

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	Raw reads were quality-trimmed using fastp v0.20.1 (Chen et al., 2018) and assembled with metaSPAdes v3.13 (Nurk et al., 2017). Core UCES were extracted using PHYLUCE v1.6.6 (Faircloth, 2016) and LASTZ (Harris, 2007) with the termite-specific bait set targeting 50,616 loci from Hellemans et al. (2022). UCE loci were aligned with MAFFT (Katoh & Standley, 2013) implemented in phyluce_align_seqcap_align. Alignments were either trimmed internally using phyluce_align_get_gblocks_trimmed_alignments_from_untrimmed that implements Gblocks (Castresana, 2000; Talavera & Castresana, 2007) with default parameters or edge-trimmed with phyluce_align_seqcap_align.
Data analysis	Loci absent in more than 30% of taxa were filtered out using phyluce_align_get_only_loci_with_min_taxa. We assessed the compositional bias of each locus (after either internal- and edge-trimming) to identify low-quality and rogue loci using indices and tests implemented in RCFV_Reader v1 (Fleming & Struck, 2023) and BaCoCa v1.105 (Kück & Struck, 2014). Three supermatrices, the unfiltered, cogenic, and intergenic supermatrices, were reconstructed from internally-trimmed alignments. Four supermatrices were reconstructed from protein-coding UCES and the corresponding amino acid sequences were identified and retrieved using the transeq function of EMBOSS v6.6.0 (Rice et al., 2000). To format the data to codon and protein from obtained UCES, we used the GFF annotation file (NCBI Annotation Release 100) of Zootermopsis nevadensis (Terrapon et al., 2014) with BLASTn v2.10.0+ and bedtools v2.29.2 (Quinlan & Hall, 2010). Protein sequences were aligned with MAFFT v7.305 (Katoh & Standley, 2013), and back-translated the protein alignments into codon alignments with PAL2NAL v14 (Suyama et al., 2006). We used these seven supermatrices to reconstruct a total of 30 maximum-likelihood phylogenetic trees using IQ-TREE v1.6.12 with 1,000 ultrafast bootstrap replicates (UFB) to assess branch supports (Nguyen et al., 2015; Hoang et al., 2018). Additionally, we carried 21 reconstructions using multi-species coalescence in ASTRAL-III v5.7.7 (Zhang et al., 2018), with local posterior probabilities (LPPs) as branch support measure (Sayyari & Mirarab, 2016). For rogue taxa detection, we used RogueNaRok v1.0.1-3 (Aberer et al., 2013) on analyses from concatenated data, and TreeShrink v1.3.9 (Mai & Mirarab, 2018) for coalescence analyses. The phylogenetic trees were combined into a 50% majority rule consensus summary tree using phytools (Revell, 2012) in R v4.0.2 (R Core Team, 2020), separately for concatenation and

coalescence analyses. Site concordance factors (sCFs) are provided as an alternative branch support metric (Minh et al., 2020). We assessed alternative topologies for unresolved interfamilial nodes using the approximately unbiased (AU) test (Shimodaira, 2002) with 10,000 RELL replicates.

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](#)

Samples used in this study are described in Supplementary Data 1. The raw sequence reads newly generated in this study have been deposited in the NCBI Sequence Read Archive under BioProject accession code PRJNA1122986[<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1122986>]. The raw data from prior publications used in this study are available under BioProject accession codes PRJNA560101[<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA560101>], PRJNA1122307 [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1122307] and PRJNA1122966[https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1122966]; and the genome assemblies are available under accession codes GCF_000696155 [https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000696155.1/], GCF_002891405 [https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_002891405.2/], GCA_013340265 [https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_013340265.1/] and GigaDB100057[https://doi.org/10.5524/100057]. The raw sequence data can also be accessed using individual accession numbers given in Supplementary Data 1. The UCE data from prior publications used in this study are available in the Dryad Digital Repository at <https://doi.org/10.5061/dryad.x0k6djhn0>; <https://doi.org/10.5061/dryad.tmpg4f53w>. Source data for this study (UCE sequences, supermatrices, and trees) are available from the Dryad Digital Repository at <https://doi.org/10.5061/dryad.02v6wwqbm>. This study is registered under the ZooBank LSID: <https://zoobank.org/urn:lsid:zoobank.org:pub:B8A8464B-2FB2-4CA2-8330-A907111FD62E>.

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

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

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

Research sample

Sampling strategy

	arose shortly after the origin of the first Termitinae. Specific taxa were chosen based on sample availability in the collections of this study's co-authors.
Data collection	SH, GJ, YR, EC, MR, JŠ, RHS, and TB provided samples from collections. MW and EK performed lab experiments and generated data. SH assembled, processed, and analysed the data.
Timing and spatial scale	This project is the result of a global effort, with samples coming from all continents. Sequence data was obtained in 2018-2023, duration which made it possible to acquire all the data necessary to complete this study's goal.
Data exclusions	No data was excluded.
Reproducibility	All raw data, UCE data, alignments, and phylogenetic trees were either submitted to NCBI or made available in Dryad repository. We therefore believe the study to be fully reproducible.
Randomization	Phylogenies were carried out using different supermatrices and a diverse range of models.
Blinding	Blinding design was not necessary for the phylogenetic analyses reported herein as there are no differential treatment which may affect participants.

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	No laboratory animals were involved.
Wild animals	We used specimens from insect collections. The specimens were preserved in RNA-later or 70-100% ethanol. No living animal was used in this study.
Reporting on sex	Both males and females were used in the study. This data does not impact the phylogenetic reconstructions reported here.
Field-collected samples	No new field collection was performed for this study. All samples were housed in collections.
Ethics oversight	Insect specimens of this study (termites) are not regulated under IUCN or other ethical committees.

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

Plants

Seed stocks

No plant material was used in the study.

Novel plant genotypes

No plant material was used in the study.

Authentication

No plant material was used in the study.