nature portfolio

Corresponding author(s): Ava Ghezelayagh

Last updated by author(s): Apr 1, 2022

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	Whole genomic DNA was downloaded from NCBI and converted to 2bit format using faToTwoBit from the UCSC Genome Browser Downloads (https://hgdownload.cse.ucsc.edu/admin/exe/). UCE loci were located from each genome assembly using the publicly available PHYLUCE v.1.7.1 script run_multiple_lastzs_sqlite.py with default parameters and the FASTA file containing the UCE bait sequences. The program phyluce_probe_slice_sequence_from_genomes.py was used (with default parameters) to extract the identified loci.
Data analysis	No custom code needed for reproducibility, but the following open source programs were used for data analysis: the pipeline PHYLUCE v.1.7.1, MAFFT v7.130b, trimAL v1.4.rev, IQTREE v.1.7, IQTREE v.2, PartitionFinder2 v2.1.1, RAxML-ng v0.9.0, TOPD_v4.6, BEAST v.1.8.4 , BEAST v2.5, ASTRAL-III v5.6.3, BUCKy, BAMM v.2.5.0, and the following R packages: TESS, PHYTOOLS, GEIGER. Citations to all software used are given in the manuscript. Code availability Analyses relied on open source programs and scripts used for data analysis are available on the Dryad Digital Repository: https://datadryad.org/stash/share/-vfd5XqnNuJ1BHG7s2nBDw2nRRyK80Rc4BAtkrAkkoU . Data availability NCBI BioSample Accession numbers corresponding to sequence data are listed in Supplementary Table 1. New raw sequence data is available for download from the NCBI Sequence Read Archive (SRA), under BioProject ID PRJNA758064. Sequence alignments, partitioning schemes, phylogenetic trees, and other related data files are available on the corresponding Dryad Digital Repository: https://datadryad.org/stash/share/-vfd5XqnNuJ1BHG7s2nBDw2nRRyK80Rc4BAtkrAkkoU.

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

Data availability
NCBI BioSample Accession numbers corresponding to sequence data are listed in Supplementary Table 1. New raw sequence data is available for download from the NCBI Sequence Read Archive (SRA), under BioProject ID PRJNA758064. Sequence alignments, partitioning schemes, phylogenetic trees, phenotypic trait data, and other related data files are available on the corresponding Dryad Digital Repository: https://datadryad.org/stash/share/-vfd5XqnNuJ1BHG7s2nBDw2nRRyK80Rc4BAtkrAkkoU.

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	A phylogenetic analysis of 1,084 species of teleosts fishes (1,109 individual specimens; duplicates for species are noted in Supplementary Table 1). The phylogeny is time-calibrated using information from the fossil record. The pattern of body shape diversification from 608 teleost species (one individual per species) is explored using the phylogeny.
Research sample	We include genomic data sampled from 1,084 species of teleost fishes targeting 91% of the taxonomic families. The taxon sampling, which used museum accessions of fish tissues, was designed to capture the major lineages of acanthomorph fishes. Some of the species were already sequenced (e.g. Alfaro et al. (2018)) and we gathered the appropriate data from NCBI.
Sampling strategy	We were interested in relationships among major lineages, so we typically only included one specimen per species sampled.
Data collection	All new DNA sequencing was performed by Ava Ghezelayagh, Peter Cowman, Richard Harrington, and Jessica Glass.
Timing and spatial scale	All new DNA sequence data was collected in 2018 and 2019.
Data exclusions	No sequence data was excluded. For diversification rate analyses, trees were pruned to exclude outgroup and duplicate that would be irrelevant to the question and/or drive misleading results.
Reproducibility	Some species were sequenced with multiple individuals to test the veracity of our data work flow. These duplicates can be found in Supplementary Table 1. We report all protocols for the generation of phylogenomic data in the Methods section of the Supplementary Information, allowing for the reproduction of alignments and downstream phylogenetic analyses using open source, previously published code.
Randomization	This is not particularly relevant to phylogenetic analyses concerning historical evolution. The only relevant randomization that occurred was the random selection of time trees outputted by BEAST to produce a summary Maximum Clade Credibility tree with TreeAnnotator. This randomization was performed using the shuf command in Unix.
Blinding	This is not relevant to phylogenetics.

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

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	NA
Wild animals	NA
Field-collected samples	Tissue samples of animals used in this study were obtained from several research museum collections, listed in the Acknowledgments.
Ethics oversight	The Peabody Museum of Natural History is under the ethics mandate of IACUC, so any tissues used from the Peabody collection followed IACUC approval. However, no animal or paleontological materials were collected for the purpose of this study and the study itself did not require IACUC approval.

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