nature ecology & evolution

Article

https://doi.org/10.1038/s41559-026-03110-5

Rates of passerine body plan evolution in time and space

Received: 9 September 2025

Accepted: 8 May 2026

Published online: 20 July 2026

Check for updates

Jacob S. Berv[1,2] ✉, Charlotte M. Probst[1], Santiago Claramunt[3,4], J. Ryan Shipley[5], Matt Friedman[2,6,7], Stephen A. Smith[8], David F. Fouhey[9,10] & Brian C. Weeks[1] ✉

The rates and patterns of phenotypic evolution are known to vary dramatically across time and space, yet the drivers of this variation remain poorly understood. Using over 170,000 skeletal measurements across 2,057 species of Passeriformes—the most diverse order of living birds—we reconstruct the evolutionary history of passerine body plan diversification over the past ~50 Myr. We find that the history of passerine body plan evolution is dominated by rare, hierarchically nested bursts of phenotypic innovation near the origins of major taxonomic groups, indicating that such radiations have played a disproportionately large role in shaping global passerine diversity. These bursts are temporally aligned with intervals of increased climatic instability in Earth's history and are reflected in contemporary latitudinal gradients, where assemblages at higher latitudes host species with elevated mean rates of phenotypic evolution and more modular trait covariation. Together, our findings support the hypothesis that climatic variability may act as a recurring driver of phenotypic diversification, linking deep-time pulses of innovation to broad-scale spatial patterns in evolutionary tempo.

A central challenge in evolution is determining whether large-scale patterns of biodiversity are the result of ecological and evolutionary processes or the legacy of Earth's historical contingencies. Viewed through a macroevolutionary lens, contemporary biodiversity gradients reflect the cumulative impact of repeated radiations and extinctions, modulated by environmental variability[1–3]. These dynamics are exemplified in the fossil record, leading George Gaylord Simpson to propose that much of life's diversification resulted from deep-time 'early bursts' of adaptive radiation[4,5]. During these events, ecological opportunities arising from environmental disturbance or evolutionary innovation can trigger simultaneous divergence of lineages and phenotypes[6]. Just as environmental changes over geological time likely catalysed bursts of phenotypic evolution, spatial variation in ecological opportunity may generate and structure phenotypic disparity across contemporary gradients[7]. At the same time, present-day patterns of phenotypic variation may be shaped by the footprints of extinction and dispersal, particularly in regions affected by glaciation, aridification, or habitat fragmentation[7–11]. Together, these perspectives suggest that environmental variability—acting both as a source of opportunity and as a filter through extinction—may generate recurring patterns of morphological disparity and evolutionary potential across time and space.

[1]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA. [2]Michigan Institute for Data Science and AI in Society, University of Michigan, Ann Arbor, MI, USA. [3]Department of Biological Sciences, University of New Orleans, New Orleans, LA, USA. [4]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada. [5]Department of Fish Ecology and Evolution, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Kastanienbaum, Switzerland. [6]Museum of Paleontology, University of Michigan, Ann Arbor, MI, USA. [7]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA. [8]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA. [9]Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, New York, NY, USA. [10]Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, New York, NY, USA. ✉e-mail: jacob.berv@gmail.com; bcweeks@umich.edu

To test this prediction, we examine variation within the avian bauplan, or body plan, across 2,057 species of passerine birds, using a 12-dimensional dataset summarizing >170,000 linear measurements of skeletal elements[12,13]. These data capture the primary axes of avian body plan variation, including limb proportions, the flight and perching apparatuses, and cranial and sensory structure, reflecting broad patterns of morphological organization (Methods).

Passeriformes, or perching birds, include over 6,500 species distributed across nearly 140 families[14]. Their global radiation spans a remarkable range of ecological niches, morphologies and behaviours, making them a powerful model for investigating how spatial and temporal patterns of morphological evolution are connected across lineages, environments, and geologic time. Although the passerine fossil record is sparse[15–17], recent studies have produced consistent estimates of the group's temporal dynamics[14,18,19], providing a foundation for evaluating the forces shaping their diversification—including ecological opportunity, resource heterogeneity, sexual selection, competition, and climate-driven environmental change[6,20–24]. At the same time, recent analyses of the avian skeleton suggest that passerine skeletal evolution may be conservative, with lower skeletal disparity and simpler appendicular skeletons than those of other bird clades[25,26], creating an apparent paradox between their extraordinary species richness and limited skeletal diversity. Understanding how this large fraction of global bird diversity assembled its phenotypic diversity is central to interpreting passerine evolution and to situating their radiation within the wider landscape of avian body plan diversification.

To characterize body plan evolution, we developed a new approach for inferring histories of evolutionary processes from multivariate phenotypic data (Methods). By operating in a latent space that captures the proportional organization of the skeleton, our method enables inference of macroevolutionary dynamics shaped by changes in developmental and functional constraints[27–30]. Specifically, we (1) extend multivariate evolutionary models with a greedy heuristic search procedure[31–38], (2) apply this framework to infer phenotypic shifts across the passerine phylogeny and (3), evaluate how these shifts align with episodes of climatic instability over the past ~50 Myr. To assess whether deep-time dynamics shape contemporary patterns, we then (4) model how rates of body plan evolution vary across global environmental gradients, with a focus on latitude and temperature seasonality.

Results and discussion

Historical patterns of body plan evolution

Across ~50 Myr of passerine evolution, our focal analysis infers 82 shifts in the mean rate of body plan evolution across the passerine phylogeny (Fig. 1a). Here, 'rate' is the mean of the diagonal variances (σ^2) from the evolutionary variance–covariance matrix (Methods; also Supplementary Fig. 4). Under this definition, the root branch separating oscines and suboscines (Eupasseres) has the fastest inferred rate; it is ~3 times faster than the next fastest rate (the nine-primaried oscines (Emberizoidea and Fringillidae); 'J' in Fig. 1) and almost 900 times faster than the slowest, a subclade of Palearctic emberizid buntings. Most shifts were statistically well supported; in our focal analysis, ~90% had generalized information criterion (GIC) weights >95% (Methods). Of all inferred shifts, only 18 (~20%) represent increases in rate (modal rate increase of ~130%), while all other shifts represent rate decreases (modal rate decrease of ~64%). Across multiple analyses, the magnitudes of inferred rate increases and decreases differed; a typical rate increase was ~2.75 times greater than a typical rate decrease ($D = 0.24$, $P < 0.0001$; Supplementary Fig. 5).

The distribution of inferred mean rates across all lineages (Fig. 1a) is best described by a Gumbel distribution (maximum likelihood estimate (MLE): $\mu = -8.84$, $\sigma = 0.83$; Kullback–Leibler divergence of ~0.03). Used in extreme value theory to model the distribution of rare, high-magnitude events[39], the Gumbel distribution is consistent with an evolutionary process shaped by infrequent but consequential shifts in rate, as expected under models of pulsed macroevolution[40–42]. Considering each lineage separately, accelerations are separated by a median waiting time of 20.5 Myr (or ~1 acceleration per 0.9 billion years of summed branch length), broadly consistent with ref. 41. For all shifts, 86% of tip-to-root paths have waiting times among shift events best fit by constant-rate exponential models (MLE: $\lambda \approx 0.2$, mean ≈ 5 Myr, 95% highest density interval (HDI) = 2.5–12.6 Myr), consistent with memoryless dynamics. Across the entire phylogeny, waiting times are best fit by a Weibull distribution (MLE: $k = 0.77$, $\lambda = 0.45$), indicating that the instantaneous probability of a shift occurring anywhere in the tree is highest immediately after a previous shift. Shifts are therefore temporally clustered and define episodic bursts (Figs. 1a and 2).

Together, these dynamics have produced a pattern of hierarchically nested bursts, in which numerous clade originations coincide with dramatic increases in the average rate of body plan evolution (Fig. 1a). Theoretical models suggest that such reorganizations, particularly those that increase modularity and reduce pleiotropy, can enhance evolvability by enabling modular responses to selection while preserving overall organismal integration[30,43]. Researchers have previously proposed that these reorganizations underlie the origins of higher taxa, where early bursts of disparity reflect rapid shifts in trait integration rather than gradual modification of individual features[29,44]. Post hoc analyses of regime-specific phenotypic correlation structure (Supplementary Fig. 4a–d) support this interpretation and show that inferred shifts align with changes in overall patterns of phenotypic integration. Specifically, shifts fall along a dominant axis that partitions a distal wing/flight module from hindlimb and cranial traits, suggesting variable coupling between the flight apparatus and terrestrial locomotion/feeding structures[45] (Supplementary Text).

Consistent with this interpretation, rate increases in passerines are concentrated near the origins of clades historically recognized for their phenotypic distinctiveness (Fig. 1a and Supplementary Table 7), supporting the idea that evolutionary re-patterning of body plan architecture enables access to distinct regions of ecomorphospace[44,46]. Viewed alongside pan-avian analyses that portray passerines as morphologically conservative[25], our results show that passerine skeletal diversity has been assembled through striking dynamism within their occupied region of morphospace, via rare but consequential changes in body plan architecture.

Passerine body plan evolution thus aligns with features of classic macroevolutionary models: repeated accelerations followed by slowdowns resemble punctuated equilibrium and evolutionary constraint[47]; the concentration of accelerations at clade origins is consistent with Simpsonian quantum evolution[5]; and the alternation of rapid and gradual change resembles punctuated gradualism[48]. Similar to ref. 49, we find that body plan diversification follows a process in which rare accelerations punctuate a background diffusion process. Importantly, our results show that classic macroevolutionary patterns, largely derived from the fossil record, can be observed even when phenotypic data are available only for extant taxa. Despite our dense sampling (2,057 species spanning ~89% of passerine families), many recently diverged species remain unsampled, so the overall pattern of variation should be interpreted as a lower bound on the complexity of passerine body plan diversification (see the sensitivity analysis in the Supplementary Information).

More broadly, the burst–slowdown dynamics inferred here are compatible with the view that some avian adaptive radiations may slow as niche space fills. In the broader avian literature, classic island and regional radiations such as Darwin's finches, Hawaiian honeycreepers, Madagascan vangas, Himalayan songbirds and Neotropical tanagers show either early pulses of morphological or lineage diversification, or patterns consistent with diversification constrained by niche filling and ecological saturation[20,23,50–54]. Still, other studies examining a wide range of species have not detected early bursts or diversification slowdowns[55–58]. These differences may reflect real variation in

