nature portfolio

Corresponding author(s): Joshua S. Weitz

Last updated by author(s): Oct 7, 2021

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	The water-column sampling was achieved using a 24 x 12 L Niskin bottle rosette attached to a conductivity-temperature-depth (CTD) package (SBE 911Plus, SeaBird) with additional fluorescence, oxygen, and transmissometer sensors. Metabolites were measured with a Waters Xevo TQ-S triple quadrupole and a Thermo Scientific Q-Exactive Orbitrap HF with both reversed-phase and hydrophilic interaction liquid chromatography (HILIC). Metabolite peaks were integrated with Skyline for small molecules 69, followed by quality control and normalization. Details of the data acquisition and processing have been previously reported in Boyesen et al 2020. The total lipid extract was analyzed on an Agilent 1200 high performance liquid chromatography (HPLC) system coupled to a ThermoFisher Exactive Plus Orbitrap high resolution mass spectrometer (HRMS; ThermoFisher, Waltham, MA, USA) equipped with an electrospray ion source. For the identification and quantification of lipids, we used LOBSTAHS, an open-source lipidomics software workflow based on adduct ion abundances and several other orthogonal criteria. Lipids identified using the LOBSTAHS software were quantified from MS data after pre-processing with XCMS and CAMERA. For the >0.2 micron transcriptome, generated reads were trimmed of adapter sequences with Trimmomatic v 0.27, end-joined with PandaSeq v2.4, and filtered for quality using Sickle v1.33. Reads containing ribosomal RNA sequences were removed in silico using sortmerna v2.1. Spiked-in RNA internal standard sequences were identified using lastal v756, quantified, and then removed. For the >5 micron transcriptome, mapping was conducted using the Burrows–Wheeler Aligner (BWA-MEM, parameters –k 10 –aM;) against a reference database constructed from MMETSP transcriptomes. Resulting alignments were counted using the HTSeq 0.6.1 package (options –a 0, --m intersection-strict, -s no). Read counts were then filtered for contigs with average read counts ≥ 10 across the time series and then DESeq2's variance stabilizing normalization was implemented on remaining data. KEGG Orthologs were assigned with UProC and putative taxonomic assignments at the phylum level were assigned from MMETSP taxon designations.
Data analysis	All code and feature/abundance tables used to complete this analysis are available at https://github.com/WeitzGroup/community_scale_metabolism_NPSG archived under https://zenodo.org/badge/latestdoi/262179139. This analysis was conducted using R version 3.6, for the versions of specific packages implemented please see Methods.

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

Sequence data for the >0.2 µm metatranscriptome are deposited in the Sequence Read Archive through the National Center for Biotechnology Information under BioProject ID PRJNA358725. The Station ALOHA gene catalogue data are available under Bioproject no. PRJNA352737, and iMicrobe (http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/263/ALOHAgene cat_v1_nonredundant.annot). Sequence data for the >5 µm metatranscriptomes are available in the Sequence Read Archive through the National Center for Biotechnology Information under accession no. SRP136571, BioProject no. PRJNA437978. Raw files for metabolomics data are available in Metabolomics Workbench under project ID PR000797. Processed data will be available soon in Metabolomics Workbench under Project ID PR000926. Lipidomics mass spectral raw data are available from authors upon request. All code and feature/abundance tables used to complete this analysis are available at https://github.com/WeitzGroup/community_scale_metabolism_NPSG archived under https://zenodo.org/badge/latestdoi/262179139.

