nature portfolio

Corresponding author(s): Manli Wang, Zhihong Hu
Last updated by author(s): Sep 11, 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	Ubuntu v22.04.3 Python v3.10.14 LocalColabFold v 1.5.0 MMseqs2 v14.7e284 Foldseek v6.29e2557 orfipy v0.0.3 Pymol v2.5.0 IQ-TREE v2.2.5 MAFFT v7.505 trimAl v1.4.1 BEAST v2.7.6 FigTree v1.4.4 Bowtie2 v2.5.3 samtools v1.20 iVar v1.4.2 TB-Tools v2.085
Data analysis	R version 4.2.3 - all analysis code is available in Figshare file Anvi'o version 8

GraphPad Prism version 8.3.0

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

Databases used in our study include:
NCBI RefSeq (https://ftp.ncbi.nlm.nih.gov/refseq/);
NCBI WGS (Whole Genome Shotgun) genomes (https://ftp.ncbi.nlm.nih.gov/genbank/wgs/);
NCBI TSA (Transcriptome Shotgun Assemblies) (https://ftp.ncbi.nlm.nih.gov/genbank/tsa/);
UniRef90 (https://ftp.ebi.ac.uk/pub/databases/uniprot/uniref/uniref90/);
National Genomics Data Center (https://ngdc.cncb.ac.cn/);
AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/);
SpiderDB 1.0 (https://spider.bioinfotoolkits.net).

The high-throughput sequencing raw data of A. ventricosus genomic DNA were deposited into the Genome Sequence Archive (GSA) of the National Genomics Data Center and are available through accession number: CRA017441.

The original data generated in our study, including phylogenetic trees, predicted protein structures, custom codes, etc., have been made publicly available at Figshare.

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	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

Life sciences study design

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

Sample size	The samples included all the accessible Refseq genomes of NALDVs, WGS data for Arachnida and TSA data for Araneae at the start of our study. The list of samples could be found in the supplementary tables.
Data exclusions	The WGS data for Arachnida and TSA data for Araneae were obtained at June 2023. All the sequencing data that generated after our accession time were excluded.
Replication	All computational pipelines were run successfully at least twice. The qRT-PCR samples was performed from three female spiders.
Randomization	This study did not perform experiments that require sample randomization.
Blinding	This study did not perform experiments that require blinding.

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	N/A
Wild animals	The adult spiders (A. ventricosus and T. clavata) captured in Nanyang, Henan Province, China, were purchased via Taobao (https://shop104568564.taobao.com). For A. ventricosus, the spiders were first immersed in liquid nitrogen and macerated prior to DNA extraction. For T. clavata, the fresh tissues of spiders were dissected in PBS and prior to RNA extraction.
Reporting on sex	For A. ventricosus, individuals of both sexes were used to determine whether the spider nrEVEs was present in both sexes. For T. clavata, only females were used for RNA extraction, because females are larger and easier to dissect, and have a greater variety of silk glands.
Field-collected samples	N/A
Ethics oversight	N/A

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

Plants

Seed stocks	NA/
Novel plant genotypes	N/A
Authentication	N/A