

Genomic divergence between species identifies taxonomic outliers and establishes calibrated molecular clocks

Natalie M. Allen¹, Mekala Sundaram², Avril M. Harder³, Jong Yoon Jeon^{1,4}, Gina F. Lamka¹, Janna R. Willoughby⁵, and J. Andrew DeWoody^{1*}

¹Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA

²Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, GA 30602, USA

³HudsonAlpha Genome Sequencing Center, Huntsville, AL 35806, USA

⁴Harvard University Center for the Environment, Harvard University, 26 Oxford St, Cambridge, MA 02138, USA

⁵College of Forestry, Wildlife, and Environment, Auburn University, Auburn, AL 36849, USA

***Email:** dewoody@purdue.edu

Author Contributions:

Conceptualization: JAD; Methodology: NMA, JRW, MS, AMH; Investigation: NMA, MS, JYJ, GFL, JRW; Visualization: NMA; Funding acquisition: JAD, NMA; Project administration: JAD; Supervision: JAD; Writing – original draft: NMA; Writing – review & editing: NMA, MS, AMH, JYJ, GFL, JRW, JAD.

Competing Interest Statement: Authors declare that they have no competing interests.

Keywords: Systematics, genome-wide divergence, species delimitation, substitution rate

25 **Abstract**

26 For decades, molecular clocks that rely on one or a few genes have been employed to
27 quantify evolutionary rates (e.g., anagenesis) or to help date important phylogeographic
28 events (e.g., vicariance). Genome-wide clocks remain largely unexplored, and the degree
29 to which genome-wide divergence scales with time is still unknown in most vertebrates.
30 Here, we expand the concept of the molecular clock to entire genomes by calibrating both
31 nuclear and mitochondrial clocks. We do so in ~700 congeneric tetrapod species pairs
32 using distance estimates from pairwise genome alignments and best estimates of time
33 since divergence. We find that divergence time is the dominant and most consistent
34 predictor of genome-wide divergence, with both mitochondrial and nuclear clocks scaling
35 log-linearly with time. Because our genomic clocks average across entire genomes, they
36 have the ability to smooth over locus-specific variance. The two clocks complement each
37 other temporally, with nuclear clocks being most reliable at deeper divergence times (>5
38 MYA) and mitochondrial clocks effectively resolving shallower divergence times (<5
39 MYA). Calibration at narrower taxonomic scales improves accuracy, as illustrated by our
40 Order-level clocks that reflect known biological differences among clades (i.e., life
41 history pace). When applied as diagnostic tools, the clocks recover known cryptic species
42 complexes (e.g., mustached bats) and biological phenomena such as mitochondrial
43 capture (e.g., rock lizards), establishing their utility in flagging candidate species groups
44 for additional evolutionary analyses and/or taxonomic reevaluation.

45 **Significance Statement**

46 Genome sequencing rates have surpassed the rate at which systematists can designate
47 taxonomies, yet biologists have not yet developed and embraced robust methods to
48 evaluate genomic similarity and dissimilarity among species pairs in the context of their
49 taxonomic relatives. By leveraging publicly available DNA sequence data from both
50 nuclear and mitochondrial genomes, we calibrate general-purpose molecular clocks for
51 most vertebrates. The two clocks tick at complementary rates (the mitochondrial clock

faster, the nuclear clock slower) and we use them to demonstrate their potential to systematically flag cryptic, misclassified, and/or unusual species.

Main Text

Introduction

The comparison of both protein and DNA sequence divergence has long been used to estimate rates, timing, and drivers of evolutionary change. The molecular clock hypothesis posits that sequence differences accumulate in a roughly clock-like fashion over time (1). However, the rate of sequence divergence is heterogeneous among taxonomic groups, and the use of sequence divergence as a proxy for evolutionary divergence must be evaluated with lineage-specific molecular clocks (2). Mitochondrial genes were the basis of most original molecular clocks: with short sequences, rapid evolution due to small effective population sizes, and limited recombination, they formed the backbone of eukaryotic phylogenetics and species delineation (2, 3, 4). In 1998, Johns and Avise (5) used the mitochondrial cytochrome *b* gene to establish a “yardstick” of genetic dissimilarity for comparing taxonomic groups across the vertebrate tree, revealing relatively slow rates of sequence evolution in birds relative to reptiles and amphibians. Such foundational clocks were built from one or a few mitochondrial genes, but it remains unclear how genetic distance scales with time across the entire mitochondrial or nuclear genomes, how the rates of evolution actually differ between genomes with different effective population sizes, and whether the noise associated with locus-specific clock variance can be meaningfully reduced with genome-wide clocks.

The last decade or two has ushered a rapid shift towards whole-genome sequence datasets for resolving phylogenetic relationships (6, 7) and for estimating lineage divergence times (8). The establishment of comprehensive molecular clocks for both nuclear and mitochondrial genomes would enable comparisons of divergence between the two genomes, but also among lineages where a comprehensive clock would help circumvent distortions that can occur when using single genes (9). With advances in both sequencing technologies and bioinformatics, whole-genome clocks are now within reach. To date,

there are over 14,000 vertebrate nuclear genomes available from the NCBI, and rapid whole-genome alignment has made it possible to calculate genetic distance as genome-wide nucleotide divergence (10, 11). The development of genome-wide clocks presents the opportunity for population genomics to inform theoretical population genetics with empirical estimates of evolutionary rates contrasted between the nuclear and mitochondrial genomes (where effective population sizes differ four-fold; 12). Additionally, this work enables systematic efforts by identifying taxonomic groupings that—based on genome sequence distances—are relatively over- or under-split. Such cases may be a) bookkeeping anomalies that require taxonomic revision and/or b) evolutionary anomalies that warrant further research to better understand the underlying biology (e.g., cryptic species).

Interpreting genome-wide sequence divergence requires differentiating between the mutation rate and substitution rate. Mutation rate (μ) is the rate at which heritable changes emerge within a generation such that $\mu = \frac{\text{mutations}}{\text{sites} \times \text{generations}}$. Importantly, most mutations never become fixed; the substitution rate (k) is the rate at which mutations become fixed within a population such that $k = \frac{\text{fixed differences}}{\text{site} \times \text{time}}$. Under neutral conditions, mutation rate and substitution rate are equal and independent of population size within a given molecule (i.e., nuclear or mitochondrial genome). While more mutations accumulate in larger populations, this is offset by the lower probability of fixation. Consequently, neutral substitutions accumulate linearly over time, forming the basis of the molecular clock. The accumulation of substitutions in two distinct lineages can be calculated as pairwise nucleotide distance (D). Under a neutral clock, $D = 2kt$, where t is the time since lineage divergence and the factor of 2 accounts for the independent substitutions accumulating in each lineage. Departures from the neutral clock arise due to molecule- and lineage-specific variations in germline mutation rates, generation times, and effective population size (N_e), which mediates the efficiency of selection versus drift (13, 14, 15).

Our clocks are calculated as pairwise whole-genome nucleotide distance regressed against divergence time. We consider them to be calibrated tools that estimate the molecule-specific and clade-specific relationship between sequence divergence and time.

There are many reasons why nucleotide substitutions form the basis of excellent clocks (e.g., the absolute number of available sites, ease of assessment, etc.), but they represent only one axis of genomic divergence. Structural variants (SVs) can also reshape nuclear genomes by elevating nucleotide divergence in rearranged or repetitive regions and sometimes by driving reproductive isolation (16, 17, 18). Because pairwise whole-genome alignment measures substitutions only within homologous, collinear blocks, our nuclear genome divergence estimates are largely insensitive to this structural variation. We view the SV agnosticism of our method as a feature that isolates the roughly clock-like signal of nucleotide substitution for three reasons. First, genome-wide clocks can quantify the variance in temporal resolution within and between each molecule. Second, they can evaluate taxa that differ in evolutionary age and substitution rate on a common scale (i.e. standard normal distribution), allowing for the direct comparison of realized genomic divergence across taxonomic groups. Third, they can estimate the expected pairwise divergence within a group such that taxa that fall outside the expected range can be flagged as candidates for closer inspection.

