nature research

Corresponding author(s): Tsuyoshi Komiya

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

▸ Experimental design

1. Sample size

Describe how sample size was determined.	We collected 156 sedimentary rocks from the Saglek area, and found graphite in 54 rock samples. We selected 28 samples with a larger amount of the graphite to cover all of lithologies and studied areas for geochemical works. We analyzed the whole rock TOC contents and carbon isotopes of 20 pelitic rocks, four conglomerates, three carbonate rocks and two chert nodules in carbonate rocks. We conducted in-situ analyses of graphite in two samples of a pelitic rock (LAF491) and carbonate rock (LAF760). We did not conducted statistical application.

2. Data exclusions

Describe any data exclusions.	We used all of data.

3. Replication

Describe whether the experimental findings were reliably reproduced.	We analyzed the compositions of graphite and whole-rocks using well-established analytical methods, and did not conduct duplicate analyses. But, the reproducibilities are described in the main text as well as Extended Data.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.	The experiments were not randomized.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.	The investigators were not blinded to allocation during experiments and outcome assessment.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ A statement indicating how many times each experiment was replicated
☒	☐ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☒	☐ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☒	☐ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	We used a software of "PeakFit 4.12 (SeaSolve Software Inc.) for the Laser Raman micro-spectrometer analysis. The software is also described in "Method" as well as "Extended Data."

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	Anyone can collect the samples in Labrador. They are also stored at the University of Tokyo.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	We did not use antibodies.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	We did not use eukaryotic cells.
b. Describe the method of cell line authentication used.	We did not use eukaryotic cells.
c. Report whether the cell lines were tested for mycoplasma contamination.	We did not use eukaryotic cells.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	We did not use eukaryotic cells.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	We did not use eukaryotic cells.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	We did not use eukaryotic cells.

nature research

Corresponding author(s): Tsuyoshi Komiya

☐ Initial submission ☐ Revised version ☒ Final submission

ChIP-seq Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data deposition

1. For all ChIP-seq data:

☒ a. Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ b. Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

2. Provide all necessary reviewer access links. *The entry may remain private before publication.*	Now, we do not the reviewers' names yet. When we know, we will show the reviewers' name on acknowledgment.
3. Provide a list of all files available in the database submission.	Figures 1-3, Extended Data Figures 1-9, Extended Data Tables 1-6.
4. If available, provide a link to an anonymized genome browser session (e.g. UCSC).	We do not study genome.

Methodological details

5. Describe the experimental replicates.	We collected 156 sedimentary rocks from the Saglek area, and found graphite in 54 rock samples. We selected 28 samples with a larger amount of the graphite to cover all of lithologies and studied areas for geochemical works. We analyzed the whole rock TOC contents and carbon isotopes of 20 pelitic rocks, four conglomerates, three carbonate rocks and two chert nodules in carbonate rocks. We conducted in-situ analyses of graphite in two samples of a pelitic rock (LAF491) and carbonate rock (LAF760).
6. Describe the sequencing depth for each experiment.	We did not performed PCR analysis.
7. Describe the antibodies used for the ChIP-seq experiments.	We did not performed PCR analysis, and did not study genome.
8. Describe the peak calling parameters.	We did not performed PCR analysis, and did not study genome.
9. Describe the methods used to ensure data quality.	We did not performed PCR analysis, and did not study genome.
10. Describe the software used to collect and analyze the ChIP-seq data.	We did not performed PCR analysis, and did not study genome.

nature research

Corresponding author(s): Tsuyoshi Komiya

☐ Initial submission ☐ Revised version ☒ Final submission

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data presentation

For all flow cytometry data, confirm that:

☐ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ 3. All plots are contour plots with outliers or pseudocolor plots.

☐ 4. A numerical value for number of cells or percentage (with statistics) is provided.

Methodological details

5. Describe the sample preparation.	We did not performed Flow Cytometry.
6. Identify the instrument used for data collection.	We did not performed Flow Cytometry.
7. Describe the software used to collect and analyze the flow cytometry data.	We did not performed Flow Cytometry.
8. Describe the abundance of the relevant cell populations within post-sort fractions.	We did not performed Flow Cytometry.
9. Describe the gating strategy used.	We did not performed Flow Cytometry.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☐

nature research

Corresponding author(s): Tusyohi Komiya

☐ Initial submission ☐ Revised version ☒ Final submission

MRI Studies Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Experimental design

1. Describe the experimental design.	We did not perform MRI studies.
2. Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.	We did not perform MRI studies.
3. Describe how behavioral performance was measured.	We did not perform MRI studies.

Acquisition

4. Imaging	
a. Specify the type(s) of imaging.	We did not perform MRI studies.
b. Specify the field strength (in Tesla).	We did not perform MRI studies.
c. Provide the essential sequence imaging parameters.	We did not perform MRI studies.
d. For diffusion MRI, provide full details of imaging parameters.	We did not perform MRI studies.
5. State area of acquisition.	We did not perform MRI studies.

Preprocessing

6. Describe the software used for preprocessing.	We did not perform MRI studies.
7. Normalization	
a. If data were normalized/standardized, describe the approach(es).	We did not perform MRI studies.
b. Describe the template used for normalization/ transformation.	We did not perform MRI studies.
8. Describe your procedure for artifact and structured noise removal.	We did not perform MRI studies.
9. Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.	We did not perform MRI studies.

Statistical modeling & inference

10. Define your model type and settings.	We did not perform MRI studies.
11. Specify the precise effect tested.	We did not perform MRI studies.

12. Analysis

a. Specify whether analysis is whole brain or ROI-based.	We did not perform MRI studies.
b. If ROI-based, describe how anatomical locations were determined.	We did not perform MRI studies.
13. State the statistic type for inference. (See Eklund et al. 2016.)	We did not perform MRI studies.
14. Describe the type of correction and how it is obtained for multiple comparisons.	We did not perform MRI studies.
15. Connectivity	
a. For functional and/or effective connectivity, report the measures of dependence used and the model details.	We did not perform MRI studies.
b. For graph analysis, report the dependent variable and functional connectivity measure.	We did not perform MRI studies.
16. For multivariate modeling and predictive analysis, specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.	We did not perform MRI studies.