Marine metabolomics: measurement of metabolites in seawater by gas chromatography mass spectrometry

Authors: E. Maggie Sogin*, Erik Puskas, Nicole Dubilier, Manuel Liebeke*

Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, Bremen, Germany

*Corresponding authors

Maggie Sogin, email: esogin@mpi-bremen.de, tel: +49 421 2028-823

Manuel Liebeke email: mliebeke@mpi-bremen.de, tel: +49 421 2028-822

Abstract

One of the major bottlenecks in describing marine ecosystems is characterizing the metabolome of seawater, as salt prevents metabolite analysis. We present SeaMet, a method that can detect hundreds of metabolites in less than one mL of seawater and quantifies them down to nano-molar levels using gas chromatography-mass spectrometry. Our method provides a major advance in marine metabolomics by enabling the unbiased analysis of primary metabolites across saltwater habitats.

Main

Gas chromatography-mass spectrometry (GC-MS) analysis enables the detection of primary metabolites, or small molecules that occur in central metabolic pathways across biological systems[1, 2]. High reproducibility coupled to the widespread availability of annotation resources make GC-MS the "workhorse" of analytical chemistry facilities. GC-MS has allowed for identification of metabolites associated with human disease[3], detection of compounds that serve as environmental cues in foraging[4], and description of metabolic fluxes within and between cells[5]. GC-MS is also an important analytical platform for environmental profiling of terrestrial soils and microbial activity on land[6, 7].

In contrast to terrestrial or limnic systems, the defining characteristic of marine habitats - high salt concentration - inhibits MS-based metabolomic analysis. Consequently, our knowledge of metabolite composition of the oceans is severely limited. This, in turn, restricts our understanding of global biogeochemical cycles, as the oceans contain the largest organic carbon pool on Earth[8]. While metagenomic and metatranscriptomic studies of the ocean, driven by low sequencing costs and projects such as Tara Oceans[9], have deepened our knowledge of the identity and activity of marine microbes, current metabolomic techniques do not permit equivalent surveys of the ocean metabolome. This is largely due to high instrumentation costs, large sample volume requirements, and the long amount of time needed for metabolome analysis. Available methods for measuring metabolites in marine samples, such as solid phase extraction (SPE), efficiently remove salt in preparation for either high resolution-MS[10] or liquid chromatography-MS approaches[11]. However, while SPE measures up to 50% of the dissolved organic carbon of a marine sample, it is inefficient in retaining small polar compounds[12], and thus excludes the majority of metabolites underlying primary metabolism. To more efficiently decipher ocean metabolism, it is necessary to develop cost-effective, high-throughput, and untargeted workflows that can readily identify and quantify molecules from high salinity environments.

Here, we present SeaMet, a marine metabolomics workflow that uses sample derivatization to enable metabolite detection in seawater. Our method modifies the well-established two-step metabolite derivatization procedure involving methoximation followed by trimethylsilylation[13], which permits the detection of non-volatile compounds using GC-MS. Like routine GC-MS sample preparation techniques[14, 15], SeaMet removes liquid through vacuum drying prior to derivatization – a process that results in a salt pellet when working with marine samples, which restricts MS analysis. Our initial results suggested that water locked within the dried salt crystals hindered the chemical reactions needed for GC-MS analysis (**Supplementary Fig. 1a, b**). Our method overcomes this limitation by first eliminating residual water within the salt crystals and then extracting small compounds into the derivatization reagents (**Fig. 1a**).