https://doi.org/10.1038/s41559-026-03110-5

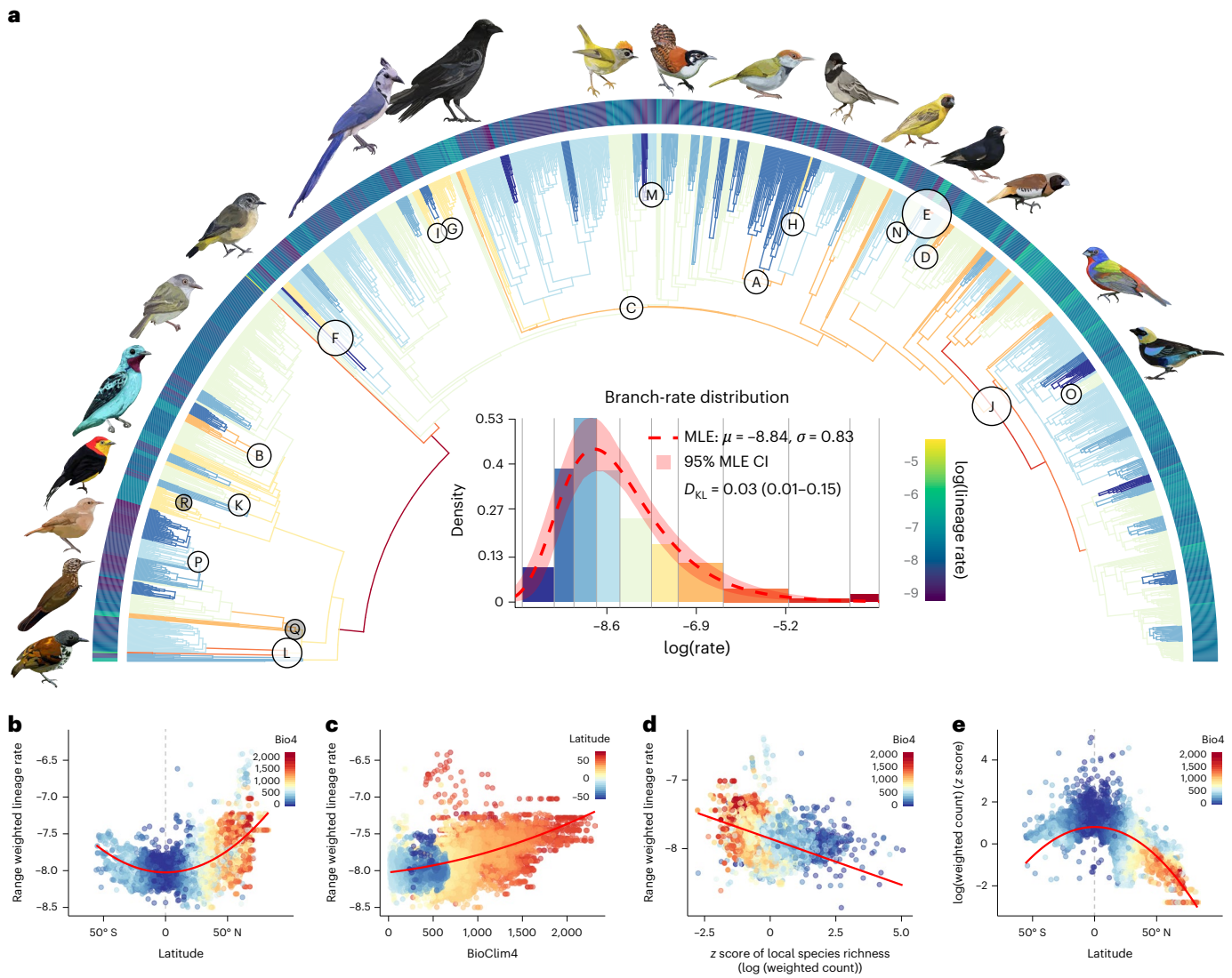

Fig. 1 | Phylogenetic and spatial patterns of passerine body plan evolution. **a**, Passerine phylogeny[18] showing the results of our focal analysis (Methods). We infer 82 shifts in the process of multivariate phenotypic evolution (summarized as rate variation). Letters at highlighted clades (white circles) indicate nodes where we observe multivariate rate increases (18 total), while all other branch colour shifts show inferred rate decreases (64 total). Of these highlighted clades, 16 have GIC weights of ~100%, whereas two (R, Q; in grey) have GIC weights <90%. These rate increases align with recognized taxa, for example, Passerida (C), Nine-Primaried Oscines (J), Furnariida (Q, L), Corvinae (G), Cyanocoracinae (I), the 'core' Cotingas (K), Acanthizidae (F) and others (Supplementary Table 7). Circle sizes are proportional to the percentage rate increases on those nodes. MLE of a Gumbel distribution, fit to inferred mean branch rates, with Kullback–Leibler divergence (D_{KL}) of 0.03, indicating a close fit to the empirical distribution of rates. Bars at the tree tips summarize lineage rates (inset legend, Methods). **b**–**d**, Associations between range-weighted mean lineage rates and latitude (**b**), between range-weighted mean lineage rates and temperature seasonality (BioClim4, °C × 100) (**c**) and between range-weighted mean lineage rates and species richness (**d**). **e**, The association between local species richness and latitude, peaking at the equator. Scatterplot point colours correspond to seasonality or latitude values. Simplified quadratic or linear curves overlaid for reference. Credit: Illustrations of representative taxa from labelled clades (not to scale) in **a** by Olivia Stein; scaled graphics provided in Supplementary Fig. 18.

evolutionary dynamics or, alternatively, may arise from methodological challenges, including differences in scale, clade selection, or model adequacy[59,60]. Traditional comparative approaches often assume a homogeneous evolutionary process across lineages[61] and rely on ordination[62], thereby limiting their sensitivity. Our approach addresses these challenges by providing a framework that can identify shifts in phenotypic diversification across multidimensional trait spaces with fewer assumptions.

When compared to a continuous paleoclimate proxy for mean global temperature (benthic $\delta^{18}O$; ref. 63), we find that periods of increased climatic variability are aligned with increased global frequencies of inferred model shifts (Fig. 2). This pattern is clearest during the Eocene–Oligocene transition (EOT, ~34 Ma) and the Mid-Miocene climatic optimum/transition (MMCO/MMCT, ~15 Ma)—intervals characterized by dramatic global cooling and warming, respectively[63–66]. While the overall frequency of shifts appears elevated across both intervals, their directionality differs: rate increases are relatively more frequent near the EOT, whereas rate decreases peak closer to the MMCO/MMCT (Fig. 2).

The EOT was characterized by global restructuring of ecosystems, including, in many regions, the contraction of warm, closed-canopy forests, the expansion or reorganization of more open, seasonal habitats (for example, temperate woodlands, shrublands, and grassland-dominated systems), and shifts in marine productivity[65]. Such environmental changes would have displaced incumbent lineages and created new ecological opportunities—a dynamic long proposed as a catalyst

https://doi.org/10.1038/s41559-026-03110-5

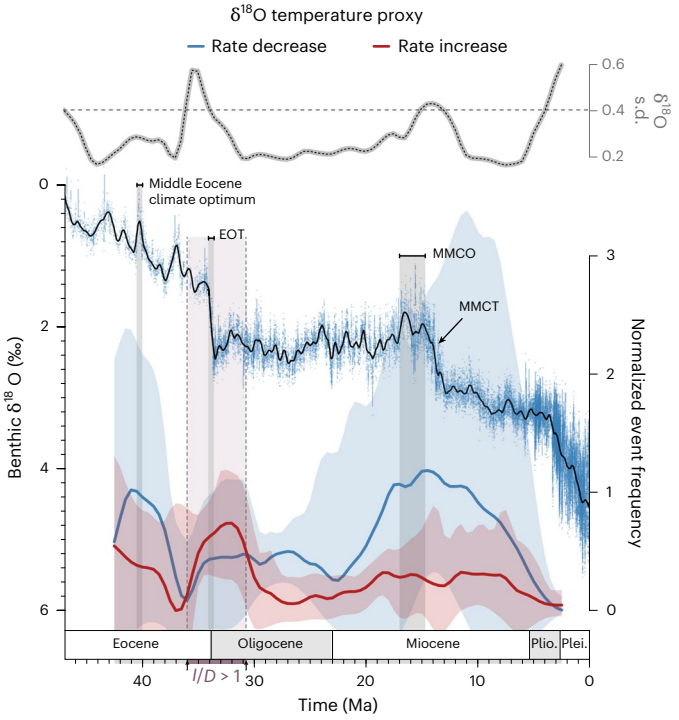

Fig. 2 | Climatic variability and temporal clustering of evolutionary rate shifts. Black curve: Benthic $\delta^{18}O$ global temperature proxy[63] across the Cenozoic. Top: we overlay a curve (grey) showing a moving average of Benthic $\delta^{18}O$ standard deviation in a 5 Myr sliding window. Maximum temperature variability peaks at the EOT and the MMCO/MMCT. Below: we summarize the normalized frequency of inferred model shifts across the history of Passerines using a 5 Myr sliding window (Methods); these curves are separated into inferred increases and inferred decreases in multivariate rates. The frequency of increases (I) peaks near the EOT and exceeds that of decreases (D) in one interval across the EOT boundary (~36–31 Myr, I/D > 1). By contrast, rate decreases peak near the MMCO/MMCT. Shaded bands represent the 90% density interval for shift frequency detection across multiple analyses (Methods). Because the heuristic search is conditional on a minimum clade size of 10, very recent shift frequencies are subject to a soft temporal horizon near the present.

for adaptive radiation[4,6,9,67]. These patterns are consistent with the hypothesis that cooling-associated habitat turnover at the EOT generated ecological opportunity and promoted accelerated body plan evolution[67]. In some cases, this broader interval of ecological opportunity may have been reinforced by lineage-specific biogeographic expansion. For example, the acceleration at Passerida (C in Fig. 1a) coincides with the expansion of this radiation beyond Australasia (the temporally later acceleration at nine-primaried oscines (J in Fig. 1a) may likewise reflect diversification associated with colonization of new regions in the New World[68])[14]. Other inferred accelerations of similar age, including Furnariida (L and Q in Fig. 1a), do not coincide with comparably clear biogeographic transitions, suggesting that the broader clustering of rate increases around the EOT is more consistent with a shared influence of climatic instability. By the MMCO/MMCT, when many passerine lineages were relatively mature and had already diversified extensively within these environments, decelerations in the rate of body plan evolution appear more common, as would be expected under increasing ecological saturation and stabilization[67]. Our observations of nested phenotypic bursts near the origins of major clades, and the broader pattern of accelerations around the EOT followed by decelerations near the MMCO/MMCT, point to a common burst–slowdown dynamic repeated across scales, from subclade-level body plan evolution to the temporal trajectory of the radiation as a whole.

Temperature variability links body plan evolution across time and space

To better understand potential linkages between climatic instability and phenotypic evolution, we modelled contemporary spatial variation, leveraging the high resolution of available climate data to test whether recent evolutionary patterns recapitulate those inferred across time. Latitudinal diversity gradients are among the most striking patterns in nature; competing hypotheses attribute these gradients to variation in speciation and extinction rates, ecological and energetic constraints and asymmetric dispersal across latitudes[3,69–72]. Range-weighted passerine species richness tracks this pattern, peaking near the equator and declining toward the poles (Fig. 1e). While some studies report increased lineage diversification rates toward the poles[2,70,71,73–76], researchers have not yet determined whether patterns of phenotypic diversification broadly mirror gradients of lineage diversification[7,77–79]. Although diversification and phenotypic evolution may be linked[60], they can also vary independently[80]. Such decoupling can arise through cryptic speciation[81], non-ecological speciation[82], drift in geographically structured populations[83] or bursts of morphological divergence without speciation[84]. More broadly, it is unclear whether contemporary biodiversity gradients and deep-time evolutionary dynamics share a common underlying cause.

We find a strong signal of elevated mean rates of multivariate body plan evolution at higher latitudes (Fig. 1b). We focus on a useful summary statistic—grid-cell mean rates that account for both recent lineage effects and variation in range sizes (Fig. 1a and Methods). A parsimonious spatial autoregressive lag (SAR-Lag) model including a quadratic term for latitude and a linear control for species sampling confirms a strong statistical signal of increasing mean phenotypic rates with latitude (total effect latitude z-score = 2.78, $P \approx 0.005$; total effect sampling fraction z-score = 1.6, $P \approx 0.1$; Supplementary Fig. 13 and Supplementary Table 4). In this and other SAR-Lag models, Moran's I, a statistic of spatial autocorrelation, was significantly reduced in model residuals (for example, an ~84–86% reduction compared to ordinary least squares (OLS)), indicating that our results are robust to the majority of spatial autocorrelation in these data (see Fig. 3c). Additional SAR-Lag models incorporating spatial variation in temperature seasonality (positive effect; Supplementary Fig. 15 and Supplementary Table 6) and local species richness (negative effect; Supplementary Fig. 14 and Supplementary Table 5) confirm these factors as significant covariates of spatial variation in mean rates (Fig. 3).