Because molecular clocks describe the expected pairwise divergence of an operational taxonomic unit (OTU) given the time since a common ancestor, deviations from the clock occur when the underlying substitution rate changes. The rate of genomic divergence may shift dramatically between molecules (due to differences in mutation and recombination rates) or subtly with life history traits. For example, species with shorter generation times experience more germline mutations per unit time (19). Nuclear genome size is thought to be positively associated with DNA sequence divergence (20, 21), and body size is often negatively associated with sequence divergence because it is mechanistically tied to metabolic rate (22, 23). Consistent with these expectations, the per-generation germline mutation rate varies roughly 40-fold across vertebrates and is predicted by generation time, age at maturity, and fecundity, with species of larger N_e having lower mutation rates (14). Additionally, mitochondrial DNA evolves more rapidly due to higher mutation rates.

The difficulty for any comparative test of life history traits is that they are not independent. Body size, generation time, metabolic rate, genome size, and N_e all covary,

and each has strong phylogenetic signal in that closely related OTUs tend to share these traits (24). We therefore expect that after accounting for shared ancestry, lineage divergence time will be the most consistent predictor of genomic sequence divergence, and that the predictive power of individual traits will map to phylogenetic structure. In addition to establishing genome-wide clocks, quantifying the relative predictive power of different traits in different Classes will provide context for variation within, or deviations from, each clock.

Tetrapods are a well-suited system in which to build and test genomic clocks. Many reference-quality genome assemblies exist thanks to coordinated efforts including the Vertebrate Genomes Project (25), making genome-wide divergence estimates possible for many pairs of molecules and OTUs. These contrasts exhibit variation in body size, generation time, genome size, metabolic strategy, and germline mutation rate, and taxonomic traits are well catalogued by multiple comprehensive databases. Additionally, tetrapods vary in genome architecture (e.g., birds have conserved karyotypes compared to mammals) (26, 27). By contrasting clocks, we critically test whether a nucleotide-level measure of divergence behaves consistently across molecules and lineages. Finally, time-calibrated phylogenetic trees are available for many tetrapods, making it possible to account for shared ancestry when examining the relative predictive capacity of different traits.

Here, we extend the concept of molecular clocks by examining whole-genome divergence in over 700 pairs of tetrapods. For each species pair, we aligned the nuclear genomes and calculated pairwise Kimura-2-parameter (k2p) nucleotide divergence, a key metric that corrects the observed proportions of differing sites for unobserved multiple substitutions and for the differing rates of transitions versus transversions (28). We then regressed $\log(k2p)$ on $\log(\text{divergence time})$ obtained from TimeTree (29) to estimate Class-specific genomic clocks for birds, mammals, and reptiles. TimeTree derives divergence time estimates from thousands of independent, fossil-calibrated phylogenies, typically generated from a small number of loci. Amphibians were included in a pooled nuclear clock, and where sample sizes allowed, we calibrated Order-specific clocks as well. Using standardized z-scores, we examined deviations from our clocks as outlier

case studies, demonstrating the utility of our clocks as diagnostic tools. In the best-studied tetrapod Classes, birds and mammals, we compared the relative predictive power of different traits (divergence time, body size, generation time, genome size, clutch size, population density, geographic range overlap) and data characteristics (genome N50 and genome GC content) in explaining nuclear divergence using phylogenetic mixed models and variance partitioning. Finally, we repeated each analysis with an overlapping dataset of mitochondrial genomes, comparing the utility of nuclear and mitochondrial clocks across molecules, timescales, and taxonomy.

Results and Discussion

Nucleotide divergence accrues in a roughly clock-like fashion, yet its rate varies among molecules and lineages. By fitting the relationship between lineage divergence time and genomic sequence divergence, we present genome-wide clocks which function as a common scale for comparing divergence across molecules, taxa, and as a taxon-specific expectation that can be used to flag anomalous species pairs. Our central finding is that lineage divergence time is by far the strongest predictor of both nuclear and mitochondrial sequence divergence and that the effect of individual life history traits on genomic divergence is largely inseparable from the effect of shared ancestry.

Genome-wide clocks recover a time-dependent signal in tetrapods

After filtering, the nuclear clock was calibrated with 409 bird, 245 mammal, and 44 reptile pairs. The mitochondrial clock was calibrated with 234 bird, 415 mammal, and 28 reptile pairs; the two datasets overlap but are not identical (*Supplementary Fig. S1*). Amphibians were included only in pooled tetrapod clocks as there were too few species pairs in both datasets (7 nuclear and 10 mitochondrial; *Supplementary Tables S1 and S2*).

Regressing $\log(k2p)$ on $\log(\text{divergence time})$ yielded nuclear slopes of 0.34 for birds, 0.57 for mammals, and 0.69 for reptiles, explaining 22%, 41%, and 36% of the variance, respectively (Table 1, Fig. 1A). The mitochondrial clocks were faster, with steeper slopes (0.68 in birds and mammals, 0.80 in reptiles; Table 1, Fig. 1B). These steeper slopes are unsurprising, as mtDNA is a single, non-recombining molecule that evolves nearly

neutrally (30). Thus, its divergence is largely determined by the germline mutation rate (which scales closely with time) whereas nuclear divergence averages over millions of loci that vary quantitatively with respect to selective regimes, linkage disequilibrium, and effective population sizes.

Beyond rate of divergence, nuclear and mitochondrial clocks differ in the divergence time range at which they are most accurate. We found that divergence is inflated at recent splits (less than ~2 MYA), where it scales more weakly with time in birds and mammals (the reptile dataset did not include recently diverged pairs; Fig. 1C and *Supplementary Table S3*). This may be due to inflated nucleotide divergence in recent lineage splits, where genomes reflect variation already segregating in the common ancestor at the time of lineage coalescence (31). We do not observe this “floor” of nucleotide divergence in the mitochondrial genome, where the effective population size is only $\frac{1}{4}$ that of the nuclear genome. Thus, the mitochondrial genome harbors less ancestral polymorphism at recent splits. Instead, mitochondrial nucleotide distance tracks time closely below 5 MYA (slope ≈ 0.77 – 0.9 ; *Supplementary Table S4*), but plateaus at deep divergence times (~7–10 MYA), as expected given its high substitution rate leading to saturation (32). Therefore, the two molecular clocks are best used in tandem to accurately capture the widest range of divergence times.

Contemporary and antique clocks agree on divergence rates

Relative to the nuclear genome, contemporary mitochondrial clocks reveal genomic distances between two- and seven-fold higher across Classes and timepoints (*Supplementary Fig. S2*). Mutation rate differences between nuclear and mitochondrial genomes are known to be highly variable across genes and across taxa (32, 33). On longer timescales, our data are in surprising agreement with the classic finding that vertebrate mtDNA evolves roughly tenfold faster than single copy nuclear DNA. Brown et al. (34) first estimated divergence rates of primate mitochondrial genomes to approximate 2% divergence per nucleotide site by 1 million years. Our 13-gene mitochondrial clock shows a divergence rate per million years of 1.7% in birds and 2.7% in mammals. These results show that our “modern” clock is well aligned with decades of “antique” clocks and that mammals have higher mitochondrial substitution rates than

birds (35). By contrast, our nuclear clocks diverge about an order of magnitude more slowly, reaching similar divergence (~1.8%; 1.7% in birds, 1.9% in mammals, 1.8% in reptiles) by 10 million years.

Nuclear clocks are also taxon-specific

Biologists have long known that mitochondrial clocks tick at different rates in different taxa (32, 34) and thus local clocks are warranted. Our novel nuclear clocks exhibit this same pattern. As expected, the slope and intercept of each nuclear genomic clock also varied among Classes and their Orders, indicating that taxon-specific clocks are most precise. At the deepest divergence times for which our nuclear clock is most accurate, birds show the lowest genome-wide divergence per unit time. This may seem counterintuitive given the short generation times of birds, which allow for more germline replications in the same unit time. However, the low divergence of birds is congruent with the "avian slowdown" phenomenon in which birds evolve more slowly than mammals across nuclear and mitochondrial genomes due to low germline mutation rates and efficient purifying selection (14, 36). Our mitochondrial clocks corroborate this concept; birds exhibit less nucleotide divergence than mammals across lineage divergence times (32), consistent with conservative avian genome evolution that manifests as low germline mutation rates and well-conserved avian karyotypes (27). The Class-specific genomic clocks therefore reflect real biological differences among taxonomic groups and demonstrate that taxonomy-specific clocks capture systematic biological differences among groups rather than a single vertebrate-wide rate of genomic divergence.