We used a synthetic mixture of 45 metabolites (**Supplementary Table 1**) dissolved in artificial seawater to document the performance in metabolite detection of our method. Overall, SeaMet increased total signal intensity on average by 42% and up to 89% for high salinity samples in comparison to the standard protocol for GC-MS sample preparation (**Fig. 1b**). We first replaced the most commonly used trimethylsilylation reagent, N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA)[15], with one that is less susceptible to inhibition by water, N, O-Bistrifluoroacetamide (BSTFA), which resulted in higher metabolite signals (**Supplementary Fig. 1b**). To eliminate water from the samples, we increased the speed-vacuum drying time from four to eight hours, and integrated a toluene drying step that is used in urine-based metabolomic analyses[14]. We further enhanced metabolite signals by treating the salt pellet to a combination of ultrasonication and vortexing after the additions of toluene and derivatization reagents, and following completion of the trimethylsilylation reaction. These steps break apart the salt crystals and release water into the toluene to enhance salt drying and metabolite extraction. Finally, following a recently described method for improving metabolite detection regardless of sample type (Liebeke et al. *in review*), we dried the samples between the methoximation and trimethylsilylation derivatization reactions under a gentle flow of N_2 gas (see **Fig. 1b** for signal improvements of each protocol step).

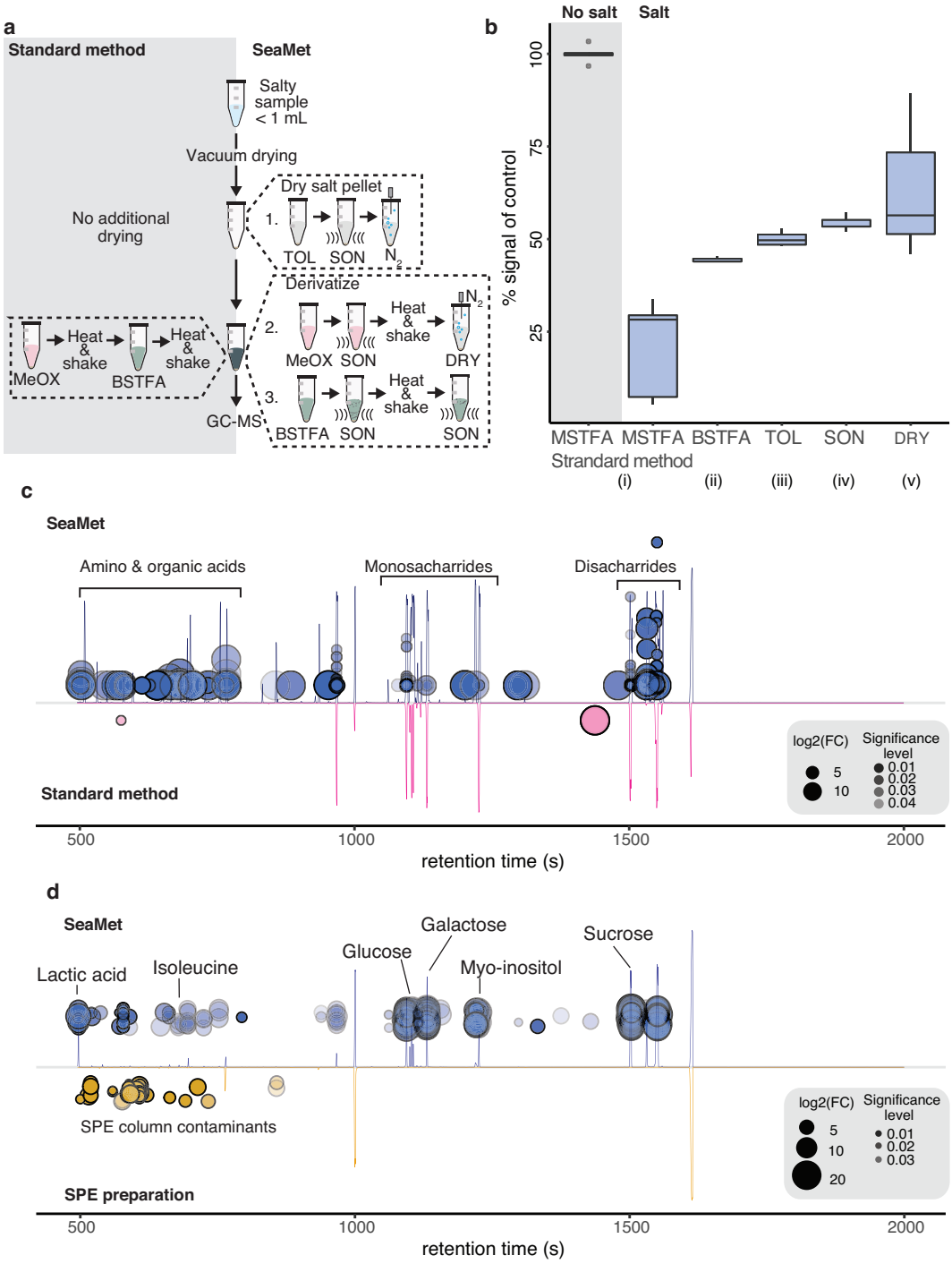

Fig. 1| SeaMet enhances the detection of metabolites in marine samples. a, Proposed modifications (right hand side outlined in white) to the standard (left hand side outlined in gray) two-step methoximation (MeOX)-trimethylsilylation (TMS) derivatization protocol includes key steps that strongly enhance metabolite detection in seawater as depicted in **b.** Steps modified from the standard method (i) include (ii) a switch in derivatization reagents from the water prone N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA) to *N*,*O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA), (iii) further drying of the salt pellet including the addition of toluene (TOL) to remove water azeotrophically, (iv) ultrasonication (SON) after the addition of TOL, MeOX , BSTFA and

