doi:10.1038/nature11631

1. SUPPLEMENTARY METHODS

1.1 Phylogenetic tree construction

1.1.1 Overview of tree inference and uncertainty

Genetic data is now available for at least 6,663 of a total of 9,993 bird species recognized in our taxonomy (see below). Many studies of diversification, trait evolution, community phylogenetics and phylogenetic diversity require resolved and dated trees. The most advanced relaxed clock methods for simultaneously inferring tree topology and estimating divergence times (e.g. BEAST, MrBayes 3.2) are well-suited to smaller datasets (typically <500 tips) but computational limits preclude their use for large trees of 1000's of taxa. Dating methods for very large trees (e.g. PathD8, Britton et al. 2007) do not infer topology, place tight restrictions on calibration priors and can yield biased divergence time estimates (Brown et al. 2007, Svennblad 2008). Moreover, although it is technically possible to infer a point estimate additive tree with our data (using, e.g. RaxML), data matrices of this size are typically sparse and may lack sufficient information to resolve many topological combinations, seriously impeding tree inference (Sanderson et al. 2011). The proportion of taxon combinations for which large data matrices are indecisive can be reduced substantially by analyzing subsets of the data that represent established well-supported clades. The locus by species coverage for a concatenated supermatrix would be 0.31 and decisiveness ~0.6. By using smaller matrices that correspond to well-supported monophyletic clades both coverage and decisiveness are improved. For the clades used in our tree-building approach (see point 2 below and section 1.1.3) the average locus by species coverage is significantly greater (mean 0.67; range 0.42-1.0) and average decisiveness (Sanderson et al. 2011) substantially higher (average 0.92 across clades) than obtained using the supermatrix approach. Consequently, our approach to building a complete tree of extant birds is to combine time-calibrated trees of deep avian relationships with separately inferred trees of well-supported constituent clades.

Importantly for our approach, avian phylogenetics has advanced tremendously over the past 10-20 years and molecular studies of both deeper relationships and core clade memberships of almost all of the 2,091 avian genera are now available. In particular, a comprehensive recent phylogenomic study of key higher-level relationships in birds (Hackett et al. 2008) together with several other assessments of core clades (e.g. Barker et al. 2004, Jonsson & Fjeldsa 2006, Jonsson et al. 2008a, Ohlson et al. 2007) now offer reasonable, if not undisputed, resolution of the main relationships in the avian tree. Rather than re-assessing these topologies, we here accept them as the current state-of-the-art evidence for higher level avian relationships. Consequently, we constructed three sets of component trees as follows:

1. We inferred and dated distributions of "backbone" trees using information from existing higher-level topologies. These trees contained 158 tips. Each tip represents a crown clade in the complete avian tree. A portion of the topology of these backbone trees is highly constrained on the basis of strong support (posterior probabilities => 0.95) from previous studies. The remaining interior nodes with comparatively poor support were left unconstrained (41 nodes for our

primary backbone based on Hackett et al. 2008, and 37 in the case of an alternative backbone based on Ericson et al 2006).

2. The backbone determines a total of 129 avian crown clades with > 4 species (varying from 5 to 460 species in size, see clade_summary.csv) and 29 smaller clades with 4 or fewer species. We used a careful evaluation of the existing avian phylogenetic literature to assign each avian genus to one of the 129 crown clades (the members of ten passerine genera were split among clades). We inferred distributions of relaxed clock (ultrametric) trees including all species with genetic data for each of the 129 crown clades. These trees were wholly unconstrained and we refer to them as "Stage 1 trees".

3. We built an additional set of trees that included the remaining 3,323 species that are members of the 129 crown clades but lack genetic data. 2,997 of them have member(s) of their genus represented in the Stage 1 trees, and we formulated constraints that attach these to their congeners as situated in the Stage 1 consensus trees (i.e. in the case of non-monophyletic Stage 1 congeners these species are constrained deeper in the tree, allowing placement across multiple genera). 326 species (9.9 %) lacked congeners with genetic data, and for the 214 genera they belonged to we formulated constraints based on taxonomic evidence (e.g. recognized sub-family relationships). We inferred distributions of these topologically constrained relaxed clock (ultrametric) trees for each crown clade, and refer to them as "Stage 2 trees".

To form distributions of class-wide species level trees we combined the distribution of dated backbones trees with the distributions of crown clade trees into a distribution of full trees. These then constituted the class-level Stage 1 (genetic data only) trees, consisting of 6,670 species (which includes 7 backbone species without genetic data), and the class-level Stage 2 trees including all 9,993 recognized species.

Our phylogenetic inference involves topology constraints as informed priors as follows: In the class-level Stage 1 trees 3.7% o of the 6,669 nodes (245 nodes when using the Hackett backbone and 249 when using the Ericson backbone) are fully constrained (this includes the 129 constraints due to our assignments of genera to clades and the remainder from well-supported nodes in the backbone tree as outlined above). This excluded a portion of exceedingly unlikely possible topologies from our tree search. The remaining search space on topologies alone is $>>10^{226}$, or effectively infinite.

Class-level Stage 2 trees included additional constraints: i) the Stage 1 tree position of species with genetic data informed the placement of congeners without genetic data, ii) the position of the 214 genera without genetic data within a crown clade excluded highly unlikely topologies, and iii) the 57 non-monotypic of these 214 genera were assumed to be monophyletic. This resulted in a total of 1,525 such constraints across the entire avian tree, meaning the class-level (9993-tip) Stage 2 tree had <18% nodes with some constraints associated (1,770 and 1,774 nodes for the class-level Stage 2 trees based on the 'Hackett' and 'Ericson' backbone tree, respectively). This is an upper bound as a large number of these are very weak constraints. For example, additional Stage 2 species belonging to non-monophyletic Stage 1 genera are allowed

to attach to all nodes encompassed by the genus polyphyly (see below). Further, the 214 genera with no genetic data were allowed to associate with on average 68% (median 94%) of genera within their crown clade and 103 of them (48%) with all genera.

Our class-level Stage 2 tree distributions thus encompass a very large portion of the currently remaining phylogenetic uncertainty in the extant avian tree of life. Future fossil evidence will likely lead to continued changes to dating and branch lengths of the presented tree. Future topological rearrangements in the backbone tree may affect some clades and ultimately change the assignment of several genera to their constituent crown clade. Finally, select species currently without genetic data may ultimately be placed outside the groupings currently defined by their congeners with genetic data; and in very few cases our very liberal constraints for the genera without genetic data may turn out not to be liberal enough. But these cases excepting, the class-level Stage 2 tree distribution represent a pragmatic capture of current day phylogenetic uncertainty as presented by molecular and expert evidence that offers a robust (and updatable) basis for analysis in avian diversification, trait evolution, community phylogenetics, phylogenetic diversity and comparative analyses as we continue to collect data for a more complete molecular treatment.

1.1.2 Backbone tree construction

We used BEAST v.1.5.1 (Drummond and Rambaut 2007) to construct and date backbone trees using genetic data, topological constraints, and fossil constraints. BEAST simultaneously estimates the tree topology and time-calibrated branch lengths using a Bayesian uncorrelated relaxed molecular clock (Drummond et al. 2006). BEAST also allows several alternative prior distributions on calibrations and allows the inherent uncertainty in fossil and biogeographic dating to be explicitly incorporated (Ho and Phillips, 2009).

(i) Genetic data, alignment and partitioning

For the backbone tree we collated data for 151 key species and 15 genes (19 loci), including nuclear introns from eight genes (MUSK intron 4, myoglobin intron 2, ornithine decarboxylase introns 6 and 7, rhodopsin introns 1-3, transforming growth factor beta 2 intron 5, tropomyosin 1 alpha intron 6, glyceraldehyde-3-phosphate dehydrogenase intron 11, zinc finger protein (ZENK) gene, 3' untranslated region), nuclear exons from four genes (recombination activating genes 1 and 2, proto-oncogene myc exon 3, zinc finger protein (ZENK) exon 2), and four protein coding mitochrondrial genes (cytochrome b, cytochrome oxidase I, NADH dehydrogenase subunits 2 and 3) from GenBank using Geneious 4.04 (Drummond et al. 2007). The 151 species included one species from each of 129 clades of size >= 4, 17 monotypic species and four species found across genera of size two or three. Genetic data were unavailable for seven species in the smallest clades (monotypic genera *Chunga burmeisteri*, *Xenicus gilviventris*, *Urocynchramus pylzowi*, as well as *Mesitornis variegata*, *Dasyornis brachypterus*, *Dasyornis longirostris*, *Daphoenositta miranda*) and were added to our alignments as empty sequences (see Section (ii) *Topological constraints*). Otherwise, we chose species to maximize the availability of the 15

genes for each of the 129 clades. GenBank accession numbers for each taxon and each sequence are available in Supplemental File 'backbone_tree_accessions.csv'.

Mitochondrial genes and nuclear exons were initially aligned using the local alignment tool MAFFT (Kathoh et al. 2009) and adjusted by eye where necessary using Se-Al v2.0a11 (Rambaut 1996). Nuclear introns were aligned using the Probabilistic Alignment Kit (PRANK) v.081202 (Loytynoja and Goldman 2005, 2008). PRANK treats insertions and deletions as distinct evolutionary events and avoids systematic biases associated with large numbers of indels in traditional multiple sequence alignment methods (Loytynoja and Goldman 2008). Ambiguously aligned regions in the nuclear intron alignments were removed prior to analysis. Alignments were carried out and cross-checked by GHT and JBJ. Our complete aligned dataset consisted of 13 data partitions; we utilized jModeltest 0.1.1 (Posada 2008, Guindon and Gascuel 2003) to determine the most appropriate nucleotide substitution model for each partition (Supplemental Methods Table 1). For protein coding regions (mtDNA and nuclear exons) we used codon position models (Shapiro et al. 2006) by further partitioning the data into codon positions 1+2 and codon position 3.

Supplementary Methods Table 1. Substitution models for the loci used to produce the backbone phylogeny.

Gene	Partition	Type	Model	b freq	bp	start	end
cytochrome b cytochrome oxidase 1 NADH dehydrogenase (subunit 2) NADH dehydrogenase (subunit 3)	mtDNA	mtDNA (coding)	(1+2,3) GTR+G+I	unequal	4062	1	4062
RAG1	RAG1	exon	(1+2,3) SYM+I+G	equal	2901	4063	6963
RAG2	RAG2	exon	(1+2,3) TIM3+G	unequal	1152	6964	8115
Proto-oncogene myc (exon 3)	Cmyc	exon	(1+2,3) HKY+I+G	unequal	525	8116	8640
ZENK	Zenk	exon	(1+2,3) TVM+G	unqeual	1227	8641	9867
ZENK 3'UTR	ZENK3UTR	intron	TIM2+G	unequal	502	9868	10369
Glyceraldehyde (intron 11)	Glycer-aldehyde	intron	HKY+G	unequal	217	10370	10586
MUSK intron 4	MUSK	intron	K80+G	equal	338	10587	10924
Myoglobin intron 2	Myoglobin	intron	K80+G	equal	500	10925	11424
Ornithine decarboxylase (intron 6) Ornithine decarboxylase (intron 7)	Ornithine	intron	K80+G	equal	260	11425	11684
Rhodopsin (intron 1) Rhodopsin (intron 2) Rhodopsin (intron 3)	Rhodopsin	intron	SYM+G	equal	765	11685	12449
TGFB2 (intron 5)	TGFB2	intron	TIM3+G	unequal	450	12450	12899
Tropomyosin (intron 6)	Tropomyosin	intron	GTR+G	unequal	378	12900	13277

(ii) Topological constraints

The higher-level phylogeny of birds continues to be subject to intensive study (Cracraft and Donoghue 2004; Fain and Houde 2004; Poe and Chubb 2004; Edwards et al. 2005; Ericson et al. 2006; Brown et al. 2007; Hackett et al. 2008). Recent studies, including the most comprehensive gene sequencing yet undertaken for all birds, propose that Neoaves consists of two major clades, the Metaves and Coronaves (Fain and Houde 2004; Hackett et al. 2008). However, monophyly of the Metaves is dependent on the inclusion of the β-fibrinogen gene (Ericson et al. 2006; Hackett et al. 2008). Studies using complete mitochrondial genomes (Gibb et al. 2007; Morgan-Richards et al. 2008) or multiple nuclear genes excluding β-fibrinogen (Ericson et al. 2006) have failed to recover Metaves. Morgan-Richards et al. (2008) suggest that the high number of indels in β-fibrinogen may present problems in sequence alignment leading to artifacts during analysis. Analyses of nuclear DNA excluding β-fibrinogen (e.g. supplementary figure 6 in Ericson et al. 2006) are highly congruent with those based on mitochondrial DNA (Morgan-Richards et al. 2008).

It is clear that β-fibrinogen has a major and not yet fully understood impact on higher-level avian relationships, particularly among the non-passerines. We therefore constructed and dated backbone trees using topological constraints derived from two main sources, which either include (Hackett et al. 2008) or exclude (supplementary figure 6 in Ericson et al. 2006) β-fibrinogen. We supplemented the Hackett et al. and Ericson et al. topologies with data from a range of sources (Supplementary Methods Table 2); in particular we derived topological constraints for the Passeriformes mainly from Barker et al. (2004). Constrained topologies (with no branch-length information) for "Ericson" and "Hackett" trees are shown in Supplementary Methods Figure 1 where black circles indicates nodes that were constrained and nodes labeled A-J correspond to the fossil priors detailed in Supplementary Methods Table 3 and the next section. We use the "Hackett" backbone for our primary analyses due to the more extensive genomic scope of loci used in the study. However and importantly, analyses based on the Ericson backbone do not alter our conclusions.