Order-level clocks offer additional clarity (*Supplementary* Tables S5 and S6). Within-Order nuclear regressions explained 32–92% of divergence variance, versus 22–41% at the Class level, and their slopes tracked clade biology. In mammals, for example, Rodentia and Primates experience faster and slower genomic divergence, respectively, compared to the Class average (Fig. 1D). In line with this, rodents have short generations and small body sizes, raising substitution rates. By contrast, primates have longer generations and larger bodies (19, 23). The ability of Order-level clocks to reflect known

biology suggests that Family- or Genus-level clocks may further improve resolution in speciose groups.

Traits are inseparable from ancestry as drivers of divergence

We used dyadic phylogenetic mixed models to compare the relative predictive power of nine traits (divergence time, body size, generation time, genome size, clutch size, population density, geographic range overlap, genome N50, and genome GC content) for explaining molecular divergence in bird (278 pairs) and mammal (162 pairs) nuclear genomes. Divergence time was by far the strongest and most consistent predictor of genomic distance (birds: $\beta = 0.26$; mammals: $\beta = 0.35$; 100% posterior probability; Fig. 2). Only two secondary predictors were credible, albeit weakly: assembly N50 in birds ($\beta = 0.08$), likely a technical signal due to better alignment of divergent regions in high-quality assemblies, and population density in mammals ($\beta = 0.17$), a possible proxy for N_e that may reflect inflated nucleotide divergence at shallow timescales due to increased ancestral polymorphism (31). Because population density is notoriously difficult to estimate (37), this result should be interpreted with caution. Importantly, none of our primary traits of interest (body size as a proxy for metabolic rate, generation time, and genome size) were credible predictors of genomic divergence, and the variance partitioning analysis revealed why. Trait variation independent of phylogeny was indistinguishable from zero in both Classes, while phylogenetically structured trait variation accounted for 36.0% (birds) and 57.5% (mammals) of explained variance (Fig. 3A). The phylogenetic PCA corroborated this pattern, assigning body size and generation time to a fast-to-slow life history axis (pPC1) that was positively associated with divergence in both Classes, consistent with faster life histories elevating substitution rates (14, 19; Fig. 3B and C; *Supplementary Table S7*). However, even these axes retained phylogenetic autocorrelation and should be interpreted as descriptive associations.

In the mitochondrial models, divergence time was again the dominant predictor (birds $\beta = 0.36$, mammals $\beta = 0.54$). Unlike in the nuclear models, generation time credibly predicted mammalian mitochondrial divergence ($\beta = -0.31$), in line with the theoretical expectation that short generation times increase per-year mitochondrial mutation rate (32, 34). Traits also explained substantially more mitochondrial than nuclear variance

independent of phylogeny (17% in birds and 19% in mammals; Fig. 3A), and the life-history axis (pPC1) was again credibly associated with nucleotide divergence (Supplementary Fig. S3 and Table S8). The non-recombining mitochondrial genome can therefore resolve trait effects such as generation time separately from shared ancestry, which remains difficult with the nuclear clock. By contrast, the much larger nuclear molecule is buffered against mitochondrial capture and locus-specific effects; thus the two clocks are complementary tools to examine different timescales and evolutionary forces.

Clock outliers as case studies

To demonstrate the utility of our calculated clocks, we examine case studies from our nuclear outlier pairs (Supplementary Tables S9–S12; Fig. 4A). Excess nucleotide divergence over a given timeframe could indicate the presence of cryptic species or other unrecognized discontinuities in a putative gene pool. For example, the mustached bats *Pteronotus mesoamericanus* and *P. parnellii* were significant outliers as gauged by the mammalian nuclear clock (Fig. 4A). This is an independently documented cryptic species complex in which *P. parnellii* is likely to encompass multiple reproductively isolated species with high mitochondrial divergence (38). The identification of this pair as an outlier serves as a positive control for cryptic speciation that can be recovered by our clock. As another example, the primates *Presbytis melalophos* and *P. potenziani* were also significant outliers on the mammalian nuclear clock. These species belong to an island radiation with a long history of taxonomic revisions, including frequent elevations from subspecies to species status (39). As these congeneric species exhibit greater nuclear distances than expected for a given divergence time, our results suggest their TimeTree age may be under-dated. Conversely, nucleotide divergence deficits may indicate over-split taxonomy or recent divergence as exemplified by the parrot genus *Pyrrhura*, which had multiple low-divergence outlier pairs in our avian nuclear clock. The taxonomy of this group has been frequently contested, presumably because of a rapid radiation, incomplete lineage sorting, and introgression (40). Based on low z-scores in our avian nuclear clock, the Genus may be over-split.

We found few species pairs that were outliers with respect to both the nuclear and mitochondrial clocks, likely due to incomplete overlap in the two molecular datasets. However, the howler monkeys *Alouatta juara* and *A. macconnelli* exhibited excess nucleotide divergence in both mammalian clocks. This genus has extensive karyotype variation and within-species mitochondrial structure, congruent with the increased nucleotide divergence we observed (41). Outliers in the mammalian mitochondrial clock were often nucleotide divergence deficits (*Supplementary* Tables S13–S16; Fig. 4B). One viable explanation for low nucleotide divergence among mitochondrial genomes is mitochondrial capture, where one species inherits the mitogenome of another following hybridization. A clear example in our dataset is found in rock lizards *Darevskia armeniaca* and *D. mixta*. The former is a parthenogenetic species formed through hybridization and whose mitogenome was inherited from *D. mixta* (42), resulting in an FDR significant outlier in our reptile mitochondrial clock. Mitochondrial capture through introgression may be responsible for many of the extreme deficit outliers in the mitochondrial dataset including the baboons *Papio anubis* and *P. cynocephalus*, whose mitochondrial lineages are paraphyletic due to introgression (43), and the deer *Odocoileus hemionus* and *O. virginianus*, which have shared mitochondrial haplotypes across a hybrid zone (44). These case studies offer concrete examples of clock outliers representing biological reality rather than simply noise. Because the nuclear and mitochondrial genome evolve differently (e.g., the mitochondrial genome is maternally inherited and non-recombining), identifying outliers using both clocks may help to reduce false positives and diagnose interesting phenomena such as mitochondrial capture.

Implications and future directions

Our results establish genome-wide molecular clocks for tetrapods, recapitulate the idea that mitochondrial genomes evolve 5-10x faster than nuclear genomes, and demonstrate that lineage divergence time is the dominant predictor of genomic sequence divergence. We show that Class and Order-specific clocks are necessary due to inherent differences among OTUs, and that individual life-history effects are largely inseparable from shared ancestry. Mitochondrial clocks have particular utility at recent lineage divergence or coalescent timescales due to more accurate nucleotide divergence estimates. The

mitochondrial clock is more sensitive to generation time (19, 34), while the nuclear clock is more robust at deep time scales and for genome-wide comparisons. Both clocks are informative and complementary when it comes to identifying taxonomic outliers.

We see this work as a proof-of-concept that provides an initial estimate of taxon-specific genomic clocks. Because our dataset only includes genome pairs that properly aligned, meaning highly diverged genomes were disproportionately excluded from the analysis, and because TimeTree divergence time estimates inevitably contain errors, we imagine regular clock updates (i.e., recalibrations) as data availability and accuracy improve. A key advantage of the genome-wide approach is that averaging across the entire genome should buffer clock estimates against locus-specific heterogeneity that impacts single-locus clocks. Whole-genome clocks may become invaluable as rapid and reliable tools to screen for taxonomic disparities (e.g., cryptic species and over- or under-split taxonomies) across the tree of life. Future work should focus on expanding genome-wide clocks to other taxonomic groups, including new Order and Family-specific clocks, to help illuminate interesting branches on the tree of life.

