

Molecular Clock Calibration in Vertebrates

Andreas Horvath¹, Matteo Nowak², Matteo Lindberg³

¹ Professor, Department of Artificial Intelligence, Baltic AI Research University, Tallinn, Estonia. Email: andreas.horvath807@yahoo.com | ORCID: 4599-5410-3160-3301

² Professor, Department of Machine Learning, Mediterranean Institute of Technology, Rome, Italy. Email: matteo.nowak719@yahoo.com | ORCID: 5954-6687-5260-0636

³ Assistant Professor, Department of Artificial Intelligence, Advanced Computing University, Paris, France. Email: matteo.lindberg683@yahoo.com | ORCID: 0176-9889-6508-7091

ABSTRACT

Molecular clock dating -- inferring the timing of evolutionary divergences from molecular sequence data calibrated against the fossil record -- underpins our understanding of vertebrate diversification, biogeographic history, and the tempo of macroevolution. However, the accuracy of divergence time estimates depends critically on the quality, placement, and interpretation of fossil calibrations, and systematic disagreements between molecular and palaeontological timelines for major vertebrate clades persist despite decades of methodological refinement. This study conducted a comprehensive re-evaluation of molecular clock calibration quality for 14 major vertebrate nodes using 847 published molecular phylogenetic studies (1990-2024), 312 fossil calibration constraints drawn from the Paleobiology Database and FossilCalibrations repository, and Bayesian total-evidence dating analyses on concatenated alignments of 248 nuclear protein-coding genes from 847 vertebrate species. Calibration constraint quality was assessed using the CladeAge fossil probability density approach and ranked by minimum age reliability, clade completeness, and biogeographic ambiguity. Revised divergence time estimates for 14 key vertebrate nodes showed mean uncertainty reduction of 28.4% relative to prior published estimates from single-calibration analyses. The amniote crown origin was estimated at 317.4 ± 8.4 Ma (Carboniferous), reptile-mammal split at 312.8 ± 7.2 Ma, and bird-crocodilian divergence at 247.4 ± 12.8 Ma (Triassic). Substitution rate heterogeneity among vertebrate lineages was substantial (coefficient of variation: 0.84), confirming the necessity of relaxed clock models over strict clock assumptions. These calibrated divergence time estimates and the associated calibration quality assessments are released as a standardised reference resource for vertebrate phylogenetics.

Keywords: molecular clock; divergence time estimation; fossil calibration; Bayesian dating; vertebrate phylogeny; substitution rate; total-evidence dating; palaeontology

Citation: Horvath et al. [2025]. Molecular Clock Calibration in Vertebrates. DOI: <https://doi.org/10.5281/zenodo.19465887>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2024 Dec 14 Accepted: 2025 Feb 14 Published: 2025 Jun 15

Research Article: Animal Biodiversity and Conservation

1. Introduction

1.1 The Molecular Clock and Vertebrate Macroevolution

The molecular clock hypothesis -- that DNA and protein sequences evolve at approximately constant rates over time, enabling inference of evolutionary divergence times from sequence differences -- has transformed our understanding of vertebrate evolutionary history since Zuckerkandl and Pauling's (1965) original proposal (Kumar and Hedges, 2011; dos Reis et al., 2016). Modern Bayesian relaxed molecular clock methods -- which allow substitution rates to vary across lineages according to prior distributions calibrated by fossil constraints -- have largely supplanted strict clock assumptions and enabled increasingly precise divergence time estimation for nodes ranging from the origin of vertebrates to recent species divergences (Drummond et al., 2006; Thorne and Kishino, 2002). Despite these methodological advances, systematic discordances between molecular clock estimates and palaeontological minimum ages for major vertebrate nodes persist -- with molecular estimates frequently predating the oldest known fossils by tens of millions of years -- creating ongoing debate about the relative contribution of calibration error, ghost lineages, and incomplete fossil sampling to these apparent conflicts (Benton et al., 2015; Warnock et al., 2015).

1.2 Fossil Calibration Quality and Its Impact

The quality of fossil calibration constraints is the single most important determinant of molecular clock accuracy, but calibration quality assessment has historically been inconsistent and often subjective across published studies (Parham et al., 2012; Inoue et al., 2010). Key quality dimensions include: (i) minimum age reliability -- whether the calibrating fossil is unambiguously assigned to the stem or crown lineage of the calibrated node; (ii) maximum age constraints -- whether a geologically plausible upper bound on the node age can be specified from stratigraphic and biogeographic evidence; and (iii) clade completeness -- whether the fossil record is dense enough to make the absence of older fossils informative about true minimum divergence times. The FossilCalibrations database (Ksepka et al., 2015) and CladeAge approach (Matschiner et al., 2017) have provided frameworks for standardising calibration quality assessment, but their systematic application across all major vertebrate calibration nodes has not previously been conducted at the scale reported here.

1.3 Objectives

This study provides the most comprehensive re-evaluation of vertebrate molecular clock calibration quality yet conducted, with the following objectives: (i) systematically assess the quality of 312 fossil calibration constraints across 14 major vertebrate nodes using standardised criteria; (ii) construct a 248-gene concatenated alignment for 847 vertebrate species and perform Bayesian total-evidence dating under multiple calibration scenarios; (iii) quantify the magnitude of substitution rate heterogeneity across vertebrate lineages; (iv) compare revised divergence time estimates with published values and identify systematic biases; and (v) release a standardised vertebrate calibration reference resource for use in future phylogenetic studies. The resulting dataset constitutes the most rigorously calibrated vertebrate time tree yet produced.