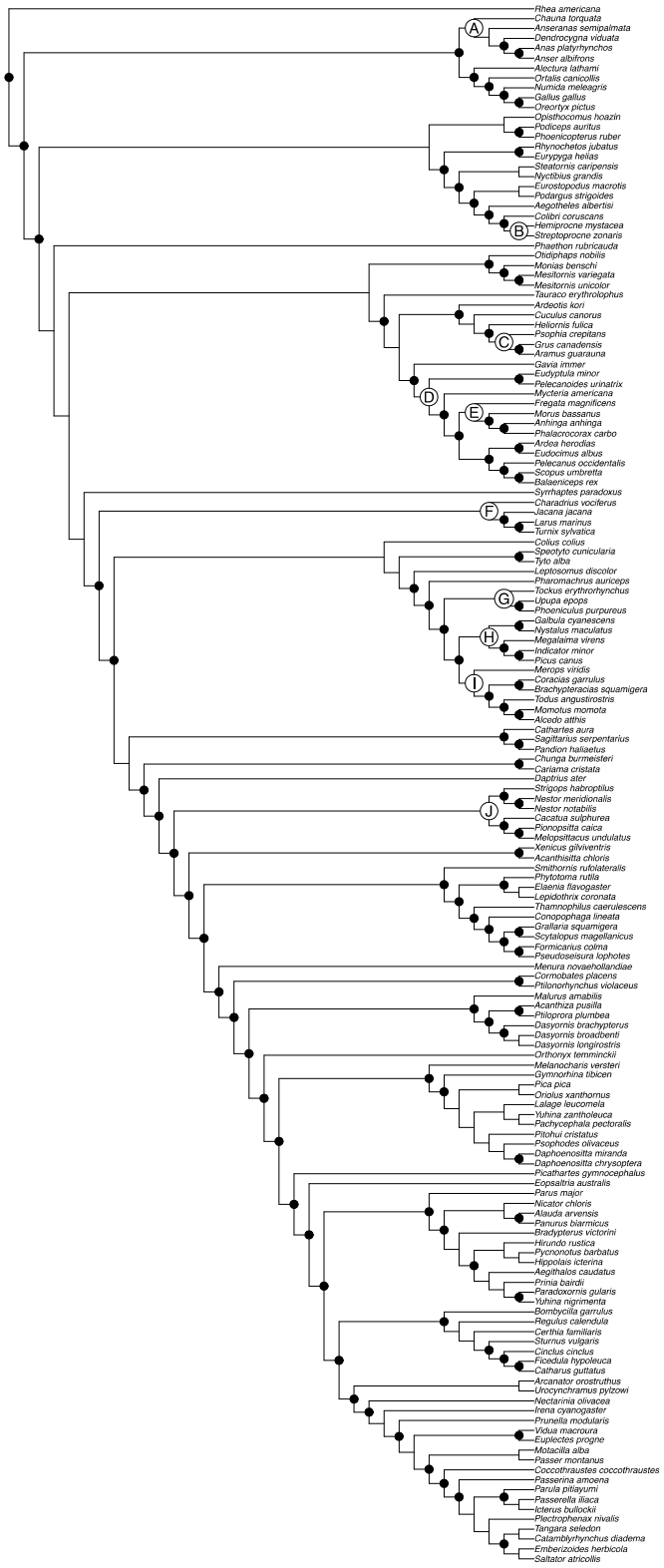
All additional sources for topological constraints are documented in Supplementary Methods Table 2. We only used constraints for nodes that are well-supported (posterior probability >0.95 or bootstrap >70%) in the respective source trees. This allows us to explore and potentially resolve topologically uncertain parts of the published trees.

We then screened our data matrix for data decisiveness, a potential problem where some species have such low data overlap with others that their movement across candidate trees does not affect the tree likelihood, such that tree inference can be seriously compromised (Sanderson et al. 2010). We measured data decisiveness as the proportion of branches that the data could distinguish across 1000 random trees consistent with our constraints (Sanderson et al., 2011): average decisiveness was >0.98 for both backbones.

Supplementary Methods Table 2. Sources for backbone constraints. In addition to the three main sources (Hackett et al. 2008 or Ericson et al. 2006 for the non-passerines; Barker et al. 2004 for passerines) several other sources were used to add further resolution. The column "Figure in ref" refers to the figure in the source reference that we used to inform constraints for the "Clade" highlighted in column 4.

Backbone	Reference	Figure in ref	Clade
NP, Hackett	Hackett et al. (2008)	Fig. 2,4	Non-pass (main source)
NP, Hackett	Wright et al. (2008)	Fig. 2	Parrots
NP, Hackett	Crowe et al. (2006)	Fig. 4	Galliformes
NP, Hackett	Donne-Goussé et al. (2002)	Fig. 5b	Anseriformes
NP, Ericson	Ericson et al. (2006)	Suppl. Fig. 6	Non-pass (main source)
NP, Ericson	Wright et al. (2008)	Fig. 2	Parrots
NP, Ericson	Crowe et al. (2006)	Fig. 4	Galliformes
NP, Ericson	Donne-Goussé et al. (2002)	Fig. 5b	Anseriformes
NP, Ericson	Baker et al. (2007)	Fig. 1	Charadriiformes
Passerine	Barker et al. (2004)	Fig. 1	Pass (main source)
Passerine	Ohlson et al. (2007)	Fig. 2	Tyrannidae
Passerine	Moyle et al. (2009)	Fig. 2	Furnariida
Passerine	Driskell et al. (2004)	Fig. 3	Basal Corvida
Passerine	Jonsson et al. (2008a)	Fig. 2	Core Corvida
Passerine	Norman et al. (2009)	Fig. 4	Core Corvida
Passerine	Dumbacher et al. (2008)	Fig. 1	Core Corvida
Passerine	Jonsson et al. (2008b)	Fig. 1	Core Corvida
Passerine	Fuchs et al. (2006)	Fig. 4	Core Corvida
Passerine	Ericson et al. (2005)	Fig. 5, 3a	Core Corvida
Passerine	Beresford et al. (2005)	Fig. 1	Muscicapida
Passerine	Beresford et al. (2005)	Fig. 1	Sylvoidea
Passerine	Johansson et al. (2008)	Fig. 2	Sylvoidea
Passerine	Beresford et al. (2005)	Fig. 1	Passerida
Passerine	Klicka et al. (2007)	Fig. 1	Passerida
Passerine	Yuri & Mindell (2002)	Fig. 3b	*Catamblyrhychus*
Passerine	Groth (2000)	Fig. 1	*Urocynchramus*

Ericson constraint



Hackett constraint

Supplementary Methods Figure 1 (two pages). (top panel) The 'Ericson' and (bottom panel) 'Hackett' constraint trees used for the backbone tree analyses, with nodes used for time calibration indicated with letters (see Supplementary Methods Table 3). Branch lengths are not to scale and are displayed so that all nodes can be clearly viewed.

(iii) Fossil constraints

We used ten stratigraphically and phylogenetically well-known fossils to calibrate the backbone trees, outlined in Supplementary Methods Table 3.

Supplementary Methods Table 3. Fossil calibration dates, sources and parameters for the lognormal prior distribution. Node refers to the backbone constraint trees. The Fossil age is used to define a hard minimum age constraint. The soft maximum age of each constraint is given by *Gansus yumenensis*, from the Early Cretaceous Xiagou formation (~110 MYA) of northwestern China (You et al. 2006).

Taxon	Node	Fossil age (min)	95% quant.	97.5% quant.	Calibration density (sd)	Calibration density (mean)	Ref
Root	root	66	110	117.5	0.5	85.33	1
Stem Pici (Picidae, Ramphastidae, Indicataoridae)	H	30	88.38	110	1	41.27	2
Stem Upupidae + Phoeniculidae	G	47.5	92.92	110	1	56.27	3,4
Stem Coraciidae + Brachypteraciidae	I	47.5	92.92	110	1	56.27	4,5
Stem Anatidae	A	66	98.11	110	1	72.20	6
Stem Fregatidae	E	52.5	94.46	110	1	60.60	7,8
Stem Sphenisciformes	D	60.5	96.62	110	1	67.47	9
Stem Apodidae	B	47.5	92.92	110	1	56.27	10, 11
Stem Grues	C	30	88.38	110	1	41.27	12
Stem Psittacidae + Cacatuidae	J	54	94.86	110	1	61.89	13, 14, 15
Crown Charadriiformes	F	55.8	95.35	110	1	63.43	16, 17

References: 1) You et al. 2006; 2) Mayr 2005a; 3) Mayr 2000; 4) Franzen 2005; 5) Mayr & Mourer-Chauvire 2000; 6) Clarke et al. 2005; 7) Olson 1977; 8) Simmons et al 2008; 9) Slack et al. 2006; 10) Mayr 2003; 11) Harrison 1984; 12) Mayr 2005b; 13) Waterhouse et al. 2008; 14) Kundu et al. 2011; 15) Pacheco et al. 2011; 16) Bertelli et al. 2010; 17) Dyke & van Tuinen 2004.

Following the recommendation of Ho and Phillips (2009) we avoid the common but flawed approach of using fossils as minimum age constraints (e.g. Ericson et al. 2006; Brown et al. 2007), but rather used them to set calibration densities on node ages. Specifically, we used a lognormal calibration density on each fossil date. The lognormal calibration density is particularly well-suited to calibrations based on palaeontological data because in addition to a hard minimum bound, it also allows the mean (the highest point probability) to be set slightly

older than the fossil date reflecting the expected age underestimation associated with fossil calibrations (Ho and Phillips, 2009). Furthermore, the tail of the lognormal distribution can be set as a soft maximum bound (Benton and Donoghue 2006, Yang and Rannala 2006) on node ages. Benton and Donoghue (2006) propose the use of the clades Ichthyornithiformes and Hesperornithiformes (that together with the Neornithians form the Ornithurae) of the Niobrara Chalk Formation to suggest a soft maximum age of 86.5 MYA for the divergence of palaeognaths and neognaths. However, the discovery of an older member of the Ornithurae, *Gansus yumenensis*, from the Early Cretaceous Xiagou formation of northwestern China (You et al. 2006) suggests that a soft maximum of ~110 MYA is more appropriate. For calibration densities on the ten calibrated internal nodes, we set the mean and standard deviation of the lognormal distribution so that 97.5% of the probability distribution lies between the hard minimum and soft maximum bound. For the root node we set the hard minimum bound at 66 MYA on the basis of the oldest robust fossil member of the Neornithes (*Vegavis iaai*; Clarke et al. 2005), and set the mean and standard deviation of the lognormal distribution so that 95% of the probability distribution lies between the hard minimum and soft maximum bound.

(iv) Relaxed clock dating with BEAST

We used the topological and fossil constraints described above, along with the uncorrelated lognormal relaxed clock prior on molecular rates (in which rate variation among adjacent branches is assumed to follow an uncorrelated lognormal distribution) and a Yule speciation process prior on tree shapes to estimate and date the higher-level relationships of birds. We ran two sets of analyses, the first with "Ericson" topological constraints (supplementary figure 6 in Ericson et al. 2006, without β-fibrinogen) and the second with "Hackett" topological constraints (figure 2 in Hackett et al., 2008). For each set we ran four independent chains of 100 million generations (~5760 hours of CPU time). We assessed convergence, mixing and burnin by visual examination of plots of the parameters and on the basis of effective sample size (ESS, Drummond et al. 2006) using Tracer v1.4.1 (Rambaut & Drummond 2008). An ESS >200 for a continuous parameter indicates that the chain has run for an adequate number of generations (Drummond et al. 2006). We discarded the first 10 million generations from each chain as burnin. Note that since the majority of nodes are constrained, we were also able to examine convergence and mixing of node ages using Tracer. After discarding burnin and combining the output from the four independent chains the ESS for all parameters was >200.

We report results from the Hackett backbone for our main analyses, but results from the two full tree distributions (Hackett and Ericson) are almost identical: the Spearman Rank correlation of the mean species-level lineage diversification rate (see section 1.2.2 for explanation) across the two distributions of trees exceeds 0.999. The maximum clade credibility tree with posterior values and 95% c.i. on the edge lengths for each of the two sets of constraints (as a nexus file) can be found in Supplementary Methods Archive 'MCC_trees.zip.'

1.1.3 Phylogenetic inference for each crown clade

(i) Recognized species and clade assignment

We first assigned species to genera and genera to clades. Our treatment of species follows the Birdlife V3 world list (June 2010, 9,895 extant species recognized). Of Birdlife V3 we did not recognize nine species (*Anodorhynchus glaucus, Gallinula pacifica, Gallirallus lafresnayanus, Oceanodroma macrodactyla, Ophrysia superciliosa, Rhodonessa caryophyllacea, Siphonorhis americana, Tadorna cristata, Vanellus macropterus*) that are widely considered extinct and three (*Heliangelus zusii, Atlapetes blancae, Upupa marginata*) that are considered not valid by most authorities (resulting in 9,882 accepted species). In addition to Birdlife V3 we recognize 111 species that are considered valid by Handbook of the Birds of the World (del Hoyo et al. 1992-2011) and/or Birds of the Western Palearctic (Cramp et al. 1978-1994) and/or Birds of Africa (Urban et al. 1986-2000) and that are also recognized by IOC world list V2.7 (DEC 29, 2010), resulting in a total of 9,993 recognized species. Our taxonomy is found in Supplementary Methods File 'mastertaxonomy.csv'.

The backbone determines a total of 129 avian crown clades with > 4 species (varying from 5 to 460 species in size, see clade_summary.csv). We used a careful evaluation of the existing avian phylogenetic literature (including a total of 179 references from the primary literature, available from authors upon request) to assign each avian genus to one of the 129 clades. For ten genera we found equivocal evidence for clade membership and used a combination of molecular and literature evidence and expert judgment to assign the members of such genera to one single (*Chamaeza*, *Pteroptochos*, *Yuhina*), two (*Alcippe*, *Bradypterus*, *Neomixis*, *Orthotomus*, *Saltator*) or in one case, three (*Pitohui*) clades.