Materials and Methods

Nuclear Divergence Data

Extended methods can be found in the Supplementary Information. For each tetrapod Class (Mammalia, Aves, Amphibia, and Reptilia), we identified congeneric genome assemblies using the NCBI Genomes database, requiring that at least two species per genus had available assemblies and TimeTree divergence time estimates (29). We aligned each possible congeneric pair using minimap2 v. 2.26 (11) and calculated Kimura-2-parameter (k2p) divergence using a custom python script. For each assembly, we recorded genome size (bp), GC content (%), scaffold N50 (bp; contig N50 for unscaffolded assemblies), and repeat content (%) via WindowMasker (45).

Genome-wide clock analyses

We excluded domesticated or taxonomically unreliable species and required both nuclear genomes in each pair to achieve $\geq 80\%$ alignment coverage (this threshold was selected

because pairs below this value showed median alignment percentages of approximately 25–35%; pairs above showed median alignment of 96.9% in the query and 96.4% in the reference). Two overrepresented bird genera (Falco, N = 169; Pyrrhura, N = 53) were capped at 25 pairs per genus; no other Class was affected. To calculate clocks, we modeled the relationship between $\log(k2p)$ and $\log(\text{divergence time})$ for each Class using ordinary least squares regression, producing Class-specific prediction intervals and standardized (z-score) residuals. Cross-Class comparisons used predicted k2p values at standard divergence times (1, 2, 5, 10, and 15 MYA). Outliers were flagged in R v.4.5.2 at $|z| > 1.96$ with Benjamini–Hochberg FDR correction, and we tested for excess outliers using DHARMA simulation-based tests (simulateResiduals/testOutliers). Amphibians were included only in a pooled all-Class nuclear clock due to small sample size. Where Order-level sample sizes were sufficient (≥ 15 pairs), we fit Order-specific regressions, retaining only time-informative clocks ($R^2 \geq 0.28$, the 50th percentile of the Order R^2 distribution; chosen to avoid uninterpretable fits in Orders with narrow divergence time ranges).

Trait associations with nuclear sequence divergence

We compiled body size (kg), generation time (years), clutch size, and population density (individuals/km²) for each species. Sources included COMBINE (46) for mammals and AVONET (47) for birds. Population density gaps were filled from TetraDENSITY (48), and generation time and geographic range overlap estimates were obtained from IUCN. Remaining gaps were filled via literature searches; final sources for each datapoint can be found in Purdue University’s Research Repository (PURR).

To test whether species traits predict nuclear genomic divergence while accounting for shared evolutionary history, we fit dyadic Bayesian phylogenetic mixed models separately for birds and mammals using brms v.2.23.0 (49) with log-transformed k2p divergence as the response. Amphibia and Reptilia were excluded due to insufficient sample sizes, and divergence times were capped at 15 MYA to allow for direct comparisons across Classes. Each species pair was treated as a single dyadic observation. Fixed effects included pairwise trait averages (divergence time, body size, generation time, clutch size, population density, genome size, GC content, and assembly N50) and a

binary geographic overlap term (sympatric: overlap > 0 km²; allopatric: overlap = 0 km²). In the case of one species missing a trait datapoint, the other species was used in the “average” predictors; pairs with both species lacking data were excluded. All continuous predictors were log-transformed where appropriate and z-score standardized. We assessed multicollinearity using variance inflation factors prior to model fitting. We included a dyadic phylogenetic random effect modeled via multi-membership syntax to control for taxonomic structure and shared ancestry. This approach follows de Villemereuil & Nakagawa (50) and controls for pseudoreplication arising from shared ancestry among pairs that share a species (51). We used weakly informative priors (Normal(0, 1) for fixed effects; Exponential(1) for variance components) and verified convergence via $\hat{R}$, effective sample size ratios, and divergent transitions. Models were fit across $N = 100$ posterior phylogenetic trees from VertLife (mammals: DNAonly 4098sp FBD tree set, 52; birds: Hackett Stage 1, 53). Species absent from the phylogenies due to taxonomic revisions were matched with synonyms identified through GBIF. Posterior draws were combined across trees to propagate phylogenetic uncertainty. We assessed phylogenetic signal using Pagel's λ (54) and Blomberg's K (55) via phytools v.2.5.2 (56).

Variance Partitioning

Because life history traits are often phylogenetically structured, variance attributable to traits versus shared ancestry is difficult to parse with a phylogenetic mixed model alone. To examine the relationship between traits and genomic divergence more closely, we partitioned variance following Desdevices et al. (57). For each Class, we fit full (traits + phylogeny), traits-only, and phylogeny-only models on the maximum clade credibility consensus tree and decomposed the Bayesian R^2 (58) into four components: pure trait effects, pure phylogenetic, the shared (phylogenetically structured) trait variation, and unexplained residual. Uncertainty was propagated by combining 4,000 posterior draws. We complemented the variance partitioning with a phylogenetic principal components analysis (pPCA; 59) to summarize covarying traits into composite axes, following Sundaram et al. (60). The pPCA was computed using Abouheif's method in the adephylo v. 1.1.17 R package, retaining the first two global axes and the first local axis. We

recomputed the pPCA on each of the 100 posterior trees, converted species scores to pair-level predictors (pairwise mean), and regressed log(k2p) on the retained axes, divergence time, and sympatry, pooling coefficients across trees with Rubin's rules.

Mitochondrial genome analyses

We assembled the mitochondrial dataset by searching the same pool of candidate congeneric pairs for complete mitogenomes. For each species, we searched for a complete mitogenome from NCBI, prioritizing RefSeq records. Pairs were retained only if a mitogenome was available for both members; the mitochondrial and nuclear datasets therefore overlap but are not identical. We extracted the 13 protein-coding genes, aligned them with MAFFT v.7.525 (61), concatenated the alignments, and computed k2p distance. Pairs with fewer than 8 genes or fewer than 3,000 aligned coding sites were excluded. We replicated all nuclear analyses (clock regressions, mixed models, variance partitioning, and pPCA) on this dataset using identical specifications, substituting mitochondrial GC content for nuclear GC content and excluding genome size and assembly N50, which have no mitochondrial analog.

Acknowledgments

Supported in part by the National Institute of Food and Agriculture, a Fulbright U.S. Scholar fellowship to JAD, by U.S. Fish and Wildlife Service Award F22AF01479, and Texas Parks and Wildlife Department Contract CA-0004020. NMA is supported by a Graduate Research Fellowship with the Rob & Bessie Welder Wildlife Foundation.

Data Availability

All data used in this study were obtained from publicly available sources. Nuclear and mitochondrial genome assemblies were downloaded from the NCBI Genomes database, divergence-time estimates from TimeTree, phylogenetic trees from VertLife, and life-history and trait data from published databases (COMBINE, AVONET, and TetraDENSITY) and the IUCN. A complete list of the specific accessions, source databases, and per-datapoint citations used in our analyses is archived in the Purdue

University Research Repository (PURR) at DOI: 10.4231/GYDX-9711. All code is available at https://github.com/nataliemallen/genomic_divergence.

References

1. E. Zuckerkandl, L. Pauling, “Evolutionary divergence and convergence in proteins” in *Evolving Genes and Proteins*, V. Bryson, H. J. Vogel, Eds. (Academic Press, 1965), pp. 97–166.
2. W. M. Brown, E. M. Prager, A. Wang, A. C. Wilson, Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* 18, 225–239 (1982).
3. A. C. Wilson et al., Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26, 375–400 (1985).
4. P. D. N. Hebert, S. Ratnasingham, J. R. de Waard, Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B Biol. Sci.* 270, S96–S99 (2003).
5. G. C. Johns, J. C. Avise, A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Mol. Biol. Evol.* 15, 1481–1490 (1998).
6. R. Allio, C. Scornavacca, B. Nabholz, A.-L. Clamens, F. A. Sperling, F. L. Condamine, Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail butterfly evolution. *Syst. Biol.* 69, 38–60 (2020).
7. G. E. Sims, S.-R. Jun, G. A. Wu, S.-H. Kim, Whole-genome phylogeny of mammals: Evolutionary information in genic and nongenic regions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17077–17082 (2009).
8. M. dos Reis, P. C. J. Donoghue, Z. Yang, Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* 17, 71–80 (2016).
9. J. B. W. Wolf, H. Ellegren, Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* 18, 87–100 (2017).