These results show that increased temperature seasonality across space correlates with increased mean rates of body plan evolution (Figs. 1c and 3b, Supplementary Fig. 15 and Supplementary Table 6), corroborating patterns we find across time (total effects: linear effect z-score = 3.17, $P \approx 0.0015$; quadratic effect z-score = 3.53, $P \approx 0.0004$, after accounting for latitude, interaction effects, and species sampling). Mean phenotypic rates are highest where local species richness is lowest (Fig. 1d, Supplementary Fig. 14 and Supplementary Table 5), suggestive of an effect of ecological interactions and/or regional variation in diversification patterns (total linear effect z-score = −2.1, $P \approx 0.03$, after accounting for latitude, interaction effects, and sampling fraction). A more parameter-rich, aggregated SAR-Lag model included local species richness, temperature seasonality, latitude, interactions between latitude and the two focal predictors, and controls for speciation rate (λ_{DR}) and sampling fraction (19 parameters) (Figs. 1b and 3a–c and Supplementary Table 3). This aggregate model recapitulated the primary trends in the empirical data, and suggests that a small number of parameters explain much of the variation in recent mean phenotypic rates across >10,000 grid cells (Nagelkerke pseudo-R^2 = 0.74); with temperature seasonality still significant after accounting for the other predictors (total effects: linear effect z-score = 2.45, $P \approx 0.01$, quadratic effect z-score= 4.8, $P \approx 1.6 \times 10^{-6}$). Sampling intensity itself was not a significant predictor of spatial variation in mean phenotypic rates (for example,

https://doi.org/10.1038/s41559-026-03110-5

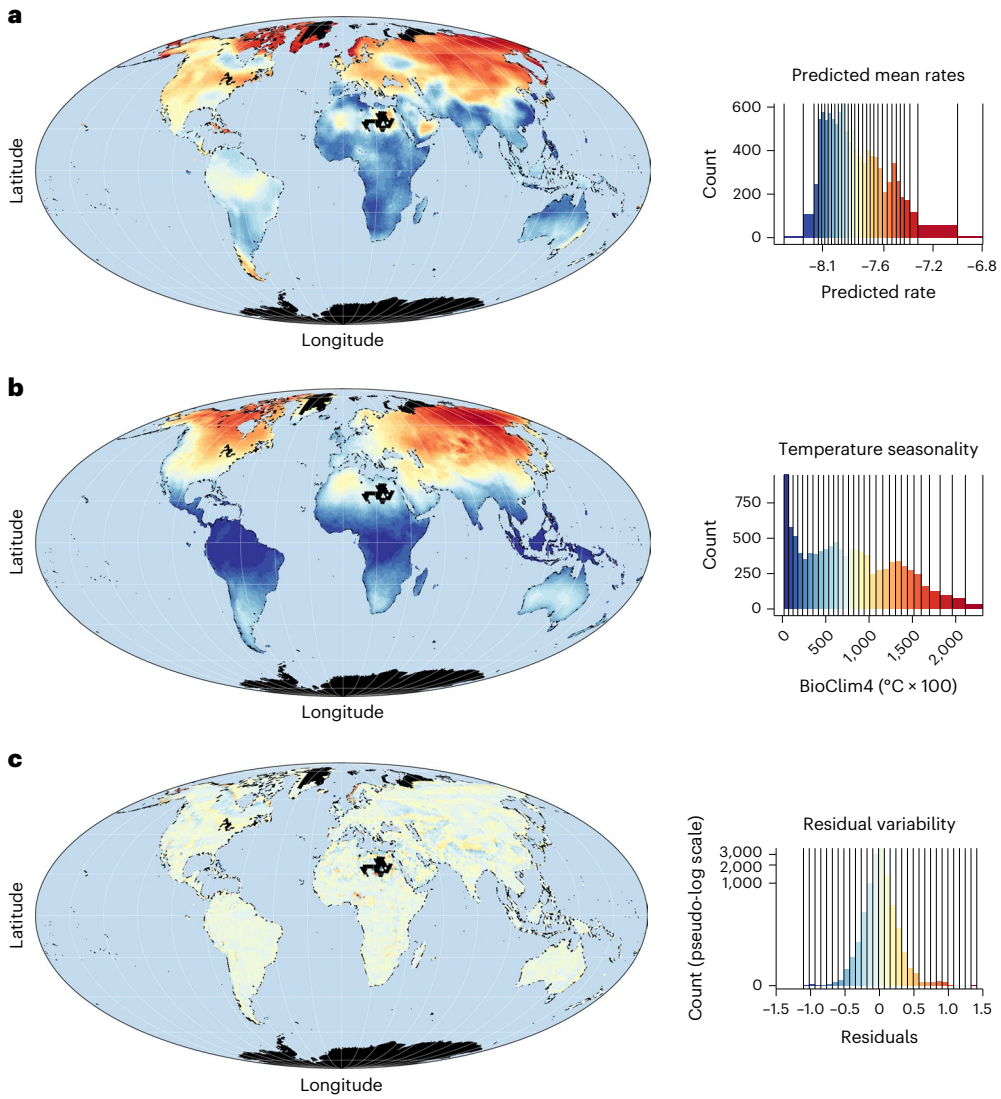

Fig. 3 | Spatial predictors of phenotypic evolution across global assemblages. **a**, Predicted grid-cell mean multivariate phenotypic rates (inverse range weighted; Methods), from an aggregated SAR-Lag model including key predictors: temperature seasonality (BioClim4, °C × 100), species richness, and latitude, as well as additional controls for sampling and speciation rates (Methods). Estimated (Fig. 1a and Supplementary Figs. 8 and 9) and predicted mean rates are typically higher at higher latitudes and are well approximated by a simple quadratic relationship with latitude (Supplementary Fig. 13). **b**, Temperature seasonality (BioClim4) shows spatial patterns similar to **a**. **c**, Aggregated SAR-Lag model (in **a**) residuals projected across space, showing minimal residual structure after accounting for spatial autocorrelation.

Supplementary Fig. 13b). In Fig. 3c, residuals are small relative to the fitted model: predicted rates (Fig. 3a) vary by ~3- to 4-fold, but 95% of residuals lie within ±0.25 ln units (~25%) of the predictions, in line with the pseudo-R^2 estimate of 0.74.

In sum, these analyses reveal a global pattern of body plan evolution: high-latitude and seasonal environments tend to host assemblages with elevated mean rates of body plan evolution, a feature often associated with increased evolvability[30,43,85]. These phenotypic patterns broadly parallel observed increases in lineage diversification rates at higher latitudes[2,70,71,73–76] and faster body-shape evolution in fishes[79]. Means can obscure variation within assemblages, so we also summarized lineage mixtures within assemblages and found that our assemblage-level mean rates are well correlated with measures of compositional enrichment for globally fast or slow lineages (Supplementary Fig. 17a–h). Overall, these patterns are consistent with the hypothesis that climatic variability has contributed to phenotypic innovation both across ~50 Myr of passerine evolution and in more recent history, reflected in strong relationships among seasonal temperature variation, latitude, and recent rates of phenotypic evolution. While deep-time climatic instability and modern temperature seasonality operate on very different timescales, both represent large-amplitude thermal variation; we therefore treat them as conceptually related axes of environmental fluctuation.

Although these spatial signals may suggest local adaptive evolution, we considered several nonadaptive explanations (Methods). First, differential extinction may bias mean rates: high turnover in unstable regions could filter out slow-evolving lineages (for example, ref. 86), while lower extinction risk in the tropics may preserve lineages shaped by slower niche-partitioning dynamics. Although we view this as a plausible explanation, a recent analysis did not support the hypothesis that avian extinction rates are systematically higher at higher latitudes[18]. Alternatively, elevated rates might reflect genetic drift during population bottlenecks associated with climatic stress or reduced population density[87], where reduced effective population sizes, N_e, are expected to increase phenotypic variation in proportion to heritability $-h^2-(\sigma^2 \sim h^2/N_e)$[32].

https://doi.org/10.1038/s41559-026-03110-5

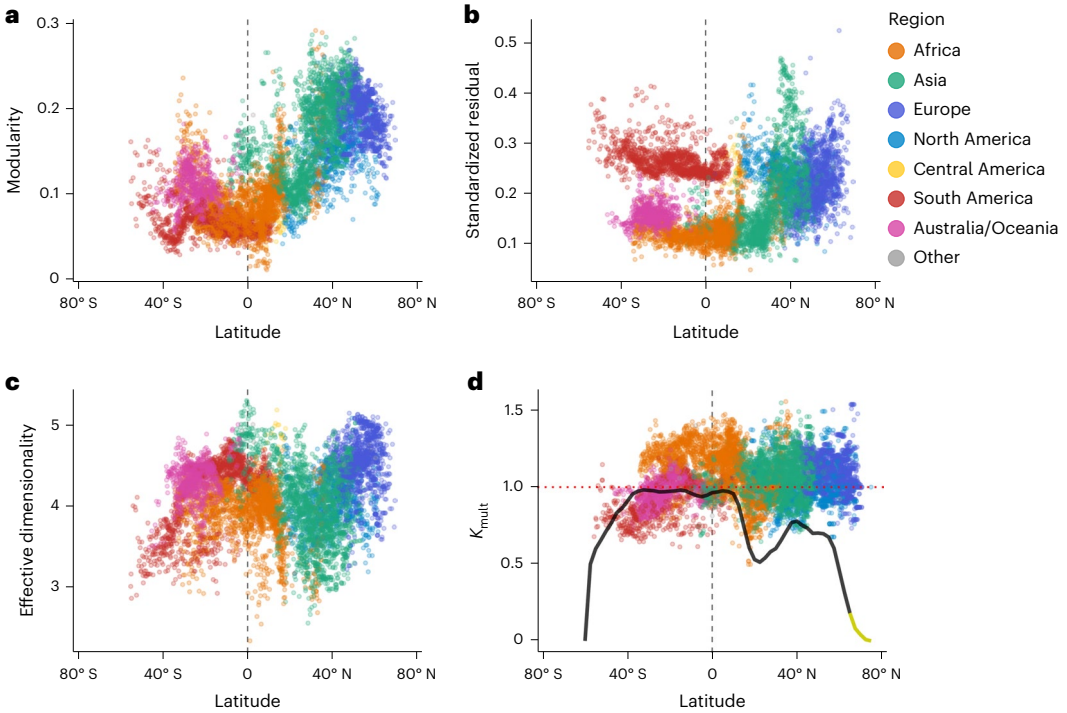

Fig. 4 | Latitudinal variation in phenotypic structure. Mid-to-high latitude assemblages, characterized by greater temperature seasonality, show faster, more modular, and less phylogenetically constrained patterns of phenotypic variation. **a**, Local estimates of modularity increase with increasing latitude, similar to patterns of mean multivariate rates (Fig. 1b). **b**, Standardized phenotypic residuals across grid cells suggest higher-latitude assemblages host lineages with phenotypic dispersion greater than expected given their species' divergence times, with strong biogeographic effects primarily separating New World (North and South America) and Old World (Africa, Europe, Asia and Australasia) assemblages. **c**, Effective dimensionality, approximating the number of orthogonal axes of variation, is relatively stable across latitude (3–5 axes). **d**, K_{mult}, a descriptive metric of multivariate phylogenetic signal within each grid cell. We overlay a curve (black) showing the proportion of grid cells with significant departures from phylogenetic independence (10° sliding window). This proportion approaches 0 beyond ±40°, suggesting a weakening of evidence for phylogenetic structuring of trait variation; the yellow segment of the curve denotes cells with <10 species (approximately >68° N).

To investigate these possibilities, we independently assessed assemblage-level phenotypic variance, dispersion, and dimensionality (Methods)—aspects of variation that we expect may be eroded by extinction[88–90] or potentially inflated by drift in small populations[32]. We also estimated the modularity of local trait covariance matrices using Louvain community detection[91], which quantifies the degree of phenotypic compartmentalization and may be shaped by ecological, drift or extinction processes[92–94].

Standardized mean phenotypic residuals increase sharply at higher latitudes (Methods), suggesting disproportionate phenotypic divergence relative to evolutionary distances (Fig. 4b). Continental bioregions appear to strongly influence these patterns; New World lineages show greater local phenotypic dispersion than Old World localities. Estimates of local effective dimensionality (Fig. 4c) and total variance (Supplementary Fig. 16) are relatively stable across latitudes. Estimates of modularity within local assemblages also increase at mid-latitudes, suggesting greater clustering of trait covariation, consistent with regime-level patterns (Supplementary Fig. 4). Further, the proportion of grid cells showing significant departures from phylogenetic independence declines around ±40° latitude, consistent with reduced phylogenetic structuring of phenotypic variation at high latitudes (Fig. 4d). This weakening of evidence for phylogenetic signal may align with patterns of phenotypic overdispersion and modular organization observed at high latitudes, although some of these declines may reflect statistical limitations in smaller species pools.

These patterns are difficult to reconcile with simple drift or extinction-filtering scenarios, suggesting more complex dynamics. At higher latitudes, species assemblages have disproportionate phenotypic disparity, stable dimensionality and variance, and stronger modular organization, features that neither drift nor extinction alone are expected to produce. While drift can increase phenotypic variance in small populations, it is unlikely to generate or preserve structured trait integration[43,94,95]. Extinction can erode morphological diversity[88] and, under some conditions, favour modular architectures that promote resilience[93–95], but the stable variance and dimensionality in our data argue against extinction filtering as the sole driver of the patterns we observe[18]. The spatial convergence of high mean rates, high modularity of trait covariation, and overdispersed traits, aligns with theoretical expectations for evolvable systems, suggesting that environmental instability fosters the enrichment of phenotypically labile lineages and/or selects for their persistence.

