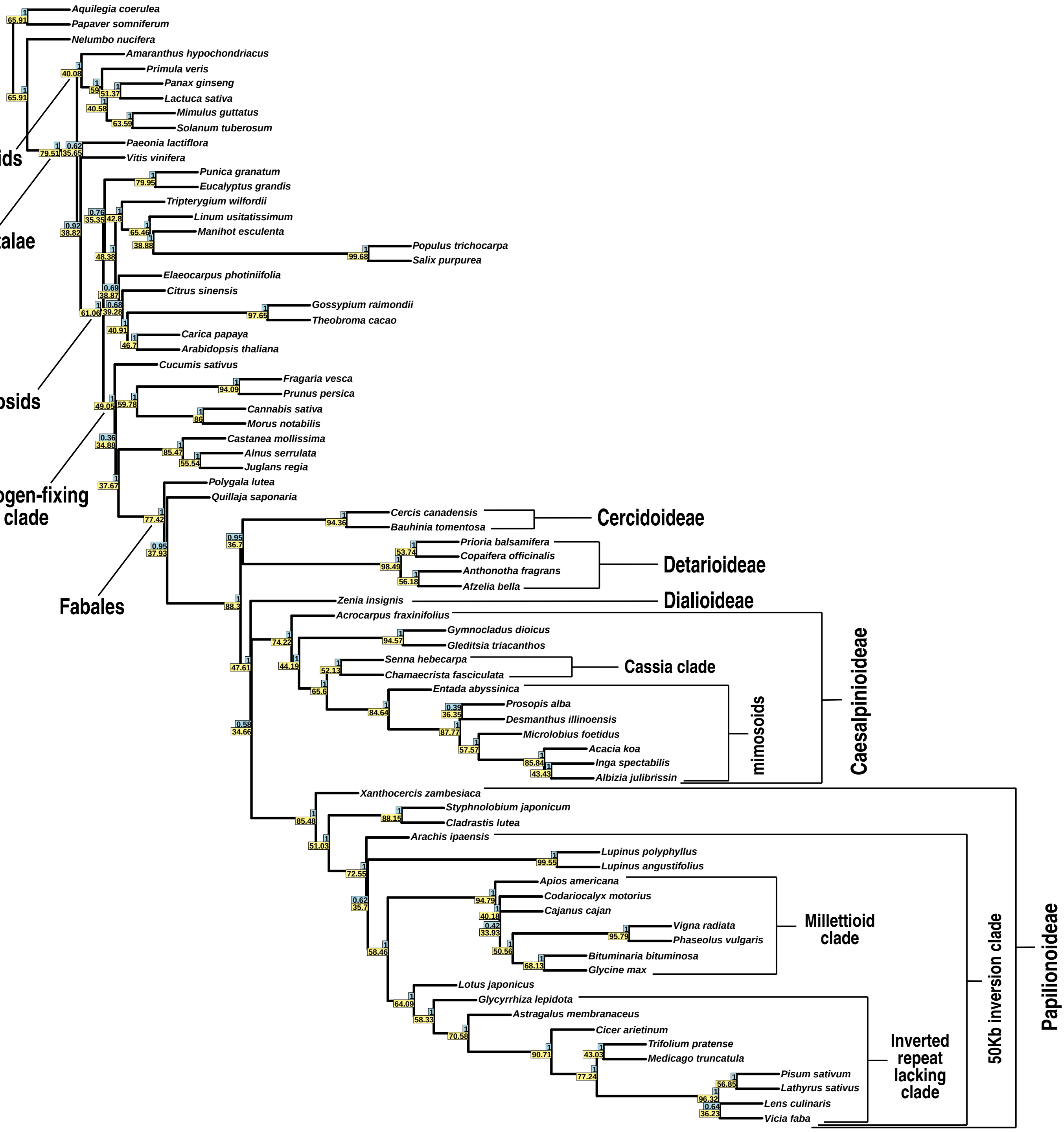

Figure S7. Phylogeny estimated under the multi-species coalescent with ASTRAL. Support values indicated represent local posterior probability (blue rectangles) and quartet support (yellow rectangles).