- 496 10. J. Armstrong et al., Progressive Cactus is a multiple-genome aligner for the thousand-
497 genome era. *Nature* 587, 246–251 (2020).
- 498 11. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* 34,
499 3094–3100 (2018).
- 500 12. W. S. Moore, Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees
501 versus nuclear-gene trees. *Evolution* 49, 718–726 (1995).
- 502 13. S. Kumar, Molecular clocks: Four decades of evolution. *Nat. Rev. Genet.* 6, 654–662
503 (2005).
- 504 14. L. A. Bergeron et al., Evolution of the germline mutation rate across vertebrates.
505 *Nature* 615, 285–291 (2023).
- 506 15. B. Charlesworth, Effective population size and patterns of molecular evolution and
507 variation. *Nat. Rev. Genet.* 10, 195–205 (2009).
- 508 16. O. S. Harringmeyer, H. E. Hoekstra, Chromosomal inversion polymorphisms shape
509 the genomic landscape of deer mice. *Nat. Ecol. Evol.* 6, 1965–1979 (2022).
- 510 17. E. K. H. Ho, E. S. Bellis, J. Calkins, J. R. Adrion, L. C. Latta IV, S. Schaack, Engines
511 of change: Transposable element mutation rates are high and variable within
512 *Daphnia magna*. *PLoS Genet.* 17, e1009827 (2021).
- 513 18. J. M. de Vos, H. Augustijnen, L. Bäscher, K. Lucek, Speciation through
514 chromosomal fusion and fission in Lepidoptera. *Philos. Trans. R. Soc. Lond. B*
515 *Biol. Sci.* 375, 20190539 (2020).
- 516 19. J. A. Thomas, J. J. Welch, R. Lanfear, L. Bromham, A generation time effect on the
517 rate of molecular evolution in invertebrates. *Mol. Biol. Evol.* 27, 1173–1180
518 (2010).
- 519 20. J. Ma, J. L. Bennetzen, Rapid recent growth and divergence of rice nuclear genomes.
520 *Proc. Natl. Acad. Sci. U.S.A.* 101, 12404–12410 (2004).

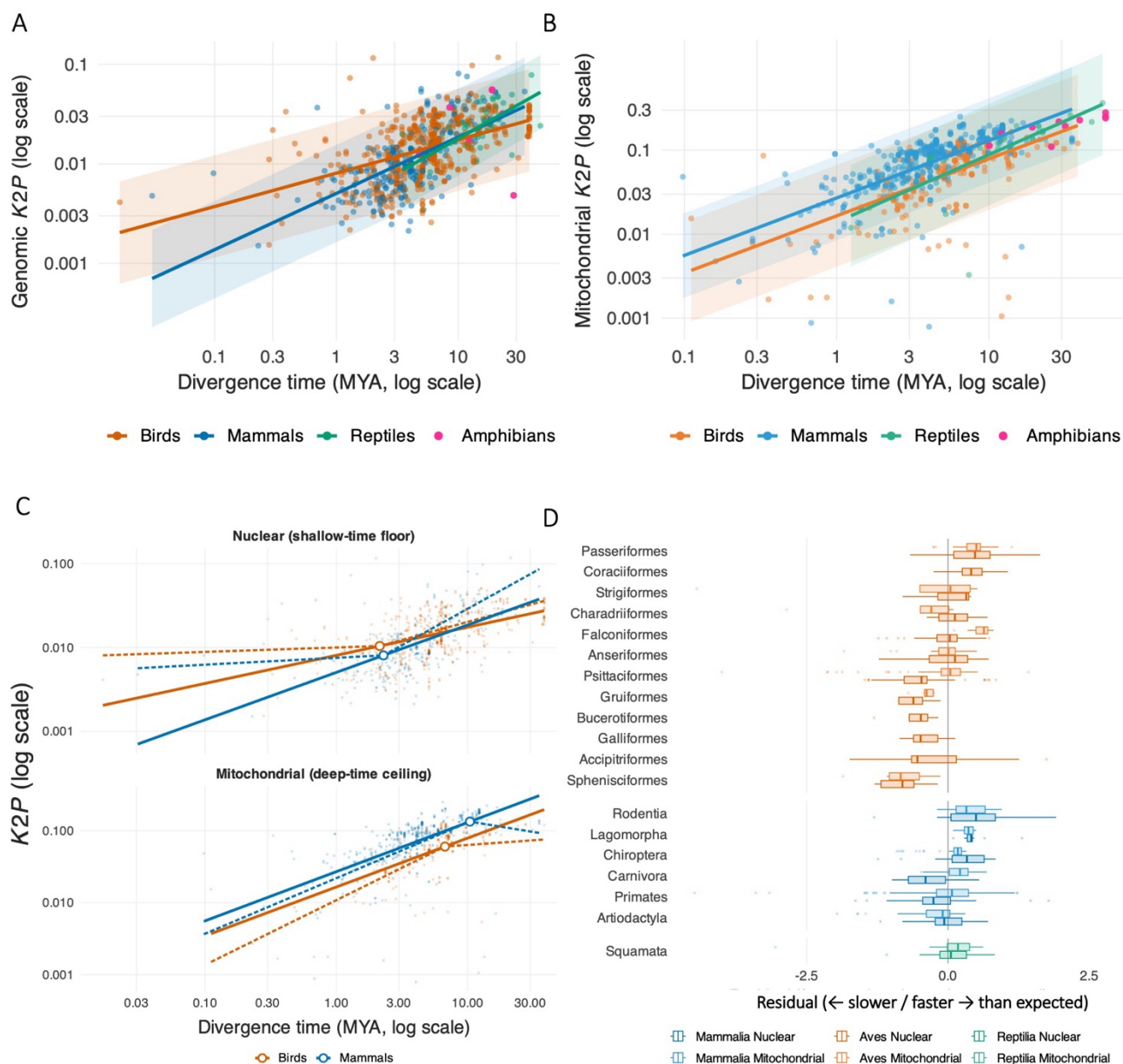
- 521 21. J. H. Nadeau, D. Sankoff, Comparable rates of gene loss and functional divergence
522 after genome duplications early in vertebrate evolution. *Genetics* 147, 1259–1266
523 (1997).
- 524 22. J. F. Gillooly, A. P. Allen, G. B. West, J. H. Brown, The rate of DNA evolution:
525 Effects of body size and temperature on the molecular clock. *Proc. Natl. Acad.*
526 *Sci. U.S.A.* 102, 140–145 (2005).
- 527 23. A. P. Martin, S. R. Palumbi, Body size, metabolic rate, generation time, and the
528 molecular clock. *Proc. Natl. Acad. Sci. U.S.A.* 90, 4087–4091 (1993).
- 529 24. J. M. Kamilar, N. Cooper, Phylogenetic signal in primate behaviour, ecology and life
530 history. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120341 (2013).
- 531 25. A. Rhie et al., Towards complete and error-free genome assemblies of all vertebrate
532 species. *Nature* 592, 737–746 (2021).
- 533 26. H. Ellegren, Evolutionary stasis: The stable chromosomes of birds. *Trends Ecol. Evol.*
534 25, 283–291 (2010).
- 535 27. M. A. Ferguson-Smith, V. Trifonov, Mammalian karyotype evolution. *Nat. Rev.*
536 *Genet.* 8, 950–962 (2007).
- 537 28. M. Kimura, A simple method for estimating evolutionary rates of base substitutions
538 through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120
539 (1980).
- 540 29. S. Kumar, M. Suleski, J. M. Craig, A. E. Kasprowicz, M. Sanderford, M. Li, G.
541 Stecher, S. B. Hedges, TimeTree 5: An expanded resource for species divergence
542 times. *Mol. Biol. Evol.* 39, msac174 (2022).
- 543 30. N. Galtier, B. Nabholz, S. Glémin, G. D. D. Hurst, Mitochondrial DNA as a marker
544 of molecular diversity: A reappraisal. *Mol. Ecol.* 18, 4541–4550 (2009).

