

## ***The benchmark species trees used in this analysis***

The inference method and input data for each of the species trees for each of the 12 datasets used in this analysis.

### ***Birds (Jarvis et al. 2014)***

**Data:** exons from 8251 protein coding genes, introns from 2516 genes, 2769 ultraconserved elements x 48 species. Total ~48.1 million base pairs. GTR + GAMMA.

**Method:** CMSA nucleotides, partitioned by data type. Third nucleotide of codons excluded

### ***Diptera (Wiegmann et al. 2011)***

**Data:** Tier 1: 42 species x (14 nuclear genes + mitochondrial genome + 371 morphological features). Tier 2: 202 taxa x 5 nuclear genes. GTR+GAMMA

**Method:** CMSA nucleotides. Likelihood & Bayesian. Data partitioned by type. Nucleotide with 3<sup>rd</sup> codon removed

### ***Fish (Betancur et al. 2013)***

**Data:** 16 nuclear gene exons + 1 intron + 4 mitochondrial genes x 1184 taxa

**Method:** CMSA, nucleotide, partitioned by codon position. GTRGAMMA & GTRCAT

### ***Fungi (James et al. 2006)***

**Data:** 3 protein coding genes, 3 RNA genes x 214 taxa

**CMSA** of AA (JTT+I+GAMMA) and nucleotides (GTR+I+GAMMA/GTR+GAMMA), partitioned by gene. Bayesian. Assessment of conflict.

### ***Hymenoptera (Mao et al. 2015)***

**Data:** nucleotide for the 13 protein coding + 2 rRNA genes of the mitochondrial genome x 48 species

**Method:** CMSA, best partitioning & model by BIC. With/without 3rd codon. ML & Bayesian.

### ***Kinetoplastids (Kelly et al. 2017)***

**Data:** AA for 375 genes x 19 species

**Method:** CMSA, 47,625 AA using equal sampling from each gene. ML tree PROTGAMMAAUTO

### ***Laurasiatheria (Meredith et al. 2011)***

**Data:** 171 species x 26 genes (35,603bp/11,010 AA)

**Method:** CMSA, amino acids and nucleotide. ML & Bayesian. Model selection by AIC. Coalescence with STAR & STEAC.

### ***Metazoa (Dunn et al. 2008)***

**Data:** 77 taxa x 150 genes

**Method:** CMSA, amino acids ML (WAG selected by MyBayes), Bayesian (CAT & WAG)

### ***Nematoda (Liu et al. 2013)***

**Data:** AA for 12 mitochondrial protein-coding genes x 65 species

**Method:** MP, ML & Bayesian. AIC to select best model. MSA, SSU rRNA

### ***Primates (Perelman et al. 2011)***

**Data:** 191 taxa x 34,972 bp: nucleotide, 6 partitions (X-chromosome, Y-chromosome, autosome, intron, exon and UTR segments)

**Method:** CMSA of nucleotides, MP, ML & Bayesian. Partitions analysed separately & combined. Model selection by AIC.

### ***Rodents (Meredith et al. 2011)***

**Data:** 171 species x 26 genes (35,603bp/11,010 AA)

**Method:** CMSA, amino acids and nucleotide. ML & Bayesian. Model selection by AIC. Coalescence with STAR & STEAC.

### ***Plants (Ruhfel et al. 2014; Byng et al. 2016)***

**Data:** Expert review of latest molecular studies including 17 nuclear and plastid genes x 640 taxa (Byng) & 78 protein-coding plastid genes x 360 taxa (Ruhfel)

**Method:** Ruhfel: separate and partitioned, concatenated alignments of. 78 partitioned, concatenated plastid genes including AA; nucleotide; 1st & 2nd codon; and RY-coding. Soltis: with and without mitochondrial genes, ML, GTRGAMMA. Individual & concatenated alignments.

### ***References***

Betancur RR, Broughton RE, Wiley EO, Carpenter K, Lopez JA, Li C, Holcroft NI, Arcila D, Sanciangco M, Cureton JC et al. 2013. The tree of life and a new classification of bony fishes. *PLoS Curr* 5.

Byng JW, Chase MW, Christenhusz MJM, Fay MF, Judd WS, Mabberley DJ, Sennikov AN, Soltis DE, Soltis PS, Stevens PF et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc* 181(1): 1-20.