Finally, while we have attempted to account for assemblage structure (for example, via range-weighting, richness, and sampling covariates), our framework does not explicitly separate in situ evolutionary change from the effects of dispersal and local extirpation, particularly in species-poor cells at extreme latitudes. In these cases, quantities such as modularity and local trait disparity may be influenced by small species pools and noisier estimates of phenotypic structure. Metrics that summarize mean lineage or trait histories can also, in principle, generate spatial patterns even under constant diversification, if range shifts and local turnover are structured in space; our approach does not model these dynamics directly. We therefore view our spatial analyses as approximations of recent evolutionary dynamics in local assemblages. Spatially explicit models and simulations that jointly track range dynamics, speciation and trait evolution will be valuable for further disentangling these effects.

Conclusion

By linking the temporal and spatial dimensions of passerine bauplan evolution, our findings support the hypothesis that climatic variability structures the exploration of ecomorphospace[67]. Just as deep-time episodes of environmental disruption coincide with changes in patterns of phenotypic evolution, contemporary high-latitude or highly seasonal environments show similar patterns and host assemblages with elevated mean rates and trait architectures capable of generating variation under fluctuating conditions[95–97]. Together with our finding that rapidly evolving regimes tend to show reduced phenotypic integration and distinctive axes of covariance structure (Supplementary Fig. 4a–d), these patterns suggest that modular reorganization is a key pathway by which climatic variability is translated into morphological diversification in passerines. Environmental variability may thus act as a common cause of both temporal bursts of phenotypic change and spatial gradients in evolutionary tempo.

Methods

Empirical dataset and phylogenetic framework

Our study relies on a large phenotypic dataset described in ref. 12. In brief, we developed and applied a computer vision pipeline (Skelevision)[13] that leverages U-Net[98] and Mask R-CNN[99] approaches to measure traits across thousands of images of avian skeletal specimens. This approach identifies and measures the longest linear dimension of bones in the legs, torso, wings, hand, and skull with low error[13]. The initial Skelevision dataset spans 2,057 species (Aves: Passeriformes) and includes data from 12 skeletal elements across 14,419 voucher specimens[12]. These traits comprise linear measurements of major hindlimb (femur, tibiotarsus, tarsometatarsus), forelimb (humerus, ulna, radius, carpometacarpus, second digit first phalanx), and axial/cranial (keel, furcula, sclerotic ring diameter, skull–bill length) elements. The species in our dataset encompass 34% of passerine species and 89% of passerine families[56], enabling us to examine a wide taxonomic sample across the most diverse avian order. Acanthisittidae (New Zealand wrens), sister to other passerines, were not included because suitable skeletal specimens were unavailable. The geographic breadth of these species is summarized in Supplementary Fig. 11, which also highlights opportunities to expand future coverage.

In addition to the trait values directly measured from photographs, Skelevision includes estimates of missing values, imputed using a multivariate evolutionary model that accounts for both within- and between-species variation[12,100]. We previously showed that this approach for estimating missing values has low root mean square error (r.m.s.e.) and percentage bias[12]. Following the procedure described in ref. 12, we derived species-level body mass estimates from associated specimen metadata; 54.2% of individual skeletons had body mass recorded on their museum labels (archived in VertNet)[12]. Across these specimens, 69.2% of species had mass data (on average, $n \approx 5.5$ per species), providing a reasonable basis for estimating missing mass values within and across species. In alignment with this expectation, cross-validation of our approach against known specimen masses indicates very low imputation error (mean r.m.s.e. ≈ 0.16 log(grams); mean P bias ≈ 0.45%) at the specimen level. Therefore, we used the species means (for both skeletal traits and body mass) returned by our approach[12,100] (best linear unbiased predictions) as covariates for phylogenetic comparative analysis.

All downstream analyses use the phylogenetic hypothesis of Passeriformes from ref. 18. This phylogeny was generated using a sequential approach, integrating a fossil-calibrated family-level backbone phylogeny with high-quality and nearly complete phylogenetic trees of bird groups (for example, refs. 75,101–108). In our sample, 101 species did not cleanly map to the Passeriform phylogeny from ref. 18. To include these additional species (~5% of the dataset), we used matrix representation with parsimony[109,110], implemented in phangorn[111] to graft them into the ref. 18 topology. We used an alternative consensus phylogeny from refs. 56,112 as a reference and constrained existing relationships from ref. 18 by applying 1,000:1 weighting in favour of ref. 18. We then restored the exact node ages from ref. 18, using congruification[113] with treePL[114]. Although this procedure may introduce noise or bias when placements are not perfectly correct (for example, ref. 19), we observed no added species far outside expected groups. Therefore, we believe that the increase in power from adding these 101 species outweighs the potential error in estimated parameters (see the sensitivity analysis in the Supplementary Information).

Although the topology and timescale of avian diversification are active areas of research, the phylogeny we use is consistent with the fossil record[115–117] and aligns with other studies using subgenomic or genomic datasets[14,19,118–120]. These studies and others support the hypothesis that Passeriformes originated in the Eocene, 45-47 Ma[14,18].

Fitting time-heterogeneous models

A variety of statistical approaches have been developed to characterize the accumulation of phenotypic variation through stochastic processes, including Brownian motion (BM)[31,121], Ornstein–Uhlenbeck process (OU)[122,123], and accelerating-decelerating (ACDC) models[55,124]. Extending these models, researchers now apply the concept of 'macroevolutionary regimes'[125] to evaluate how broad-scale patterns of phenotypic variation may be explained by shifts in a background evolutionary process (for example, refs. 126,127). This model-based approach is widely recognized as a powerful way to test hypotheses about macroevolution[128].

Some researchers have also recognized that the process generating phenotypic variation may itself evolve along the branches of a phylogenetic tree[126], such that variation may be best partitioned on the basis of a mixture of distinct evolutionary models across the tree of life[37,38]. This idea has spurred a renaissance of methodologies for identifying the phylogenetic positions of evolutionary events that partition phenotypic variation under a consistent statistical criterion[60,129–138]. A few approaches also allow researchers to extend these ideas to multivariate datasets of several dimensions[38,62,135,139–141], although, as far as we are aware, there is only one existing likelihood framework that can be used to estimate a full multivariate evolutionary process for high-dimensional phenotypes[34].

Although phenotype is an inherently multivariate phenomenon[32,33,62,142], theoretical or computational constraints have typically required that researchers reduce the dimensionality of their datasets through techniques such as principal component analysis (PCA)[62,143]. However, PCA can bias evolutionary analyses by distorting variation across trait dimensions such that the first few principal components artifactually appear to have evolved via an 'early burst' process, similar to expectations of adaptive radiation[62]. More generally, PCA's objective of minimizing the sum of squared errors has long been recognized to have the potential to suppress task-relevant details if they do not dominate the overall variance of the data[144]. True multivariate methods that can capture the evolution of high-dimensional phenotypes in association with large phylogenies are therefore required to identify statistically consistent partitions of biological diversity, enabling robust tests of macroevolutionary hypotheses.

We developed an efficient approach for detecting such multivariate shifts in phenotypic evolution. The underlying model-fitting machinery we use builds on the mvgls framework from the R package mvMORPH, which enables high-dimensional phylogenetic generalized least squares[34,145]. We extend this framework with a model selection procedure inspired by recent work on phylogenetic shift detection[36–38,146] to infer multi-rate multivariate Brownian motion models. In these models, evolutionary shifts are represented as scalar transformations of the phenotypic variance–covariance matrix, corresponding to proportional matrices in the 'Flury hierarchy' with shared eigenvectors[147,148]. This proportionality assumption balances model complexity, computational feasibility and statistical

identifiability in high-dimensional trait spaces, reducing the number of free parameters (p) from $\frac{p(p+1)}{2}$ per matrix to a single scalar multiplier.

Although our approach does not explicitly model more complex evolutionary dynamics—such as changing covariance structures or deviations from Brownian motion (although see post hoc analysis in the Supplementary Information)[38,131,135,139,140,148,149]—it enables automated inference of multivariate evolution across thousands of species without relying on lossy procedures such as PCA[62]. The stepwise greedy search algorithm we apply (Supplementary Figs. 6 and 7) compares alternative configurations of model shifts using information criteria, the generalized information criterion (GIC)[150] or Bayesian information criterion (BIC)[151]. Similar to refs. 36–38,136,146, our approach requires a rooted phylogenetic tree and a matching phenotype matrix and applies the following algorithm:

1. Estimate a globally uniform mvBM model with mvgls[34], using a user-specified generalized least squares regression structure (for example, **Y** ~ 1 for intercept-only or **Y** ~ X to include numeric and/or factor covariates).
2. Generate a pool of all candidate shift points; this includes all nodes with a specified minimum number of descendant tips. From this pool, use branch painting (for example, using paintSubTree from phytools[152]) to represent each candidate shift as a 'simmap' tree object. Each candidate shift is thus encoded as a 'painted' tree with two states: a uniform ancestral state and a derived state representing the candidate shift.
3. Each painted tree is evaluated in mvgls[34] under the 'BMM' model by constructing a regime-scaled multivariate covariance matrix and computing the corresponding multivariate normal log-likelihood.
4. Generate a ranked list of candidate shifts, according to GIC or BIC score, representing the improvement in model fit relative to the initial global model.
5. Initialize the final model configuration with the global model. The first candidate shift is added to the final configuration (an 'aggregated' model) from the ranked candidate list, and the likelihood is re-estimated. If the updated aggregated model has an information criterion (IC) score lower than the current aggregated model's by a specified threshold, the candidate shift is accepted, and the algorithm moves to the next candidate in the ranked list. At this stage, if adding a particular candidate shift does not reduce the IC score of the aggregated model by the specified threshold or more, we discard it and move to the next candidate. This loop continues until all candidates have been evaluated. The intention of this strategy is to efficiently add shifts to the aggregated model in the order that we expect will maximally improve model fit[36–38].
6. Lastly, reassess the statistical support for each accepted shift using IC weights. This step compares the final aggregated model (for example, with the best IC score) to reduced models in which each shift is removed sequentially. In this context, IC weights provide an interpretable metric of how much each shift contributes to the final configuration, when assessed independently[153,154], and are computed as follows:

$$w = \frac{1}{1 + e^{\frac{1}{2}\Delta IC}}, \text{ where } \Delta IC = IC_{\text{with shift}} - IC_{\text{without shift}}$$

Higher weights (that is, closer to 1) indicate a large increase in IC when the shift is excluded, suggesting that including the shift improves overall model fit. Weights near 0.5 suggest more equivocal support, while values below 0.5 imply that excluding the shift leads to a better-fitting model.

Our model selection procedure is implemented as a set of new R functions; all code is available as supplementary material and on GitHub (https://github.com/jakeberv/passerine-bodyplan-evolution), Zenodo (https://doi.org/10.5281/zenodo.19211076)[155] and CRAN (https://doi.org/10.32614/CRAN.package.bifrost). R code development was assisted by OpenAI large language models (GPT-4, GPT-4o and GPT-5); all generated outputs were reviewed and verified by the authors.

Here we restrict our analyses and simulations to restricted maximum likelihood optimization, which we show performs well for our empirical data (Supplementary Information), although ML can be used when fixed effects vary. For linear measurement data, our approach also relaxes the need for phylogenetic size correction as a data preprocessing step[28,143]; because it is built around a phylogenetic generalized least squares framework[145], we include log-body-mass as a covariate to account for phylogenetic size variation, consistent with isometric expectations for linear dimensions[156].

Analysis of Passeriform skeleton evolution

We applied our heuristic search procedure to our passerine skeletal dataset that includes data from 12 skeletal elements measured on 14,419 individuals spanning 2,057 species of passerines ('Empirical dataset and phylogenetic framework'). We ran 12 searches that varied the search parameters to identify optimized configurations of model shifts. We varied the minimum clade size (10, 20, 30 tips), the IC acceptance threshold (20, 40 units) and the information criterion (GIC or BIC) across replicated analyses (Supplementary Figs. 5–7). All analyses specify a likelihood optimization method (in our case, 'method = "LL" '; log-likelihood, to use default restricted maximum likelihood estimation) and estimate a measurement error term ('error = *T*') from the data as a nuisance parameter, which is recommended for empirical data[34]. In Fig. 1a, we highlight results from the best-performing model (minimum clade size of 10 and GIC threshold of 20 units; Supplementary Fig. 6).

We summarized our results using several approaches. First, for each heuristic search, we partitioned inferred shifts into rate-increase or rate-decrease categories. Here, a 'rate' is operationally defined as the arithmetic mean of the variance terms (σ^2) for each regime's estimated phenotypic variance–covariance matrix. We counted model shifts in each category, summarized the frequency of increases and decreases for each analysis, and compared these counts using a Wilcoxon signed-rank test with continuity correction[157] (Supplementary Fig. 5). Then, we compared the magnitudes of observed rate increases to observed rate decreases using a Monte Carlo two-sample Kolmogorov–Smirnov test (Supplementary Fig. 5)[158–160]. Next, we estimated waiting times between shift events and compared the fit of alternative parametric distributions to the observed distribution of log rates using the model_select function in the R package 'univariateML'[161]. We highlight key nodes representing rare rate increases in Fig. 1a.

