nature research

Corresponding author(s): Ferrari

☐ 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.

No statistical methods were used to predetermine sample sizes. Three biologically independent soil samples were collected from the two sampling sites, Robinson Ridge (December 2005) and Adams Flat (January 2014). For both sites, samples were collected along a spatially explicit sampling design comprised of three 300 m long transects separated by two meter distances from each other. This sample size was sufficient to perform the in-depth metagenome analyses and supporting biochemical assays needed to confirm that novel primary production processes operating in these soils. However, the relatively limited sample size meant that we were unable to confirm whether trace gas scavenging is a widespread mechanism operating in Antarctic soils. Due to the very limited amounts of Robinson Ridge soils that remained following metagenome and physicochemical analysis, we used sample sizes lower than three for some gas chromatography and carbon fixation experiments.

2. Data exclusions

Describe any data exclusions.

No exclusions.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful. For carbon fixation studies (Figure 3b, Extended Data Figure 4), we observed that there was a large and unexplained variation in basal levels of CO2 that was fixed between each of the five soils tested. However, the core finding that H2 addition stimulated carbon fixation was reliably reproduced. H2 addition stimulated all 17 sample pairs in the dark condition and 15 of the 17 sample pairs under light illumination.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Samples were not randomized for the experiments.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not applicable to use in this study as we were investigating a new process and only two sites were examined.

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.	ARC GIS 10.1, Trimmomatic 0,36, CLC Genomics Workbench v8, BamM 1.4.1, GroopM 0.3, MetaBAT 0.25.2, CheckM v0.9.4, RefineM 0.0.20, CommunityM 1.0.6,FastTree 2.1, iTOL v3.6, dbCAN 5.0, ClustalX 2.0, MEGA7, Antismash 2.0, Graphpad Prism 7, BBMap 35.66, Prokka 1.12.

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.	All materials are available. However, only limited quantities remain of the Antarctic soil samples used for this study.

9. Antibodies

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

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	Not applicable.
b. Describe the method of cell line authentication used.	Not applicable.
c. Report whether the cell lines were tested for mycoplasma contamination.	Not applicable.
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.	Not applicable.

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.	Not applicable.

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.	Not applicable.