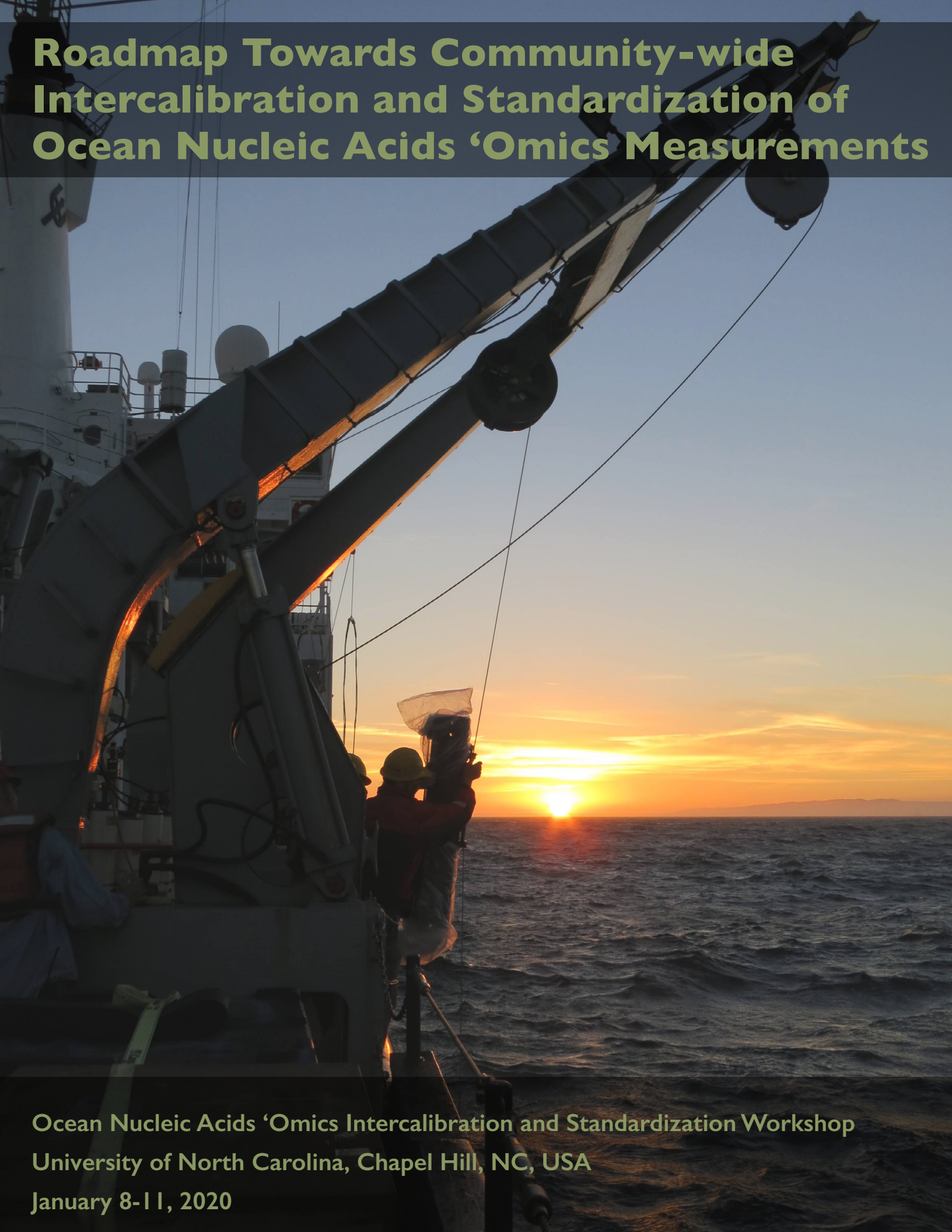
Roadmap Towards Community-wide Intercalibration and Standardization of Ocean Nucleic Acids 'Omics Measurements