For downstream spatial analyses, we summarize tip-level multivariate rates using a lineage-based statistic that approximates recent rates of multivariate phenotype evolution. This metric calculates a weighted mean of the inferred multivariate phenotypic variance along each tip-to-root path:

$$\bar{\sigma}^2_{\text{WPR}} = \frac{1}{Z}\sum_{i=1}^{L} 2^{-a_i/T} \cdot \bar{\sigma}^2_i$$

where

- $\bar{\sigma}^2_{\text{WPR}}$ is the weighted mean phenotypic rate along the root-to-tip path.
- $\bar{\sigma}^2_i$ is the mean variance at node i, computed across multiple trait dimensions.
- a_i is the age of node i, measured as time since the present.
- T is the half-life parameter controlling the rate of decay in weighting.

- $Z = \sum_{j=1}^{L} 2^{-a_j/T}$ is the normalization constant ensuring weights sum to 1.
- L is the total number of nodes along the root-to-tip path, including both shift and non-shift nodes.

Here, each node's variance contribution σ_i^2 is down-weighted by its age a_i using an exponential kernel with half-life T and normalized by Z to ensure the weights sum to 1. Under a model of piecewise-constant evolutionary rate variation—where $\sigma^2(t)$ changes occur discretely at shift points—this decay-weighted estimator accumulates variance across entire lineages but gives more weight to recent evolutionary history. Because the exponential decay drops off with age, the estimator behaves like a soft time filter, concentrating signal within the most recent T million years ago. We set T to 5 Myr, which partitions tip-to-root paths into approximately 10 stages of declining half-life, given a root age of ~45 Ma.

This lineage-rate statistic is intended as a present-weighted summary of each lineage's inferred regime history, rather than an estimate of an instantaneous rate over a fixed recent time window. In this sense, the statistic is best interpreted as a tunable, temporally filtered average of multivariate evolutionary tempo along a root-to-tip path, with the half-life parameter (here, $T = 5$ Myr) controlling the timescale over which past regimes remain influential (for example, how rapidly the 'memory' of ancestral regimes is overwritten by recent cladogenesis). Our approach differs from a recently proposed estimator that summarizes tip-specific rates by summing standardized phylogenetic contrasts under a constant-rate Brownian motion model[162]. While that contrast-based method offers a computationally efficient summary of cumulative phenotypic divergence along lineages, our lineage-weighted metric integrates rate heterogeneity across a fitted multi-regime, multivariate Brownian motion model, enabling comparisons that reflect macroevolutionary variation across both lineages and evolutionary regimes.

For comparison, we also compiled a simpler 'tip phenotype rate', defined as the rate of the regime currently covering each terminal branch (that is, using only the σ^2 of each tip's assigned regime and ignoring ancestral nodes). This tip-level measure was highly correlated with our lineage-rate statistic ($r \approx 0.8$) and produced very similar spatial patterns (Supplementary Fig. 8), indicating that the spatial patterns we report do not depend on how our proposed statistic aggregates rates along root-to-tip paths.

Temporal and spatial patterns of rate variation

Statistical analysis of waiting times between shift events shows that shifts conform to an episodic burst model. To further investigate links between inferred macroevolutionary shifts and Earth's climate history, we summarized the temporal frequency of inferred rate shift events across the ~50 Myr evolutionary history of Passeriformes. Using a window size of 5 Myr and a step size of 0.5 Ma, we estimated a moving average capturing the frequency of shift events per node and visualized these patterns with spline fitting (Fig. 2). To visualize temporal intervals of extreme climatic variability, we downloaded the continuous composite of benthic foraminifer isotope records ($\delta^{18}O$) from ref. 63 (the 'benthic d18O smoothLoess10' dataset) and estimated a similar moving average of standard deviation. We then visually assessed alignment between periods of high shift-event frequency and periods of peak climatic variability (Fig. 2).

Next, to investigate how climatic variability may influence recent phenotypic evolution, we examined the spatial distribution of evolutionary rates across gradients of temperature seasonality and latitude. Similar to the approaches of refs. 56,74, we processed year-round species range data to generate a global gridded biodiversity dataset using range maps from BirdLife International[163]. Our approach partially accounts for the complex biogeographic affinities of migratory species by focusing on species with clearly defined year-round spatial affinities (for example, effectively removing species with very distinct wintering and breeding ranges). Using the sf R package[164,165], we standardized species distribution polygons and assigned species occurrences to the H3 hexagonal geospatial indexing system[166,167] based on the 2023 GBIF Taxonomy Backbone speciesKeys[168]. After cleaning, merging or removing invalid shapefiles, we were able to include 1,863 out of 2,057 species from our dataset in the spatial analyses. Although our spatial coverage is substantial, key areas for improvements in sampling include East and Southeast Asia, Indonesia, and Middle and Southern Africa.

Within each H3 grid cell, we computed several summary metrics, including the mean log rates of phenotypic evolution, the mean log-speciation rate (λ_{DR})[56,169], local species richness, and the sampling fraction (capturing the cell-level sampling relative to the expected species counts). We then computed range-weighted metrics to account for the contributions of geographically widespread taxa to grid-cell-level means. Similar to refs. 74,56, we computed species weights as the inverse of the species polygon range area. Our focal models use range-weighted means (although see our analyses of spatial variation in within-assemblage compositional variation in the Supplementary Information).

Following the approach of ref. 74, we explored SAR-Lag models[170–172] to assess the relationships among grid-cell means and environmental and ecological predictors. We first fit SAR-Lag models to estimate the independent effects of latitude, mean temperature seasonality (BioClim4; ref. 173), and species richness (log of range-weighted species counts; Supplementary Fig. 12c), on cell-mean lineage rates. For these cases, we first generated parsimonious models that controlled only for sampling and spatial effects, thereby isolating the effects of focal predictors. For the latitude model (Supplementary Fig. 13 and Supplementary Table 4), we included a quadratic term for latitude, and a control term for the sampling fraction (arcsine-square-root-transformed sampling fraction), to account for sampling coverage. The temperature seasonality model included second-degree polynomial terms for temperature seasonality, latitude, their interactions, and sampling fraction (Supplementary Fig. 15 and Supplementary Table 6). Lastly, the species richness model included a linear term for local species richness, a quadratic term for latitude, and linear terms for their interactions and sampling intensity (Supplementary Fig. 14 and Supplementary Table 5). To account for spatial autocorrelation in cell-mean lineage rates, we used an inverse distance-squared ($1/d^2$) spatial weighting scheme[174], defining neighbours within 1,000 km using great-circle distances. We evaluated the degree to which SAR-Lag models reduced spatial structure in the response by comparing estimates of Moran's I (ref. 175) in model residuals from baseline OLS to SAR-Lag models.

We then assembled an aggregated SAR-Lag model to assess the joint influence of temperature seasonality (BioClim4), species richness, and latitude on variation in grid-cell-mean lineage rates across space. This model included second-degree polynomials for temperature seasonality, species richness, and latitude. We also included interactions between latitude and the two focal predictors (temperature seasonality and richness), and additional linear terms to account for sampling (via an arcsine square-root transform of sampling fraction) and variation in speciation rate, λ_{DR} (ref. 56). We include λ_{DR} as a control term to account for the possibility that our search procedure, which assigns multivariate phenotypic rate shifts to cladogenic events, might induce a correlation between patterns of lineage splitting and patterns of phenotypic evolution.

We quantified the effect sizes of temperature seasonality, latitude, and species richness, by estimating spatially adjusted impact measures using the 'impacts' function from the 'spatialreg' R package. In SAR-Lag models, changes in a predictor can influence both the focal unit and neighbouring units because of spatial autocorrelation, making raw regression coefficients difficult to interpret. The 'impacts' function addresses this by decomposing each predictor's effect into

direct impacts (within-unit effects accounting for feedback), indirect impacts (spillover effects on neighbouring units), and total impacts (the sum of both). We approximated these values using a trace-based method that estimates spatial feedback from the traces of powers of the spatial weights matrix. To quantify uncertainty, we performed Monte Carlo simulations (R = 1,000) by drawing from the multivariate normal distribution defined by the model's coefficient estimates and their covariance matrix. These impact measures are derived from the full model, where each effect size is estimated while accounting for all other predictors. For each SAR-Lag model, we further assessed the relative importance of each predictor by implementing a leave-one-out procedure, in which individual terms (excluding controls) were sequentially removed from each full model. We refit each reduced model and calculated the resulting changes in AIC and BIC scores, and use these values to estimate model weights that reflect each term's contribution to overall model fit (Supplementary Tables 3–6).

Testing global patterns of trait variation

We considered the possibility that mean phenotypic rates at higher latitudes could be biased by non-random patterns of extinction, particularly in regions with strong seasonality, low local species richness, and a history of glacial cycling. Pleistocene glaciations are thought to have driven repeated bottlenecks in avian populations, reducing effective population sizes, and increasing extinction risk, especially in temperate and boreal zones with harsh climatic oscillations[176,177]. Some studies have suggested these dynamics have imparted a demographic imprint on extant bird lineages[86,178], while others suggest that glacial cycles have catalysed speciation through geographic isolation and ecological opportunity[179].

In one scenario, selective extinction of slow-evolving or ecologically specialized lineages could inflate estimates of mean macroevolutionary rates, producing a spurious signal of elevated rates at high latitudes. Alternatively, the loss of such lineages may have opened ecological space, facilitating subsequent diversification among surviving taxa[180]. At the same time, we expect that selective pruning of slow-evolving lineages could reduce within-assemblage trait disparity or dimensionality, as the ecological diversity of lineages is diminished[88–90]. To help distinguish between signatures of extinction-driven filtering and those more consistent with recent ecological diversification, we quantified spatial variation in trait disparity, dimensionality, and modularity across global assemblages. If elevated phenotypic rates at high latitudes reflect ongoing ecological diversification rather than extinction bias, we expect patterns consistent with local phenotypic overdispersion, indicative of expansion into a broader range of functional forms than expected, given evolutionary history.

To test for overdispersion in local phenotypic trait space relative to evolutionary age, we used OLS regressions of pairwise Mahalanobis distances against pairwise divergence times, constrained to pass through the origin, reflecting the expectation that zero evolutionary time should yield zero phenotypic divergence (Fig. 4b). Mahalanobis distances quantify multivariate dissimilarity between species while accounting for correlations among traits, making them well suited for measuring multivariate phenotypic disparity in our context[181,182] (Supplementary Fig. 16a). This nonparametric framework provides a simple test of whether empirical trait disparity in a grid cell, on average, exceeds or falls below expectations under a simple null model of proportional divergence[32]. For each spatial grid cell, we calculated pairwise Mahalanobis distances between size-corrected species residual trait vectors[143] using the inverse of the covariance matrix estimated via standard matrix inversion. We also extracted pairwise phylogenetic divergence-time estimates between species from the dated phylogeny.

Residuals from these OLS regressions approximate deviations from proportional expectations (that is, residuals with a mean of 0), with positive average residuals indicating greater-than-expected disparity (for example, consistent with adaptive divergence or ecological opportunity) and negative average residuals indicating constrained or filtered evolution. To account for uncertainty and pseudoreplication in OLS model residuals, we applied a nonparametric bootstrap procedure within each grid cell, resampling with replacement and refitting each model to generate a distribution of bootstrapped residuals. For comparisons across grid cells, we normalized for sample size and variance by averaging standardized mean residuals, where each was defined as the mean residual from a bootstrap replicate divided by its standard deviation. We also examined the proportions of positive and extremely positive residuals (>2 s.d.) within each grid cell across pooled bootstrap replicates (not shown).

Next, we explored several model-based alternatives to approximate the structure of multivariate phenotypic covariance within local assemblages. As a basis, we fit multivariate Brownian motion models to local phylogenetic subtrees using mvgls[34], yielding phylogenetically informed trait-covariance matrices for each grid cell. Then, we computed three metrics: the trace (Supplementary Fig. 16b), representing the total phenotypic variance; effective dimensionality (Fig. 4c), which estimates the number of orthogonal axes of variation; and modularity, reflecting the extent to which traits cluster into covarying groups (Fig. 4a).