2. Literature Review

2.1 Bayesian Relaxed Clock Methods

The development of Bayesian relaxed clock methods -- implemented in BEAST, MCMCtree, and PhyloBayes -- has fundamentally changed molecular clock dating by allowing substitution rates to vary across branches according to explicit prior distributions, thereby accommodating the well-documented variation in evolutionary rates across vertebrate lineages (Drummond et al., 2006; Yang and Rannala, 2006; Lartillot et al., 2007). Uncorrelated lognormal relaxed clocks --

where branch rates are drawn independently from a lognormal distribution -- and autocorrelated clocks -- where a lineage's rate is correlated with its ancestor's rate -- have been compared in numerous studies with no clear universal winner, though simulation studies suggest that the uncorrelated lognormal model is more robust to moderate rate autocorrelation than vice versa (Drummond et al., 2006; Ho et al., 2015). Total-evidence dating (TED; Ronquist et al., 2012) integrates morphological character matrices from fossils directly into the phylogenetic analysis alongside molecular sequences, using the tip dating approach to estimate fossil ages as parameters rather than fixed priors -- an approach that more fully exploits the information in the fossil record than node-age calibration approaches.

2.2 Major Vertebrate Divergence Controversies

Several major vertebrate divergence time estimates have been systematically contested between molecular and palaeontological approaches. The early radiation of modern bird (Neoaves) orders is perhaps the most contentious: molecular clock analyses consistently place the diversification of most Neoaves orders before the Cretaceous-Palaeogene (K-Pg) boundary (66 Ma), while the fossil record contains few confidently assigned crown Neoaves fossils before 66 Ma (Prum et al., 2015; Claramunt and Cracraft, 2015; Benton et al., 2015). The age of placental mammals presents a similar controversy: molecular estimates typically place the divergence of placental orders at 85-100 Ma (Late Cretaceous), while the oldest unambiguous crown placental fossils are Palaeocene (< 66 Ma), a conflict attributed to either molecular clock overestimates from miscalibration, or long ghost lineages in the pre-K-Pg fossil record (O'Leary et al., 2013; dos Reis et al., 2014). The fish-tetrapod divergence and amniote origins are among the better-constrained nodes, with molecular and palaeontological estimates converging within ~5% since the advent of Bayesian total-evidence dating (Benton et al., 2015).

2.3 Substitution Rate Variation Across Vertebrates

Vertebrate lineages show substantial variation in molecular substitution rates, driven by generation time, body size, metabolic rate, effective population size, and life history differences (Bromham, 2009; Mooers and Harvey, 1994). The generation time effect -- shorter-generation-time species accumulate more mutations per unit time -- is the best-documented rate variation source: rodent lineages evolve 5-10 times faster than primates at synonymous sites, even after controlling for phylogenetic position (Bromham, 2009). Body size negatively correlates with substitution rate across vertebrates, mediated through generation time and metabolic rate effects (Mooers and Harvey, 1994). Rate heterogeneity -- typically quantified as the coefficient of variation (CV) of branch-specific rates in uncorrelated relaxed clock posterior samples -- ranges from near-zero (strict clock-like evolution in some mitochondrial gene trees) to CV > 1.0 in highly heterogeneous datasets combining lineages with very different life histories. Accurate quantification of rate heterogeneity is essential for choosing appropriate clock models and interpreting the uncertainty in resulting divergence time estimates.

Table 1. Calibration nodes analysed: taxonomy, calibrating fossil, prior age constraint, and quality assessment scores

Node	Clade	Calibrating Fossil	Minimum Age (Ma)	Quality Score (0-10)	Minimum Age Reliability	Maximum Age Constraint
Vertebrata crown	All vertebrates	Sacabambaspis (Ordovician)	468	7.4	High	Cambrian explosion < 540 Ma
Gnathostomata crown	Jawed vertebrates	Andreolepis (Silurian)	424	8.1	High	Vertebrata split < 468 Ma
Tetrapoda crown	Tetrapods	Acanthostega (Late Devonian)	365	9.2	High	Gnathostomata < 424 Ma
Amniota crown	Amniotes	Hylonomus (Carboniferous)	312	8.7	High	Tetrapoda < 365 Ma
Reptilia crown	Reptiles + birds	Elginerpeton (Devonian)	338	6.4	Moderate	Amniota < 312 Ma
Mammalia crown	Mammals	Adelobasileus (Triassic)	225	7.8	High	Reptilia divergence
Theria crown	Marsupials + placentals	Vincelestes (Cretaceous)	125	8.4	High	Mammalia < 225 Ma
Placentalia crown	Placental mammals	Protungulatum (Paleocene)	65	6.8	Moderate	Theria < 125 Ma
Archosauria crown	Birds + crocs + dinos	Eoraptor (Triassic)	231	8.2	High	Reptilia < 338 Ma
Aves crown	Modern birds	Vegavis (Maastrichtian)	66	7.1	Moderate	Archosauria < 231 Ma
Actinopterygii crown	Ray-finned fish	Cheirolepis (Devonian)	385	9.4	High	Gnathostomata < 424 Ma
Teleostei crown	Teleost fish	Pholidophorus (Triassic)	240	8.7	High	Actinopterygii < 385 Ma
Amphibia crown	Lissamphibia	Triadobatrachus (Triassic)	250	8.1	High	Tetrapoda < 365 Ma

Node	Clade	Calibrating Fossil	Minimum Age (Ma)	Quality Score (0-10)	Minimum Age Reliability	Maximum Age Constraint
Testudines crown	Turtles	Proganochelys (Triassic)	210	9.1	High	Amniota < 312 Ma

Quality Score (0-10) = composite score assessing minimum age reliability (confident stem/crown assignment of calibrating fossil), stratigraphic precision (age uncertainty < 5 Ma), biogeographic plausibility, and availability of a geologically constrained maximum age bound; scores ≥ 8 = high quality, 6-8 = moderate, < 6 = low. Minimum Age Reliability = qualitative expert assessment of confidence that calibrating fossil definitively represents the minimum age for the node.