(ii) Genetic data, alignment and partitioning

We collated genetic data for 6,641 of the 9,993 species in these 129 clades and ten loci to construct clade trees (data for an additional 22 species were included only in the backbone trees). Data were obtained from GenBank using Geneious 4.04 (Drummond et al. 2007) with additional data for the family Cuculidae supplied by M. Sorenson (Sorenson and Payne, 2005) and consist of: four protein coding mitochrondrial genes (cytochrome b, 4902 species; cytochrome oxidase I, 2335 species; NADH dehydrogenase subunit 2, 4308 species; and NADH dehydrogenase subunit 3, 1232 species), and six nuclear loci (recombination activating protein 1 [rag-1], 1528 species; beta-fibrinogen intron 5 [bfib5] 5, 1089 species; beta-fibrinogen intron 7 [bfib7], 1460 species; glyceraldehyde 3-phosphate dehydrogenase [gapdh], 967 species; myoglobin [myo], 1867 species; and ornithine decarboxylase [odc], 1405 species). Final updates on the alignment were completed on April 13th, 2011. All sequences were aligned using MAFFT and inspected by eye in Se-Al v2.0a11 (as above) and were independently crosschecked by GHT and JBJ. GenBank accession numbers are available in Supplementary Methods File 'clade_tree_accessions.csv'.

We treated the mitochondrial loci as a single gene, partitioned by codon position. We treated each nuclear loci as a separate partition and determined the nucleotide substitution model for each partition for each clade separately in mrmodeltest 2.3 (Nylander, 2004). We further

partitioned the four coding genes (rag1, gapdh, myo, and odc) into codon positions 1+2 and codon position 3 (Shapiro et al. 2006).

We initially included all partitions in a focal clade with data for at least two species. We analysed each clade with two outgroup species, one from each of two nearby outgroup clades identified using the Maximum Clade Credibility backbone tree ('Ericson' or 'Hackett'), and with these outgroup clades being chosen to be paraphyletic with respect to the focal clade. In step one, we first identified the most common partition for nuclear coding genes in the focal clade (from rag1, myo, odc, and gapdh). We then identified all clades in the sister clade to a focal clade on the backbone, and all species in this pool which had data for this most common nuclear coding partition in the focal clade were considered candidate species. Then, in a second step, the individual overlap for each candidate species was compared with the second most common nuclear coding gene partition in the focal clade, and the species with the most overlap was chosen as the first outgroup. In order to increase the likelihood of using rag-1 to root trees, if rag-1 was not the most common partition identified in the first step, it was used preferentially to identify outgroup species in the second step. If no candidate species in the candidate clades had rag-1 sequences present (or if rag-1 was not present in the focal clade) the second-most common nuclear coding gene partition was used. The Maximum Clade Credibility backbone tree was then traversed down a node toward the root and the process was repeated, but focusing on the partitions identified for the first outgroup. In the cases where we could not use two nuclear coding genes, we used the most common partition from the rest of the candidate gene partitions.

As with the backbone trees, we screened our data matrices for data decisiveness. For each of 129 monophyletic clades in our tree with >= 4 taxa, we measured triplet coverage, and then iteratively dropped species until every possible triplet of species was represented by at least one locus (a necessary but not sufficient condition for decisiveness). Fifty-seven clades failed this decisiveness tests and thus required pruning (24 clades were pruned of one species, 10 clades were pruned of 2 species, 7 clades were pruned of three species, and 16 clades were pruned of >3 species). We then measured the average data decisiveness on these trimmed datasets as the proportion of branches that the data could distinguish across fully random trees (Sanderson et al., 2011). For 1000 random trees on the trimmed datasets, this was always 100%. We note that showing a dataset is not decisive for all trees (or a random set of trees) does not mean it is indecisive for trees near the true tree. We rebuilt clade trees for these trimmed data sets, and recreated the distribution of full bird trees. The Spearman rank correlation of the mean species-level lineage diversification rate (DR) values (see section 1.2.2) from the original and fully decisive tree distributions exceeded 0.997.

(iii) Stage 1 tree inference

We inferred initial posterior distributions of phylogenies for each clade using the substitution models determined above, a relaxed clock model (independent branch rates - ibr prior), with the default (exponential) prior on the distribution of branching rates of the ibr prior, a Yule prior on

branch lengths, an all-trees-equiprobable prior on topology, and with the two outgroup species constrained as outgroups. A sample tree inference nexus block for a single run would be:

```
prset brlenspr=clock:birthdeath;
prset Extinctionpr = Fixed(0);
prset Speciationpr=exponential(1);
prset clockvarpr=ibr;
prset ibrvarpr=exponential(10);
mcmcp nruns=1 nchains=1 ngen=50000000 samplefreq=1000;
propset ParsSPRClock(Tau{all},V{all})$warp=0.01;
```

As with the backbone tree, we assessed convergence, mixing and burnin by visual examination of parameter plots and on the basis of effective sample size (ESS, Drummond et al. 2006) using both Tracer v1.4.1 (Rambaut & Drummond 2008) and AWTY (Nylander et al. 2008). We kept four runs that converged on the same tree likelihood distributions. For a handful of clades, it was necessary to modify the substitution model to allow variable rates (prset ratepr=variable;) to achieve convergence. We used the hashcs program (Sul & Williams, 2009) to produce consensus trees by combining the post burn-in trees from each of the four runs. The Maximum Clade Credibility trees for each clade can be found in Supplemental Methods Archive 'MCC_trees.zip' and distributions of these Stage 1 trees can be found at birdtree.org.

(iv) Topological constraints for Stage 2 trees

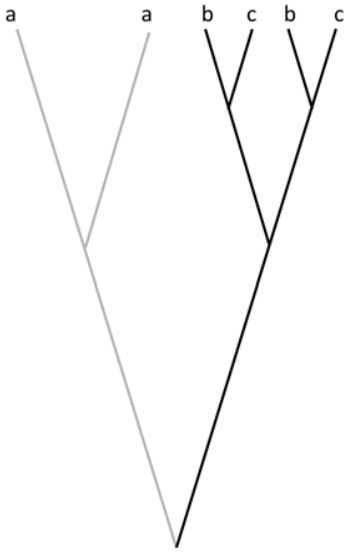
Inference about evolutionary patterns across time and space require complete dated trees. Such complete trees are almost always restricted to small clades and small spatial scales. The motivation for this study was to provide a methodology to overcome this limitation by integrating all available systematic information. We develop a general approach that is suited to any taxonomic group with missing phylogenetic information.

For birds the most taxonomically inclusive published tree remains Sibley and Ahlquist's (1990) tapestry, which contains ca. 15% of extant species sampled across the Class. More recent studies have focused either on small subsets of taxa chosen to illuminate deeper avian relationships (e.g. Hackett et al. 2008) or on portions of the clade Aves such as Passerines (e.g. Hugall and Stuart-Fox 2012). Our specific objective here was to produce a reasonable pseudo-posterior distribution of trees including all extant species in a way that is consistent with raw sequence data and taxonomic information. This is achieved using a method that samples trees in a MCMC framework such that they are consistent with the genetic data, and include all species in a consistent fashion using placement constraints and a birth-death model of diversification. Placement constraints are statements dictating where species can and cannot be placed while inferring a phylogenetic tree (see, e.g., Lanfear and Bromham 2011).

A given clade now consists of species that fall into one of three categories: Type 1 species have genetic information and are found in a stage 1 tree topology; type 2 species are congeners of a species found in a stage 1 tree topology; and type 3 species are members of a genus not represented in that preexisting topology (214 genera out of 2091 total). For each clade, we formulated the simplest possible placement constraints that combine taxonomic (type 2, 3) and

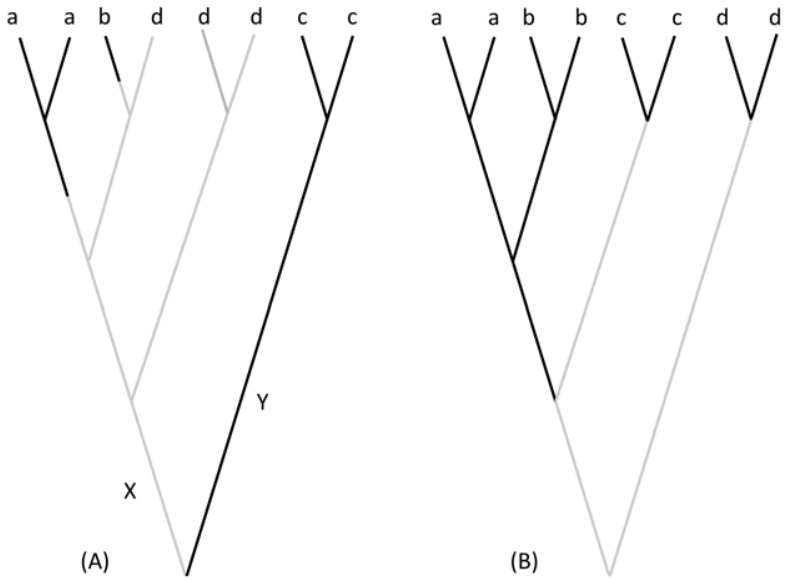
genetic (type 1) data species as follows. First, type 1 species were anchored relative to one another using the consensus trees from the unconstrained relaxed clock searches (the Stage 1 trees). For the great majority of clades, we used the 95% consensus tree from our posteriors as our Stage 1 tree (created using hashcs from Sul and Williams 2009), but in six cases, we used the 75% consensus tree to allow for more precise placement of missing species, and in one case (the 'Titytyrannidae' clade), we used the 50% majority-rules consensus tree. Supplementary Methods file 'clade_summary.csv' lists these, and Supplementary Methods file 'ConstraintTrees.zip' contains these trees as nexus files. It is important to note that the constraints imposed by the use of these Stage 1 trees will be found in every tree in the stage 2 distributions.

Second, we restricted each type 2 species to its genus or its 'supragenus' as described in Supplementary Methods Figure 2. Supragenera are defined to accommodate named genera with evidence of non-monophyly in the stage 1 trees.



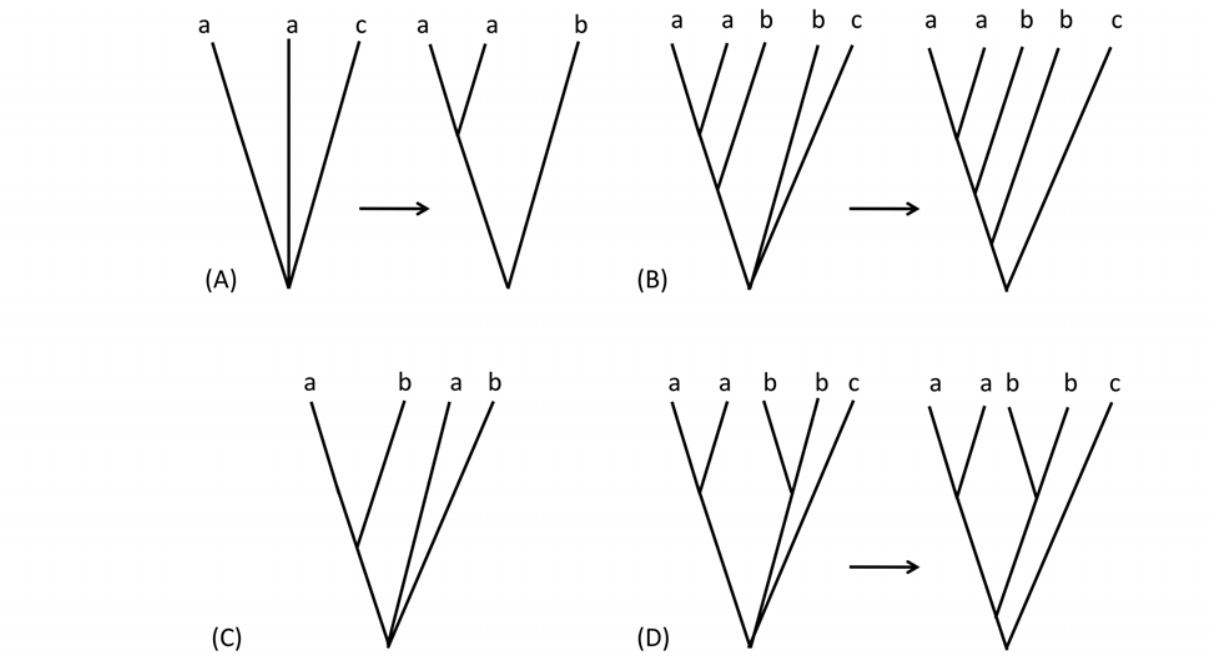
Supplementary Methods Figure 2. Placement of congeners of a species found in astage 1 tree topology (type 2 species). A type 2 species of genus *a* is confined to the area of the tree containing genus *a* (highlighted in grey). A type 2 species of genus *b* is confined to its supragenus as genus *b* is not monophyletic (highlighted in black).

Third, each type 3 species was restricted to its named genus, which in turn was constrained in its placement among other genera using additional taxonomic information (Supplemental Material 'missing_genera.csv'). Supplementary Methods Figure 3 illustrates how these various placement constraints interact.



Supplementary Methods Figure 3. Here we consider the interaction among placement constraints for species without genetic data (type 2, 3). (A) A type 2 species of genus *a, b* or *c* will be confined to its respective genus (areas in black). Genus *d* is not monophyletic, hence a type 2 species of genus *d* is confined to the supragenus including genera *a, b* and *d*. However, these type 2 species may not enter any genus within that supragenus and so can only enter the grey areas (i.e., below all *a* species and all *b* species, represented as edges that start black and become grey). A type 3 species that groups with genera *a* and *b* will be restricted to the edge *X* as it may not enter the smallest supragenus containing both genera *a* and *b*. A type 3 species that groups with genus *c* is restricted to the edge *Y* because it may not enter genus *c*. (B) Taxonomic information may suggest that a type 3 species is aligned with genera *c* and *d*. It will then be restricted to the grey edges (i.e. below all *c* and *d* species).

