nature research

Corresponding author(s): Marie-Eve Monchamp

☐ 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 attempted to collect sediment cores from all major peri-alpine lakes impacted by eutrophication. We chose our sample of lakes based on availability of long-term lake data, accessibility for sampling and diversity in morphometric variables (to cover different lake sizes and morphometries). Only one core was sampled from every lake since we have shown in previous work that DNA from one sediment core was sufficient for confidently describing community composition over centennial time scales [see Monchamp et al. 2016, Appl and Environ Microbiol].

2. Data exclusions

Describe any data exclusions.

Next generation sequencing: the 16S sequence reads which did not pass the pre-established quality check and filtering steps are discarded. In the analyses requiring rarefication of samples to account for variation in sequencing depth (e.g. for estimating richness), the samples that contain less reads than the established threshold are discarded.
Statistical analyses: no outlier was found and no data points have been excluded from analyses.

3. Replication

Describe whether the experimental findings were reliably reproduced.

This study did not include any experiment. Replication was achieved in the survey by sampling 10 lakes. Two DNA extractions were pooled for each sample, and in PCR amplification steps, 3 separate PCR reaction on the same DNA extract were performed and pooled.

4. Randomization

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

Our study was a survey, in which no randomisation was applied to sample collection or analysis.

5. Blinding

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

Blinding was not possible due to the nature of the samples. Because we work with time-series of DNA samples collected in sediments of various age, we must follow strict procedures to avoid contamination of old samples with modern DNA. Therefore, it is advised to process the old and recent samples separately to help reduce the risks of cross-contamination, especially during DNA extractions. For the PCR step, all samples are treated the same way, i.e., the plates are prepared with both old and recent samples together.

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.

R version 3.3.2 using publicly available R packages all listed in the methods section CIMminer, open access Software development of the Genomics and Bioinformatics Group, Developmental Therapeutics Branch (DTB), Developmental Therapeutics Program (DTP), Center for Cancer Research (CCR), National Cancer Institute (NCI). (https://discover.nci.nih.gov/cimminer/)

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.

No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used

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.

No commonly misidentified cell lines were used

## 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. | No animals were used |

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. | The study did not involve human research participants. |