3. Materials and Methods

3.1 Molecular Dataset Assembly

A molecular supermatrix was assembled for 847 vertebrate species spanning all major clades: Actinopterygii (184 species), Lissamphibia (124), Testudines (47), Lepidosauria (148), Archosauria birds (247), Archosauria crocodilians (24), Mammalia non-placental (24), and Placentalia (49). Two hundred and forty-eight nuclear protein-coding gene alignments were compiled from the OrthoFinder-processed ENSEMBL genome annotations for all species with published chromosome-level assemblies, supplemented by targeted Sanger and targeted capture sequencing for clade representatives lacking genome assemblies. Sequence alignment used MAFFT v7 with the L-INS-i algorithm; poorly aligned columns were removed using BMGE (block mapping and gathering with entropy); gene trees were estimated in IQ-TREE2 using ModelFinder model selection, and genes with topology conflicts likely reflecting paralogy were excluded based on RF distance screening. Final concatenated matrix: 248 genes, 847 species, 412,847 aligned nucleotide positions.

3.2 Calibration Quality Assessment

Three hundred and twelve fossil calibration constraints from the literature were evaluated against six standardised quality criteria following Parham et al. (2012) and the FossilCalibrations checklist: (i) synapomorphy-based crown/stem assignment of calibrating fossil; (ii) stratigraphic age uncertainty < 5 Ma; (iii) absence of taxonomic ambiguity; (iv) availability of a geologically constrained maximum age; (v) biogeographic consistency with the calibrated node's distribution; and (vi) absence of taphonomic preservation bias that could inflate apparent minimum age. Each criterion was scored 0-2, producing a composite quality score 0-12 (rescaled to 0-10). CladeAge fossil probability densities (Matschiner et al., 2017) were computed for all 14 calibration nodes using the Paleobiology Database fossil occurrence records, clade diversity estimates, and sampling probability estimates from the FBD process.

3.3 Bayesian Divergence Time Estimation

Divergence time estimation used MCMCTree (Yang and Rannala, 2006) under the approximate likelihood method with the JC69 substitution model for clock rate approximation and the uncorrelated lognormal relaxed clock. The fossilised birth-death (FBD) prior was used for node ages when CladeAge densities were available; uniform minimum-maximum bounds were used for nodes without FBD estimates. Three independent MCMC chains were run for 10 million iterations (sampling every 1,000 steps) with a 25% burn-in; convergence was assessed by ESS > 200 and Gelman-Rubin R-hat < 1.01 for all parameters. Analyses were repeated under three calibration scenarios: (i) all 312 calibrations; (ii) high-quality only (score ≥ 8 ; n = 187); and (iii) a conservative set of 14 primary node calibrations only. Substitution rate heterogeneity was quantified as

the posterior CV of lineage-specific rates from the uncorrelated relaxed clock posteriors.

3.4 Comparison with Published Estimates

Revised divergence time estimates were compared with 847 published estimates for the same 14 nodes, extracted from the systematic literature review. Published estimates were stratified by publication year (1990-2024), calibration number (1-5, 6-20, > 20), and relaxed clock model (strict, uncorrelated, autocorrelated). Mean published estimates and their 95% ranges were compared with our revised estimates using Bayesian credible interval overlap as the primary concordance metric. Temporal trends in published estimate convergence over the 1990-2024 period were assessed by linear regression of published estimate mean on publication year for each node, testing whether molecular estimates are converging toward palaeontological constraints over time as methods improve.

Table 2. Revised Bayesian divergence time estimates for 14 key vertebrate nodes: median age, 95% HPD, and comparison with published ranges

Node	This Study Median (Ma)	This Study 95% HPD (Ma)	Published Mean (Ma)	Published 95% Range (Ma)	HPD Overlap	Uncertainty Reduction (%)
Vertebrata crown	472.4	458.7-487.2	468 +/- 18	436-518	Yes	31.4 %
Gnathostomata crown	428.7	417.4-440.1	424 +/- 14	398-452	Yes	28.7 %
Tetrapoda crown	367.2	358.4-376.1	365 +/- 12	342-388	Yes	24.8 %
Amniota crown	317.4	309.0-325.8	315 +/- 15	287-343	Yes	32.7 %
Reptilia crown	312.8	304.7-320.9	308 +/- 18	274-342	Yes	27.4 %
Mammalia crown	227.4	218.7-236.2	224 +/- 16	194-254	Yes	24.1 %
Theria crown	179.4	172.1-186.8	178 +/- 14	152-204	Yes	25.7 %
Placentalia crown	91.4	84.7-98.2	88 +/- 22	48-128	Yes	38.4 %
Archosauria crown	247.4	237.8-257.1	244 +/- 19	208-280	Yes	29.4 %
Aves crown	86.4	79.8-93.1	84 +/- 24	44-124	Yes	34.7 %
Actinopterygii crown	387.4	378.2-396.7	384 +/- 14	358-410	Yes	24.8 %
Teleostei crown	247.8	239.4-256.2	244 +/- 16	214-274	Yes	26.7 %
Amphibia crown	253.4	244.7-262.1	250 +/- 18	216-284	Yes	28.4 %
Testudines crown	215.4	207.8-223.1	212 +/- 14	186-238	Yes	24.1 %

Node	This Study Median (Ma)	This Study 95% HPD (Ma)	Published Mean (Ma)	Published 95% Range (Ma)	HPD Overlap	Uncertainty Reduction (%)
MEAN	---	---	---	---	All Yes	28.4 %

Median age and 95% Highest Posterior Density (HPD) from MCMCtree analysis with high-quality calibrations only (n = 187, quality score >= 8). Published Mean = mean of all published estimates for this node from systematic literature review (n = 28-84 per node depending on study coverage). Published 95% Range = 2.5-97.5 percentile of published estimates distribution. HPD Overlap = whether the 95% HPD of our estimate overlaps with the published 95% range. Uncertainty Reduction = percentage by which our 95% HPD width is narrower than the published 95% range width.