- 545 31. S. Y. W. Ho, R. Lanfear, L. Bromham, M. J. Phillips, J. Soubrier, A. G. Rodrigo, A.
546 Cooper, Time-dependent rates of molecular evolution. *Mol. Ecol.* 20, 3087–3101
547 (2011).
- 548 32. B. Nabholz, S. Glémin, N. Galtier, Strong variations of mitochondrial mutation rate
549 across mammals—the longevity hypothesis. *Mol. Biol. Evol.* 25, 120–130 (2008).
- 550 33. R. Allio, S. Donega, N. Galtier, B. Nabholz, Large variation in the ratio of
551 mitochondrial to nuclear mutation rate across animals: Implications for genetic
552 diversity and the use of mitochondrial DNA as a molecular marker. *Mol. Biol.*
553 *Evol.* 34, 2762–2772 (2017).
- 554 34. W. M. Brown, M. George, A. C. Wilson, Rapid evolution of animal mitochondrial
555 DNA. *Proc. Natl. Acad. Sci. U.S.A.* 76, 1967–1971 (1979).
- 556 35. B. Nabholz, S. Glémin, N. Galtier, The erratic mitochondrial clock: Variations of
557 mutation rate, not population size, affect mtDNA diversity across birds and
558 mammals. *BMC Evol. Biol.* 9, 54 (2009).
- 559 36. S. E. Stanley, R. G. Harrison, Cytochrome *b* evolution in birds and mammals: An
560 evaluation of the avian constraint hypothesis. *Mol. Biol. Evol.* 16, 1575–1585
561 (1999).
- 562 37. J. P. Twining, C. McFarlane, D. O’Meara, C. O’Reilly, M. Reyne, W. I.
563 Montgomery, S. Helyar, D. G. Tosh, B. C. Augustine, A comparison of density
564 estimation methods for monitoring marked and unmarked animal populations.
565 *Ecosphere* 13, e4165 (2022).
- 566 38. E. L. Clare, A. M. Adams, A. Z. Maya-Simões, J. L. Eger, P. D. Hebert, M. B.
567 Fenton, Diversification and reproductive isolation: Cryptic species in the only
568 New World high-duty cycle bat, *Pteronotus parnellii*. *BMC Evol. Biol.* 13, 26
569 (2013).
- 570 39. D. Meyer, I. D. Rinaldi, H. Ramlee, D. Perwitasari-Farajallah, J. K. Hodges, C. Roos,
571 Mitochondrial phylogeny of leaf monkeys (genus *Presbytis*, Eschscholtz, 1821)

- 572 with implications for taxonomy and conservation. *Mol. Phylogenet. Evol.* 59,
573 311–319 (2011).
- 574 40. C. C. Ribas, L. Joseph, C. Y. Miyaki, Molecular systematics and patterns of
575 diversification in *Pyrrhura* (Psittacidae), with special reference to the *picta-*
576 *leucotis* complex. *Auk* 123, 660–680 (2006).
- 577 41. L. Cortés-Ortiz, E. Bermingham, C. Rico, E. Rodríguez-Luna, I. Sampaio, M. Ruiz-
578 García, Molecular systematics and biogeography of the Neotropical monkey
579 genus, *Alouatta*. *Mol. Phylogenet. Evol.* 26, 64–81 (2003).
- 580 42. S. Freitas, A. M. Westram, T. Schwander, M. Arakelyan, Ç. Ilgaz, Y. Kumlutas, D. J.
581 Harris, M. A. Carretero, R. K. Butlin, Parthenogenesis in *Darevskia* lizards: A
582 rare outcome of common hybridization, not a common outcome of rare
583 hybridization. *Evolution* 76, 899–914 (2022).
- 584 43. D. Zinner, L. F. Groeneveld, C. Keller, C. Roos, Mitochondrial phylogeography of
585 baboons (*Papio* spp.): Indication for introgressive hybridization? *BMC Evol. Biol.*
586 9, 83 (2009).
- 587 44. S. M. Carr, S. W. Ballinger, J. N. Derr, L. H. Blankenship, J. W. Bickham,
588 Mitochondrial DNA analysis of hybridization between sympatric white-tailed
589 deer and mule deer in west Texas. *Proc. Natl. Acad. Sci. U.S.A.* 83, 9576–9580
590 (1986).
- 591 45. A. Morgulis, E. M. Gertz, A. A. Schäffer, R. Agarwala, WindowMasker: Window-
592 based masker for sequenced genomes. *Bioinformatics* 22, 134–141 (2006).
- 593 46. C. D. Soria, M. Pacifici, M. Di Marco, S. M. Stephen, C. Rondinini, COMBINE: A
594 coalesced mammal database of intrinsic and extrinsic traits. *Ecology* 102, e03344
595 (2021).
- 596 47. J. A. Tobias et al., AVONET: Morphological, ecological and geographical data for all
597 birds. *Ecol. Lett.* 25, 581–597 (2022).

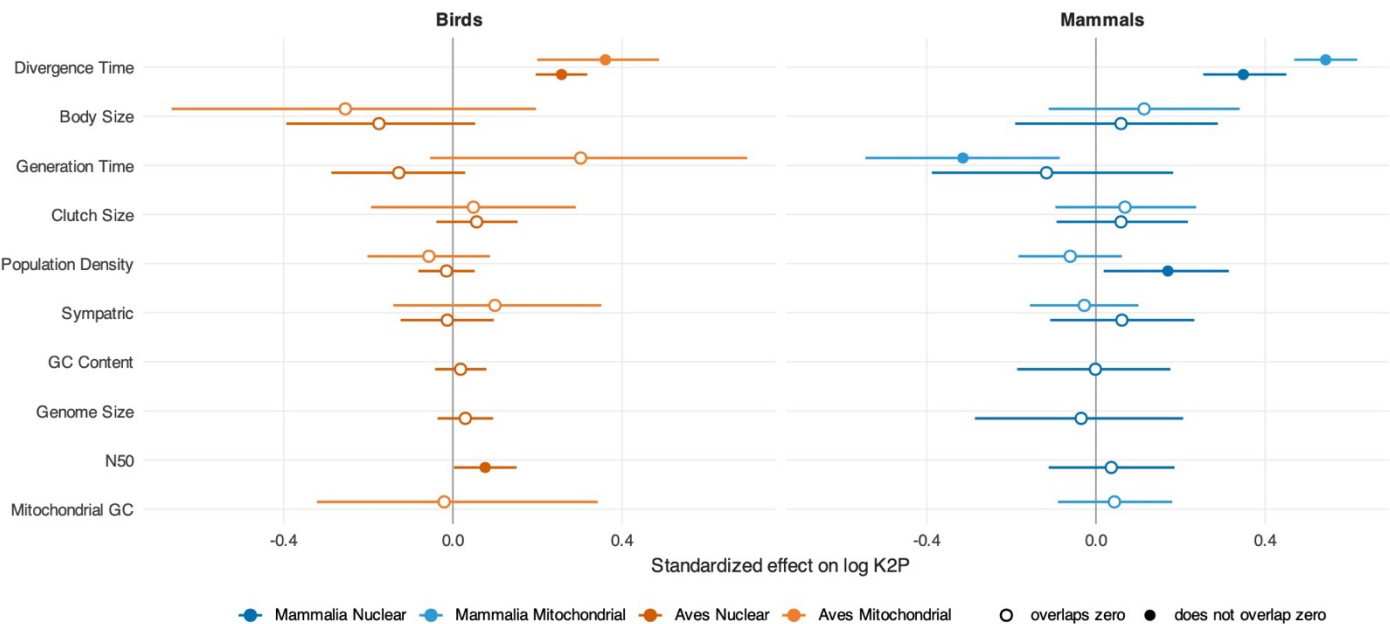
- 598 48. L. Santini, N. J. B. Isaac, G. F. Ficetola, TetraDENSITY: A database of population
599 density estimates in terrestrial vertebrates. *Glob. Ecol. Biogeogr.* 27, 787–791
600 (2018).
- 601 49. P.-C. Bürkner, brms: An R package for Bayesian multilevel models using Stan. *J.*
602 *Stat. Softw.* 80, 1–28 (2017).
- 603 50. P. de Villemereuil, S. Nakagawa, “General quantitative genetic methods for
604 comparative biology” in *Modern Phylogenetic Comparative Methods and Their*
605 *Application in Evolutionary Biology: Concepts and Practice*, L. Z. Garamszegi,
606 Ed. (Springer, 2014), pp. 287–303.
- 607 51. J. D. Hadfield, S. Nakagawa, General quantitative genetic methods for comparative
608 biology: Phylogenies, taxonomies and multi-trait models for continuous and
609 categorical characters. *J. Evol. Biol.* 23, 494–508 (2010).
- 610 52. N. S. Upham, J. A. Esselstyn, W. Jetz, Inferring the mammal tree: Species-level sets
611 of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biol.*
612 17, e3000494 (2019).
- 613 53. W. Jetz, G. H. Thomas, J. B. Joy, K. Hartmann, A. O. Mooers, The global diversity of
614 birds in space and time. *Nature* 491, 444–448 (2012).
- 615 54. M. Pagel, Inferring the historical patterns of biological evolution. *Nature* 401, 877–
616 884 (1999).
- 617 55. S. P. Blomberg, T. Garland, A. R. Ives, Testing for phylogenetic signal in
618 comparative data: Behavioral traits are more labile. *Evolution* 57, 717–745
619 (2003).
- 620 56. L. J. Revell, phytools: An R package for phylogenetic comparative biology (and other
621 things). *Methods Ecol. Evol.* 3, 217–223 (2012).
- 622 57. Y. Desdevises, P. Legendre, L. Azouzi, S. Morand, Quantifying phylogenetically
623 structured environmental variation. *Evolution* 57, 2647–2652 (2003).