Ocean Nucleic Acids 'Omics Intercalibration and Standardization Workshop

University of North Carolina, Chapel Hill, NC, USA

January 8-11, 2020

AUTHORS

Paul Berube
Civil and Environmental Engineering Department, Massachusetts Institute of Technology, Boston, MA

Scott Gifford
Department of Earth, Marine and Environmental Sciences, University of North Carolina, Chapel Hill, NC

Bonnie Hurwitz
Department of Biosystems Engineering, University of Arizona, Tucson, AZ

Bethany Jenkins
Graduate School of Oceanography, University of Rhode Island, RI

Adrian Marchetti
Department of Earth, Marine and Environmental Sciences, University of North Carolina, Chapel Hill, NC

Alyson E. Santoro
Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA

EDITORS

Mai Maheigan
Woods Hole Oceanographic Institution

Heather Benway
Woods Hole Oceanographic Institution

BIBLIOGRAPHIC CITATION

Berube, P., S. Gifford, B. Hurwitz, B. Jenkins, A. Marchetti, A. E. Santoro. 2022. Roadmap Towards Community-wide Intercalibration and Standardization of Ocean Nucleic Acids 'Omics Measurements. 50pp. DOI 10.1575/1912/28054

ACKNOWLEDGMENTS

The Ocean Nucleic Acids 'Omics Intercalibration and Standardization Workshop was supported by grants from the Ocean Carbon & Biogeochemistry Program (OCB) – funding provided by the National Science Foundation (NSF) and the National Aeronautics and Space Administration (NASA) – and the Simons Foundation. We thank the University of North Carolina, Chapel Hill for logistics and meeting space. A list of meeting participants and

COVER IMAGES

FRONT: This photo was taken from the IRNBRU cruise in 2014 on the R/V Revelle of Ken Bruland performing his last Go Flo cast, taken by Adrian Marchetti.
BACK: Workshop participants at UNC

SIMONS FOUNDATION

NSF

NASA

OCB Ocean Carbon & Biogeochemistry

Table of Contents

I Summary

In January 2020, the US Ocean Carbon & Biogeochemistry (OCB) Project Office funded the Ocean Nucleic Acids 'omics Intercalibration and Standardization workshop held at the University of North Carolina in Chapel Hill. Thirty-two participants from across the US, along with guests from Canada and France, met to develop a framework for standardization and intercalibration (S&I) of ocean nucleic acid 'omics (na'omics) approaches (i.e., amplicon sequencing, metagenomics and metatranscriptomics). During the three-day workshop, participants discussed numerous topics, including: a) sample biomass collection and nucleic acid preservation for downstream analysis, b) extraction protocols for nucleic acids, c) addition of standard reference material to nucleic acid isolation protocols, d) isolation methods unique to RNA, e) sequence library construction, and f) integration of bioinformatic considerations. This report provides a summary of these and other topics covered during the workshop and a series of recommendations for future S&I activities for na'omics approaches.

2 Workshop overview

Ocean science has now entered the molecular era. Advances in nucleic acid (i.e., DNA and RNA) sequencing over the past few decades have greatly expanded our knowledge on the diversity of marine microbes (i.e., bacteria, archaea and protists). These advances have provided great insights into the critical roles these microbes play in collective ocean metabolism and their profound impact on global biogeochemistry. But despite this surge in the use of na'omics approaches by members of the oceanographic community, there has yet to be a coordinated S&I initiative towards the development of community-wide best practices and to ensure intercomparability across future large-scale field campaigns. Currently, there are no reference materials or standardized protocols for the generation of sequence-based data, and intercalibration efforts have only occurred between a few select labs. All of this uncertainty has generated debate on protocols, introduced skepticism in the interpretation of na'omics data across projects, and stifled international collaboration and data integration. Such S&I efforts have been, and remain critical to the success of international basin-scale programs (e.g., GEOTRACES).