Finally, where possible, polytomies in the Stage 1 consensus topologies were resolved prior to the sampling approach to ensure that: (i) genera are monophyletic if possible, (ii) supragenera are as small as possible; and (iii) Type 3 species are placed among the smallest group of genera consistent with the taxonomic information in Supplementary Methods File 'missing_genera.csv'. This step uses taxonomic information and monophyly of named genera to resolve the Stage 1 topologies. Supplementary Methods Figure 4 provides an illustration of this method.



Supplementary Methods Figure 4. Placement of type 3 species (members of a genus not represented in the stage 1 topology) where there are polytomies. (A) The polytomy is resolved by enforcing the monophyly of genus *a*. (B) The polytomy is resolved as the supragenus containing genera *a* and *b* is required to be as small as possible. (C) This polytomy is not resolved as the monophyly requirements of genera *a* and *b* conflict. (D) This polytomy is not resolved as there is no taxonomic information to support this (genera a and b are already monophyletic). However if a genus containing Type 3 species is known to group with genera a and b, this will break the polytomy.

(v) Stage 2 tree inference

We inferred initial posterior distributions of phylogenies for each clade using the substitution models inferred above, a relaxed clock model (independent branch rates - ibr prior), with the default (exponential) prior on the distribution of branching rates of the ibr prior, a birth-death prior on topologies, and with the two outgroup species constrained as outgroups. These prior birth-death models can be used to sample topologies and edge lengths given the topological constraints presented above and genetic data. We used MrBayes v3.2 to sample trees from birth-death model priors when no sequence data were available. For all the trees we report here, we set extinction = 0. This is not because we believe there has been negligible extinction over the course of the bird radiation, but because recent comparative work on the shapes of molecular-clock trees suggests that there is no signal of *homogeneous* extinction rates on the resulting tree shapes (Phillimore & Price 2008; Rüber & Zardoya 2005; Weir 2006; McPeek 2008; Morlon et al. 2010). Any signal of extinction in our final tree inference will subsequently be conservative by this choice of prior.

A sample tree inference nexus block would be:

```
prset topologypr=constraints(…);
prset brlenspr=clock:birthdeath;
```

```
prset Extinctionpr = Fixed(0);
prset Speciationpr=exponential(1);
prset clockvarpr=ibr;
prset ibrvarpr=exponential(10);
mcmcp nruns=1 nchains=1 ngen=50000000 samplefreq=1000;
propset ParsSPRClock(Tau{all},V{all})$warp=0.01;
mcmc;
```

We produced four runs for each clade. As with the initial tree production, we assessed convergence, mixing and burnin by visual examination of plots of the parameters and on the basis of effective sample size (ESS, Drummond et al. 2006) using Tracer v1.4.1 (Rambaut & Drummond 2008). The resulting ultrametric trees were then dated using the backbone tree.

1.1.4 Backbone-clade grafting - an interim class-level tree

We produced a pseudo-posterior distribution of class-level complete bird trees by sampling backbone trees and clade trees from their respective posterior distributions using a simple R script. A backbone tree was sampled randomly without replacement from the final posterior distribution, and then a tree was drawn (randomly without replacement) from each clade distribution to be grafted into that backbone tree. We scaled the depth of each clade tree from the depth of the node on the sampled backbone tree corresponding to the most recent common ancestor of the clade and its first outgroup. We dropped the two outgroups from the clade tree and then substituted it into the backbone tree in place of the single species that had represented it. So, for example, if the depth of the ingroup node in clade A (i.e. its first split) is 0.60, and the depth of the split between clade A and its first outgroup is 0.80 (noting that we always scaled the depth of the tree of the clade plus its two outgroups to 1.0), then, if the depth of the node linking the clade A and the clade containing clade A's outgroup is 16 MYA, the ingroup age of the focal clade is set to 0.6/0.8*16=12 MYA.

In cases where the first outgroup for a clade was not chosen from the clade's immediate sister clade (due to the sister clade having no candidate species with data for the clade's most common nuclear partition, or due to the sampled backbone tree being different from the MCC topology used to identify outgroups) and where the inferred age of the clade was older than the stem age from the backbone (e.g., if the species representing clade A joined the tree at 10 MYA in the example above), the clade was attached to the backbone with a stem edgelength of ca. 0.0001.

Sets of 10,000 full bird trees for both the Hackett and Ericson backbones, and for both complete 'stage 2' and DNA-data only ('stage 1') trees (excepting the seven species included in the backbone with no genetic data), are downloadable from birdtree.org.

1.2 Analyses of Diversification Rates

1.2.1 Tree and clade-level analyses

To investigate the pattern of diversification across time, we visualized lineage-through time (LTT) plots (Nee et al. 1992, the cumulative number of lineages in a dated molecular phylogeny on a logarithmic scale graphed against time when the lineages arise) and then used Maximum Likelihood approaches for estimating diversification rates and fitting a priori models to trees, implemented in the R packages LASER (Rabosky 2006), TreePar (Stadler 2011, 2012), and the turboMEDUSA version (Brown et al. 2012) of MEDUSA (Alfaro et al. 2009).

We used the bd.ME.optim function in TreePar (v.2.2; Stadler, 2012) to estimate speciation (λ) and extinction (μ) rates (in units of my^{-1}) (technically, $\rho = \lambda - \mu$ and $\varepsilon = \mu/\lambda$) for each five million year time slice with >= 30 branching times (thus starting 67.5 mya) for each of 525 trees drawn from our distribution (see Main Text Figure 1). Because of the well-known problem that when speciation takes time, unfinished speciation events will not be recorded near the present (see Phillimore and Price 2008), we excluded the most recent 2.5 million years from the visualization.

We then evaluated the fit of nine models of diversification using an AIC framework. Using LASER, we assessed pure birth, constant-rate birth-death, density-dependent logistic (where extinction is set to zero and speciation rate follows a logistic curve as a function of standing diversity (Sepkowski 1978, Rabosky and Lovette 2008a)), density dependent exponential (where extinction is set to zero the speciation rate is a function of the number of extant lineages at any point in time) and two more complex models (SPVAR and EXVAR; Rabosky and Lovette 2008b).

For the logistic model, we changed the K (carrying capacity) default in LASER from 2 to 100 times standing diversity. The default setting for the fitSPVAR function is to model diversification as an exponentially *declining* speciation rate and a constant extinction rate (note that the extinction rate is estimated). We altered the default setting to also allow estimation of exponentially *increasing* speciation rates if necessary. The fitEXVAR function models diversification with exponentially increasing extinction rates and constant (estimated) speciation rates. We did not alter the fitEXVAR function. The default for fitBOTHVAR is an exponentially decreasing speciation rate and an exponentially increasing extinction rate. We altered fitBOTHVAR to allow exponentially increasing speciation. However, for most trees we were unable to optimize the fitBOTHVAR model. We note that for trees where we were able to optimize fitBOTHVAR the model fit was indistinguishable from the altered fitSPVAR model.

Using TreePar, we used the birth-death-shift model (Stadler, 2011) to assess the fit of temporal changes in diversification rates. Rate shifts were permitted at any two million year interval. We used likelihood ratio tests to infer the number of supported shifts (for us, never more than 2).

Finally, we used turboMEDUSA to optimise models with 20 clade-specific diversification rate shifts to each tree. The MEDUSA approach implemented in turboMEDUSA sequentially

identifies clades with major shifts in diversification (Alfaro et al. 2009). Statistical power to detect rate increases in MEDUSA is higher than that to detect rate decreases, though decreases are detectable (Alfaro et al., 2009). We standardized the likelihoods across the three modeling approaches (LASER, MEDUSA and TreePar) so that AIC values would be comparable using the likelihood of the fits of the constant rates birth-death model for each (Stadler, pers. comm.). To summarize and display the MEDUSA results across our sample of trees we first extracted a single representative tree from the sample. This tree is presented as Main Text Figure 2. We then assessed the correspondence between nodes in the representative tree and each of the 525 sampled trees. We first identified nodes with an exact match between a given sampled tree and the representative tree. Where there was no exact match we identified the node in the representative tree that represented the most recent common ancestor (MRCA) of all taxa within a given shift in a given sampled tree. In Main Text Figure 2 we plot node labels depicting the proportion of sampled trees with exact shifts in black and the proportion of sampled trees with MRCA shifts in grey. For simplicity we identify only those nodes where either the proportion of trees with an exact match, or the proportion of trees with an MRCA match, or the proportion of trees with summed exact and MRCA patches was >25%. This yielded a total of 25 rate shifts.

1.2.2 Individual lineage level analyses

Diversification rates have traditionally been estimated for clades rather than species. However, it is possible to assign an integrated rate of a lineage to a species, i.e. along the path from the root of a tree to an individual species. Sister species have the same history and the same value for any such measure. The only such metric we know of in the literature is the number of observed splits (or nodes) along this path (Freckleton et al., 2008). We present a related measure of such a species-level lineage diversification rate, DR, which we adapt from the literature on evolutionary isolation and which, importantly, maps onto clade-level estimates of the net diversification rate (λ-μ).

The Equal Splits (ES) measure of evolutionary isolation (Redding and Mooers 2006) uniquely apportions a phylogenetic tree among its tips, achieved by dividing the phylogenetic distance represented by an edge equally among its daughter edges. Sharing parent edges among daughter edges ensures that sum of the ES measure across the tips equals the sum of all the branch lengths of the tree (see Pauplin 2000; Semple and Steel 2003). The ES measure for a focal tip on a rooted bifurcating tree is the sum of the edge lengths from the species i to the root, with each consecutive edge discounted by a factor of ½:

$$ES_i = \sum_{j=1}^{N_i} l_j \frac{1}{2^{j-1}} \qquad \text{(Supplementary equation 1)}$$

N_i is the number of edges on the path from species i and the root, and l_j is the length of the edge j, with $j=1$ being the pendant edge leading to the species and $j=N_i$ being the edge nearest the root

on that path. For a more general equation for non-bifurcating trees, see Redding et al. (2008). Species with fewer close relatives are apportioned more of the tree and so a higher weighting as less of the path length between the tip and the root is shared by other species.

The inverse of this measure can be seen as a measure of the splitting rate of the path to a tip: species in rapidly-diversifying clades will have short edge lengths shared among many species and low ES values, while isolated species on a tree have no evidence of recent diversification and large ES values. We term the 1/ES metric for a species its species-level lineage diversification rate, or DR.

One advantage to using 1/ES as a measure of lineage-specific diversification is that it can be shown that the mean value across tips of 1/ES for a Yule tree diversifying at rate λ rapidly approaches λ. To see this, we first express the expected value of ES across tips on a Yule-shaped tree as

$$E(ES) = E(l)\sum_{k=0}^{E(N)} \frac{1}{2^k} \qquad \text{(S. eq. 2)}$$

where l denotes length of an edge, and N denotes the number of edges from a tip to a root, and the $E()$ denotes an expected value. (Note, we have changed the subscripts from j to k between Supplementary equation 1 and S. eq. 2) to facilitate the comparison of ES and λ.)

On a Yule topology, the expected number of nodes from a tip to the root on a tree with n tips is a known harmonic sum (McKenzie and Steel 2000; Mulder 2011):

$$E(N) = 2\sum_{k=2}^{n} \frac{1}{k} \qquad \text{(S. eq. 3)}$$

$E(l)$ can be also be expressed as $t/E(N)$, i.e. the depth of the tree t divided by the average number of edges that depth is divided among. (As an added intuition, recall that $E(l) = t/2\ln(n/2)$ (Steel and Mooers 2011) and that N converges to $2\ln(n)$ (Steel and McKenzie 2000): the ratio of these two expressions for $E(l)$ converges to 1.) This means we can express $E(ES)$ as:

$$E(ES) = \frac{t}{E(N)}\sum_{k=0}^{E(N)} \frac{1}{2^k} \qquad \text{(S. eq. 4)}$$

The reciprocal, $E(ES)^{-1}$ is well-represented as

$$E(ES^{-1}) = \frac{E(N)}{t}\left(\sum_{k=0}^{E(N)} \frac{1}{2^k}\right)^{-1} \qquad \text{(S. eq. 5)}$$

The second term converges very quickly to ½ from 1: at $n = 10$ tips, it is ~0.516.

