

Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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

<https://www.ncbi.nlm.nih.gov>

Data analysis

Trinity version 2.0.3
 Trimmomatic
 cd-hit-v4.6.4
 TransDecoder-3.0.0
 BLAST
www.github.com/jairly/MoSUMA_tools/
 MUSCLE v3.7
 IQTree-1.6.2
 ModelFinder
 SequenceMatrix v100
 GBLOCKS v0.91b
 OMA.2.0.0
 trimAl v1.4
 APE 5.0
 PAUP4.0a
 PhyloBayes MPI v1.5a

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

All scripts, individual gene alignments, amino acid matrices and phylogenetic trees have been deposited as a Bitbucket repository and can be accessed at https://bitbucket.org/bzxdp/lozano_fernandez_2019. The transcriptomes generated as part of our study are available at NCBI Sequence Read Archive – BioProject PRJNA438779 [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA438779>]. Individual SRA numbers for the raw read data of each species are listed in Supplementary Table 1.

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 present a phylogenomic investigation of Chelicerata, utilising both new and more complete sequence information and a robust inferential methodology.
Research sample	We compiled molecular datasets based on transcriptomic data from 95 species, mostly from Illumina transcriptomes that were largely retrieved from public repositories (NCBI), and four species that we newly sequenced. The taxonomic sample spanned 80 chelicerates, 21 of them representing Acari, and 15 outgroups. Our dataset includes representatives of all arachnid orders, with the exception of two minor lineages, Palpigradi and Schizomida, for which genomic-scale data are still missing. We include sequence information from three of four living horseshoe crab species, and most importantly we expand taxon sampling within mites and ticks, including short-branched representatives of Sarcoptiformes (Acariformes) and Mesostigmata (Parasitiformes).
Sampling strategy	We prioritized maximizing the amount of diversity within the group, trying to have an even representation of chelicerates at the order level, and particularly focusing on increasing the representation of lineages that are prone to the effects of Long Branch Attraction, like Acariformes, Parasitiformes or Pseudoscorpiones.
Data collection	Molecular data was largely retrieved from public repositories, such as NCBI. For those four newly generated transcriptomes, they were generated at the University of Maynooth and assembled by Robert Carton.
Timing and spatial scale	Transcriptomic and genomic samples were retrieved from public repositories during 2015 and 2016. We generated the four Illumina transcriptomes during the 2011 to 2013 period.
Data exclusions	No data were excluded from the analyses
Reproducibility	For the Bayesian analyses, we ran two chains until convergence. We make available all generated data so other researchers can reproduce our experiments.
Randomization	Randomization was not relevant to our study. As we have tried to maximize the inclusion of a wide diversity of chelicerates (some orders in which there are just available a few transcriptomes), performing experiments removing some of this data would have possibly exacerbated systematic errors, such as long branch attraction.
Blinding	Blinding is not relevant to this kind of study.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> 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

Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

The study did not involve laboratory animals.

Wild animals

The study involved the collection of small arthropods in ocean or terrestrial environments. The specimens were collected and determined at the genus or species level by collaborators, and stored in RNA later for posterior RNA extraction.

Field-collected samples

We generated four new transcriptomes: the sea spider *Pycnogonum* sp., the solifugid *Galeodes* sp., the pseudoscorpion *Neobisium carcinoides*, and the amblypygid *Damon* sp. The specimens were collected and determined at the genus or species level by collaborators, and stored in RNA later for posterior RNA extraction. We have complied with all relevant ethical regulations for animal testing while collecting and processing these animals.

Ethics oversight

No ethical approval was required for the handling of small arthropods.

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