With this in mind, the ocean na'omics workshop was organized to focus on marine microbial na'omics intercomparison and intercalibration efforts. Increasingly, field programs of relevance in the OCB community have major components that use high-throughput molecular barcoding, metagenomics, and metatranscriptomics to understand the functions of prokaryotic and eukaryotic organisms in the ocean. Examples of these programs include the NASA-NSF funded EXPORTS program, the international BioGEOTRACES program that couples these measurements in conjunction with GEOTRACES cruise sampling, and the next wave of potential international field programs that would couple ocean chemistry sampling with 'omics measurements. In fact, a recent OCB supported workshop was held in November 2018 to catalyze discussion within the oceanographic community to develop such a future program that is named BioGeoSCAPES.

The success of the international GEOTRACES program relies on diverse research groups producing measurements that are intercomparable. This is facilitated by a series of intercalibration efforts and standard reference materials developed by the chemical oceanography community over a period of decades. It is clear and urgent that a parallel effort needs to be undertaken by the microbial oceanography community to intercalibrate na'omics measurements with a goal of developing community benchmarks. In order for a program such as BioGEoSCAPES to be successful, a critical first step is intercalibration and standardization of methods for the na'omics measurements.

Many members of the oceanographic community have participated in OCB and EarthCube sponsored workshops where issues of intercomparability of na'omics methodologies have been discussed. Overall, there is a growing recognition across multiple subsets of the microbiome community (e.g., biological oceanography, human microbiome, agriculture, etc.) that the absence of intercomparable data sets is limiting our ability to make large and rapid advancements. Still, incremental progress has been made. Improvements in bioinformatics methods to facilitate post-sequencing data harmonization have been a useful stopgap measure. Methods used for the collection, processing and sequencing of marine microbial samples have generally converged around a common set of practices. Advances in high-throughput sequencing methods have made data generation more cost-effective. And finally, individual researchers have been employing internal and external standards for quantification and quality control. Thus, the timing is right for na'omics S&I

activities, and our community is committed to advancing such efforts.

Prior to the workshop, the organizing committee identified key issues that participants would examine in detail in the context of na'omics S&I activities. These key issues are summarized below.

2.1. Key workshop issues

2.1.1. Sample biomass collection and nucleic acid preservation for downstream analysis

Researchers have used multiple methods to filter biogenic particles from seawater. The four main filtration approaches are by vacuum pump, peristaltic pump, pressurization of head space, and large volume in-situ pumps. There are different filter matrices, construction (e.g., flat vs. capsule), and pore sizes used with variations in filtration time, pressure, and volume. Filters are preserved using numerous storage methods prior to extractions. As RNA (i.e., transcriptomic responses) can be rapidly altered as a function of time following in situ sampling, filtration methods and procedures have important consequences for RNA-based methods. Additionally, biomass in nutrient rich and nutrient poor areas of the ocean can vary widely, which also influences filtration times and yields. A goal of na'omics S&I activities would be to develop a series of recommendations for filtration protocols and times that are appropriate for different ocean environments and microbial communities.

2.1.2. Extraction protocols for nucleic acids

Nucleic acid extraction encompasses cell lysis and purification of the released material from other cellular and environmental compounds that may interfere with downstream applications. There are variations in methods for nucleic acid extraction that range from protocols with laboratory made reagents to commercially available kits that may utilize proprietary reagents. Different extraction protocols are meant to optimize extraction efficiencies across organisms. Additionally, cost and throughput of the methods are worthy considerations. Some methods use mechanical shearing to disrupt cells – these choices have downstream impacts on the suitability of the purified nucleic acids for different sequencing platforms. Sequencing technologies have been rapidly changing and current methods range from sequencing very short fragments of DNA to newer methods that can sequence long contiguous DNA molecules. Thus, extraction methods that employ mechanical shearing may impact quality of material for some sequencing platforms. Some researchers have been directly involved in comparisons of nucleic acid isolation methods and a refined set of protocols can be tested in a future intercalibration effort.