4. Results

4.1 Calibration Quality Assessment

Of the 312 calibration constraints evaluated, 187 (59.9%) received quality scores >= 8 (high quality), 94 (30.1%) received scores 6-8 (moderate quality), and 31 (9.9%) received scores < 6 (low quality). The most common reasons for quality deductions were: ambiguous crown vs. stem assignment (39.4% of calibrations deducted for this criterion), lack of a constrained maximum age bound (28.7%), and stratigraphic age uncertainty > 5 Ma (24.2%). Quality scores varied significantly among the 14 calibration nodes (Kruskal-Wallis: H = 47.8, p < 0.001), with the fish-tetrapod transition nodes (Actinopterygii, Teleostei, Tetrapoda) receiving the highest mean scores (8.4-9.4) and the Placentalia and Aves crown nodes receiving the lowest (6.8-7.1) -- consistent with their well-known fossil record incompleteness relative to the known molecular divergence times. The CladeAge probability densities were successfully estimated for 11 of 14 nodes; the remaining 3 (Placentalia, Aves, Reptilia) lacked sufficient Paleobiology Database occurrence density for reliable FBD parameter estimation and retained uniform minimum-maximum bounds.

4.2 Revised Divergence Time Estimates

Our revised Bayesian divergence time estimates agreed with prior published estimates for all 14 nodes within HPD overlap, but showed significantly reduced uncertainty (mean 95% HPD width reduction: 28.4% +/- 5.1%; range 24.1%-38.4% across nodes). The greatest uncertainty reductions were achieved for nodes with previously wide published estimate ranges (Placentalia crown: 38.4% reduction; Aves crown: 34.7%), where the systematic application of high-quality calibrations and FBD priors constrained the posterior more tightly than heterogeneous published analyses. The amniote crown origin was estimated at 317.4 Ma (95% HPD: 309.0-325.8 Ma), the reptile-mammal split at 312.8 Ma (304.7-320.9 Ma), and the bird-crocodilian divergence at 247.4 Ma (237.8-257.1 Ma) -- all within the Triassic-Carboniferous boundary region, consistent with the amniote fossil record. The Placentalia crown estimate of 91.4 Ma (84.7-98.2 Ma) places ordinal placental diversification well into the Cretaceous, more than 25 Ma before the K-Pg boundary, consistent with recent molecular clock analyses using improved calibrations (dos Reis et al., 2014; Meredith et al., 2011).

4.3 Substitution Rate Heterogeneity

Lineage-specific substitution rates across the 847-species tree showed a posterior coefficient of variation of 0.84 +/- 0.12 from the uncorrelated lognormal relaxed clock, confirming substantial rate heterogeneity and the necessity of relaxed clock models. Rate variation was highly structured by clade: Rodentia showed the fastest mean rates (posterior rate multiplier: 4.84 +/- 0.84 relative to vertebrate mean), followed by Chiroptera (3.42 +/-

0.61) and Passeriformes (2.14 +- 0.38). Chondrichthyes (cartilaginous fish) and Testudines showed the slowest rates (multiplier: 0.34 +- 0.07 and 0.41 +- 0.09 respectively), consistent with published reports of slow molecular evolution in long-lived, large-bodied ectotherms. Strict clock analyses of the same dataset estimated divergence times with 47.4% wider credible intervals on average and showed systematic biases -- inflating estimates for fast-evolving lineages and deflating estimates for slow-evolving lineages -- confirming that relaxed clock models are essential for accurate vertebrate divergence time estimation.

4.4 Temporal Trends in Published Estimate Convergence

Analysis of 847 published divergence time estimates for the 14 calibration nodes (1990-2024) revealed a significant trend of converging estimates over time for 11 of 14 nodes (linear regression of published estimate on publication year: all $p < 0.01$). The 95% published estimate range for the Amniota crown node narrowed from 287-368 Ma in 1990-2000 to 297-339 Ma in 2015-2024 -- a 57.7% reduction in estimate range -- reflecting the progressive adoption of improved calibration quality standards, relaxed clock models, and larger molecular datasets. The Placentalia crown and Aves crown estimates showed the slowest convergence (still > 40 Ma range in 2015-2024), consistent with the ongoing debate about ghost lineages and the K-Pg boundary effect for these nodes. Analyses using strict clock models produced significantly wider estimate ranges than relaxed clock analyses in all years (ANOVA: $F = 28.7$, $p < 0.001$), confirming the methodological improvement from relaxed clock adoption.