Dunn CW, Hejnal A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD et al. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452(7188): 745-U745.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443(7113): 818-822.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346(6215): 1320-1331.

Kelly S, Ivens A, Mott GA, O'Neill E, Emms D, Macleod O, Voorheis P, Tyler K, Clark M, Matthews J et al. 2017. An Alternative Strategy for Trypanosome Survival in the Mammalian Bloodstream Revealed through Genome and Transcriptome Analysis of the Ubiquitous Bovine Parasite *Trypanosoma (Megatrypanum) theileri*. *Genome Biol Evol* 9(8): 2093-2109.

Liu GH, Shao RF, Li JY, Zhou DH, Li H, Zhu XQ. 2013. The complete mitochondrial genomes of three parasitic nematodes of birds: a unique gene order and insights into nematode phylogeny. *Bmc Genomics* 14.

Mao M, Gibson T, Downton M. 2015. Higher-level phylogeny of the Hymenoptera inferred from mitochondrial genomes. *Mol Phylogenet Evol* 84: 34-43.

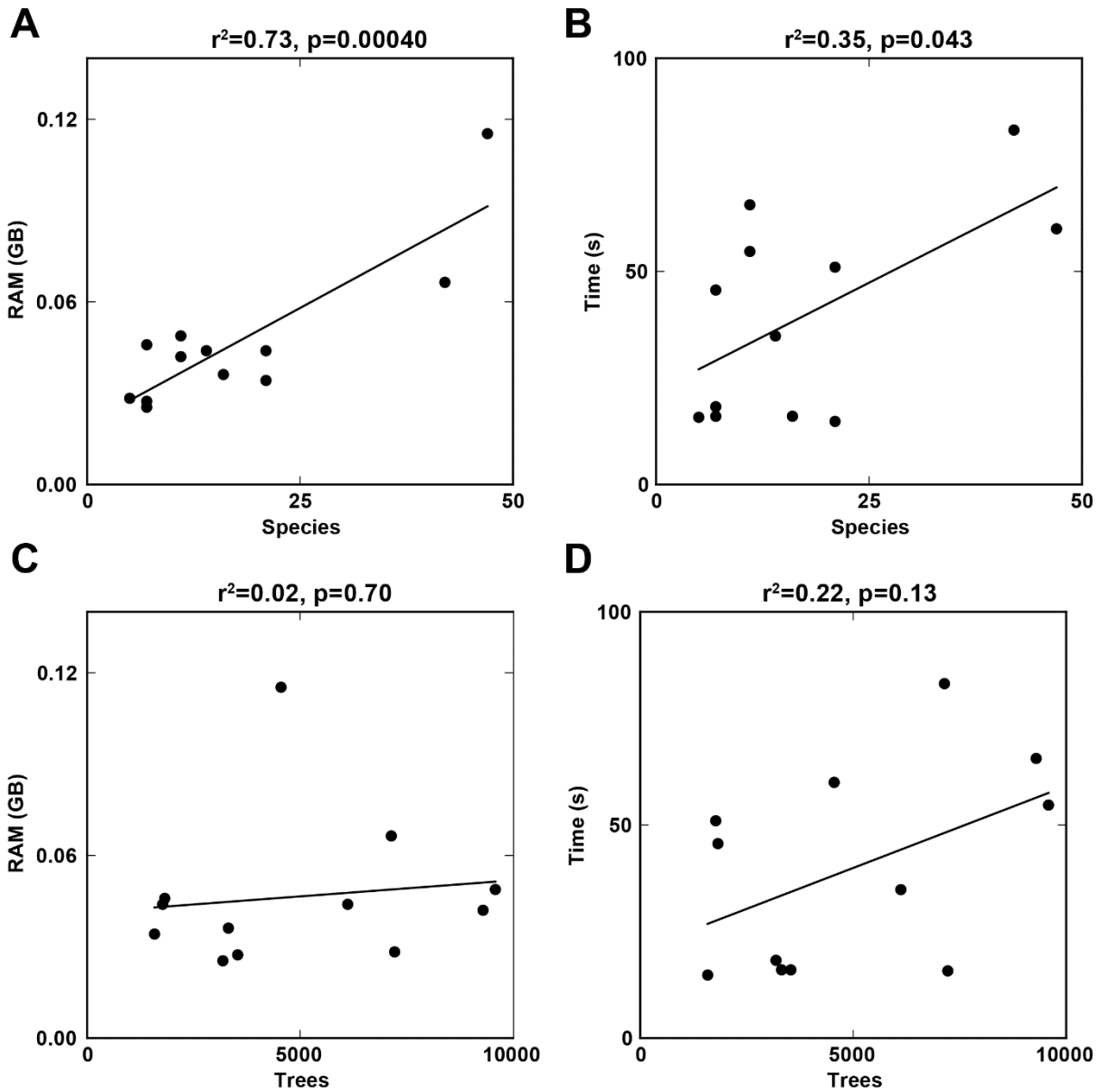
Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T et al. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg Extinction on Mammal Diversification. *Science* 334(6055): 521-524.

Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MAM, Kessing B, Pontius J, Roelke M, Rumpler Y et al. 2011. A Molecular Phylogeny of Living Primates. *Plos Genet* 7(3).

Ruhfel BR, Gitzendanner MA, Soltis PS, Soltis DE, Burleigh JG. 2014. From algae to angiosperms-inferred the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *Bmc Evol Biol* 14.

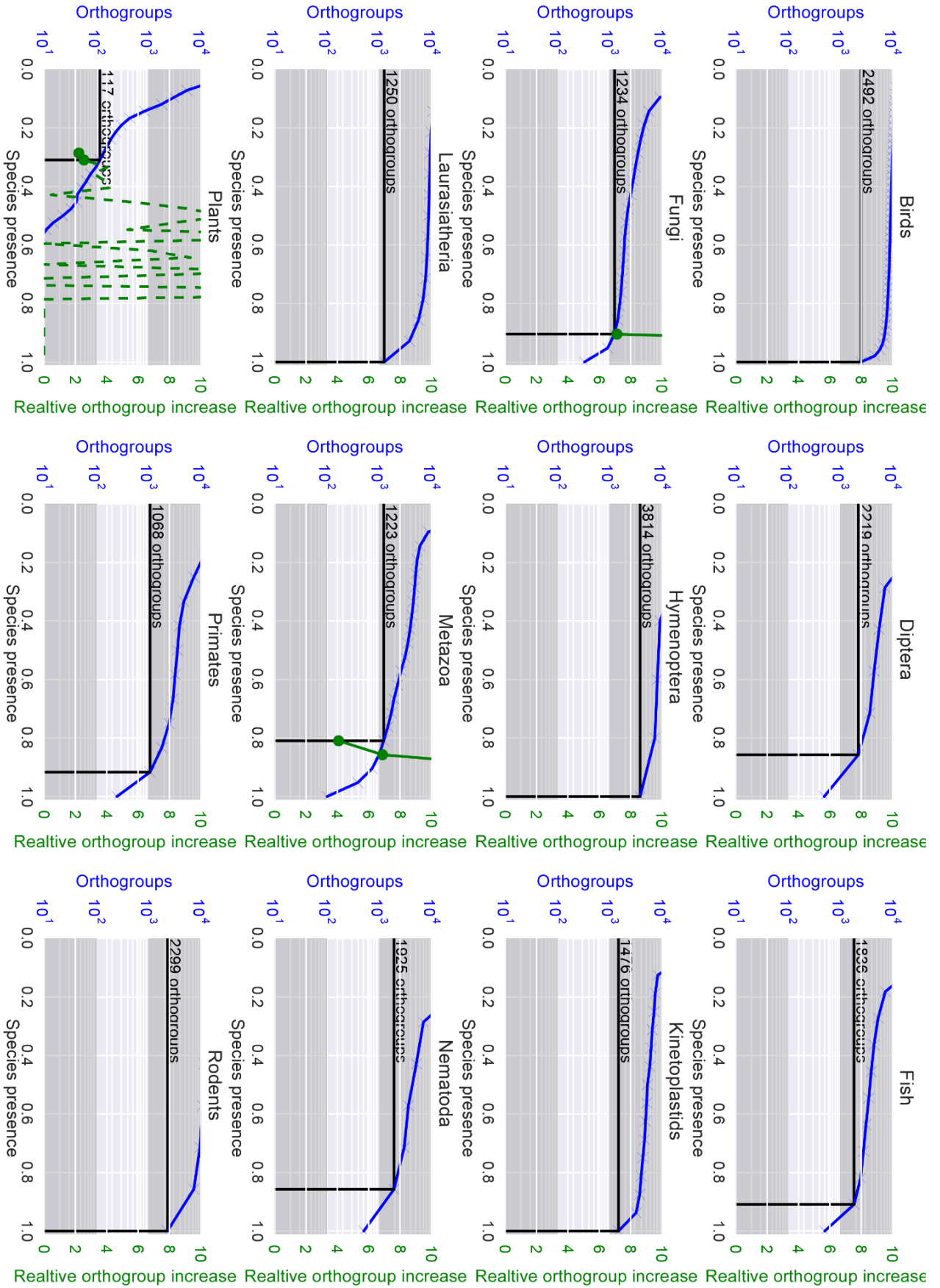
Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM et al. 2011. Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences of the United States of America* 108(14): 5690-5695.