2.1.3. Addition of standard reference material to nucleic acid isolation protocols

There are various methods that utilize the addition of known quantities of nucleic acid standards to samples prior to extraction as a way for normalizing extraction efficiency. These standards can include whole genomes from known organisms or engineered pieces of DNA that do not correspond to any living organism. Standards can be used in all na 'omics pipelines. This is a rapidly developing area for recommendations for best practices and an area where our community can learn from members of the human microbiome community that have been converging on the use of exogenous reference materials to improve the interpretation of downstream data products. This is an area in which the microbial oceanography community would benefit from discussion with members of

different microbiome research communities.

2.1.4. Isolation methods unique to RNA

Total RNA isolated from a biomass sample is a matrix of highly abundant structural molecules (ribosomal RNA) and low abundance molecules derived from protein coding genes (mRNA) – both of these RNA molecules contain metabolic information and so are often sequenced separately or in combination. There are methods that enrich for mRNA through the removal of rRNA molecules. Other methods take advantage of the presence of rRNA for assessing taxonomy. Additionally, RNA molecules from prokaryotes and eukaryotes have sequence differences that are utilized in isolation protocols making it possible to enrich for eukaryotic mRNA over prokaryotic mRNA and total RNA. Often these removal steps are part of the sequencing library preparation and conducted at a sequencing facility outside a user's laboratory. The fraction of rRNA depleted during these steps can be variable and dependent on the sequence composition of the rRNA pool. Methods and reagents are also subject to the decisions of manufacturers to discontinue kits utilized for sequence library construction (see below). One goal for the workshop was to assess the state of the art for RNA based methods and a series of recommendations that are feasible across laboratories.

2.1.5. Sequence library construction

Once nucleic acids are isolated, they are subject to downstream protocols that prepare the isolated material for sequencing on high throughput platforms (e.g., Illumina, Oxford Nanopore). Generally, these library preparation methods are performed outside an individual user's laboratory and are often contracted through the same center that conducts the sequencing. The library preparation methods are optimized for the type of sequencing selected but different centers utilize different protocols. As this landscape is changing rapidly, how these variables impact data can be tested by providing different centers with different methods on identical samples. We would envision this activity to be encompassed in a funded intercalibration effort but would utilize the workshop to make recommendations about how many center technologies to test.

2.1.6. Considerations unique to amplicon barcoding methods

Community barcoding methods typically utilize PCR amplification with primers universal to prokaryotic or eukaryotic taxa. Target regions on universal molecules can differ as can methods for standardization. This workshop developed recommendations for intercalibrating barcoding methods.

2.1.7. Integration of bioinformatic considerations

While the workshop did not focus on specific differences in bioinformatic pipelines, members of the marine 'omics bioinformatic community were invited to participate as they work with data generated by many different methods and may have recommendations for standardized pipelines by which to verify intercalibration.

An ultimate goal of the workshop was to conceptualize what a na'omics intercalibration program would entail. We developed plans for a concerted na'omics S&I effort that combines optimization and comparison for multiple field and laboratory sampling protocols. As we describe recommendations in this workshop report, it may appear that the number of methodological variables to examine is extremely high, but we feel that by leveraging our community's combined expertise and converging

agreement on methodological considerations, we can develop recommendations for an S&I framework and best practices for the na'omics approaches. We would like to highlight that our goal isn't to develop one prescribed set of methods but rather a best practices framework that could be tested in a formal intercalibration effort.

At the workshop, there was also some discussion on whether or not we want to try and calibrate methods (i.e., focus on standards to improve accuracy of measurements as much as possible) vs. harmonize methods (i.e., focus on making measurements/data consistent and compatible to facilitate intercomparison). Calibration is going after the "true" value and may be more difficult and costly. Harmonization might sacrifice some accuracy for the sake of making the data outputs as compatible with each other as possible. Calibrated measurements with high tolerances would entail expensive and or cumbersome sampling and analysis methodologies. Less accurate (but still acceptable) measurements that yield consistent data may be desirable if deployment is more widely achieved.

