

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

We used R (v4.1.3) code to programatically download and extract data from GBIF (gbif.org), which is directly used in the study's species distribution modeling pipeline.

Data analysis

Two separate pipelines are used in this study, both based in R (v4.1.3). The species distribution modeling component uses the 'glmnet' R package version 4.1-3. Code for both species distribution models and the viral sharing simulation components of the study, and the code to produce all visualizations, has been made publicly available (github.com/viralemergence/iceberg).

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](#)

No original data is used in, generated by, or produced by our study. The Data Availability statement includes a full list of open data sources we use.

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)

Life sciences study design

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

Sample size	We report the sample size of species included in the study throughout the paper, including in Extended Data Figure 10. The decision to apply point process or range bagging models is based on a sample size cutoff, informed by literature on the performance of these models with different levels of data. Species with 3 to 9 unique georeferenced points on a 25km grid were entered into the range bagging pipeline, and species with 10 or more were subject to the point process modeling pipeline. Species with fewer points were omitted from the analyses. For climate datasets used in the study, we always used every available climate model at the time of pipeline initiation.
Data exclusions	Species with data limited by sample size (above) were excluded, as were those absent from the phylogeny, marine mammals (inappropriate to model with terrestrial species), and non-eutherian mammals (due to model calibration challenges for phylogenetic regression explained in Albery et al.'s original publication). Species occurrence points were also excluded based on a Grubb outlier test to identify spurious records of their distribution.
Replication	Our study is reproduced on both WorldClim2's CMIP6 models and the CHELSA CMIP6 ensemble, which each include several global climate models (GCMs) and multiple scenarios. Throughout the text, we report key statistics with a standard deviation across GCMs. The code pipeline made available with the study makes simultaneous replication across climate products possible.
Randomization	Not applicable (no experimental trials with a control and intervention; this is a simulation study).
Blinding	Not applicable (no experimental trials of any kind; this is a simulation study, which would be impossible to design "blinded").

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