Now, the maximum likelihood estimate $MLE(l) = (n-2)/S$, where S is the sum of all the edge lengths in a bifurcating tree of size n (Nee, 2001). Because pendant and interior edges are the same mean length in expectation on a Yule tree (and on trees with mild extinction), $E(S)$ can be

expressed as *(2n-2)*E(l)*, i.e. the number of edges times their expected length (Steel and Mooers, 2011). This means that:

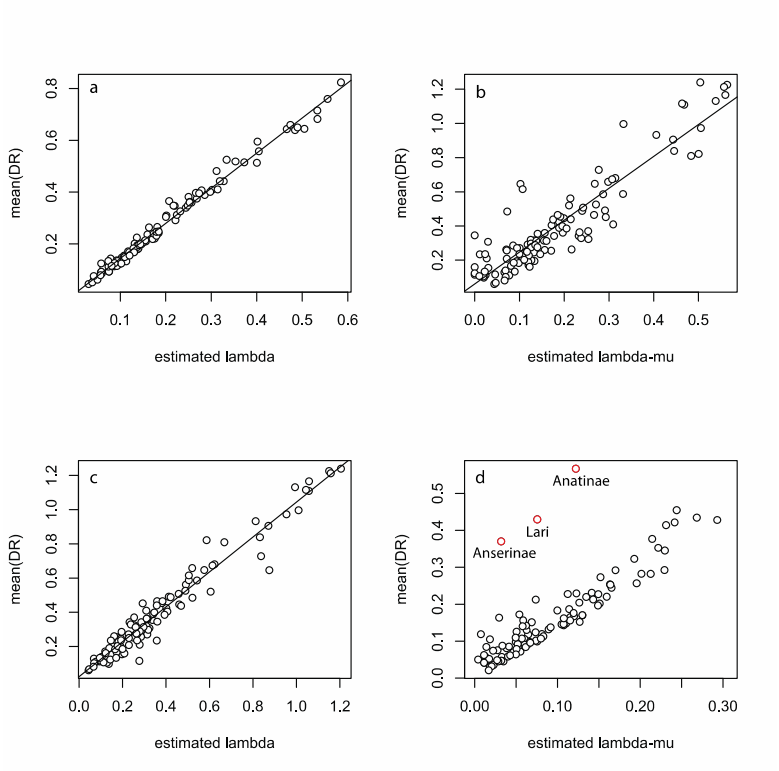
$$MLE(\lambda) = \frac{(n-2)}{E(S)} = \frac{(n-2)}{(2n-2)E(l)} = \frac{1}{2E(l)} = \frac{1}{2} * \frac{E(N)}{t} \qquad \text{(S. eq. 6)}$$

So, the maximum likelihood estimate of λ on trees grown under a Yule process with depth *t* and size *n* can be expressed as (1/2)**E(N)/t* (S. eq. 6), while the reciprocal of the expected equal splits metric $E(ES)^{-1}$ on that same tree converges to this quantity rapidly from above (S. eq. 5).

In Supplementary Methods Figure 5, we compare standard clade-wide diversification estimates with our measure of diversification rate, using the R package TreePar (v.2.2; Stadler, 2011) to simulate trees conditioned on reasonable values for *n* and *t*. For both Yule trees and for homogeneous birth-death trees with $\mu = 0.5*\lambda$ we find that the inverse of the average value of (ES) across tips increases tightly and linearly with the MLE diversification rates as estimated using the LASER package (Rabosky 2006). Under homogeneous birth-death models, DR is a better measure of the lineage speciation rate (lambda, r=0.98) than the net diversification rate (lambda-mu; r=0.92). In order to investigate how DR fares with real data, we present the same comparison for the 104 major clades in our tree with n>8 spp. The two estimates correlate strongly, except for three outliers: these outliers occur within the two clades with the most dramatic diversification increases, as identified by MEDUSA analyses (see Main Text Table 1).

We also considered the ability of mean DR to capture clade-level speciation rate under the flexible non-homogeneous Linear Density Dependence plus Extinction model (Etienne et al. 2011). We used the simulation component of Bio::Phylo (Hartmann et. al. 2010; Vos et. al. 2011) to produce trees under the DDL+E model with t=10 and carrying capacity k=64, with mu = 0.1, and with the speciation rate at the base of the radiation being twice the Yule model expected rate (i.e. 2*ln(64/2)/10=0.69); the average speciation rate under this set of parameters over the course of the diversification is ~0.44 (R. Etienne, pers. comm.). Average DR measured across these trees (c.i. 0.307-0.359) was much closer to this average rate than was lambda (0.245- 0.301) or lambda-mu (c.i. 0.240- 0.274) inferred under a homogeneous birth-death model (Rabosky, 2006). Though much more work is needed, DR may turn out to be a good integrator of changing speciation rate over time for at least some non-homogeneous diversification processes.

Supplementary Methods Figure 5. (a) The relationship between mean DR of members of a clade and the estimated λ (Nee, 2001) for that clade, for a range of clades grown to different sizes $n = \{8, 16, 32, 64, 128, 256, 512\}$ and times $t = \{10, 20, 30, 40\}$ under a Yule model with $\lambda = \ln(n/2)/t$. (b) The relationship between mean DR across species and the MLE of λ-μ (Rabosky, 2006) for the same-sized clades, but grown with $\lambda - \mu = \mu = \ln(n/2)/t$ (i.e. for trees grown under a birth-death model with $\varepsilon = 1/2$. (c) The same simulated data, but with mean DR plotted against the MLE estimate of λ. (d) The grand mean across species of DR for each of 108 avian clades with n>8 species (summarized over 10K trees) versus the MLE estimate of λ - μ. The three outliers are clades associated with significant increases in diversification rate (see Main Text Table 1).



1.3 Mapping of Diversification Rates

1.3.1 Species distributions

Species occurrence data were compiled from global expert opinion range maps of breeding distributions as provided by the best available sources for a given broad geographical region or taxonomic group. We used the same sources as in Jetz et al. (2007), but with an updated taxonomy as described above and geographic ranges updated and added using the now complete "Handbook of the Birds of the World" (del Hoyo et al. 1992-2011) and also Ridgely et al. (2003). We manually checked every distribution map for concordance with the maps in these sources and adjusted ranges with discrepancy > 200km (with Ridgely et al. 2003 taking precedence over del Hoyo et al. 1992-2011 in the New World). Recent validation indicated that these expert maps provide reliable presence/absence accuracy down to spatial resolutions of ca. 150-100km. We extracted species distributions across a 110x110 km equal area grid in a Behrman equal area projection. We excluded all non-breeding and non-native occurrence grid cells and also all cells with < 30% dry land that were not intersecting with islands.

1.3.2 Geography of Diversification Rate

We used these grid cells to calculate the latitudinal centroid of species geographic range as used in Main Text Figure 4. For mapping (Main Text Figure 2A-C) we calculated the geometric mean of species-level lineage diversification rate (DR) of all species in a grid cell assemblage with each species' value weighted by the inverse of the number of grid cells it occupies. This range-size weighting of averages limits the pseudo-replication resulting from wide-ranging species contributing to many and narrow-ranging species only to few grid cell assemblages (Jetz & Rahbek 2002; Jetz & Fine 2012) and facilitates the evolutionary interpretation of broad-scale geographic variation in species attributes. We additionally identified the 25% of species with highest DR and plotted their relative as well as absolute richness per grid cell (Main Text Figure 2D, E).

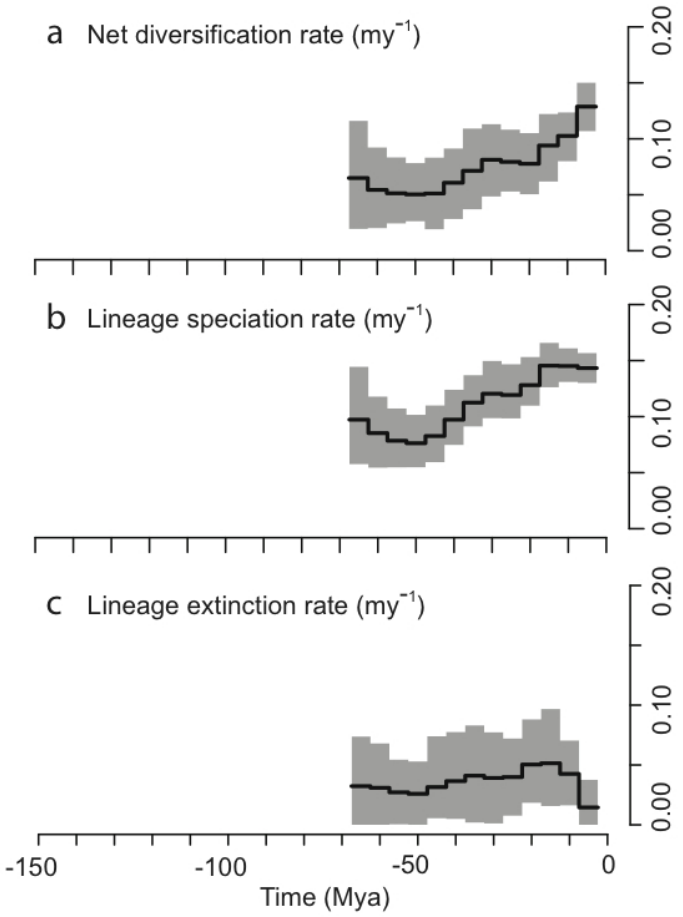
1.3.3 Comparative Analyses of Diversification Rates

We account for phylogenetic relatedness when testing hypotheses concerning differences in the diversification rate (DR) between species currently found on islands vs. mainland, tropical vs. temperate regions, along absolute latitude, and between the hemispheres using phylogenetic generalized least squares (PGLS) models (Rohlf, 2001). Running PGLS over very large trees is computationally highly intensive, and we used a recently developed fast likelihood calculation method available in the R environment (Freckleton, 2012; as used in Sol et al., 2012) to perform PGLS over a representative number of trees. In total we ran our tests across 100 trees and report average p values and frequency (f) of trees for which $p < 0.05$. For island-mainland comparisons we excluded 279 species with predominantly pelagic feeding habitats.

2. SUPPLEMENTARY DISCUSSION

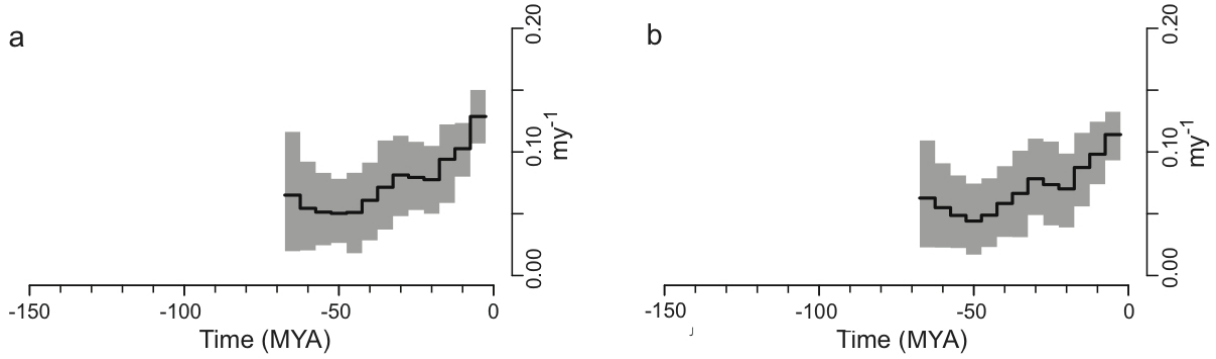
2.1 Diversification through time

One consequence of variance in diversification rates between lineages or among clades, and of the necessary time lag between initiation and completion of speciation is that extinction rate estimates derived from molecular phylogenies are often be biased downward (Morlon 2010, Rabosky 2009, Rabosky 2010); a low extinction rate implies that a clade has been growing in absolute size for much of its history. The generally very low inferred extinction rate we report here (Supplementary Discussion Figure 1) may be due to this bias, and we therefore do not analyse extinction rates.



Supplementary Discussion Figure 1: Diversification through time for all birds. Estimates of the (a) net diversification (speciation – extinction) rate, (b) tree-wide lineage speciation rate, and (c) lineage extinction rate calculated in 5 million year intervals (line segments). The shaded region represents the area between the 5th and 95th quantiles for 525 assessed trees with the mean rate traced in black. Intervals outside 67.5 and 2.5 MYA are not shown due to lack of data (≤ 30 lineages per interval) and the difficulty of accounting for ongoing speciation events, respectively.

In order to evaluate the influences of the placement of taxa with no genetic data on diversification rate results we repeated the diversification through time analyses for the class-level Stage 1 trees (i.e. the 6,670 species not including data-deficient species). Patterns are broadly consistent with those the class-level Stage 2 trees (Supplementary Discussion Figure 2(a) vs. (b)). This suggests the addition of species without genetic data to the tree had limited influence on the results of the diversification analysis.



Supplementary Discussion Figure 2. Comparison of net diversification rate (speciation – extinction) between (a) class-level Stage 2 trees (all 9,993 species; identical to Main Text Figure 1 (a) and Supplementary Discussion Figure 1 (a) above) and (b) class level Stage 1 trees (6,670 species).

2.2 Model fits

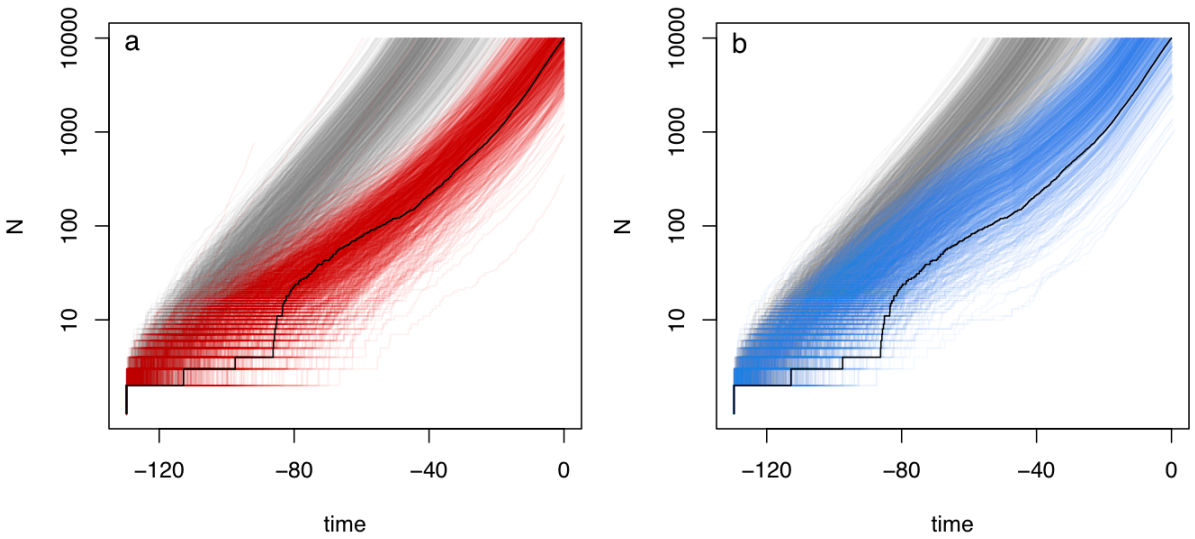
The best model for diversification rate (median delta AIC compared to all other models >> 1,700, Supplementary Discussion Table 1) is one where subsets of individual clades are assigned their own constant, and elevated, diversification rates. We note that statistical power to detect rate decreases rather than rate increases is likely to be low (but see Alfaro et al. 2009) and observing exclusively increases is perhaps not surprising. We confirmed the validity and robustness of these results using two simulation approaches. Trees were simulated with the simulation component of Bio::Phylo (Hartmann et. al. 2010; Vos et. al. 2011).