**Supplementary Figures**  
**Supplementary Figure S1**



**Supplementary Figure S1.** The RAM (A) and time (B) requirements for STAG on the 12 datasets on a single core versus the number of species and RAM (C) and time (D) requirements versus the number of trees with all species present.

Supplementary Figure S2



**Supplementary Figure S2.** Selection of single-copy orthologues with high proportion of species present for CMSA, NJst (Plants only) and ASTRAL applied to the 12 biological datasets. Each plot shows the number of orthogroups (blue line, left y-axis) in which a given fraction of species are single-copy. Single-copy orthogroup selection starts by identifying all orthogroups that are single-copy in all species. In most cases there are sufficiently many such orthogroups ( $\geq 1000$ , top grey band). If not, the requirement for all species to be present is relaxed until at least 100 orthogroups are identified (middle white band). The minimum fraction of single copy species required is then further relaxed so as to identify more orthogroups provided that the proportional increase in orthogroups is more than double the proportional decrease in minimum fraction of species (green line, where applicable) or until 1000 orthogroups are identified. The number of orthogroups selected and the minimum fraction of species in each orthogroups is shown (black line, text annotation). The number of orthogroups identified and the cut-off minimum proportion of single-copy species are given in Supplementary Table S1.

## Supplementary Tables

### Supplementary Table S1

Supplementary Table S1 contains 59134 rows and is available separately as an .xlsx file.

### Supplementary Table S2

Supplementary Table S2 contains the species tree in newick format inferred by each of the five tested methods as is available as an .xlsx file.

### Supplementary Table S3: p-values for correlation between branch length estimates across methods.

Branch length, correlation p-value				
	STAG	CMSA	ASTRAL	NJst
STAG	0	2.71E-78	7.71E-22	3.31E-10
CMSA	2.71E-78	0	5.61E-22	1.11E-09
ASTRAL	7.71E-22	5.61E-22	0	2.51E-21
NJst	3.31E-10	1.11E-09	2.51E-21	0

### Supplementary Table S4

Clade	Single-copy orthogroups, all species present				Single-copy orthogroups, most species present				
	Orthogroups	CMSA	NJs	ASTRAL	Orthogroups	Min. species present	CMSA	NJs	ASTRAL
Birds	2616	0.47	0.40	0.409	2616	100.0%	0.34	0.40	0.409
Diptera	513	0	0	0	2263	85.7%	0	0	0
Fish	518	0.12	0.12	0.25	1904	90.9%	0.12	0.12	0.125
Fungi	338	0	0	0	1182	90.9%	0.05	3	0
Hymenoptera	3947	0	0	0	3947	100.0%	0	0	0
Kinetoplastids	1476	0	0	0	1476	100.0%	0	0	0
Laurasiatheria	1386	0.18	0.18	0.182	1386	100.0%	0.18	0.18	0.182
Metazoa	102	0.27	0.11	0.167	1030	85.7%	0.22	0.16	0.167
Nematoda	517	0.25	0.25	0.25	1935	85.7%	0.05	0.20	0.25
Plants	0	-	-	-	109	35.7%	1	5	0.308
Primates	286	0	0	0	1196	91.7%	0	0	0
Rodents	2505	0.25	0.25	0.25	2505	100.0%	0.25	0.25	0.25
Mean (ex. plants)	-	0.14	0.12	0.137	-	-	0.12	0.12	0.126