We calculated the effective dimensionality as the exponential of the Shannon entropy of the normalized eigenvalue distribution of each covariance matrix (Fig. 4c). To estimate modularity, we transformed each estimated covariance matrix into a correlation matrix and constructed a weighted, undirected graph in which traits were represented as nodes and edge weights corresponded to the absolute values of pairwise correlations. Using the igraph package[183], we applied Louvain community detection[91] and quantified modularity based on the strength of the resulting structure[184,185]. Higher modularity values indicate that traits are more distinctly organized into cohesive clusters, suggesting greater functional compartmentalization (Fig. 4a). Under the hypothesis of extinction-biased filtering, we might expect declines in variance and dimensionality, as ecologically distinctive or slow-evolving lineages are lost[88]. Conversely, if high-latitude phenotypic rates reflect ecological diversification, we expect variance to be dispersed across more orthogonal axes and organized into more distinct or cohesive modules—signatures of evolutionary innovation[93–95].

Lastly, we estimated a multivariate metric of phylogenetic signal, inspired by Blomberg's K (ref. 124), K_{mult} (ref. 186) for multivariate data and its associated standardized effect size Z (ref. 187). K_{mult} captures the extent to which multivariate trait similarity tracks phylogenetic relatedness under a single-rate Brownian motion expectation, while Z provides a standardized effect-size measure based on permutations, enabling comparisons across grid cells (Fig. 4d). Within each spatial grid cell, we extracted the local subtree and, using the species-level residual trait vectors (above), computed K_{mult} and Z using the physignal.z() function from the geomorph package[188]. We assessed significant departures from phylogenetic independence by comparing the observed log-likelihood to a null distribution generated by 1,000 randomized residual permutation procedure (RRPP) permutations, after optimizing Pagel's λ on the observed data[187]. To adjust resulting P values for multiple comparisons across grid cells, we applied the Benjamini–Yekutieli procedure[189], which controls the false discovery rate under arbitrary dependence among tests. We then summarized the proportion of grid cells with statistically significant departures from phylogenetic independence in a sliding window (Fig. 4d). We emphasize that because z-scores are sensitive to sample size, the spatial patterns of significance (Fig. 4d) reflect the strength of statistical evidence, which is naturally lower in species-poor, high-latitude assemblages.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The primary dataset analysed in this work is published as part of a companion paper and is described there in detail[12]. All data objects for this manuscript are archived on Zenodo (https://doi.org/10.5281/zenodo.19198393)[190].

Code availability

Reproducible software code is available at the corresponding author's GitHub repository at https://github.com/jakeberv/passerine-body-plan-evolution and is archived at Zenodo (https://doi.org/10.5281/zenodo.19211076)[155]. Our heuristic search method is implemented in the bifrost R package[191], which is archived on CRAN (https://doi.org/10.32614/CRAN.package.bifrost).

References

1. Jablonski, D., Roy, K. & Valentine, J. W. Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science* **314**, 102–106 (2006).
2. Weir, J. T. & Schluter, D. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. *Science* **315**, 1574–1576 (2007).
3. Mittelbach, G. G. et al. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).
4. Simpson, G. G. *The Major Features of Evolution* (Columbia Univ. Press, 1953).
5. Simpson, G. G. *Tempo and Mode in Evolution* (Columbia Univ. Press, 1944).
6. Schluter, D. *The Ecology of Adaptive Radiation* (OUP Oxford, 2000).
7. Schluter, D. Speciation, ecological opportunity, and latitude. *Am. Nat.* **187**, 1–18 (2016).
8. Quintero, I. The diffused evolutionary dynamics of morphological novelty. *Proc. Natl Acad. Sci. USA* **122**, e2425573122 (2025).
9. Saupe, E. E. et al. Extinction intensity during Ordovician and Cenozoic glaciations explained by cooling and palaeogeography. *Nat. Geosci.* **13**, 65–70 (2020).
10. Eiserhardt, W. L., Borchsenius, F., Plum, C. M., Ordonez, A. & Svenning, J.-C. Climate-driven extinctions shape the phylogenetic structure of temperate tree floras. *Ecol. Lett.* **18**, 263–272 (2015).
11. Quintero, I. & Wiens, J. J. Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecol. Lett.* **16**, 1095–1103 (2013).
12. Weeks, B. C. et al. Skeletal trait measurements for thousands of bird species. *Sci. Data* **12**, 884 (2025).
13. Weeks, B. C. et al. A deep neural network for high-throughput measurement of functional traits on museum skeletal specimens. *Methods Ecol. Evol.* **14**, 347–359 (2023).
14. Oliveros, C. H. et al. Earth history and the passerine superradiation. *Proc. Natl Acad. Sci. USA* **116**, 7916–7925 (2019).
15. Mayr, G. *Paleogene Fossil Birds* 2nd edn (Springer Cham, 2022).
16. Lowi-Merri, T. M., Gjevori, M., Bochenski, Z. M., Wertz, K. & Claramunt, S. Total-evidence dating and the phylogenetic affinities of early fossil passerines. *J. Syst. Paleontol.* **22**, 2356086 (2024).
17. Steell, E. M., Nguyen, J. M. T., Benson, R. B. J. & Field, D. J. Comparative anatomy of the passerine carpometacarpus helps illuminate the early fossil record of crown Passeriformes. *J. Anat.* **242**, 495–509 (2023).
18. Claramunt, S. et al. A new time tree of birds reveals the interplay between dispersal, geographic range size, and diversification. *Curr. Biol.* **35**, 1–13 (2025).
19. Stiller, J. et al. Complexity of avian evolution revealed by family-level genomes. *Nature* **629**, 851–860 (2024).
20. Grant, P. R. & Grant, B. R. Evolution of character displacement in Darwin's finches. *Science* **313**, 224–226 (2006).
21. Prum, R. O. *The Evolution of Beauty: How Darwin's Forgotten Theory of Mate Choice Shapes the Animal World-and Us* (Anchor, 2018).
22. Ricklefs, R. E. Species richness and morphological diversity of passerine birds. *Proc. Natl Acad. Sci. USA* **109**, 14482–14487 (2012).
23. Price, T. D. et al. Niche filling slows the diversification of Himalayan songbirds. *Nature* **509**, 222–225 (2014).
24. Claramunt, S. & Cracraft, J. A new time tree reveals Earth history's imprint on the evolution of modern birds. *Sci. Adv.* **1**, e1501005 (2015).
25. Navalón, G., Bjarnason, A., Griffiths, E. & Benson, R. B. J. Environmental signal in the evolutionary diversification of bird skeletons. *Nature* **611**, 306–311 (2022).
26. Brinkworth, A. et al. Bird clades with less complex appendicular skeletons tend to have higher species richness. *Nat. Commun.* **14**, 5817 (2023).
27. Chitwood, D. H. et al. Latent developmental and evolutionary shapes embedded within the grapevine leaf. *New Phytol.* **210**, 343–355 (2016).
28. Klingenberg, C. P. Size, shape, and form: concepts of allometry in geometric morphometrics. *Dev. Genes Evol.* **226**, 113–137 (2016).
29. Wagner, G. P. *Homology, Genes, and Evolutionary Innovation* (Princeton Univ. Press, 2014).
30. Pavlicev, M. & Hansen, T. F. Genotype-phenotype maps maximizing evolvability: modularity revisited. *Evol. Biol.* **38**, 371–389 (2011).
31. Cavalli-Sforza, L. L. & Edwards, A. W. F. Phylogenetic analysis: models and estimation procedures. *Evolution* **21**, 550–570 (1967).
32. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334 (1976).
33. Felsenstein, J. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* **19**, 445–471 (1988).
34. Clavel, J., Aristide, L. & Morlon, H. A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to new-world monkeys brain evolution. *Syst. Biol.* **68**, 93–116 (2019).
35. Clavel, J., Escarguel, G. & Merceron, G. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. *Methods Ecol. Evol.* **6**, 1311–1319 (2015).
36. Berv, J. S. et al. Genome and life-history evolution link bird diversification to the end-Cretaceous mass extinction. *Sci. Adv.* **10**, eadp0114 (2024).
37. Smith, S. A., Walker-Hale, N. & Parins-Fukuchi, C. T. Compositional shifts associated with major evolutionary transitions in plants. *New Phytol.* **239**, 2404–2415 (2023).
38. Mitov, V., Bartoszek, K. & Stadler, T. Automatic generation of evolutionary hypotheses using mixed Gaussian phylogenetic models. *Proc. Natl Acad. Sci. USA* **116**, 16921–16926 (2019).
39. Coles, S. *An Introduction to Statistical Modeling of Extreme Values* (Springer, 2001).
40. Gould, S. J. & Eldredge, N. Punctuated equilibrium comes of age. *Nature* **366**, 223–227 (1993).
41. Uyeda, J. C., Hansen, T. F., Arnold, S. J. & Pienaar, J. The million-year wait for macroevolutionary bursts. *Proc. Natl Acad. Sci. USA* **108**, 15908–15913 (2011).
42. Landis, M. J. & Schraiber, J. G. Pulsed evolution shaped modern vertebrate body sizes. *Proc. Natl Acad. Sci. USA* **114**, 13224–13229 (2017).
43. Wagner, G. P. & Altenberg, L. Perspective: complex adaptations and the evolution of evolvability. *Evolution* **50**, 967–976 (1996).
44. Wagner, P. J. Early bursts of disparity and the reorganization of character integration. *Proc. R. Soc. B* **285**, 20181604 (2018).
45. Gatesy, S. M. & Dial, K. P. Locomotor modules and the evolution of avian flight. *Evolution* **50**, 331–340 (1996).

