

Reporting Summary

Nature Research 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 Research 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 All data were collected in the R software v. 3.5.3 (2019).

Data analysis All analysis were performed in the R software v. 3.5.3 (2019).

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 Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

Raw data and R-scripts developed for this research are available at <http://vertlife.org/data/discoverypotential/>

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 used species-level biological, environmental, and sociological attributes to parameterize time-to-event (Accelerated Failure Time) models of discovery probability across time for tetrapod species worldwide.
Research sample	Our dataset comprised 32,878 species of terrestrial vertebrates: 7202 species of amphibians, 10,004 of reptiles, 5679 of mammals, and 9993 of birds.
Sampling strategy	At the species-level, sample size followed the global datasets available for each of four terrestrial vertebrate classes. At the taxon- and assemblage-level, metrics of central tendency and descriptive patterns of discovery potential were presented for samples with at least five species.
Data collection	Species-level data were compiled from the literature available (e.g. data papers and taxon-related literature), and spatial data (species geographic ranges) were extracted from Map of Life database (https://mol.org/).
Timing and spatial scale	Species included in our dataset were described from 1758 to 2014. Species-level attributes based on geographic range maps were computed for grid cells at the spatial resolution of 110 x 110 km. All remaining analysis were performed at the spatial resolutions of 220, 440, and 880 km.
Data exclusions	We discarded species described after 2014 to minimize potential biases from their potentially incomplete geographic characterization. We did not modeled the discovery probability for species described in the year 1758 itself to avoid the large number of descriptions due to Linnaeus work.
Reproducibility	Raw data and R-scripts developed for this research are available at http://vertlife.org/discoverypotential/
Randomization	We analysed each class of terrestrial vertebrates (amphibians, reptiles, mammals, and birds) separately given the uniqueness of their taxonomic history and patterns of species discovery.
Blinding	Not applicable.
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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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