after the completion of the BSTFA derivatization reaction, and (v) drying (DRY) of the pyridine from the MeOX reaction prior to BSTFA addition. **b,** Boxplots showing the total detected signals from each chromatogram detected by GC-MS from a synthetic mixture of 45 metabolites (**Supplementary Table 1**) dissolved in seawater ($n = 5$) relative to the no salt controls for each modification step. **c, d** Total ion chromatogram cloud plots from metabolite mixture samples indicate significant differences (B.H. adjusted $p < 0.05$) between ion abundances when comparing **c,** SeaMet (top) to the standard (bottom) GC-MS protocol and **d,** chromatograms using SeaMet before (top) and after (bottom) solid phase extraction. Larger bubbles indicate higher log2 fold changes between groups and more intense colors represent lower t-test p-values when comparing individual feature (*m/z* values) intensities between groups. Only features with log2(fold change) scores ≥ 2 are displayed. Metabolites with higher abundances in the SeaMet method include organic acids (lactic acid, succinic acid, and fumarate), amino acids (isoleucine, leucine, threonine and valine), sugar alcohols (myo-inositol and mannitol), and sugars (fructose, glucose, cellobiose, maltose, ribose, galactose, and sucrose). While peaks with high abundances from SPE prepared were either unknown peaks or column contaminant. Representatives from each compound classes are indicated.

Overall, SeaMet allowed us to detect significant increases in metabolite abundances across molecular classes when compared to the standard method (adjusted *p*-value < 0.05; mean fold change across all ions = 323; **Fig. 1c**). This included measurement of organic acids, amino acids, and fatty acids alongside sugars (and their stereoisomers), sugar alcohols, and sterols (**Supplementary Table 1**). Our detection limits in the nano-molar range are comparable to those of targeted techniques for marine systems that were developed to quantify single compounds from specific molecular classes (**Supplementary Tables 2 and 3**). Moreover, our workflow only requires 0.5 mL to 1 mL of seawater for metabolite detection, representing at least an order of magnitude less sample volumes than previously published techniques[12, 16]. To date, we successfully measured 107 metabolite standards dissolved in artificial seawater, representing major metabolite groups involved in primary metabolic pathways (**Supplementary Table 4**) making SeaMet applicable for both targeted and untargeted applications. Our method provides reproducible quantification and gives similar linearity, dynamic range, and coefficient of variation to salt-free samples prepared with the standard GC-MS derivatization procedure (**Supplementary Table 2**; **Supplementary Fig. 2**).