As with any S&I effort, defining a reasonable range of variability (e.g., derived from unconstrained biases in sampling, storage, extraction) is an important consideration. There may be a cost vs. benefit trade-off in achieving an acceptable degree of variability (e.g., how exhaustive does sequencing need to be to achieve a certain range of variability?). Community buy-in for widespread adoption of na'omics S&I protocols will require some degree of agreement on the amount of variability that can be tolerated between measurements to identify whether or not they are intercalibrated. Work will need to be done across the broader community to assess the most appropriate metrics and cutoffs.

3 Workshop participants

Workshop participants were from 14 research and academic institutions from across the US as well as two academic institutions in Canada (See list of participants in appendix A). Also present was a representative from the Tara Oceans Program. A pre-workshop survey was conducted to obtain participant demographics and examine most commonly implemented na'omics practices across the represented research groups. Of the 32 participants that attended the workshop, there was a diversity of positions and ranks represented; approximately one-third were full professors, a third were either associate or assistant professors, and the remainder made up of research scientists, post-docs and graduate students (Fig. 1). The most common self-identified primary research expertise of the participants included microbial ecology, microbial biogeochemistry and biological oceanography. Thirteen participants identified their primary research expertise in genomics/ genetics and eight in bioinformatics (Fig. 1). Workshop participants were also closely divided among focus organisms where 18 participants primarily examined heterotrophic bacteria and 14 participants focused on either autotrophic protists or bacteria. To a lesser extent, other organisms

