nature portfolio

Corresponding author(s): DAPR NCOMMS-24-14994

Last updated by author(s): Mar 27, 2024

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 database of bee occurrences that support the findings of this study are available in Dorey at al. (2023) figshare with the identifier: https://doi.org/10.25451/flinders.21709757. The phylogenetic tree of bee (Hymenoptera: Apoidea) that supports the phylogenetic diversity analyses is available in Henríquez-Piskulich, Hugall, and Stuart-Fox (2023) Dryad Digital Repository with the identifier: https://doi-org/10.5061/dryad.80gb5mkw1. The functional trait database that supports the functional diversity analyses is not yet publicly available, since it is currently in consideration for publication. The data is currently available as a supporting material not for publication, and will be available in a published article as well as a public database. We collected all the data using R version 4.3.1, R Studio Server 2023.06.1, and packages cited in methods
Data analysis	R version 4.3.1, R Studio Server 2023.06.1, and packages cited in methods

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 database of bee occurrences that support the findings of this study are available in Dorey at al. (2023) figshare with the identifier: https://doi.org/10.25451/flinders.21709757.

The phylogenetic tree of bee (Hymenoptera: Apoidea) that supports the phylogenetic diversity analyses is available in Henríquez-Piskulich, Hugall, and Stuart-Fox (2023) Dryad Digital Repository with the identifier: https://doi-org/10.5061/dryad.80gb5mkw1.

The functional trait database that supports the functional diversity analyses is not yet publicly available, since it is currently in consideration for publication. The data is currently available as a supporting material not for publication, and will be available in a published article as well as a public database.

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	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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 evaluated the effects of urbanization, climate and agriculture on bee richness, and phylogenetic and functional diversity across Canada and the continental U.S.A. Both our response variables (bee richness, bee phylogenetic diversity, and bee functional diversity) and our fixed variables are quantitative (temperature, precipitation, solar radiation, urbanization index from population density and percent of impervious surfaces on the 10 x 10 squared-km cells, percent of croplands in the 10 x 10 squared-km cells, Shannon diversity of land uses in the 10 x 10 squared-km cells, and number of bee observations per cell). We used linear mixed models for each of our response variables, with the fixed variables, a random effect for ecoregion level II and a Gaussian spatial correlation for the cell locations.

Research sample	We included all unique observations of bees across Canada and the continental USA within 15,411 10 x 10 squared-km cells (866,247 observations of 2,553 species). The sample is meant to be used to study biogeographic patterns of bee diversity across Canada and the continental U.S.A. We choose samples in Canada and the U.S.A. since they have an appropriate sampling intensity. The raw observations (before cleaning) are included in Dorey et al. (2023) figshare with the identifier: https://doi.org/10.25451/flinders.21709757.
Sampling strategy	We selected the 10 x 10 squared-km cells that contained at least 5 observations (15,411). Data availability of bee occurrences in the U.S.A. was deemed appropriate by a previous publication (Orr, M. C. et al. Global Patterns and Drivers of Bee Distribution. Current Biology 31, 451-458.e4 (2021).).
Data collection	Data was gathered from Dorey et al. (2023) figshare with the identifier: https://doi.org/10.25451/flinders.21709757
Timing and spatial scale	The study includes all bee observations irrespective of the date of collection, more than 70% of the included observations were collected after 2000. The study includes observations within Canada and continental U.S.A.
Data exclusions	We only included all unique records by species name and 10 x 10 squared-km cell ID. We removed all records with geospatial, basis, taxonomic or metadata issues as well as the records from cells lacking climate or land use data. The cleaning process is fully described in the 'Bee occurrence records' on the methods.
Reproducibility	All data is publicly available, we used common functions included on publicly available packages to clean the variables and calculate the results. The code will be available upon acceptance of the manuscript.
Randomization	This is not relevant to the study since the study involves correlations among variables. There are no samples to assign to different treatments. We controlled for geographic correlations by adding a Gaussian correlation to the mixed model. We controlled for ecoregions by adding a random effect for ecoregions type II. Eventually, we controlled sampling intensity by considering the effect of the number of observations per cell as a fixed effect in the model.
Blinding	Blinding was not relevant to the study since the study invokes correlations among ecological and biogeographical variables relevant to each study case (15,411 cells). Additionally, the aim of the study is to explore such relationships rather than assess the model accuracy.

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	The study did not involve laboratory animals.
Wild animals	The study did not involve the gathering of wild animal samples but relied on publicly available observations from wild animals.
Reporting on sex	Sex was only considered in the study of bee length, attending to bee sexual dimorphism.
Field-collected samples	The study did not involve field sample collections.
Ethics oversight	No ethical approval was required, the study is based on publicly available observations of bee occurrences.

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*