To demonstrate that our method captures compounds commonly missed by SPE-based metabolomic approaches, we compared GC-MS profiles using SeaMet before and after SPE sample treatment. Our method detected small polar compounds that are co-removed with salt during SPE preparation, such as

sugars, sugar alcohols, amino acids, and organic acids (**Fig. 1d**). Given our results, SeaMet is complementary to current marine DOM characterization approaches that require SPE preparation and high-resolution MS for metabolite detection. Thus, SeaMet expands the possibilities for metabolite analysis in the ocean and advances marine metabolomics by recovering compounds that previously could not be analyzed by current platforms in a single analysis.

To demonstrate the power of our approach in characterizing metabolite profiles of marine bacteria, we monitored changes in the exo-cellular metabolome during growth of a heterotrophic Proteobacterium, *Marinobacter adhaerens*. Using SeaMet, we observed significant changes in the metabolite composition of marine culture medium during the bacteria's initial growth phase (adjusted p-value < 0.05; **Fig. 2a; Supplementary Fig. 3a, b).** The bacteria took up different carbon and nitrogen resources in a cascade-like fashion**,** and later in growth, began excretion of an undescribed compound (**Fig. 2a, b; Supplementary Fig. 4**). Successional dynamics in substrate use is a common energy conservation mechanism in bacteria[17] and affects central carbon and nitrogen dynamics during growth. With SeaMet, we show that like many bacteria, *M. adhaerens* participates in the release of organic carbon which is metabolized by other microorganisms or contributes to the complexity of refractory DOM. By identifying and quantifying metabolites that are consumed and excreted in cultivable marine bacteria, we may be able to expand our understanding of key primary compounds involved in the transformation of organic matter in the ocean.