First, we verified that a delta AIC of the observed magnitude for the rate-shift model would not be expected by chance alone. We first simulated 100 trees with n=9993 tips, but under a constant-rate birth death model. The 20 rate-shift MEDUSA model and a constant rate birth-death model were then fitted to each simulated tree. The delta AIC was calculated for these fitted models on each simulated tree and found to range from 44 to 107 (median=79), but in favor of the more complex and incorrect rate-shift model. This suggests that the penalty imposed by AIC to account for additional parameters is not fully sufficient to account for the multiple testing in MEDUSA on trees of this size. However, it also makes clear that the delta AIC values from our observed data (fully three order of magnitude larger) are vanishingly unlikely to arise by chance.

Second, we assessed whether our overall best model is a good model or simply the best of several poor models. We did this by simulating 1000 trees each under the best-fitting constant rate model (BD), the best-fitting temporal rate variable model (TreePar2) and the overall best-fitting model (MEDUSA). Simulated trees were parameterized from a single tree drawn from our posterior-distribution of stage 2 trees. For the MEDUSA simulations we grew trees with 20 shifts in diversification rate and parameterized the time of shift and the new diversification rates (an alternative would be to condition the simulations on the species richness of clades with shifts). The tree simulations were conditioned on age but were stopped if tree size reached 9993 species before the present. LTT plots for the simulated trees and the observed trees are shown in Supplementary Discussion Figure 3.

Supplementary Discussion Table 1. AIC values for nine diversification models competed on the 525 trees with the Hackett backbone topology. The last row in the table reports the model fit for the fixed-interval temporal shift model used in Main Text Figure 1 to visualize speciation and extinction rates through time. Likelihood values were standardized across platforms (LASER, TreePar, MEDUSA) using the likelihood of the homogeneous birth-death model at its MEDUSA-based ML parameter values.

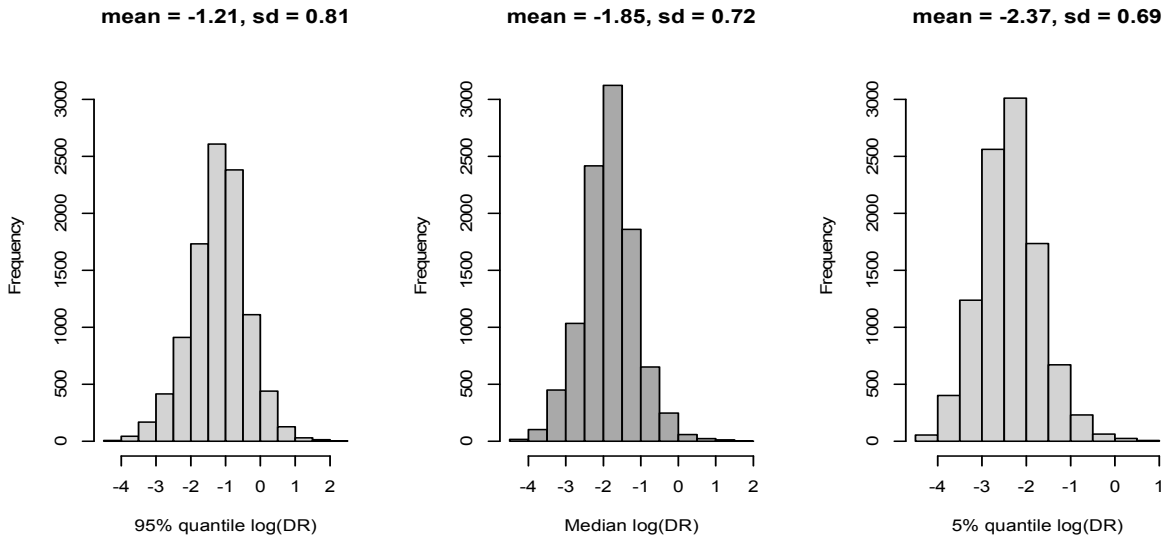
Model	Model type	median AIC	median delta AIC(range)
pureBirth	constant-rate	-101288.1	2013.4 (1463.3 - 2596.7)
BD	constant-rate	-101404.0	1896.3 (1420.9 - 2382.1)
DDX	temporal variation	-101507.3	1791.7 (1337.8 - 2268.5)
DDL	temporal variation	-101286.1	2029.1 (1465.3 - 2598.7)
SPVAR	temporal variation	-101452.3	1783.1 (1327.0 - 23590.2)
EXVAR	temporal variation	-101401.9	1898.3 (1422.9 - 2384.2)
TreePar 1	temporal variation	-101567.8	1735.8 (1283.7 - 2205.4)
TreePar 2	temporal variation	-101582.2	1725.5 (1278.0 - 2193.1)
MEDUSA	clade shifts	-103297.4	0 (0 - 0)
TreePar fixed shifts	temporal variation	-101557.7	1757.1 (1317.5 - 2229.4)



Supplementary Discussion Figure 3. Comparison of LTT plots for sets of 1000 trees simulated under model parameters estimated from a single observed tree (black line). Colored lines show the simulated trees under the constant rate birth-death model (grey in both panels), under the best-fit clade-shift MEDUSA model (red, panel a), and under the temporal rate variable model from TreePar2 (blue, panel b). The observed data falls well within the distribution of LTT plots of the MEDUSA model particularly from 80 mya to the present. The simulated constant-rate trees, and to a lesser extent the temporal rate-shift trees typically result in more rapid accumulation of diversity than the observed tree.

2.3 Tree-to-tree variation

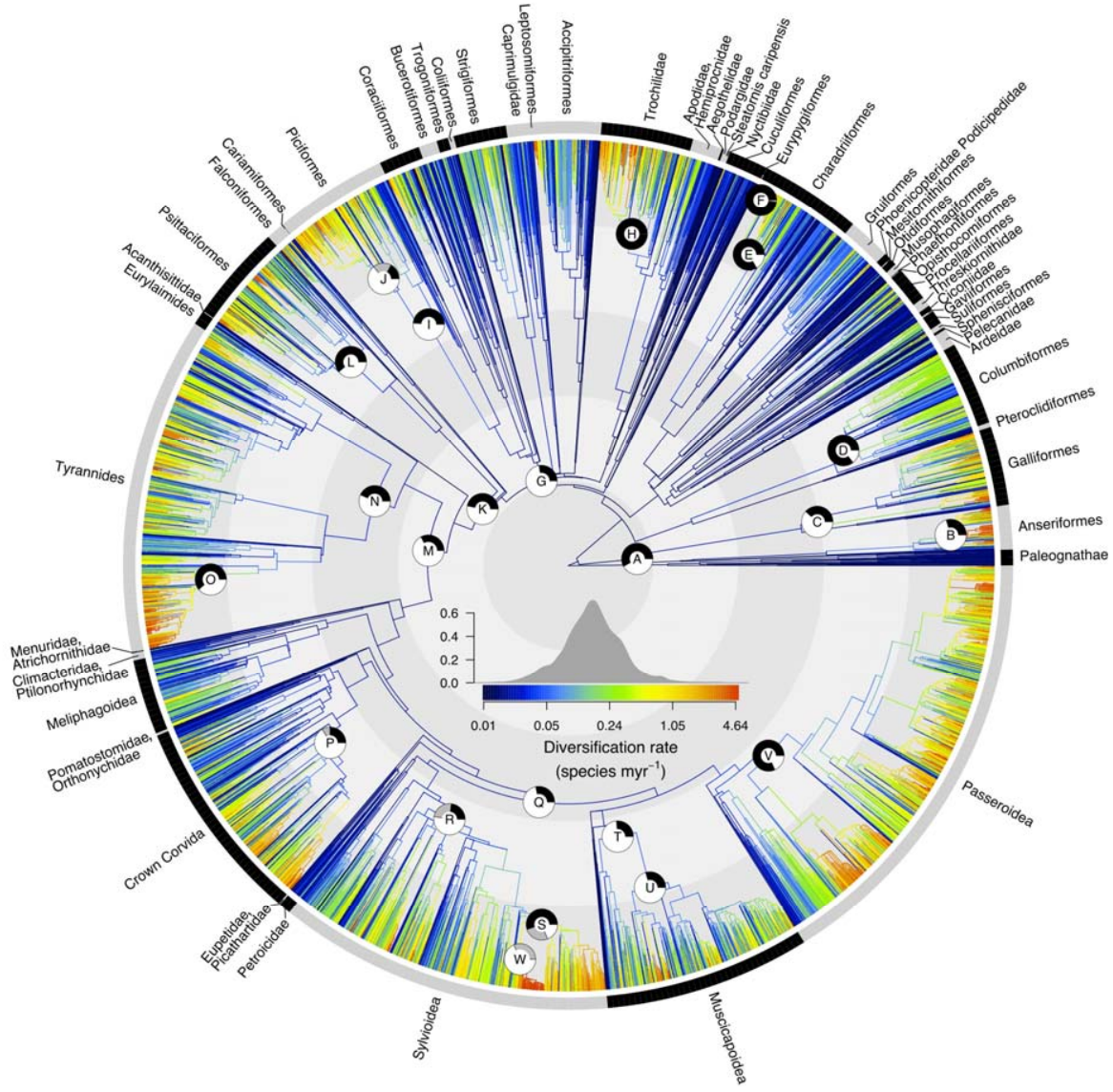
Given that many of our species are placed on our tree based only on the Yule model and generic and higher-level constraints, there could be considerable variation in their species-level lineage diversification rate (DR) scores across trees. Supplementary Discussion Figure 4 presents this source of variation. Even though mean scores vary across trees, most of the variation resides among species, not among trees.



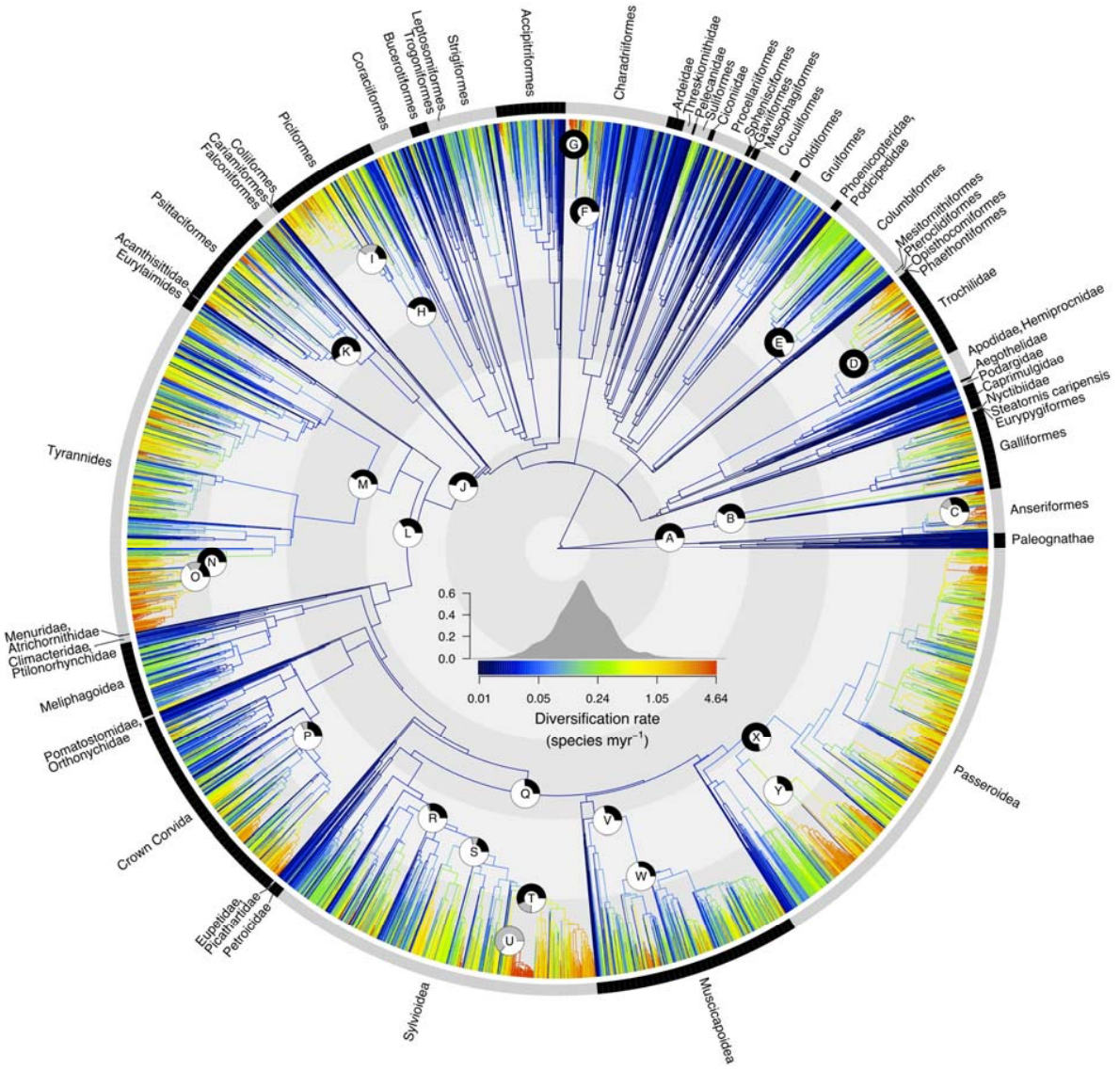
Supplementary Discussion Figure 4. Frequency histograms of the (a) 95% quantile of DR scores for each species across 10,000 trees (using the Hackett backbone tree) (b) the median DR scores across this distribution. (c) the lower 5% quantile of DR scores for each species across this distribution. Most of the variation resides among species rather than within species across trees, supporting the comparison and mapping of median DR scores across species.