- 624 58. A. Gelman, B. Goodrich, J. Gabry, A. Vehtari, R-squared for Bayesian regression
625 models. *Am. Stat.* 73, 307–309 (2019).
- 626 59. T. Jombart, F. Balloux, S. Dray, adephylo: New tools for investigating the
627 phylogenetic signal in biological traits. *Bioinformatics* 26, 1907–1909 (2010).
- 628 60. M. Sundaram, J. R. Willoughby, N. I. Lichti, M. A. Steele, R. K. Swihart,
629 Segregating the effects of seed traits and common ancestry of hardwood trees on
630 eastern gray squirrel foraging decisions. *PLoS One* 10, e0130942 (2015).
- 631 61. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7:
632 Improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780 (2013).
- 633



635
636 **Figure 1.** A genome-wide molecular clock for tetrapod divergence. (A) Genome-wide
637 Kimura-2-parameter distance against divergence time (log scales) for all pairwise
638 comparisons within Aves ($n = 600$), Mammalia ($n = 245$), Reptilia ($n = 44$), and
639 Amphibia ($n = 7$). (B) Mitochondrial genome $k2p$ distance against divergence time (log
640 scales) for all pairwise comparisons within Aves ($n = 282$), Mammalia ($n = 486$), Reptilia
641 ($n = 28$), and Amphibia ($n = 10$). For (A) and (B), points are species pairs, solid lines are
642 the fitted Class regressions, and shaded bands are 95% prediction intervals (Amphibia is

shown as points only due to sample size). Within every Class, genomic distance scales log-linearly with time, but the slope and intercept differ among Classes. (C) Nuclear (top) and mitochondrial (bottom) k2p versus divergence time (log scales) for birds and mammals, the classes with slope breakpoints. Solid lines show the clocks from (A) and (B) whereas dashed lines show the segmented two-slope fit with an open circle at the slope breakpoint. Nuclear divergence flattens below ~2 MYA (a "floor" from ancestral polymorphism); mitochondrial divergence flattens above ~7–10 MYA (a "ceiling" from saturation). (D) Residual genomic divergence (deviation from the Class expectation, in log-k2p units) for each taxonomic Order with ≥ 5 pairs, faceted by Class. Nuclear and Mitochondrial divergence are shown side-by-side in darker and lighter colors, respectively. Values near zero indicate clades that sit on the clock trendline; consistent offsets identify Orders that may diverge faster or slower than their Class average.



657

658 **Figure 2.** Standardized coefficients from the nuclear (darker) and mitochondrial (lighter)
659 dyadic phylogenetic mixed models for birds (left) and mammals (right), estimated per
660 tree and pooled over 100 posterior trees (points = posterior median, lines = 95% credible
661 interval). Filled points denote effects whose 95% interval excludes zero whereas open
662 point intervals include zero. Divergence time is the dominant predictor in all models,
663 whereas individual traits show small effects that largely overlap zero.
664

665

666

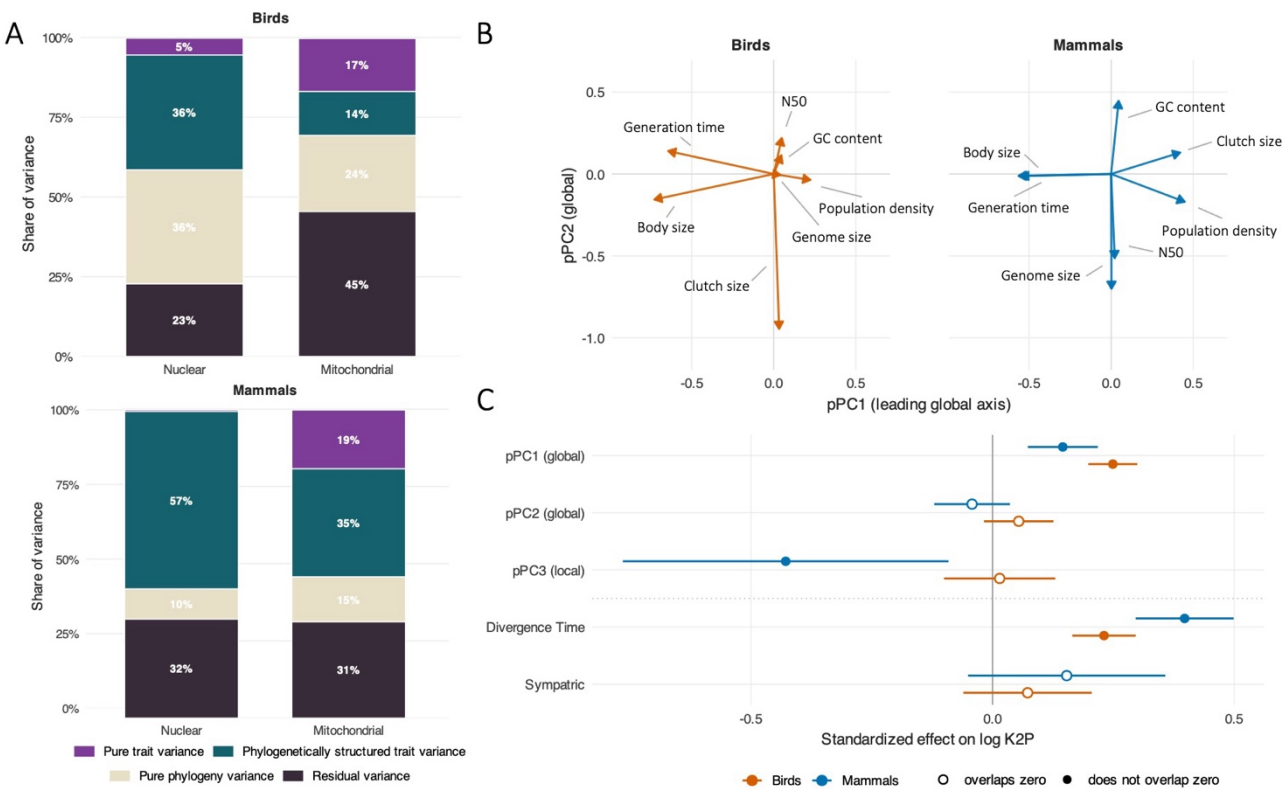


Figure 3. (A) Partition of the explained variance in nuclear and mitochondrial genomic divergence into four components for birds and mammals: pure trait variance (purple), phylogenetically structured trait (i.e. shared) variance (teal), pure phylogeny variance (beige), and residual variance (dark gray). Nuclear genomic divergence variance is dominated by phylogeny and by phylogenetically structured trait variation, not by traits acting independently. (B) Phylogenetic PCA of life-history traits and its relationship to genomic divergence. Trait loadings on the leading phylogenetic principal component axes (pPC1 $\times$ pPC2; global axes from pPCA) shown separately for birds and mammals. Arrows are individual trait contributions. Longer arrows indicate stronger loadings and similar directions indicate traits that covary along the phylogenetically structured axes. The opposite sign of the pPC2 effect between Classes is due to the arbitrary axis orientation. (C) Pooled standardized effects of the retained pPC axes on genomic divergence in the trait-averaged model ($\log K2P \sim pPC1 + pPC2 + pPC3 + \text{divergence time} + \text{sympatry}$). Points indicate the estimate while lines are the 95% interval over 100