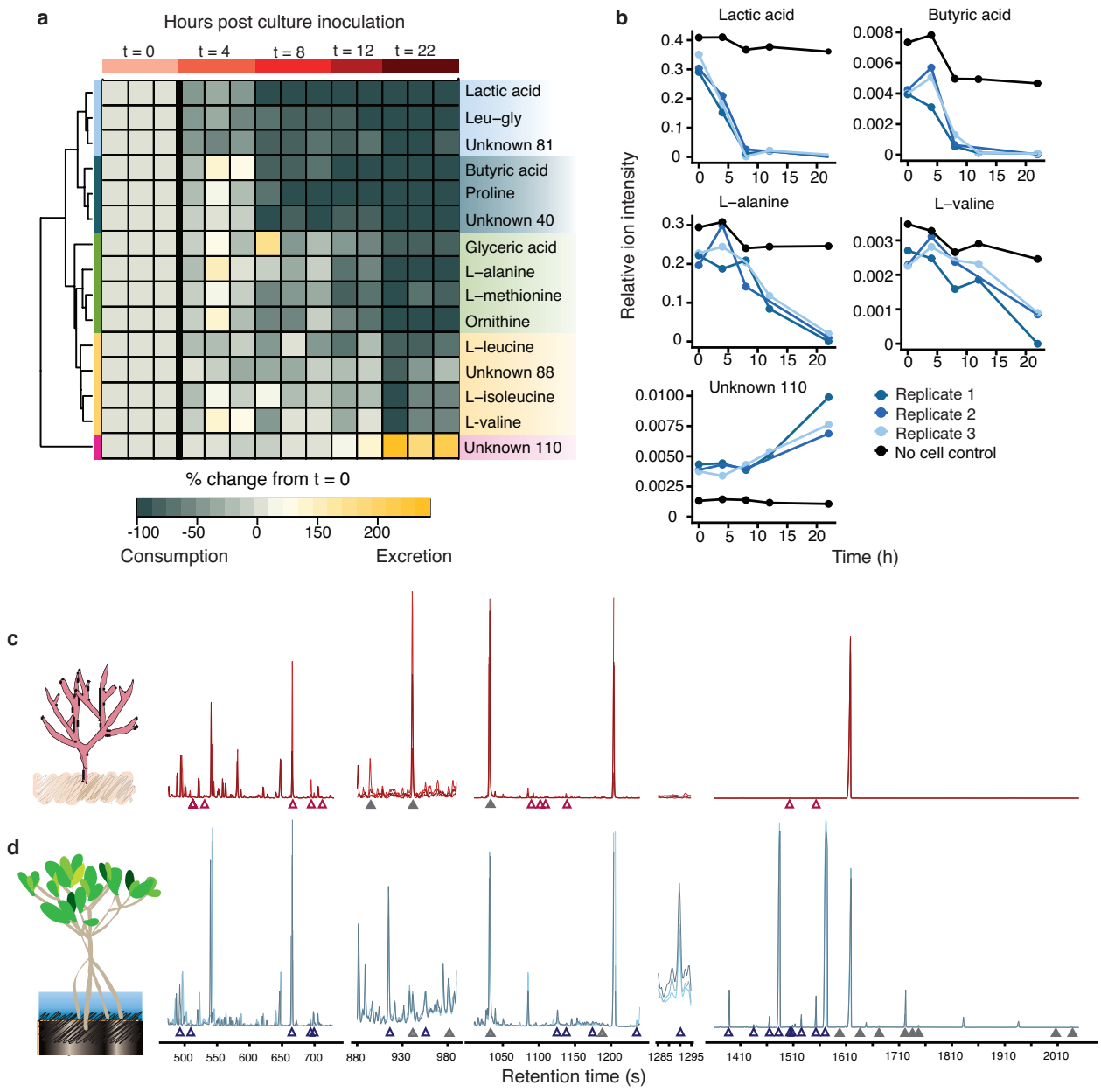

Fig. 2| Metabolite detection in marine samples. a, A heatmap of metabolite abundances relative to starting conditions during the culture of the heterotroph *Marinobacter adhaerens* in marine broth indicates some compounds, like the dipeptide leucine-glycine (leu-gly), and lactic acid are taken up before others, such as branch chain amino acids. After 12–22 hours of growth, the bacterial culture excrete an unknown compound (Unknown 110). Hierarchical clustering shows groups of metabolites that changed significantly during the experiment (left hand colored bars, B.H. adjusted p-value < 0.05; fold change > 2). These metabolite groups represent successive stages in *M. adhaerens* consumption and production of marine broth components. **b,** Relative ion abundances over time for select metabolites from each cluster group depicted in **a.** The blue lines represent biological replicate cultures while the black line shows results from a control sample with no cell addition. Low variation among biological replicates highlights the reproducibility of our marine metabolomics method. GC-MS metabolomic profiles from **c,** coralline and **d,** mangrove sediment porewaters show high concentrations of identified (open

triangles) fatty acids and sugars. Profiles also reveal unknown peaks (filled triangles) for which no matches were found in public databases (**Supplementary Table 5**).

To test the ability of our workflow to assess complex environmental metabolomes, we applied SeaMet to samples from coralline and mangrove sediment porewaters. Coral reefs and mangroves, two globally important coastal ecosystems, contain many biological compounds that remain undescribed. It is essential to characterize the metabolome of these habitats to understand the role of these ecosystems in global biogeochemical cycles. Our approach detected 295 and 428 metabolite peaks from coralline and mangrove sediment profiles (**Fig. 2 c, d**), including sugars and amino acids, organic acids, fatty acids, and background signals. Diverse and abundant sugars from coralline sediments as well as fatty acids from mangroves drove the observed significant differences between habitats (ADONIS p-value < 0.001, R^2 = 0.514; **Supplementary Fig. 5 a, b, Supplementary Table 5**). Considering corals and mangroves thrive in oligotrophic waters and their associated sediments harbor diverse and abundant microorganisms, we were surprised to measure high abundances of energy-rich metabolites that are typically consumed in primary metabolism. These data call for a reexamination of carbon sequestration in coastal sediments using techniques that can identify and quantify the accumulation of liable metabolites.

Due to the technical difficulties of detecting metabolites in seawater, a large portion of ocean chemistry remains unannotated, reflecting one of the central challenges in metabolomics research[18]. By providing a new method to measure a broad scope of the marine metabolome, we offer an avenue to identify molecules from marine environments and expand existing mass spectrometry databases that aim to characterize chemical space across ecosystems. As an example, our samples from sediment porewaters of mangroves and coral reefs revealed 11 metabolites driving variation between habitats that did not match public database entries (**Supplementary Table 5)**[19, 20].