46. Losos, J. B. Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* **175**, 623–639 (2010).
47. Eldredge, N. & Gould, S. J. in *Models in Paleobiology* (ed. Schopf, T. J. M.) 82–115 (Freeman, Cooper and Company, 1972).
48. Malmgren, B. A., Berggren, W. A. & Lohmann, G. P. Evidence for punctuated gradualism in the Late Neogene *Globorotalia tumida* lineage of planktonic foraminifera. *Paleobiology* **9**, 377–389 (1983).
49. Brennan, I. G., Chapple, D. G., Keogh, J. S. & Donnellan, S. Evolutionary bursts drive morphological novelty in the world's largest skinks. *Curr. Biol.* **34**, 3905–3916.e3905 (2024).
50. Dufour, P., Sayol, F., Steinbauer, M. J., Cooke, R. & Faurby, S. Niche filling predicts evolutionary trajectories in insular bird communities. *Funct. Ecol.* **38**, 2636–2647 (2024).
51. Rabosky, D. L. & Lovette, I. J. Density-dependent diversification in North American wood warblers. *Proc. R. Soc. B* **275**, 2363–2371 (2008).
52. Cooney, C. R. et al. Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature* **542**, 344–347 (2017).
53. Jønsson, K. A. et al. Ecological and evolutionary determinants for the adaptive radiation of the Madagascan vangas. *Proc. Natl Acad. Sci. USA* **109**, 6620–6625 (2012).
54. Vinciguerra, N. T. & Burns, K. J. Species diversification and ecomorphological evolution in the radiation of tanagers (Passeriformes: Thraupidae). *Biol. J. Linn. Soc.* **133**, 920–930 (2021).
55. Harmon, L. J. et al. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* **64**, 2385–2396 (2010).
56. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
57. Phillimore, A. B. & Price, T. D. Density-dependent cladogenesis in birds. *PLoS Biol.* **6**, e71 (2008).
58. Derryberry, E. P. et al. Lineage diversification and morphological evolution in a large-scale continental radiation: the neotropical ovenbirds and woodcreepers (Aves: Furnariidae). *Evolution* **65**, 2973–2986 (2011).
59. Pennell, M. W., FitzJohn, R. G., Cornwell, W. K. & Harmon, L. J. Model adequacy and the macroevolution of angiosperm functional traits. *Am. Nat.* **186**, E33–E50 (2015).
60. Rabosky, D. L. et al. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* **4**, 1958 (2013).
61. Slater, G. J. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods Ecol. Evol.* **4**, 734–744 (2013).
62. Uyeda, J. C., Caetano, D. S. & Pennell, M. W. Comparative analysis of principal components can be misleading. *Syst. Biol.* **64**, 677–689 (2015).
63. Westerhold, T. et al. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science* **369**, 1383–1387 (2020).
64. Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693 (2001).
65. Prothero, D. R. & Berggren, W. A. in *Princeton Series in Geology and Paleontology* 582 (Princeton Univ. Press, 1992).
66. Steinthorsdottir, M. et al. The Miocene: the future of the past. *Paleoceanogr. Paleoclimatol.* **36**, e2020PA004037 (2021).
67. Clavel, J. & Morlon, H. Accelerated body size evolution during cold climatic periods in the Cenozoic. *Proc. Natl Acad. Sci. USA* **114**, 4183–4188 (2017).
68. Selvatti, A. P., Gonzaga, L. P. & Russo, C. A.dM. A Paleogene origin for crown passerines and the diversification of the Oscines in the New World. *Mol. Phylogenet. Evol.* **88**, 1–15 (2015).
69. Rohde, K. Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* **65**, 514–527 (1992).
70. Quintero, I., Landis, M. J., Jetz, W. & Morlon, H. The build-up of the present-day tropical diversity of tetrapods. *Proc. Natl Acad. Sci. USA* **120**, e2220672120 (2023).
71. Pulido-Santacruz, P. & Weir, J. T. Extinction as a driver of avian latitudinal diversity gradients. *Evolution* **70**, 860–872 (2016).
72. Jablonski, D. et al. Out of the tropics, but how? Fossils, bridge species, and thermal ranges in the dynamics of the marine latitudinal diversity gradient. *Proc. Natl Acad. Sci. USA* **110**, 10487–10494 (2013).
73. Dimitrov, D. et al. Diversification of flowering plants in space and time. *Nat. Commun.* **14**, 7609 (2023).
74. Rabosky, D. L. et al. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).
75. Harvey, M. G. et al. The evolution of a tropical biodiversity hotspot. *Science* **370**, 1343–1348 (2020).
76. Upham, N. S., Esselstyn, J. A. & Jetz, W. Ecological causes of uneven mammal diversity. Preprint at *bioRxiv* https://doi.org/10.1101/504803 (2023).
77. Lawson, A. M. & Weir, J. T. Latitudinal gradients in climatic-niche evolution accelerate trait evolution at high latitudes. *Ecol. Lett.* **17**, 1427–1436 (2014).
78. Weir, J. T. & Wheatcroft, D. A latitudinal gradient in rates of evolution of avian syllable diversity and song length. *Proc. R. Soc. B* **278**, 1713–1720 (2011).
79. Burns, M. D. et al. High-latitude ocean habitats are a crucible of fish body shape diversification. *Evol. Lett.* **8**, 669–679 (2024).
80. Tietje, M. et al. Global variation in diversification rate and species richness are unlinked in plants. *Proc. Natl Acad. Sci. USA* **119**, e2120662119 (2022).
81. Adams, D. C., Berns, C. M., Kozak, K. H. & Wiens, J. J. Are rates of species diversification correlated with rates of morphological evolution? *Proc. R. Soc. B.* **276**, 2729–2738 (2009).
82. Cracraft, J. Historical Biogeography and Patterns of Differentiation within the South American Avifauna: Areas of Endemism. *Ornithol. Monogr.* **36**, 49–84 (1985).
83. Knowles, L. L. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Mol. Ecol.* **10**, 691–701 (2001).
84. Venditti, C., Meade, A. & Pagel, M. Multiple routes to mammalian diversity. *Nature* **479**, 393–396 (2011).
85. Hansen, T. F. Is modularity necessary for evolvability?: Remarks on the relationship between pleiotropy and evolvability. *Biosystems* **69**, 83–94 (2003).
86. Smith, B. T., Gehara, M. & Harvey, M. G. The demography of extinction in eastern North American birds. *Proc. R. Soc. B* **288**, 20201945 (2021).
87. Santini, L., Tobias, J. A., Callaghan, C., Gallego-Zamorano, J. & Benítez-López, A. Global patterns and predictors of avian population density. *Glob. Ecol. Biogeogr.* **32**, 1189–1204 (2023).
88. Puttick, M. N., Guillerme, T. & Wills, M. A. The complex effects of mass extinctions on morphological disparity. *Evolution* **74**, 2207–2220 (2020).
89. Hull, P. Life in the aftermath of mass extinctions. *Curr. Biol.* **25**, R941–R952 (2015).
90. Cole, S. R. & Hopkins, M. J. Selectivity and the effect of mass extinctions on disparity and functional ecology. *Sci. Adv.* **7**, eabf4072 (2021).
91. Clauset, A., Newman, M. E. & Moore, C. Finding community structure in very large networks. *Phys. Rev. E* **70**, 066111 (2004).
92. Cantwell-Jones, A. et al. Mapping trait versus species turnover reveals spatiotemporal variation in functional redundancy and network robustness in a plant-pollinator community. *Funct. Ecol.* **37**, 748–762 (2023).

93. Kashtan, N., Parter, M., Dekel, E., Mayo, A. E. & Alon, U. Extinctions in heterogeneous environments and the evolution of modularity. *Evolution* **63**, 1964–1975 (2009).
94. Melo, D. & Marroig, G. Directional selection can drive the evolution of modularity in complex traits. *Proc. Natl Acad. Sci. USA* **112**, 470–475 (2015).
95. Wagner, G. P., Pavlicev, M. & Cheverud, J. M. The road to modularity. *Nat. Rev. Genet.* **8**, 921–931 (2007).
96. Parter, M., Kashtan, N. & Alon, U. Environmental variability and modularity of bacterial metabolic networks. *BMC Evol. Biol.* **7**, 169 (2007).
97. Kashtan, N. & Alon, U. Spontaneous evolution of modularity and network motifs. *Proc. Natl Acad. Sci. USA* **102**, 13773–13778 (2005).
98. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. in *Medical Image Computing and Computer-Assisted Intervention* (eds Navab N., Hornegger J., Wells W. M., & Frangi A. F.). 234–241 (Springer, 2015).
99. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask R-CNN. in *2017 IEEE International Conference on Computer Vision* (eds.). 2980–2988 (2017).
100. Goolsby, E. W., Bruggeman, J. & Ané, C. Rphylopars: fast multivariate phylogenetic comparative methods for missing data and within-species variation. *Methods Ecol. Evol.* **8**, 22–27 (2017).
101. Marki, P. Z. et al. Supermatrix phylogeny and biogeography of the Australasian Meliphagides radiation (Aves: Passeriformes). *Mol. Phylogenet. Evol.* **107**, 516–529 (2017).
102. McCullough, J. M. et al. Wallacean and Melanesian Islands promote higher rates of diversification within the global passerine radiation Corvides. *Syst. Biol.* **71**, 1423–1439 (2022).
103. Cai, T. et al. Near-complete phylogeny and taxonomic revision of the world's babblers (Aves: Passeriformes). *Mol. Phylogenet. Evol.* **130**, 346–356 (2019).
104. Alström, P. et al. Systematics of the avian family Alaudidae using multilocus and genomic data. *Avian Res.* **14**, 100095 (2023).
105. Shakya, S. B. & Sheldon, F. H. The phylogeny of the world's bulbuls (Pycnonotidae) inferred using a supermatrix approach. *Ibis* **159**, 498–509 (2017).
106. De Silva, T. N., Peterson, A. T. & Perktas, U. An extensive molecular phylogeny of weaverbirds (Aves: Ploceidae) unveils broad nonmonophyly of traditional genera and new relationships. *Auk* **136**, ukz041 (2019).
107. Olsson, U. & Alström, P. A comprehensive phylogeny and taxonomic evaluation of the waxbills (Aves: Estrildidae). *Mol. Phylogenet. Evol.* **146**, 106757 (2020).
108. Barker, F. K., Burns, K. J., Klicka, J., Lanyon, S. M. & Lovette, I. J. New insights into New World biogeography: an integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. *Auk* **132**, 333–348 (2015).
109. Baum, B. R. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**, 3–10 (1992).
110. Ragan, M. A. Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet. Evol.* **1**, 53–58 (1992).
111. Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).
112. Weeks, B. C. et al. Shared morphological consequences of global warming in North American migratory birds. *Ecol. Lett.* **23**, 316–325 (2020).
113. Eastman, J. M., Harmon, L. J. & Tank, D. C. Congruification: support for time scaling large phylogenetic trees. *Methods Ecol. Evol.* **4**, 688–691 (2013).
114. Smith, S. A. & O'Meara, B. C. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 2689–2690 (2012).
115. Claramunt, S. et al. Calibrating the genomic clock of modern birds using fossils. *Proc. Natl Acad. Sci. USA* **121**, e2405887121 (2024).
116. Claramunt, S. CladeDate: calibration information generator for divergence time estimation. *Methods Ecol. Evol.* **13**, 2331–2338 (2022).
117. Parham, J. F. et al. Best practices for justifying fossil calibrations. *Syst. Biol.* **61**, 346–359 (2012).
118. Prum, R. O. et al. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569–573 (2015).
119. Jarvis, E. D. et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320–1331 (2014).
120. Kimball, R. T. et al. A phylogenomic supertree of birds. *Diversity* **11**, 109 (2019).
121. Felsenstein, J. Phylogenies and the comparative method. *Am. Nat.* **125**, 1–15 (1985).
122. Hansen, T. F. & Martins, E. P. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* **50**, 1404–1417 (1996).
123. Hansen, T. F. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**, 1341–1351 (1997).
124. Blomberg, S. P., Garland, T. Jr. & Ives, A. R. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**, 717–745 (2003).
125. Jablonski, D. Background and mass extinctions: the alternation of macroevolutionary regimes. *Science* **231**, 129–133 (1986).
126. O'Meara, B. C., Ané, C., Sanderson, M. J. & Wainwright, P. C. Testing for different rates of continuous trait evolution using likelihood. *Evolution* **60**, 922–933 (2006).
127. Rabosky, D. L. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* **9**, e89543 (2014).
128. Uyeda, J. C., Zenil-Ferguson, R. & Pennell, M. W. Rethinking phylogenetic comparative methods. *Syst. Biol.* **67**, 1091–1109 (2018).
129. Eastman, J. M., Alfaro, M. E., Joyce, P., Hipp, A. L. & Harmon, L. J. A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution* **65**, 3578–3589 (2011).
130. Revell, L. J., Mahler, D. L., Peres-Neto, P. R. & Redelings, B. D. A new phylogenetic method for identifying exceptional phenotypic diversification. *Evolution* **66**, 135–146 (2012).
131. Uyeda, J. C. & Harmon, L. J. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Syst. Biol.* **63**, 902–918 (2014).
132. Beaulieu, J. M., Jhwueng, D.-C., Boettiger, C. & O'Meara, B. C. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* **66**, 2369–2383 (2012).
133. Landis, M. J., Schraiber, J. G. & Liang, M. Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. *Syst. Biol.* **62**, 193–204 (2013).
134. Thomas, G. H. & Freckleton, R. P. MOTMOT: models of trait macroevolution on trees. *Methods Ecol. Evol.* **3**, 145–151 (2012).
135. Khabbazian, M., Kriebel, R., Rohe, K. & Ané, C. Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. *Methods Ecol. Evol.* **7**, 811–824 (2016).
136. Ingram, T. & Mahler, D. L. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike information criterion. *Methods Ecol. Evol.* **4**, 416–425 (2013).
137. Bastide, P., Mariadassou, M. & Robin, S. Detection of adaptive shifts on phylogenies by using shifted stochastic processes on a tree. *J. R. Stat. Soc. B* **79**, 1067–1093 (2017).
138. Revell, L. J. & Harmon, L. J. A discrete character evolution model for phylogenetic comparative biology with Γ-distributed rate heterogeneity among branches of the tree. Preprint at *bioRxiv* https://doi.org/10.1101/2024.05.25.595896 (2024).