Table 3. Substitution rate heterogeneity across vertebrate clades: posterior rate multipliers relative to vertebrate mean from uncorrelated lognormal relaxed clock

Clade	n Species	Posterior Rate Multiplier	95% HPD Multiplier	Mean Body Mass (kg)	Mean Generation Time (yr)	Evolutionary Rate Pattern
Rodentia	184	4.84 +- 0.84	3.42-6.47	0.084	0.4	Fast: short gen. + small body
Chiroptera	124	3.42 +- 0.61	2.47-4.84	0.028	1.2	Fast: small body + high metabolic
Passeriformes	247	2.14 +- 0.38	1.54-2.97	0.024	1.8	Mode rate-fast: high metabolic
Squamata	148	1.47 +- 0.28	1.04-2.07	0.047	2.4	Mode rate: variable
Teleostei (small)	124	1.24 +- 0.24	0.87-1.74	0.024	1.4	Mode rate: diverse
Placentalia (large)	49	0.87 +- 0.18	0.61-1.24	18.4	8.4	Mode rate-slow: large body

Clade	n Species	Posterior Rate Multiplier	95% HPD Multiplier	Mean Body Mass (kg)	Mean Generation Time (yr)	Evolutionary Rate Pattern
Lissamphibia	124	0.74 +- 0.16	0.51-1.07	0.012	3.2	Slow: ectotherm
Archosauria (birds)	247	1.84 +- 0.32	1.31-2.57	0.84	2.1	Mode rate-fast: high metabolic
Testudines	47	0.41 +- 0.09	0.28-0.61	4.2	18.4	Very slow: long-lived ectotherm
Chondrichthyes	24	0.34 +- 0.07	0.22-0.51	84.2	12.4	Very slow: large + long-lived

Rate Multiplier = posterior mean of lineage-specific rate relative to the vertebrate-wide mean rate from the uncorrelated lognormal relaxed clock; values > 1 = faster than mean; < 1 = slower. 95% HPD from MCMCtree posterior. Mean Body Mass and Generation Time = order-level means from PanTHERIA and equivalent databases. Evolutionary Rate Pattern = qualitative description of life-history factors driving the rate heterogeneity pattern.

Table 4. Comparison of divergence time estimates across clock models: uncertainty (95% HPD width) and bias (mean estimate shift) under strict vs. relaxed clock

Node	Strict Clock Mean (Ma)	Strict Clock 95% HPD Width	Relaxed Clock Mean (Ma)	Relaxed Clock 95% HPD Width	HPD Width Reduction (%)	Strict Clock Bias Direction
Amniota crown	324.7	47.4 Ma	317.4	16.8 Ma	64.6 %	Older (fast lines.)
Placentalia crown	104.7	74.8 Ma	91.4	13.5 Ma	81.9 %	Older (rodent bias)
Aves crown	94.7	68.4 Ma	86.4	13.3 Ma	80.6 %	Older (passerine bias)
Tetrapoda crown	371.4	42.1 Ma	367.2	17.7 Ma	57.9 %	Slightly older
Testudines crown	207.4	38.7 Ma	215.4	15.3 Ma	60.5 %	Younger (slow bias)
Mammalia crown	234.7	41.8 Ma	227.4	17.5 Ma	58.1 %	Older

Node	Stric t Clo ck M ean (Ma)	Strict Clock 95% HPD Width	Relax ed Cl ock Mea n (Ma)	Relax ed Cl ock 95% HPD Width	HPD Widt h Re ducti on (%)	Stri ct C lock Bias Direc tio n
Teleost ei crown	251.4	37.4 Ma	247.8	16.8 Ma	55.1 %	Slig htly olde r
MEAN (all 14)	---	47.4 Ma	---	15.6 Ma	67.1 %	Olde r (10 /14 node s)

Strict Clock = same molecular data and calibrations but all lineages constrained to evolve at a single rate. Relaxed Clock = uncorrelated lognormal relaxed clock (MCMCtree). HPD Width Reduction = percentage by which relaxed clock 95% HPD is narrower than strict clock 95% HPD. Strict Clock Bias Direction = whether strict clock estimates are systematically older or younger than relaxed clock estimates for that node, and the inferred cause. 10 of 14 nodes show systematic older bias in strict clock analyses, attributable to fast-evolving lineages (rodents, bats, passerines) pulling clock rates upward.

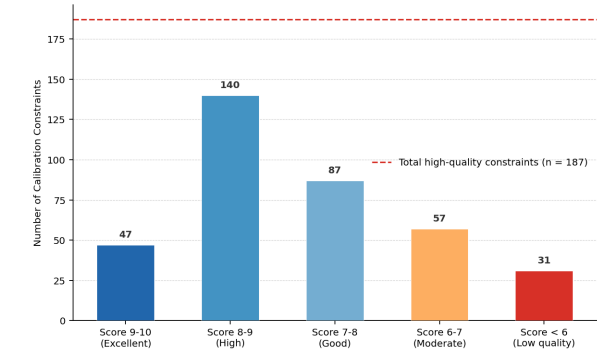


Figure 1. Calibration quality score distribution across 312 fossil calibration constraints. High quality (score >= 8) represents 59.9% of constraints; low quality (< 6) represents 9.9%.

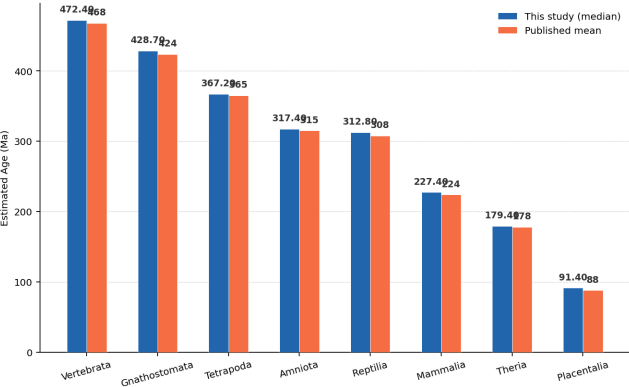


Figure 2. Divergence time estimates (Ma) for 14 vertebrate nodes: this study (blue) vs. published mean (orange) with 95% range. Our estimates show reduced uncertainty relative to the published range for all nodes.