We have presented SeaMet, a marine metabolomics workflow that enables the analysis of primary metabolism in the oceans. It is time efficient, allows the detection of diverse metabolite classes in a single

run, and expands the analytical window for molecules that can be detected within marine samples. This advance enables untargeted metabolomics for marine ecosystems using a low-cost, easy to use GC-MS platform. Moreover, SeaMet is independent of MS instrumentation, allowing it to be combined with time-of-flight MS or Orbitrap detectors to provide faster analysis time and higher mass resolving power to improve metabolite identification. We expect our marine metabolomics workflow will enable the exploratory analysis of untargeted metabolites occurring in seawater and thereby advance our understanding of the ocean's vast and largely unexplored metabolome.

Methods

Methods and associated references are available with the online version.

Acknowledgements

We thank T. Gulstad, K. Caspersen, F. Fojt, and M. Meyer (MPI-Bremen) for support with data acquisition and sample preparation, N. Böttcher (Jacobs University Bremen) for providing culture samples, and D. Michellod, B. Geier and J. Beckmann (MPI-Bremen) for valuable discussions. We acknowledge the Max-Planck Society and the Gordon and Betty Moore Foundation (Marine Microbial Initiative Investigator Award to ND, Grant #GBMF3811) for financial support. This work is contribution *XXX* from the Carrie Bow Cay Laboratory, Caribbean Coral Reef Ecosystem Program, National Museum of Natural History, Washington DC.

References

1 A, J. et al. *Anal. Chem.* **77**, 8086-8094 (2005).

2 Fiehn, O. et al. *Nat. Biotechnol.* **18**, 1157-1161 (2000).

3 Theriot, C.M. et al. *Nat. Commun.* **5**, 3114 (2014).

4 Bruce, T.J. et al. *PNAS.* **105**, 4553-4558 (2008).

5 Fischer, E. & Sauer, U. *Euro. J. Biochem.* **270**, 880-891 (2003).

6 Villas-Boas, S.G., Noel, S., Lane, G.A., Attwood, G. & Cookson, A. *Anal. Biochem.* **349**, 297-305 (2006).

7 Swenson, T.L., Jenkins, S., Bowen, B.P. & Northen, T.R. *Soil Biol. and Biochem.* **80**, 189-198 (2015).

8 Moran, M.A. et al. *PNAS.* **113**, 3143-3151 (2016).

9 Sunagawa, S. et al. *Science* **384** (2015).

10 Koch, B.P., Ludwichowski, K.-U., Kattner, G., Dittmar, T. & Witt, M. *Mar. Chem.* **111**, 233-241 (2008).

11 Petras, D. et al. *Front. Mar. Sci.* **4** (2017).

12 Dittmar, T., Koch, B., Hertkorn, N. & Kattner, G. *Limnol. Oceanogr.* **6**, 230-235 (2008)

13 Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N. & Willmitzer, L. *Plant J.* **23**, 131-142 (2000).

14 Chan, E.C., Pasikanti, K.K. & Nicholson, J.K. *Nat. Protoc.* **6**, 1483-1499 (2011).

15 Kind, T. et al. *Anal. Chem.* **81**, 10038-10048 (2009).

16 Johnson, W.M., Kido Soule, M.C. & Kujawinski, E.B. *Limnol. Oceanogr.* **15**, 417-428 (2017).

17 Behrends, V., Ebbels, T.M., Williams, H.D. & Bundy, J.G. *Appl. Environ. Microbiol.* **75**, 2453-2463 (2009).

18 Luzzatto-Knaan, T. et al. *Elife* **6** (2017).

19 Kopka, J. et al. *Bioinformatics* **21**, 1635-1638 (2005).

20 Lai, Z. et al. *Nat. Methods* **15**, 53-56 (2018).