nature portfolio

Corresponding author(s): Lucy Van Kleunen, Aaron Clauset

Last updated by author(s): 12/18/2025

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	The empirical food-web datasets used in this study are available in the GATEWAy database (GlobAl daTabasE of traits and food Web Architecture), version 3, accessible at doi.org/10.25829/idiv.283-3-756. The data was pre-processed as described in Supplementary Note 2. The pre-processed data are available at https://doi.org/10.5281/zenodo.18026669.
Data analysis	The code for this paper is available at https://doi.org/10.5281/zenodo.18026669. Data pre-processing and missing link prediction scripts were run with Python 3.6.3 on CentOS Linux 7 using High Performance Computing resources supported by BioFrontiers IT. Results visualization scripts were run on Windows 11 with Python 3.12.4. A configuration file is provided in the code repository for reproducing a Python environment with the necessary packages for running the scripts.

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

The empirical food-web datasets used in this study are available in the GATEWAy database (GlobAl daTabasE of traits and food Web Architecture), version 3, accessible at doi.org/10.25829/idiv.283-3-756. The data was pre-processed as described in Supplementary Note 2. The pre-processed data are available at https://doi.org/10.5281/zenodo.18026669. Supplementary Data 1 includes food web features across the datasets and summary statistics for food web features split by ecosystem type. Supplementary Data 2 includes details of Beta regression results.

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	Our study did not involve human participants.
Reporting on race, ethnicity, or other socially relevant groupings	Our study did not involve human participants.
Population characteristics	Our study did not involve human participants.
Recruitment	Our study did not involve human participants.
Ethics oversight	Our study did not involve human participants.

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	This study evaluates a missing link prediction method for food webs.
Research sample	290 food web datasets were evaluated from an existing global food web database, with great diversity in organism taxa, source, sex, and age range.
Sampling strategy	290 food web datasets were evaluated from an existing global food web database, with diversity in sampling strategy.
Data collection	Data was downloaded from a publicly available database.
Timing and spatial scale	290 food web datasets were evaluated from an existing global food web database, with diversity in the timing, frequency, and spatial scale of data collection.
Data exclusions	No food webs were excluded from the analysis.
Reproducibility	Code is provided to reproduce computational experiments.
Randomization	Links were randomly assigned to folds for model evaluation.
Blinding	Predictive model performance was evaluated within each fold on withheld test links not used to train the predictive model.

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

Plants

Seed stocks	Not applicable.
Novel plant genotypes	Not applicable.
Authentication	Not applicable.