139. Uyeda, J. C., Pennell, M. W., Miller, E. T., Maia, R. & McClain, C. R. The evolution of energetic scaling across the vertebrate tree of life. *Am. Nat.* **190**, 185–199 (2017).
140. Bastide, P., Ané, C., Robin, S. & Mariadassou, M. Inference of adaptive shifts for multivariate correlated traits. *Syst. Biol.* **67**, 662–680 (2018).
141. Bartoszek, K. et al. Fast mvSLOUCH: multivariate Ornstein–Uhlenbeck-based models of trait evolution on large phylogenies. *Methods Ecol. Evol.* **15**, 1507–1515 (2024).
142. Gould, S. J., Lewontin, R. C., Maynard Smith, J. & Holliday, R. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B* **205**, 581–598 (1979).
143. Revell, L. J. Size-correction and principal components for interspecific comparative studies. *Evolution* **63**, 3258–3268 (2009).
144. Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer, 2006).
145. Clavel, J. & Morlon, H. Reliable phylogenetic regressions for multivariate comparative data: illustration with the MANOVA and application to the effect of diet on mandible morphology in phyllostomid bats. *Syst. Biol.* **69**, 927–943 (2020).
146. Alfaro, M. E. et al. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl Acad. Sci. USA* **106**, 13410–13414 (2009).
147. Phillips, P. C. & Arnold, S. J. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution* **53**, 1506–1515 (1999).
148. Stepanova, N., Boyko, J. D., Lin, J., Davis Rabosky, A. R. & Rabosky, D. L. Punctuated versus gradual shifts in the multivariate evolutionary process: a test with paired radiations of scincid lizards. *Syst. Biol.* **74**, 1–17 (2025).
149. Mahler, D. L., Revell, L. J., Glor, R. E. & Losos, J. B. Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* **64**, 2731–2745 (2010).
150. Konishi, S. & Kitagawa, G. Generalised information criteria in model selection. *Biometrika* **83**, 875–890 (1996).
151. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
152. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
153. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* 2nd edn, Vol. 2 (Springer-Verlag, 2002).
154. Burnham, K. P. & Anderson, D. R. Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**, 261–304 (2004).
155. Berv, J. S. et al. Supplementary code for: Rates of passerine body plan evolution in time and space. v1.0.0. *Zenodo* https://doi.org/10.5281/zenodo.19211076 (2026).
156. Spence, A. J. Scaling in biology. *Curr. Biol.* **19**, R57–R61 (2009).
157. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1**, 80–83 (1945).
158. Kolmogorov, A. N. Sulla determinazione empírica di uma legge di distribuzione. *G. Ist. Ital. Attuari* **4**, 83–91 (1933).
159. Smirnov, N. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Stat.* **19**, 279–281, 273 (1948).
160. Marsaglia, G., Tsang, W. W. & Wang, J. Evaluating Kolmogorov's distribution. *J. Stat. Softw.* **8**, 1–4 (2003).
161. Moss, J. univariateML: An R package for maximum likelihood estimation of univariate densities. *J. Open Source Softw.* **4**, 1863 (2019).
162. Title, P. O. et al. The macroevolutionary singularity of snakes. *Science* **383**, 918–923 (2024).
163. BirdLife International & Handbook of the Birds of the World. *Bird Species Distribution Maps of the World*. Version 2021.1 edn (BirdLife International, 2021).
164. Pebesma, E. Simple features for R: standardized support for spatial vector data. *R J.* **10**, 439–446 (2018).
165. Pebesma, E. & Bivand, R. *Spatial Data Science: With Applications in R*. (Chapman and Hall/CRC, 2023).
166. *H3: A Hexagonal Hierarchical Geospatial Indexing System* (Uber Technologies, 2018).
167. O'Brien, L. *h3jsr: Access Uber's H3 Library* (CRAN, 2024).
168. GBIF Secretariat. *GBIF Backbone Taxonomy*. Checklist dataset edn (GBIF Secretariat, 2023).
169. Redding, D. W. & Mooers, A. O. Incorporating evolutionary measures into conservation prioritization. *Conserv. Biol.* **20**, 1670–1678 (2006).
170. Ord, K. Estimation methods for models of spatial interaction. *J. Am. Stat. Assoc.* **70**, 120–126 (1975).
171. Bivand, R. & Piras, G. Comparing implementations of estimation methods for spatial econometrics. *J. Stat. Softw.* **63**, 1–36 (2015).
172. Bivand, R., Hauke, J. & Kossowski, T. Computing the Jacobian in Gaussian spatial autoregressive models: an illustrated comparison of available methods. *Geogr. Anal.* **45**, 150–179 (2013).
173. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
174. Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. In *Proc. 1968 23rd ACM National Conference*, 517–524 (Association for Computing Machinery, 1968).
175. Moran, P. A. P. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).
176. Willis, K. J. et al. The tempo of avian diversification during the Quaternary. *Phil. Trans. R. Soc. Lond. B* **359**, 215–220 (2004).
177. Lovette, I. J. Glacial cycles and the tempo of avian speciation. *Trends Ecol. Evol.* **20**, 57–59 (2005).
178. Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G. & Ellegren, H. Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Curr. Biol.* **25**, 1375–1380 (2015).
179. Weir, J. T. & Schluter, D. Ice sheets promote speciation in boreal birds. *Proc. R. Soc. Lond. B* **271**, 1881–1887 (2004).
180. Jablonski, D. Lessons from the past: evolutionary impacts of mass extinctions. *Proc. Natl Acad. Sci. USA* **98**, 5393–5398 (2001).
181. Atchley, W. R. M-Statistics and morphometric divergence. *Science* **208**, 1059–1060 (1980).
182. Huttegger, S. M. & Mitteroecker, P. Invariance and meaningfulness in phenotype spaces. *Evol. Biol.* **38**, 335–351 (2011).
183. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJ. Complex Syst.* **1695**, 1–9 (2006).
184. Schuurman, T. & Bruner, E. Modularity and community detection in human brain morphology. *Anat. Rec.* **307**, 345–355 (2024).
185. Padi, M. & Quackenbush, J. Detecting phenotype-driven transitions in regulatory network structure. *npj Syst. Biol. Appl.* **4**, 16 (2018).
186. Adams, D. C. A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. *Syst. Biol.* **63**, 685–697 (2014).
187. Collyer, M. L., Baken, E. K. & Adams, D. C. A standardized effect size for evaluating and comparing the strength of phylogenetic signal. *Methods Ecol. Evol.* **13**, 367–382 (2022).
188. Baken, E. K., Collyer, M. L., Kaliontzopoulou, A. & Adams, D. C. geomorph v4.0 and gmShiny: enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods Ecol. Evol.* **12**, 2355–2363 (2021).

189. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188 (2001).
190. Berv, J. et al. Supplementary data archive for Rates of passerine body plan evolution in time and space. v1.0.0. *Zenodo* https://doi.org/10.5281/zenodo.19198393 (2026).
191. Berv, J. S. et al. bifrost: an R package for scalable inference of phylogenetic shifts in multivariate evolutionary dynamics. Preprint at *bioRxiv* https://doi.org/10.64898/2026.04.12.718036 (2026).

Acknowledgements

We thank the numerous researchers and museums whose field collection efforts over 150+ years made this research possible, including B. Winger and B. Benz at the University of Michigan Museum of Zoology, as well as J. Bates, S. Hackett and B. Marks at the Field Museum of Natural History. This work was partly supported by the facilities and staff of the University of Michigan Institute for Data Science and AI in Society, the School for Environment and Sustainability, the Department of Ecology and Evolutionary Biology and the Museum of Paleontology. We thank A. Weiss, J. Clavel, E. Troyer, L. DeHaan, R. A. Rivero-Vega, T. Carruthers, S. McHugh, J. Boyko, S. Gaynor, S. Aguillon, K. Zyskowski and R. Prum for helpful discussions. We thank O. Stein for providing illustrations. S.C. was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2018-06747). S.A.S. was supported by NSF DEB 2325835. B.C.W. was supported by the David and Lucile Packard Foundation. J.S.B. was supported by an Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship (Schmidt Sciences, LLC).

Author contributions

Conceptualization, J.S.B. and B.C.W.; methodology, J.S.B., S.C., J.R.S., S.A.S., D.F.F. and B.C.W.; software, J.S.B., B.C.W. and D.F.F.; validation, J.S.B., C.M.P., D.F.F. and B.C.W.; formal analysis, J.S.B.; investigation, J.S.B.; resources, B.C.W. and D.F.F.; data curation, J.S.B., B.C.W. and D.F.F.; writing—original draft, J.S.B.; writing—review and editing, J.S.B., C.M.P., S.C., J.R.S., M.F., S.A.S., D.F.F. and B.C.W.; visualization, J.S.B.; supervision, J.S.B., B.C.W. and D.F.F.; project administration, B.C.W.; funding acquisition, J.S.B., B.C.W., S.A.S. and S.C.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-026-03110-5.

Correspondence and requests for materials should be addressed to Jacob S. Berv or Brian C. Weeks.

Peer review information *Nature Ecology & Evolution* thanks Andrew Orkney, Elizabeth Steell and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature portfolio

Corresponding author(s): Jacob S. Berv, Brian C. Weeks

Last updated by author(s): Apr 25, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Full details are provided here at the author's github repository (also included in the Zenodo archive here doi.org/10.5281/zenodo.19211076) https://github.com/jakeberv/passerine-bodyplan-evolution/blob/main/docs/reproducibility.md
Data analysis	https://github.com/jakeberv/passerine-bodyplan-evolution

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Copy of statement in the manuscript:
The primary dataset analyzed in this work is published as part of a companion paper and is described there in detail 12. All data objects for this manuscript are

archived on Zenodo (https://doi.org/10.5281/zenodo.19198393 190).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We analyze macroevolutionary patterns of passerine body plan evolution using a high-dimensional phenotypic dataset (12 skeletal traits) across 2,057 species and a time-calibrated phylogeny. We fit multivariate Brownian motion models with time-heterogeneous rate regimes inferred via a greedy model-selection procedure, and test associations between inferred evolutionary rates and environmental variables across time (paleoclimate) and space (latitude, temperature seasonality). The design is comparative and observational (no experimental treatments), integrating phylogenetic modeling and spatial statistical analyses.
Research sample	We analyze macroevolutionary patterns of passerine body plan evolution using a high-dimensional phenotypic dataset (12 skeletal traits) across 2,057 species and a time-calibrated phylogeny. We fit multivariate Brownian motion models with time-heterogeneous rate regimes inferred via a greedy model-selection procedure, and test associations between inferred evolutionary rates and environmental variables across time (paleoclimate) and space (latitude, temperature seasonality). The design is comparative and observational (no experimental treatments), integrating phylogenetic modeling and spatial statistical analyses.
Sampling strategy	Species sampling reflects availability of museum specimens and coverage of the Skelevision dataset rather than a priori power calculations. The dataset maximizes taxonomic breadth and trait completeness, capturing ~89% of passerine families. Spatial analyses use global range maps aggregated to H3 grid cells, with range-weighting to account for uneven species distributions. Sample sizes (species, specimens, grid cells) are large relative to model complexity and were evaluated through sensitivity analyses and model diagnostics (e.g., robustness across alternative parameterizations and sampling thresholds).
Data collection	Skeletal measurements were collected from digitized images of museum specimens using a validated computer vision pipeline (U-Net + Mask R-CNN). Measurements represent the longest linear dimension of each skeletal element. Body mass data were extracted from specimen metadata (VertNet) and supplemented via phylogenetic imputation. Phylogenetic data were assembled from published sources and augmented using standard comparative methods. Spatial data were derived from BirdLife range maps and processed using geospatial workflows in R.
Timing and spatial scale	Phenotypic and phylogenetic data span ~50 million years of passerine evolution. Spatial analyses are global, using contemporary species distributions aggregated into ~10,000+ grid cells. Environmental predictors (e.g., temperature seasonality) are derived from global climate datasets at ~1 km resolution and aggregated to grid cells. Temporal analyses use sliding windows (e.g., 5 Ma) across the Cenozoic.
Data exclusions	Some species (~5%) that did not map cleanly onto the phylogeny were excluded or grafted using established methods. Spatial analyses exclude species lacking reliable year-round range data (e.g., strongly migratory species with disjunct seasonal ranges). Invalid or incomplete geographic shapefiles were removed during preprocessing. No additional exclusions were applied post hoc beyond these criteria.
Reproducibility	All analyses were conducted using reproducible pipelines implemented in R, with code publicly available (GitHub and Zenodo). Results were validated across multiple model configurations (e.g., varying minimum clade size, information criteria thresholds).

	Sensitivity analyses confirm robustness of key findings to parameter choices and sampling variation. All data and code required to reproduce results are archived and accessible.
Randomization	Not applicable. This is an observational comparative study using existing phenotypic, phylogenetic, and spatial data; no experimental allocation or randomization was performed.
Blinding	Not applicable. Data collection (automated image-based measurement) and analysis pipelines were objective and standardized; no subjective scoring or group assignment requiring blinding was involved.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	This study did not involve laboratory animals.
Wild animals	This study did not involve observation, capture, or handling of wild animals. All data were derived from existing museum specimens and associated metadata.
Reporting on sex	Sex was not a focal variable in this study. Where available, sex information was recorded in the underlying specimen metadata; however, analyses were conducted at the species level using aggregated trait values, and were not disaggregated by sex.
Field-collected samples	This study did not involve field collection of samples. All specimens were previously collected and archived in museum collections.
Ethics oversight	No ethical approval was required for this study because it did not involve live animals or experimental manipulation. All data were obtained from previously collected museum specimens and publicly available datasets.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA