nature research

Corresponding author(s): Gordon Love

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☒	☐	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No such software was used
Data analysis	No such software was used

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 upon request. 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

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Ecological, evolutionary & environmental sciences study design

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

Study description	We analyzed and quantified the lipid biomarker content of a large suite of ancient sedimentary rocks and oils, specifically targeting ancient steroid biomarker compounds sourced by demosponges of Neoproterozoic-Cambrian age.
Research sample	We targeted ancient sedimentary rocks of low thermal maturity which allowed excellent preservation of lipid biomarker constituents
Sampling strategy	64 ancient sedimentary rocks samples of Cryogenian-Cambrian age were analyzed to provide stratigraphic coverage of all formations from the Huqf Supergroup, South Oman Salt Basin. Ediacaran-Cambrian oils from Siberia and India, as well as some Phanerozoic source rocks, were also analyzed for comparison.
Data collection	Lipid biomarker analysis was performed at UC Riverside and MIT
Timing and spatial scale	Sedimentary rock and oils were obtained from research projects between 2005 and 2017, mostly supplied by oil companies from sub-surface drilling of wells. The rock and oils used are grouped by geological age and provenance.
Data exclusions	No exclusions, not applicable.
Reproducibility	Our reproducibility for molecular biomarker ratios and yields was assessed using oil hydrocarbon standards ran with each batch of analyses.
Randomization	The rocks and oil samples are grouped by geological age and provenance
Blinding	Not applicable

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	Modern sponge samples are available from the collections of two of the co-authors, Sperling and Cardenas.

Palaeontology

Specimen provenance	Sedimentary rock and oil samples were supplied by Petroleum Development Oman from wells drilled. PDO have approved publication of the data.

Specimen deposition	Rock and oil samples are held in labs of Love at UC Riverside and Summons at MIT
Dating methods	Not applicable

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.