2.4 Backbone tree variation

We used two separate backbones for constructing our distribution of trees, one based on the results and data from Hackett et al. (2008) that we present in the main text, and an alternative set based on the data from Ericson et al. (2006). The Spearman Rank correlation of mean species-level lineage diversification rate (DR, section 1.2.2) between the two tree distributions was > 0.99. In Supplementary Discussion Figures 5 and 6 we show that also the MEDUSA results were insensitive to backbone tree choice.



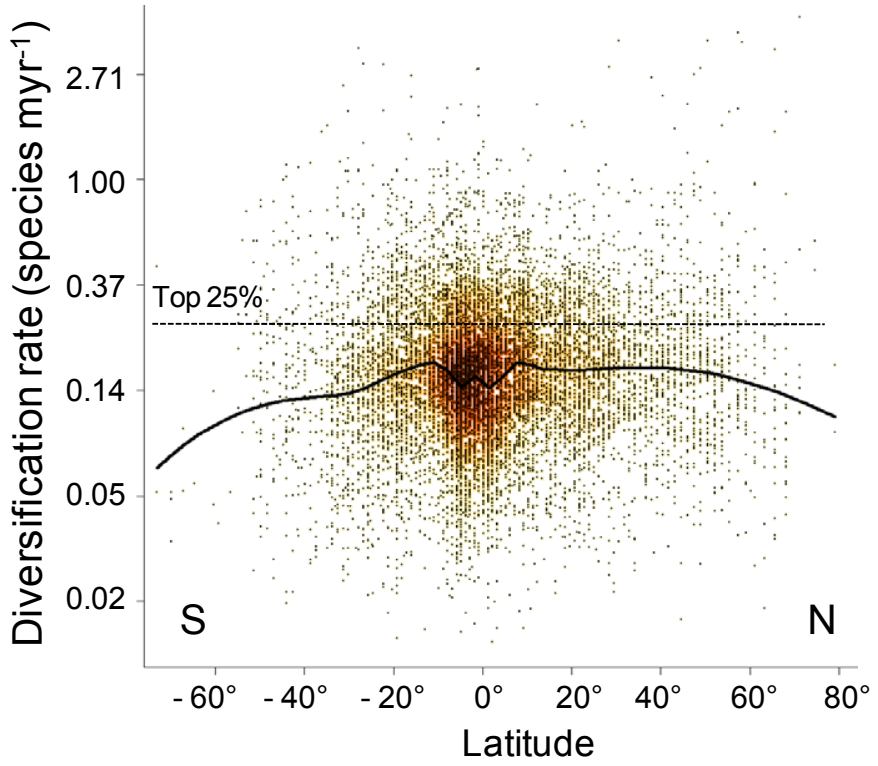
Supplementary Discussion Figure 5. MEDUSA results for 491 trees from the posterior using the **Ericson** backbone tree. All the shifts presented in Table 1 of the main text are also recovered as being the most prevalent on this alternative tree distribution (with average background diversification rate r = 0.055, versus r = 0.056 across the Hackett distribution).



Supplementary Discussion Figure 6. MEDUSA results for 525 trees from the posterior using the **Hackett** backbone tree. This tree is identical to that shown in Fig 2 of the main text, but with clades labeled as in Supplementary Discussion Figure 5. We note that the prevalence of shifts may be lowered by a "trickle-down" effect where combinations of nested shifts are rarely identified in the same tree. Notably, joint shifts at nodes A+B, H+I, J+L, and N+O are found in <5% of trees but either one or other shift occurs in >80% of trees (> 97% of trees for N+O). The trickle-down effect can extend to larger sets of nested nodes (e.g. the combination J+L+Q does not occur in any tree but at least one of the three nodes is identified by MEDUSA in >96% of trees), but the trickle-down effect does not itself preclude the possibility of nested shifts co-occurring (e.g. shifts at nodes F and G co-occur in 66% of trees).

2.5 Latitudinal patterns

Supplemental Discussion Figure 7 presents the latitudinal pattern of diversification rate (DR) for all birds, and is very similar to the Main Text Figure 4 that excludes species with the largest ranges. The overall pattern is one where DR peaks not at the equator where species richness is highest, but nearer the tropical-temperate boundary. [Note added in proof: Since our submission, Soria-Carrasco and Castresana (2012) have presented work consistent with this pattern for mammals.]



Supplementary Discussion Figure 7. Latitudinal gradient in species-level lineage diversification rate (DR) for the full dataset (N = 9,993 species). Each black point represents a single species DR at the centroid latitude of its global breeding range. There is no significant association between DR and absolute centroid latitude (p_{avg} = 0.37, f($p<0.05$) = 3/100) or for intra- (< 23° latitude) vs. extra-tropical centroid location (p_{avg} = 0.57, f($p<0.05$) = 0/100). The thin line indicates the threshold identifying the quartile of species with highest DR. Darker brown shading highlights greater density of species points.

3. INVENTORY OF SUPPLEMENTARY DATA FILES

backbone_tree_accessions.csv: Provides the genbank accession numbers for sequence data used to generate the "Hackett" and "Ericson" backbone topologies.

clade_summary.csv: Lists 129 monophyletic clades used to produce the final set of trees, their size (including the two outgroup species), decisiveness scores, the outgroup species used to root and date the clade, and which stage 1 consensus tree was used as the basis for producing the Stage 2 trees.

clade_tree_accessions.csv: Provides the genbank accession numbers for sequence data used to generate the individual clade topologies.

ConstraintTrees.zip: contains the consensus trees for each clade from our stage 1 (genes only) inference we used to place all species in our final tree inference.

MCC_trees.zip: Provides the maximum clade credibility (MCC) tree for each Stage 1 clade and for the "Hackett" and "Ericson" backbone trees. The MCC trees were extracted from the post-burnin posterior distributions (see above and clade_summary.csv supplementary file) using TreeAnnotator 1.7.1 (http://beast.bio.ed.ac.uk/TreeAnnotator). The MCC trees are nexus format files containing information on node support (posterior probabilities) and Highest Probability Densities (HPDs) for node ages. They are best viewed in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

mastertaxonomy.csv: Provides the master taxonomy for the 9,993 species in the phylogeny and analysis. *Scientific* gives the Latin binomial following Birdlife Version 3 (BL3) and IOC Version 2.7 (IOC27) taxonomies, as detailed in the field *Taxonomy*. *TipLabel* is the format in which the Latin name is listed in the tree files. *English* provides the common English name, *BLFamilyLatin* and *BLFamilyEnglish* the Latin and English family names according to the Birdlife taxonomy, together with their taxonomic sequence (*FamSequID*). The fields *Order* and *OscSubOsc* give higher level taxa names broadly following Sibley & Monroe (1990).

missing_genera.csv: Provides information on the topological constraints used for placing genera of type 3 species in stage 2 trees (see Section 1.1.3 (iv)). The *FocalGenus* field provides the Latin name of the focal genus. *Comment/Source* informs on how the constraint information was derived. *Type* indicates the way the constraint was implemented in MrBayes. In the *Type* column *include* constraints are those in which monophyly is enforced for the focal genus and all other genera on that row whereas *exclude* constraints are those in where the focal genus cannot enter a clade including the other genera on that row. *Clade* gives the name of the clade containing the genus. *Genus1* onward gives the names of genera forming a constraint.

Tree Sets: Full sets of bird trees in flat newick format for both the Hackett and Ericson backbones, and for both complete and DNA-data only trees, are downloadable from birdtree.org.

4. REFERENCES

Alfaro, M.E., Santini, F., Bock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G., and L.J. Harmon. 2008. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proceedings of the National Academy of Sciences (USA) 106: 13410-13414.

Baker, A.J, Pereira, S.L. and T.A. Paton. 2007. Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. Biology Letters 3: 205-209.

Barker, F.K., Cibois, A., Schilker, P., Feinstein, J., and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. Proceedings of the National Academy of Sciences (USA) 101: 11040-11045.

Benton, M.J. and P.C.J. Donoghue. 2006. Paleontological evidence to date the tree of life. Molecular Biology and Evolution 24: 26-53.

Beresford P., Barker F.K., Ryan P.G., and T.M. Crowe. 2005. African endemics span the tree of songbirds (Passeri): molecular systematics of several evolutionary 'enigmas'. Proceedings of the Royal Society Series B 272: 849-858

Bertelli, S., Lindow, B.E.K., Dyke, G.J., and L.M. Chiappe. 2010. A well-preserved ‘charadriiform-like’ fossil bird from the Early Eocene Fur Formation of Denmark. Palaeontology 53: 507-531.

Britton, T., Anderson, C.L., Jacquet, D., Lundqvist, S., and Bremer, K. 2007. Estimating divergence times in large phylogenetic trees. Systematic Biology 56: 741-752.

Brown, J. W., Payne, R.B., and D.P. Mindell. 2007. Nuclear DNA does not reconcile 'rocks' and 'clocks' in Neoaves: a comment on Ericson et al. Biology Letters 3: 257-259.

Brown, J.W., FitzJohn, R.G., Alfaro, M.E., & Harmon, L.J. 2012. MEDUSA: Modeling Evolutionary Diversification Using Stepwise AIC. Available from: https://github.com/josephwb/turboMEDUSA.

Soria-Carrasco, V. & Castresana, J. 2012. Diversification rates and the latitudinal gradient of diversity in mammals. Proceedings of the Royal Society B 279: 4148-4155.

Clarke, J.A., Tambussi, C.P., Noriega, J.I., Erikson, G.M., and R.A. Ketcham. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature 43: 305-308

Cracraft, J. and M. J. Donoghue, eds. 2004. Assembling the Tree of Life. Oxford University Press, Oxford.

Cramp, S., Simmons, K., Brooks, D., Collar, N., Dunn, E., Gillmor, R., Hollom, P. Hudson, R. Nicholson, E., and M. Ogilvie. 1978-1994. Handbook of the birds of Europe, the Middle East and North Africa. The birds of the Western Palearctic: Volumes 1-9.

Crowe, T.M., Bowie, R.C.K., Bloomer, P., Mandiwana, T.G., Hedderson, T.A.J., Randi, E., Pereira, S.L. and J. Wakeling (2006) Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data. Cladistics 22: 495–532.

del Hoyo, J., Elliott, A.,Sargatal, J., and D. A. Christie, editors. 1992-2011. Handbook of the Birds of the World. Lynx Editions, Barcelona.

Donne-Goussé, C., Laudet, V., and C. Hänni. 2002. A molecular phylogeny of anseriformes based on mitochondrial DNA analysis. Molecular Phylogenetics and Evolution 23: 339-356.

Driskell A.C. and L. Christidis. 2004. Phylogeny and evolution of the Australo-Papuan honeyeaters (Passeriformes, Meliphagidae). Molecular Phylogenetics and Evolution. 31: 943-960.

Drummond, A. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7: 214.

Drummond, A., Ho, S. Y. W., Phillips, M. J., and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4: e88.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., R. Moir, R., Stones-Havas, S., Thierer, T., and A. Wilson. 2007. Geneious v3.0, Available from http://www.geneious.com/

Dumbacher J.P., Deiner K., Thompson L., and R.C. Fleischer. 2008. Phylogeny of the avian genus Pitohui and the evolution of toxicity in birds. Molecular Phylogenetics and Evolution 49: 774-781.

Dyke, G. J., and M. van Tuinen. 2004. The evolutionary radiation of modern birds (Neornithes): reconciling molecules, morphology and the fossil record. Biological Journal of the Linnean Society 141: 153-177.

Edwards, S.V., Jennings, W.B., and A. M. Shedlock. 2005. Phylogenetics of modern birds in the era of genomics. Proceedings of the Royal Society of London Series B 272: 979-992.

Ericson P.G.P., Jansen A.L., Johansson U.S., and J. Ekman. 2005. Inter-generic relationships of the crows, jays, magpies and allied groups (Aves: Corvidae) based on nucleotide sequence data. Journal of Avian Biology 36: 222-234.

Ericson, P.G.P., Anderson, Britton, C.L., Eizanowski, T.A., Johansson, U.F., Källersjö, M., Ohlson, J.I., Parsons, T.J., Zuccon, D., and G. Mayr. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biology Letters 2: 543-547.

Fain, M.G. and P. Houde. 2004. Parallel radiation in the primary clades of birds. Evolution 58: 2558-2573.

Franzen, J.L. 2005. The implications of the numerical dating of the Messel fossil deposit (Eocene, Germany) for mammalian biochronology. Annales de Paléontologie 91: 329-335.

Freckleton, R.P., Phillimore, A.B., and M. Pagel. 2008. Relating traits to diversification: a simple test. American Naturalist 172: 102-115.

Freckleton, R. P. 2012. Fast likelihood calculations for comparative analyses. Methods in Ecology and Evolution: in press.

Fuchs J., Fjeldsa, J., and E. Pasquet. 2006. An ancient African radiation of corvoid birds (Aves: Passeriformes) detected by mitochondrial and nuclear sequence data. Zoologica Scripta 35: 375-385.