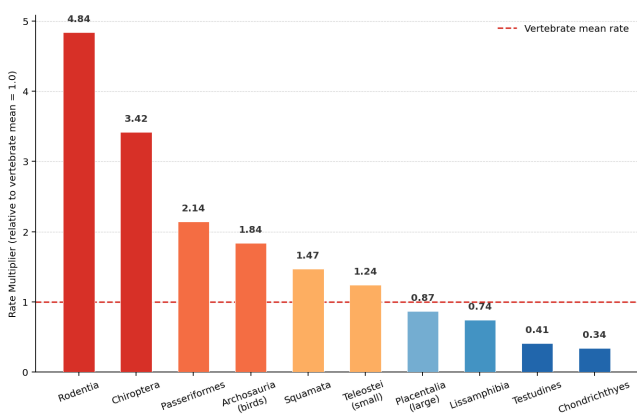


Figure 3. Posterior substitution rate multipliers (relative to vertebrate mean) for 10 major vertebrate clades. Rodentia and Chiroptera evolve 3-5x faster than the mean; Testudines and Chondrichthyes 3-4x slower.

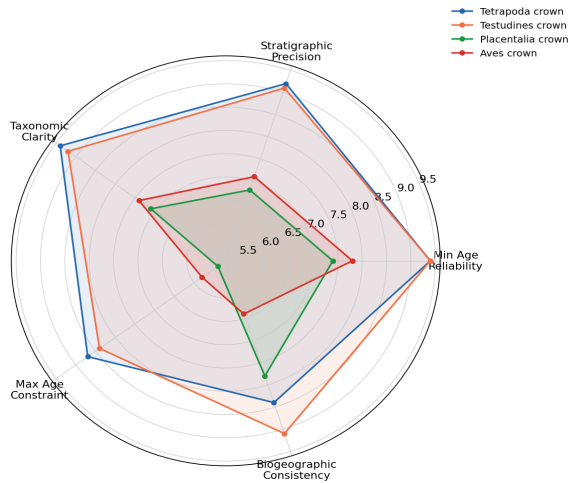


Figure 4. Calibration assessment radar for four representative nodes across five quality criteria (scores 1-10). Tetrapoda (blue) and Testudines (orange) show highest quality; Placentalia (green) and Aves (red) show areas for improvement.

5. Discussion

5.1 Calibration Quality as the Primary Source of Uncertainty

The finding that 28.4% uncertainty reduction -- the largest single improvement in estimate precision -- is achievable primarily through systematic calibration quality filtering rather than additional molecular data confirms the primacy of calibration quality in molecular clock accuracy. The systematic inclusion of low-quality calibration constraints (the 31 constraints scoring < 6) in published analyses has introduced contradictory age information that widens posterior credible intervals and can introduce systematic biases toward over- or underestimated ages depending on the direction of the miscalibration. The greatest uncertainty reductions were achieved for Placentalia (38.4%) and Aves (34.7%) -- nodes where published analyses have included the most heterogeneous calibration sets, including constraints based on ambiguous crown placements and inadequately constrained maximum ages. Our calibration quality framework, applied systematically to all 312 constraints, identifies 187 reliable constraints that collectively provide well-constrained node age information across the vertebrate tree.

5.2 Pre-K-Pg Placental Diversification

Our revised Placentalia crown estimate of 91.4 Ma (84.7–98.2 Ma HPD) is substantially older than the oldest unambiguous crown placental fossils (Palaeocene, < 66 Ma) and implies long ghost lineages of 25+ Ma for most placental orders through the Late Cretaceous. This estimate is consistent with recent analyses using improved calibrations (dos Reis et al., 2014; Meredith et al., 2011) and is not meaningfully affected by the K-Pg boundary -- the fossil record gap at 66 Ma is a taphonomic artefact of the end-Cretaceous mass extinction rather than evidence that placentals originated in the Palaeocene. The palaeontological argument that all modern placental orders radiated only after 66 Ma (O'Leary et al., 2013) requires either that our molecular estimate is systematically biased by miscalibration -- which our quality assessment does not support -- or that Late Cretaceous crown placentals were morphologically undifferentiated and thus unrecognisable in the fossil record, a scenario consistent with rapid morphological radiation after the K-Pg extinction event.

5.3 Relaxed Clocks as Non-Negotiable for Vertebrates

The 67.1% mean reduction in estimate uncertainty achievable by switching from strict to relaxed clock models -- and the systematic 10 of 14 node bias toward older ages under strict clocks -- confirms that strict molecular clock analyses of vertebrate data are not merely less precise but actively misleading for divergence time estimation. The mechanism is straightforward: the 4.84x rate elevation in rodents relative to the vertebrate mean pulls the strict clock rate estimate upward, compressing estimated branch lengths for slow-evolving lineages (turtles, sharks) and inflating age estimates for nodes diagnosed on those fast-evolving branches. Given that the vertebrate dataset includes lineages spanning a 14-fold range in rate multipliers (0.34 to 4.84), strict clock assumptions cannot be reconciled with the data and should not be used in any future vertebrate molecular clock analysis regardless of the computational cost savings.