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	This study aimed to monitor microbial ecological activity over the diel cycle via repeated sampling of surface waters (15m) on a Lagrangian cruise track over the course of several days in the North Pacific Subtropical Gyre. Sampling was conducted via Niskin bottle rosette deployment every 4 hours and seawater samples were accordingly handled for >5 micron transcriptome sequencing, >0.2 micron transcriptome sequencing, or metabolite/lipid/macromolecular extraction and mass spectrometry.
Research sample	The population intended to be studied by this sample is that of the microorganisms living in an eddy in the North Pacific Subtropical Gyre in the sunlit open ocean surface. The samples we observed were extracted via niskin bottle sampling from an anticyclonic eddy in the NPSG at a depth of 15m taken at 6 times of day - 0200 hrs, 0600 hrs (corresponding to sunrise), 1000 hrs, 1400 hrs, 1800 hrs (corresponding to sunset), and 2200 hrs. Using serial filtration for transcriptomics, we intended to separate transcripts belonging to larger eukaryotic phytoplankton from smaller cyanobacteria and heterotrophic bacteria/archaea, while chemical measurements were filtered only once to attempt to catch the bulk community's intracellular concentrations of molecules of interest.
Sampling strategy	For specifics on sampling strategy, please refer to Wilson et al 2017, Becker et al 2018, Harke et al 2018, and Boysen et al 2020 in our references for sampling procedures corresponding to the >0.2 micron transcriptome, lipidome, >5 micron transcriptome, and metabolome respectively.
Data collection	For specifics on data collection, please refer to Wilson et al 2017, Becker et al 2018, Harke et al 2018, and Boysen et al 2020 in our references for sampling procedures corresponding to the >0.2 micron transcriptome, lipidome, >5 micron transcriptome, and metabolome respectively.
Timing and spatial scale	Water-column seawater sampling for diel measurements took place every 4 h for a period of 4 days (26-30 July) and 3 days (31 July-3 August) at a depth of 15 m corresponding to the depth of a drogue used to identify the Lagrangian track. The water-column sampling was achieved using a 24 x 12 L Niskin bottle rosette attached to a conductivity-temperature-depth (CTD) package (SBE 911Plus, SeaBird) with additional fluorescence, oxygen, and transmissometer sensors. The sampling and analytical protocols for vertical profiles of nutrients, particulates, and flow-cytometry enumerated phytoplankton populations and heterotrophic bacteria were identical to those employed by the Hawaii Ocean Time-series program (http://hahana.soest.hawaii.edu/index.html).
Data exclusions	Data collected from the second 3-day long diel sampling circuit were excluded from this analysis for several reasons. Partly, the gap between the two sampling circuits for the purpose of comparative analysis constitutes data missing not at random and our analysis lacked the inferential capability to account for those missing timepoints. Furthermore, a change in wind speed between the sampling periods resulted in mixed layer shoaling causing observable ecological changes such as in the abundances of Crocosphaera observed in Wilson et al 2017. This led us to believe that the two diel sampling periods represented different ecological regimes and were not intercomparable for the purposes of a systematic analysis.
Reproducibility	All data collected for this study were observational and no experiment was conducted.
Randomization	Randomization was not relevant to our study because data collected were observational and no treatment was applied.

Blinding	Blinding was not relevant to our study because data collected were observational and no treatment was applied.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	See description in 'Location'.
Location	Fieldwork was conducted during 25 July to 5 August, 2015 in the oligotrophic North Pacific Subtropical Gyre. To maximize the signal to noise ratio in an open ocean environment, a Lagrangian sampling strategy was implemented whereby World Ocean Circulation Experiment Surface Velocity Profile (WOCE SVP) drifters from Pacific Gyre, Inc. were deployed within the center of a mesoscale anticyclonic eddy. The mesoscale eddy fields were identified using Archiving, Validation, and Interpretation of Satellite Oceanographic data (AVISO) and when the field sampling occurred, the target anticyclonic eddy was located north of the Hawaiian Islands at 24.4 N and 156.5 W, with a diameter of ~100 km. Over the 12-day sampling period, the shipboard measurements were conducted alongside the drifters as they performed an almost complete circular pattern with a diameter of ~44 km (Figure 1a). Water-column seawater sampling for diel measurements took place every 4 h for a period of 4 days (26-30 July) and 3 days (31 July-3 August) at a depth of 15 m corresponding to the depth of the drogue. The water-column sampling was achieved using a 24 x 12 L Niskin bottle rosette attached to a conductivity-temperature-depth (CTD) package (SBE 911Plus, SeaBird) with additional fluorescence, oxygen, and transmissometer sensors. The sampling and analytical protocols for vertical profiles of nutrients, particulates, and flow-cytometry enumerated phytoplankton populations and heterotrophic bacteria were identical to those employed by the Hawaii Ocean Time-series program (http://hahana.soest.hawaii.edu/index.html).
Access & import/export	The research expedition departed and returned from the port of Honolulu, Hawaii. All sampling was conducted in international oceanic waters using routine oceanographic operations that predominantly involved seawater filtration and subsequent freezing of the filters. Samples consisted of nucleic acids and other biological macromolecules such as proteins and lipids. No live biological samples were collected as part of the research activities. No permits were needed for this work and the samples were not subject to export controls.
Disturbance	Seawater was sampled using routine oceanographic methods in the open ocean with minimum disturbance to the surrounding waters.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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