Gibb, G.C., Kardailsky, O., Kimball, R.T., Braun, E., and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. Molecular Biology and Evolution 24: 269-280.

Groth, J.G. 2000. Molecular evidence for the systematic position of *Urocynchramus pylzowi*. The Auk 117:787-791

Guindon, S and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52: 696-704.

Hackett, S.J., Kimball, R.T., Reddy, S., Bowie, R.C.K. Braun, E.L., Braun, M.J., Chojnowski, J.L, Cox, W.A., Han, K.-L., Harshman, J., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320: 1763-1768.

Harmon, L. J., J. Weir, C. Brock, Glor, R. E., and W. Challenger. 2008. GEIGER: Investigating evolutionary radiations. Bioinformatics 24: 129-131.

Harrison, C.J.O. 1984. A revision of the fossil swifts (Vertebrata, Aves, suborder Apodi), with descriptions of three new genera and two new species. Mededelingen van de Werkgroep voor Tertiaire en Kwartaire Geologie 21: 157–177.

Hartmann, K., Wong, D., and T. Stadler. 2010. Sampling trees from evolutionary models. Systematic Biology 59:465-476.

Ho, S.Y.W. and M.J. Phillips. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58: 367-380.

Hugall, A.F. and D. Stuart-Fox. 2012. Accelerated speciation in colour-polymorphic birds. Nature doi:10.1038/nature11050.

Hurlbert, A.H. and W. Jetz. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proceedings of the National Academy of Sciences (USA) 104: 13384-13389.

Jetz, W., Wilcove, D.S., and A.P. Dobson. 2007. Projected Impacts of Climate and Land-Use Change on the Global Diversity of Birds. PLoS Biology 5:1211-1219.

Jetz, W. and C. Rahbek. 2002. Geographic range size and determinants of avian species richness. Science 297:1548-1551.Jetz, W. and P. V. A. Fine. 2012. Area and productivity of the world's biomes integrated over geological time predicts global patterns of vertebrate diversity. PLoS Biology 10:e1001292.

Johansson U.S., Fjeldsa J., and R.C.K. Bowie. 2008. Phylogenetic relationships within Passerida (Aves: Passeriformes): A review and a new molecular phylogeny based on three nuclear intron markers. Molecular Phylogenetics and Evolution 48: 858-876.

Jones, G.R. 2011. Tree models for macroevolution and phylogenetic analysis. Systematic Biology 60: 735-746.

Jonsson, K.A., Irestedt, M., Fuchs, J., Ericson, P.G.P., Christidis, L., Bowie, R.C.K., Norman, J.A., Pasquet. E., and J. Fjeldsa. 2008a. Explosive avian radiations and multi-directional dispersal across Wallacea: Evidence from the Campephagidae and other Crown Corvida (Aves). Molecular Phylogenetics and Evolution 47: 221-236.

Jonsson K.A., Bowie R.C.K., Norman J.A., Christidis, L., and J. Fjeldsa. 2008b. Polyphyletic origin of toxic Pitohui birds suggests widespread occurrence of toxicity in corvoid birds. Biology Letters 4: 71-74.

Kathoh K., Asimenos, G., and H. Toh (2009) Multiple alignment of DNA sequences with MAFFT. Methods in Molecular Biology 537: 39-64.

Klicka, J., Burns, K., and G.M. Spellman. 2007. Defining a monophyletic Cardinalini: A molecular perspective. Molecular Phylogenetics and Evolution 45: 1014-1032

Kundu, S., Jones, C.G., Prys-Jones, R.P., and J.J. Groombridge. 2012. The evolution of the Indian Ocean parrots (Psittaciformes): Extinction, adaptive radiation and eustacy. Molecular Phylogenetics and Evolution 62: 296-305.

Lanfear, R. and L. Bronham. 2011. Estimating phylogenies for species assemblages: A complete phylogeny for the past and present native birds of New Zealand. Molecular Phylogenetics and Evolution. 61: 958-963.

Loytynoja, A. and N. Goldman. 2005. An algorithm for progressive multiple alignment of sequences with insertions. Proceedings of the National Academy of Sciences (USA) 102: 10557-10562.

Loytynoja, A. and N. Goldman. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science 320: 1632-1635.

Mayr, G. 2000. Tiny hoopoe-like birds from the Middle Eocene of Messel (Germany). The Auk 117: 968-974.

Mayr, G. 2003. Phylogeny of early Tertiary swifts and hummingbirds (Aves: Apodiformes). The Auk 120: 145-151.

Mayr, G. 2005a. A tiny barbet-like bird from the lower Oligocene of Germany: the smallest species and earliest substantial fossil record of the Pici (woodpeckers and allies). The Auk 122: 1055-1063.

Mayr, G. 2005b. A chicken-sized crane precursor from the early Oligocene of France. Naturwissenschaften 92: 389-393.

Mayr, G. and C. Mourer-Chauvire. 2000. Rollers (Aves: Coraciiformes s.s.) from the Middle Eocene of Messel (Germany) and the Upper Eocene of the Quercy (France). Journal of Vertebrate Paleontology 20: 533-546.

McPeek M.A. 2008. The ecological dynamics of clade diversification and community assembly. American Naturalist 172:E270–E284.

Morgan-Richards, M., Trewick, S.A., Bartosch-Harlid, A., Kardailsky, O., Phillips, M.J., McLenachan, P.A., and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evolutionary Biology 8: 20.

Morlon H., Potts M.D., and J.B. Plotkin. 2010. Inferring the dynamics of diversification: a coalescent approach. PLoS Biology 8: e1000493.

Moyle R.G., Chesser, R.T., Brumfield, R.T., Tello J.G., Marchese, D.J., and J. Cracraft. 2009. Phylogeny and phylogenetic classification of the antbirds, ovenbirds, woodcreepers, and allies (Aves: Passeriformes: infraorder Furnariides). Cladistics 25: 1-20.

Mulder, F. 2011. Probability distributions of ancestries and genealogical distances on stochastically generated rooted binary trees. Journal of Theoretical Biology 280: 139-145.

Nee, S. 2001. Inferring speciation rates from phylogenies. Evolution 55: 661-668.

Nee, S., Mooers, A.O., and P.H. Harvey. 1992. Tempo and mode of evolution revealed from molecular phylogenies. Proceedings of the National Academy of Sciences (USA) 89: 8322-8326.

Norman, J.A., Ericson, P.G.P., Jonsson, K.A., Fjeldsa, J., and L. Christidis. 2009. A multi-gene phylogeny reveals novel relationships for aberrant genera of Australo-Papuan core Corvoidea and polyphyly of the Pachycephalidae and Psophodidae (Aves: Passeriformes). Molecular Phylogenetics and Evolution 58: 488-497.

Nylander, J.A.A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander J.A., Wilgenbusch, J.C., Warren, D.L., and D.L. Swofford (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24: 581-583.

Ohlson, J.I., Prum, R.O., and P.G.P. Ericson. 2007. A molecular phylogeny of the cotingas (Aves: Cotingidae). Molecular Phylogenetics and Evolution 42: 25–37.

Olson, S.L. 1977. A lower Eocene frigatebird from the Green River Formation of Wyoming (Pelecaniformes: Fregatidae). Smithsonian Contributions to Paleobiology 35: 1-3.

Pacheco, M.A., Battistuzzi, F.U., Lentino, M., Aguilar, R.F., Kumar, S., and A.A. Escalante. 2011. Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. Molecular Biology and Evolution 28: 1927-1942.

Paradis E. 2006. Analysis of Phylogenetics and Evolution with R. New York, Springer.

Pauplin, Y. 2000. Direct calculation of a tree length using a distance matrix. Journal of Molecular Evolution 51:41-47.

Phillimore, A.B. and T.D. Price. 2008. Density dependent cladogenesis in birds. PLoS Biology 6: e71.

Poe, S., and A.L. Chubb. 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution 58: 404-415.

Posada D. 2008. jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution 25: 1253-1256.

Rabosky, D. 2006. LASER: A maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. Evolutionary Bioinformatics 2: 247-250.

Rabosky, D.L. and I.J. Lovette. 2008a. Density dependent diversification in North American wood warblers. Proceedings of the Royal Society Series B 275: 2363–2371.

Rabosky, D.L. and I.J. Lovette. 2008b. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution 62: 1866–1875.

Rabosky, D.L. 2009. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. Systematic Biology 58:629-640.

Rabosky, D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution 64:1816-1824.

Rambaut A. 1996. Se-Al: Sequence alignment editor. Available at http://evolve.zoo.ox.ac.uk/.

Rambaut, A. and A. Drummond. 2007 Tracer v1.4.1, Available from http://beast.bio.ed.ac.uk/Tracer.

Redding, D. W. and A.O. Mooers 2006. Incorporating evolutionary measures into conservation prioritization. Conservation Biology 20: 1670-1678.

Redding, D.W., Mimoto, A., Hartmann, K., Bokal, D., Devos, M., and A.Ø. Mooers. 2008. Evolutionarily distinct species capture more phylogenetic diversity than expected. Journal of Theoretical Biology 251: 606-615.

Ridgely, R. S., Allnutt, T.F., Brooks, T., McNicol, D.K., Mehlman, D.W., Young, B.E., and J. R. Zook. 2003. Digital Distribution Maps of the Birds of the Western Hemisphere. Version 1.0. NatureServe, Arlington, Virginia, USA.

Rohlf, F.J. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. Evolution 55: 2143-2160.

Rüber, L. and R. Zardoya. 2005. Rapid cladogenesis in marine fishes revisited. Evolution 59: 1119-1127.

Sanderson, M.J., Mahon, M.C., and M. Steel. 2010. Phylogenomics with incomplete taxon coverage: the limits to inference. BMC Evolutionary Biology 10: 155.

Sanderson, M.J., Mahon, M.C. and M. Steel. 2011. Terraces in phylogenetic tree space. Science 333: 448-450.

Semple, C. and M. Steel. 2003. Phylogenetics. Oxford University Press, Oxford, UK.

Shapiro B., Rambaut A., and A.J. Drummond. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Molecular Biology and Evolution 23: 7-9.

Sibley, C.G. and J.E. Ahlquist. 1990. Phylogeny and classification of birds: a study in molecular evolution. Yale University Press, New Haven.

Sibley, C.G. and B.L. Monroe. 1990. Distribution and Taxonomy of the Birds of the World. Yale University Press, New Haven, USA.

Simmons, N.B., Seymour, K.L., Habersetzer, J., and G.F. Gunnell. 2008. Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation. Nature 451: 818-821.

Slack, K.E., Jones, C.M., Ando, T., Harrison, G.L., Fordyce, R.E., and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Molecular Biology and Evolution 23: 1144-1155.

Sol, D., Maspons, J., Vall-llosera, M., Bartomeus, I., García- Peña, G.E, Piñol, J., and R.P. Freckleton. 2012. Unravelling the life history of successful invaders. Science 337:580-583.

Sorenson, M.D. and R.B. Payne. 2005. Molecular systematics: cuckoo phylogeny inferred from mitochondrial DNA sequences. Pp. 68-94 in R.B. Payne. Bird Families of the World: Cuckoos. Oxford University Press.

Stadler, T. 2011. Looking at the present to learn about the past: inferring speciation and extinction processes from extant species data. Proceedings of the National Academy of Sciences (USA) 108: 16145-16146.

Stadler, T. 2012. Simulating trees on a fixed number of extant species. Systematic Biology, in press.

Steel, M. and A.O. Mooers 2010. The expected length of pendant and interior edges of a Yule tree. Applied Mathematics Letters 23:1315-1319.

Steel, M. and A. McKenzie 2001. Properties of phylogenetic trees generated by Yule-type speciation models. Mathematical Biosciences 170:90-112.

Sul, S.-J., and T.L. Williams, 2009. An experimental analysis of consensus tree algorithms for large-scale tree collections. pages 100–111 in Proceedings of the 5th International Symposium on Bioinformatics Research and Applications (ISBRA'09*)*, Berlin, Heidelberg, Springer-Verlag.

Svennblad, B. 2008. Consistent estimation of divergence times in phylogenetic trees with local molecular clocks. Systematic Biology 57: 947–954.

Urban, E.K., Fry, C.H., and S. Keith, editors. 1986 - 2000. The Birds of Africa. Academic Press, London.

Vos, R.A., Caravas, J., Hartmann, K., Jensen, M.A., and C. Miller. 2011. BIO::Phylo-phyloinformatic analysis using Perl. BMC Bioinformatics 12: 63.

Waterhouse, D.M., Lindow, B.E.K., Zelenkov, N.V., and G.J. Dyke. 2008. Two new parrots (Psittaciformes) from the lower Eocene Fur formation of Denmark. Palaeontology 51: 575-582.

Wright, T.F., Schirtzinger, E.E., Matsumoto, T., Eberhard, J.R., Graves, G.R., Sanchez, J.J., Capelli, S., Mueller, H., Scharpegge, J., Chambers, G.K., and R.C. Fleischer. 2008. A multi locus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the Cretaceous. Molecular Biology and Evolution 25: 2141-2156.

Yang, Z. and B. Rannala. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Molecular Biology and Evolution 23: 212-226.

Yuri, T. and D. Mindell. 2002. Molecular phylogenetic analysis of Fringillidae, ''New World nine-primaried oscines'' (Aves: Passeriformes). Molecular Phylogenetics and Evolution 23:229-243

You, H., Lamanna, M.C., Harris, J.D., Chiappe, L.M., O'Connor, J., Ji, S., Lu, J., Yuan, C., Li, D., Zhang, X., Lacovara, K.J., Dodson, P., and Q. Ji. 2006. A nearly modern amphibious bird from the Early Cretaceous of northwestern China. Science 312: 1640-1643.