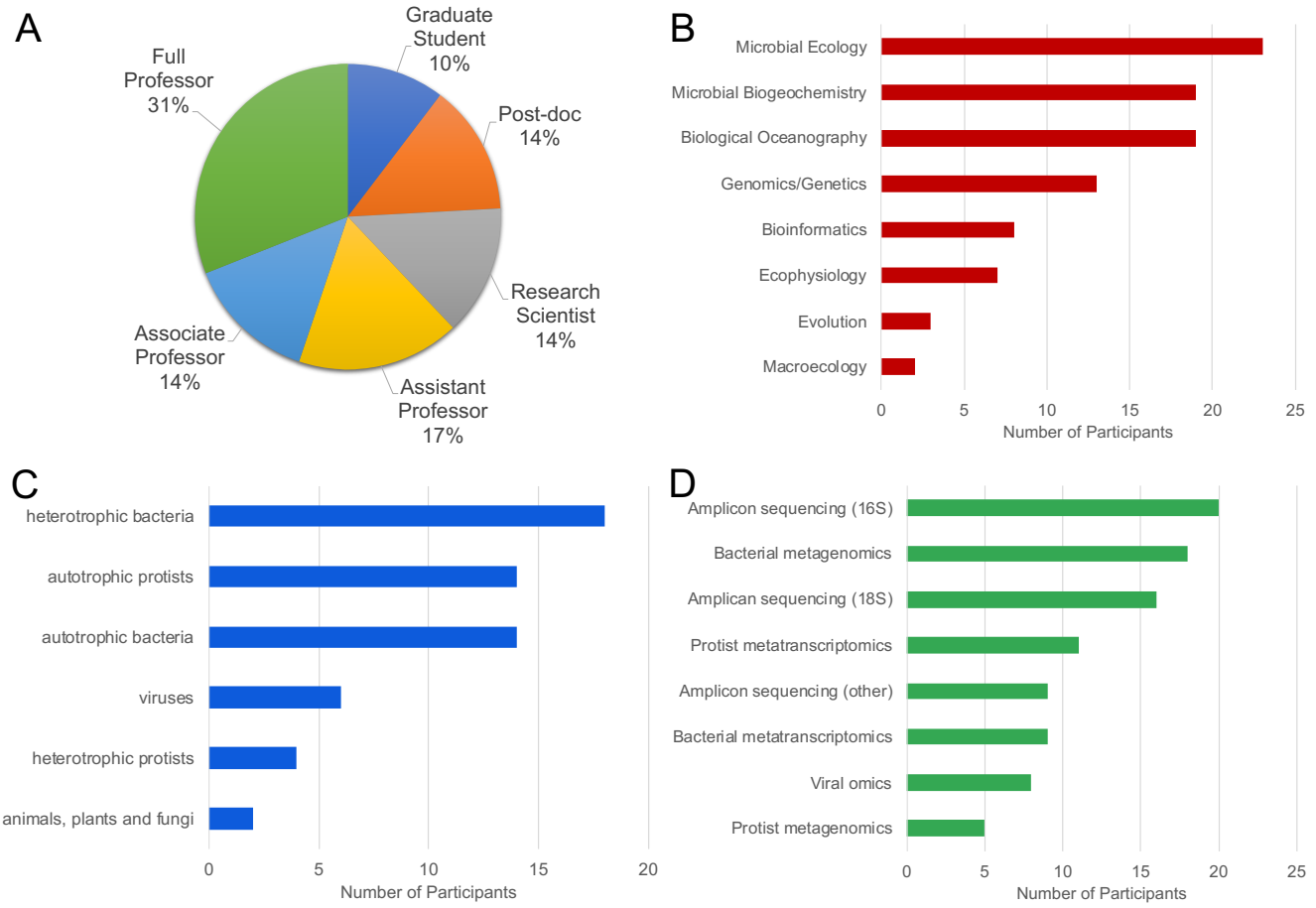


Figure 1. Workshop participant demographics including A) professional positions, B) primary research expertise, C) primary focus research organisms, and D) commonly employed molecular approaches in natural marine environments (n=25).

of interest by participants included heterotrophic protists, animals, plants, fungi and viruses. The majority of participants identified their primary research region as either open ocean or coastal ocean environments, where 16 participants focused on the euphotic zone and 12 participants on mesopelagic environments. A fewer number of participants identified their primary research locations as either estuaries or benthic habitats.

Regarding na'omics approaches, 20 and 16 participants had experience with DNA amplicon sequencing for 18S rRNA or 16S rRNA genes, respectively. Eighteen participants had experience with bacterial metagenomic sequencing and 11 and 9 participants had experience using metatranscriptomics approaches for protists and bacteria, respectively (Figure 1). A smaller number of workshop participants had experience with either protist metagenomics or viral 'omics. Many of the workshop participants have conducted some level of research and development of protocols for na'omics sample collection and preservation. The majority of participants performed na'omic extractions in their laboratories and then used high throughput sequencing facilities for sequence library preparations and/or sequencing. The most common sequencing platform used was Illumina.

4 Prior intercalibration and intercomparison efforts

The oceanography community has embarked on several historical intercomparison efforts, mostly in chemical oceanography. To name a few, significant effort went into establishing best practices for both dissolved organic carbon (DOC; Sharp et al. 2002), dissolved inorganic carbon (DIC; Dickson et al. 2003), and macronutrients (Becker et al. 2020). Perhaps the most significant intercomparison program in recent years has been that of the trace metal community, an effort necessary for launching the international GEOTRACES program (www.geotraces.org), designed to map the ocean for its trace metal distributions to understand sources and cycling of these nutrients in the ocean. Both the scale and success of GEOTRACES suggest its approach to intercomparison merits a closer look and was a major reference point during our workshop.