681 trees and filled circles are credible. The leading global axis (pPC1) is credibly associated
682 with divergence in both Classes.
683

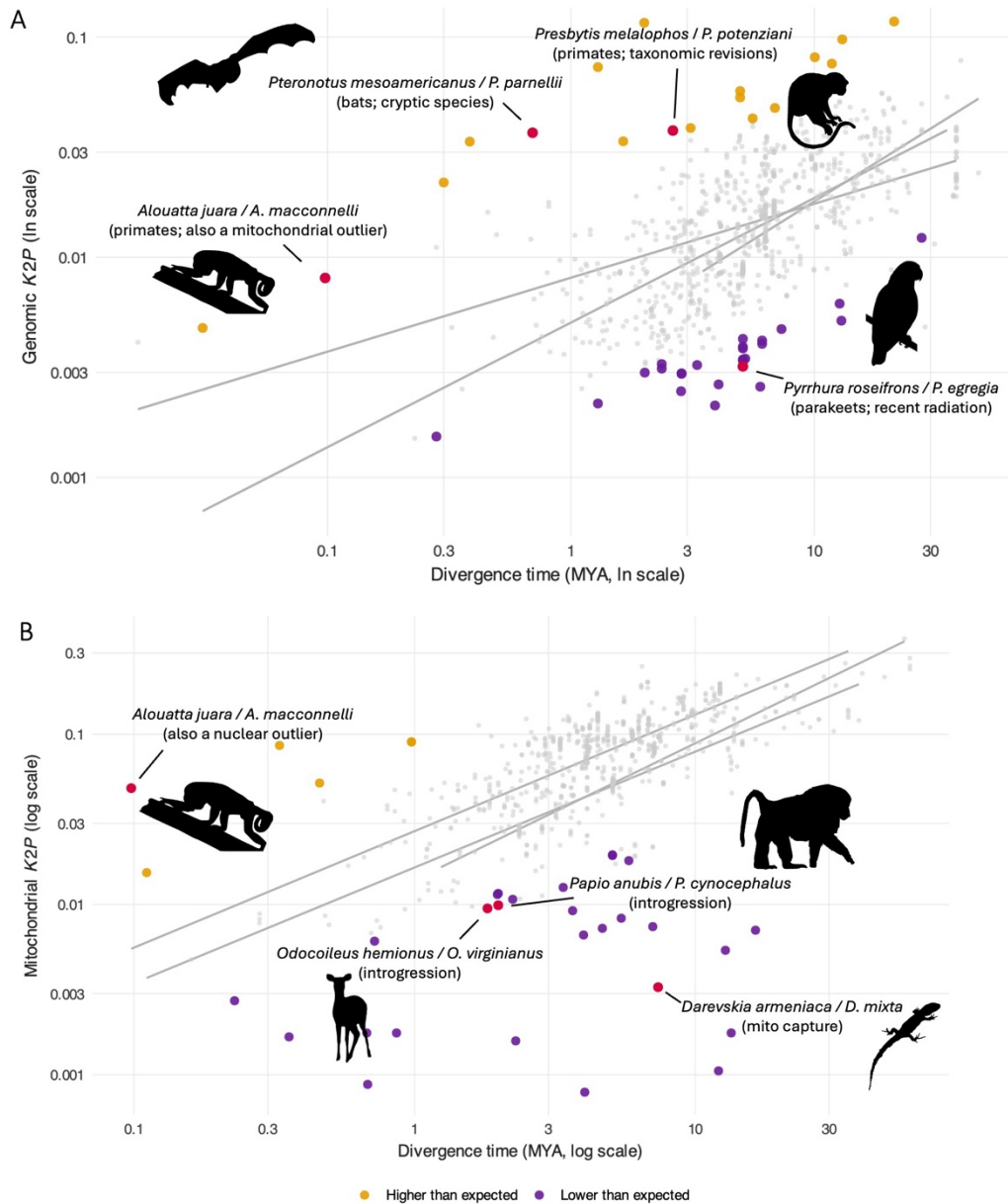


Figure 4. (A) Nuclear clock outliers (clocks as in Fig. 1A). Species pairs consistent with the expectation are shown in gray while those faster-than-average are yellow and slower-than-average are purple. Four outlier case studies are shown in red dots and illustrate the clock's ability to flag biologically meaningful outliers: *Pteronotus mesoamericanus* / *P. parnellii*, whose large genomic divergence is consistent with known cryptic speciation; *Presbytis melalophos* / *P. potenziani*, primates in a genus with frequent taxonomic revisions; *Pyrrhura roseifrons* / *P. egregia*, a recent parakeet radiation with less divergence than expected for its age; and *Alouatta juara* / *A. macconnelli*, primates with

extensive karyotype variation that are also excess outliers in the mitochondrial clock. (B)
Mitochondrial clock outliers (clocks as in Fig. 1B). Unremarkable species are shown in
gray whereas those statistically faster-than-average are yellow and slower-than-average
are purple. Notable outlier case studies are labeled with red dots: *Alouatta juara* / *A.*
macconnelli, which are also nuclear outliers; examples of mitochondrial capture
including lizards *Darevskia armeniaca* / *D. mixta*, baboons *Papio anubis* / *P.*
cynocephalus, and deer *Odocoileus hemionus* / *O. virginianus*.

Table 1. Class-level clock regressions of $\log(k2p)$ genomic distance on $\log(\text{divergence time})$, where $\log$ denotes the natural logarithm. Clock regressions were calculated as $\log(k2p) = a + b \cdot \log(t)$, where t is divergence time in millions of years. For each clock we report the intercept (a), slope (b), residual standard deviation (σ), coefficient of determination (R^2), and the divergence-time range (MYA) over which the clock was calibrated. The pooled model includes Aves, Mammalia, Reptilia, and Amphibia. Outlier pairs can be flagged by computing the standardized residual $z = [\log(k2p) - (a + b \cdot \log t)] / \sigma$. Clocks should be applied only within the calibrated divergence-time range.

Genome	Class	# of pairs (fit)	# of pairs (outlier test)	Intercept	Slope	Residual SD	R^2	Divergence time range (MYA)
Nuclear	Aves	409	600	-4.82	0.336	0.603	0.216	0.02 – 38.27
Nuclear	Mammalia	245	245	-5.29	0.566	0.574	0.413	0.03 – 35.10
Nuclear	Reptilia	44	44	-5.61	0.692	0.438	0.364	3.47 – 47.25
Nuclear	Pooled	705	889	-5.03	0.439	0.599	0.327	0.02 – 47.25
Mitochondrial	Aves	234	282	-4.10	0.680	0.712	0.317	0.11 – 38.27
Mitochondrial	Mammalia	415	486	-3.61	0.684	0.590	0.479	0.10 – 35.10
Mitochondrial	Reptilia	28	28	-4.27	0.803	0.709	0.494	1.24 – 55.68
Mitochondrial	Pooled	687	796	-3.70	0.620	0.675	0.381	0.10 – 58.22