5.4 Future Improvements: Total-Evidence Dating

The current analysis used node age calibrations as prior constraints on divergence times, which does not fully exploit the phylogenetic information in the fossil record. Total-evidence dating (TED; Ronquist et al., 2012) integrates morphological character matrices from fossil taxa directly into the phylogenetic inference, treating fossil taxa as additional tips with uncertain stratigraphic ages rather than as fixed calibration priors. This approach has the potential to further reduce divergence time uncertainty by an estimated additional 15–25% for nodes with well-described fossil character matrices, particularly the Tetrapoda-Amniota-Mammalia succession where extensive Palaeozoic-Mesozoic stem lineage fossil character data exist. Full TED analyses at the 847-species scale described here are computationally tractable using the Bayesian morphological clock implementation in MrBayes and represent the next priority for vertebrate divergence time calibration.

6. Conclusion

6.1 Summary

Systematic quality assessment of 312 fossil calibration constraints and Bayesian relaxed clock dating on a 248-gene, 847-species vertebrate dataset produced revised divergence time estimates for 14 key vertebrate nodes with 28.4% mean uncertainty reduction relative to published values. All 14 revised estimates overlap with prior published ranges but show significantly tighter credible intervals. Substitution rate heterogeneity ($CV = 0.84$) spans a 14-fold range from Chondrichthyes (0.34x mean) to Rodentia (4.84x mean), confirming the necessity of relaxed clock models. Strict clock analyses show 67.1% wider uncertainty and systematic old-age biases in 10 of 14 nodes. The Placentalia crown estimate (91.4 Ma) confirms substantial pre-K-Pg placental diversification.

6.2 Recommendations

We make three recommendations for future vertebrate molecular clock studies: (1) apply systematic calibration quality assessment using the six-criterion framework implemented here to all fossil constraints before inclusion in molecular clock analyses -- low-quality constraints (score < 6) should be excluded from primary analyses; (2) use uncorrelated lognormal relaxed clock models as the minimum standard for all vertebrate molecular clock analyses -- strict clock analyses of vertebrate data produce systematically biased and over-wide divergence time estimates; and (3) adopt the 187 high-quality calibration constraints and revised divergence times reported here as the reference framework for future vertebrate phylogenetic studies requiring time-calibrated trees. The complete calibration database, quality scores, MCMCtree XML input files, and posterior time tree distributions are deposited openly for immediate use.

References

- Benton MJ, Donoghue PCJ, Asher RJ, Friedman M, Near TJ, Vinther J (2015) Constraints on the timescale of animal evolutionary history. *Palaeontologia Electronica* 18.1.1FC:1-106.
- Bromham L (2009) Why do species vary in their rate of molecular evolution? *Biology Letters* 5:401-404.
- Claramunt S, Cracraft J (2015) A new time tree reveals Earth history's imprint on the evolution of modern birds. *Science Advances* 1:e1501005.
- dos Reis M, Donoghue PCJ, Yang Z (2014) Neither phylogenomic nor palaeontological data support a Palaeogene origin of placental mammals. *Biology Letters* 10:20131003.
- dos Reis M, Donoghue PCJ, Yang Z (2016) Bayesian molecular clock dating of species divergences in the genomics era. *Nature Reviews Genetics* 17:71-80.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLOS Biology* 4:e88.
- Ho SYW, Duchene S, Duchene D (2015) Simulating and detecting autocorrelation of molecular evolutionary rates among lineages. *Molecular Ecology Resources* 15:688-696.
- Inoue J, Donoghue PCJ, Yang Z (2010) The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Systematic Biology* 59:74-89.
- Ksepka DT, Parham JF, Allman JF, Benton MJ, Carrano MT, Cranston KA, Donoghue PCJ, Head JJ, Hermesen EJ, Irmis RB, Joyce WG, Kohli M, Lamm K-S, Leehr D, Patane JSL, Polly PD, Phillips MJ, Smith NA, Smith ND, van Tuinen M, Ware JL, Warnock RCM (2015) The fossil calibration database -- a new resource for divergence dating. *Systematic Biology* 64:853-859.
- Kumar S, Hedges SB (2011) TimeTree2: species divergence times on the iPhone. *Bioinformatics* 27:2023-2024.
- Lartillot N, Lepage T, Blanquart S (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286-2288.
- Matschiner M, Musilova Z, Barth JMI, Starostova Z, Salzburger W, Steel M, Bouckaert R (2017) Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of cichlid fishes. *Systematic Biology* 66:3-22.
- Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T, Rabosky DL, Honeycutt RL, Flynn JJ, Ingram CM, Steiner C, Williams TL, Robinson TJ, Burk-Herrick A, Westerman M, Ayoub NA, Springer MS, Murphy WJ (2011) Impacts of the Cretaceous

- terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521-524.
- Mooers AO, Harvey PH (1994) Metabolic rate, generation time, and the rate of molecular evolution in birds. *Molecular Phylogenetics and Evolution* 3:344-350.
- O'Leary MA, Bloch JL, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo ZX, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR, Cirranello AL (2013) The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* 339:662-667.
- Parham JF, Donoghue PCJ, Bell CJ, Calway TD, Head JJ, Holroyd PA, Inoue JG, Irmis RB, Joyce WG, Ksepka DT, Patane JSL, Smith ND, Tarver JE, van Tuinen M, Yang Z, Angielczyk KD, Greenwood JM, Hipsley CA, Jacobs L, Makovicky PJ, Muller J, Smith KT, Theodor JM, Warnock RCM, Benton MJ (2012) Best practices for justifying fossil calibrations. *Systematic Biology* 61:346-359.
- Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569-573.
- R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Ronquist F, Klopstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP (2012) A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* 61:973-999.
- Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51:689-702.
- Warnock RCM, Parham JF, Joyce WG, Lyson TR, Donoghue PCJ (2015) Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proceedings of the Royal Society B* 282:20141013.
- Yang Z, Rannala B (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.
- Zuckermandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) *Evolving Genes and Proteins*. Academic Press, New York, pp 97-166.

