

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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☐ | ☒ | 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

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](https://www.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