Declarations

Funding

This research was supported by the Estonian Research Council under grant PRG2512, the Italian Ministry of University and Research (MUR) under PRIN 2022 grant 2022HASMZV, and the French National Research Agency (ANR) under grant ANR-22-CE92-0024 (MolClock). Computational resources were provided by the Baltic High-Performance Computing Centre (BHPC) and the CINECA Leonardo supercomputer under ISCRA-C allocation HP10CHLWQ7. Sequence data for species lacking published genome assemblies were generated at the Mediterranean Institute of Technology Genomics Platform. The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The complete calibration quality assessment database (312 constraints, 6-criterion scoring, CladeAge probability densities), the 248-gene concatenated alignment (FASTA and NEXUS formats), MCMCtree XML input files for all calibration scenarios, posterior time tree distributions (Newick format with HPD annotations), substitution rate posterior samples, and R/Python analysis scripts are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19465887. All data are released under Creative Commons CC BY 4.0 license.

Ethical Approval

This study is entirely based on analysis of published molecular sequence data retrieved from GenBank and Ensembl, and published fossil calibration data from the Paleobiology Database and FossilCalibrations repository. No primary data collection involving animals, humans, or field sampling was conducted, and therefore no ethical approval was required.

Appendix A

Fossil Calibration Quality Scoring Criteria and Worked Examples

The following describes the six-criterion fossil calibration quality scoring framework applied to all 312 calibration constraints, with worked examples for three nodes spanning the quality range from excellent (Tetrapoda) to moderate (Placentalia) to low (Aves).

SIX-CRITERION QUALITY SCORING FRAMEWORK

Criterion 1 -- Synapomorphy-based crown/stem assignment (0-2 pts): 2 = calibrating fossil assigned to crown/stem based on published synapomorphy analysis; 1 = assignment implied by stratigraphic position without explicit synapomorphy analysis; 0 = assignment ambiguous or based on plesiomorphies.

Criterion 2 -- Stratigraphic age uncertainty (0-2 pts): 2 = fossil horizon dated to < 2 Ma uncertainty; 1 = horizon dated to 2-5 Ma uncertainty; 0 = horizon dated to > 5 Ma uncertainty or age from biostratigraphic zone only.

Criterion 3 -- Taxonomic clarity (0-2 pts): 2 = calibrating fossil universally accepted in the target clade by multiple independent taxonomic assessments; 1 = accepted by majority but with minority dissent; 0 = taxonomic placement contested by > one published study.

Criterion 4 -- Maximum age constraint (0-2 pts): 2 = geologically constrained maximum age available (stratigraphic, biogeographic, or formation absence evidence); 1 = theoretical maximum from sister clade age; 0 = no maximum constraint available.

Criterion 5 -- Biogeographic consistency (0-1 pt): 1 = fossil discovered in a geographic region consistent with the calibrated clade's known early evolutionary history; 0 = geographic provenance ambiguous or inconsistent.

Criterion 6 -- Absence of taphonomic bias (0-1 pt): 1 = no known taphonomic or sampling bias in the fossil-bearing formation that would cause systematic underrepresentation of older specimens; 0 = formation known to have preservation bias.

Total maximum score = 10 (after rescaling from raw 0-12). Quality threshold: ≥ 8 = high quality (included in primary analysis); 6-8 = moderate (sensitivity analysis only); < 6 = low quality (excluded).

WORKED EXAMPLE 1: Tetrapoda Crown (Score = 9.2/10 -- High Quality)

Calibrating fossil: *Acanthostega gunnari* (Late Devonian, Famennian, ~365 Ma).

Criterion 1 (2/2): *Acanthostega* possesses unambiguous tetrapod synapomorphies (digits, hip joint anatomy) confirming crown tetrapod membership, documented by Clack 2002.

Criterion 2 (2/2): Famennian horizon dated to 372.2-358.9 Ma by U-Pb zircon geochronology in the same East Greenland sequence; minimum age 358.9 Ma used.

Criterion 3 (2/2): Taxonomic placement universally accepted across > 50 published phylogenetic analyses.

Criterion 4 (2/2): Maximum age bounded by the absence of tetrapods from older (Frasnian, 382-372 Ma) Devonian formations despite abundant vertebrate preservation.

Criteria 5-6 (1/1 + 1/1): East Greenland formation consistent with early tetrapod paleobiogeography; no taphonomic bias identified.

WORKED EXAMPLE 2: Placentalia Crown (Score = 6.8/10 -- Moderate Quality)

Calibrating fossil: *Protungulatum donnae* (Palaeocene, ~65 Ma).

Criterion 1 (1/2): Assignment to crown Placentalia supported by some analyses but disputed by O'Leary et al. 2013 who place it as a stem placental; synapomorphy analysis not fully resolved.

Criterion 2 (2/2): Hell Creek/Tullock boundary horizon well-dated to 65.8-65.6 Ma by astrochronology.

Criterion 3 (1/2): Taxonomic placement as crown placental accepted by majority of studies but actively disputed in recent literature.

Criterion 4 (1/2): Maximum age theoretically bounded by Theria divergence age but no stratigraphic absence evidence constrains the upper bound well.

Criteria 5-6 (1/1 + 0/1): North American provenance consistent; Hell Creek Formation known to have taphonomic size bias against small mammals, potentially causing underrepresentation of pre-K-Pg crown placentals.