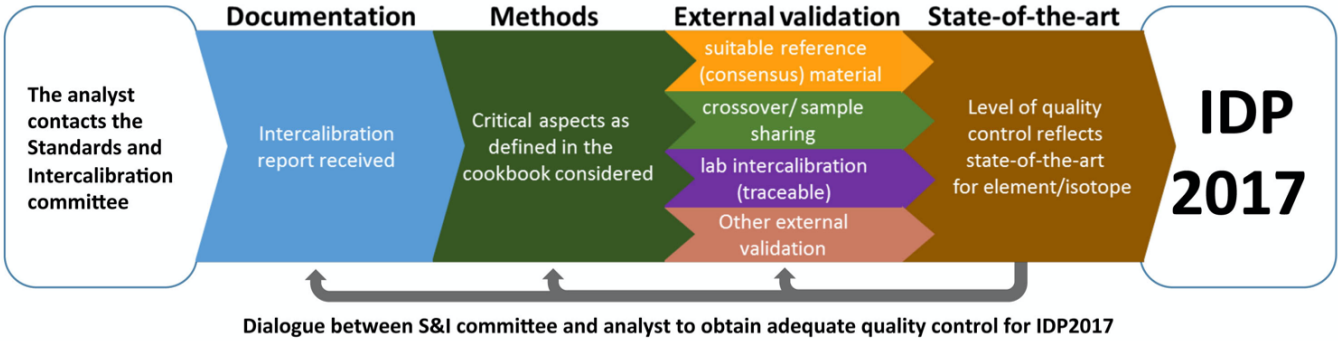


Figure 2. An example of a standardization and intercalibration pipeline as developed for the Intermediate Data Product (IDP) for the GEOTRACES program (Schlitzer et al. 2018).

4.1. Intercomparison and International GEOTRACES

The International GEOTRACES program is one of the best examples of an oceanographic community committed to the intercomparison and validation of sampling and analysis methods (Cutter 2013). The goal of GEOTRACES is "To identify processes and quantify fluxes that control the distributions of key trace elements and isotopes (TEIs) in the ocean, and to establish the sensitivity of these distributions to changing environmental conditions." Recognizing that a global mapping of trace elements in the ocean could not be accomplished by any single laboratory or nation, GEOTRACES embarked on an extensive intercalibration effort to ensure that measurements taken by any intercalibrated lab could be reliably compared to other intercalibrated labs and then included in a public database of measurements.

The key elements of the GEOTRACES intercomparison strategy were to: establish reference materials, designate elemental coordinators for individual measurement parameters, and establish a "cookbook" of approved methods. Reference materials were obtained with large volume collections of seawater presumed uniform in dissolved constituents and the deployment of large volume in-situ pumping arrays to obtain large amounts of particulate matter on filters. These collections were

done on two oceanographic cruises – one in the Atlantic and one in the Pacific – that each obtained materials from a coastal and an open-ocean location. Several formal intercalibration studies were published from these initial cruises for both dissolved (e.g., Anderson et al. 2012, Grasse et al. 2017, Boyle et al. 2012) and particulate TEIs. Scientists from the international community serve as elemental coordinators (https://www.geotraces.org/elemental-coordinators/) for distributing reference materials and collating data from labs wishing to intercompare their analytical methods that did not participate in the initial intercalibration studies. The cookbook (https://www.geotraces.org/methods-cookbook/), currently in its third edition, is reviewed and updated periodically.

Intercomparison and data quality assurance is an ongoing process within GEOTRACES. All GEOTRACES section cruises have at least one overlapping "crossover" station at the same coordinates of a station on a prior cruise to allow for intercomparison of measurements between sampling platforms and analysts. An international Standards and Intercalibration (S&I) Committee reviews, in person, an S&I report prepared by any investigator that wishes to have their data included in one of the GEOTRACES Intermediate Data Products (IDP). The S&I committee works with investigators in an iterative process (Figure 2) to ensure that data were generated and analyzed using protocols established in the cookbook, are referenced to certified or other reference materials, and that data from crossover stations are within measurement error. The entire S&I committee then votes on the inclusion of the data in the IDP.

4.2. National Institute of Standards and Technology (NIST) Microbiome Standardization

The human microbiome community has already recognized the need for reference materials to ensure intercomparability of studies from different research groups (Stulberg et al. 2016) and several current efforts serve as potential models for a future marine microbiome intercomparison study. In addition, the lack of certified reference materials limits the ability to conduct clinical trials to test the efficacy of potential microbiome-based treatments and diagnostics for human health. NIST has partnered with the Mosaic community platform (https://mosaicbiome.com) to launch the Mosaic standards challenge (Jackson 2019). Participating laboratories receive five homogenized fecal samples and two synthetic microbial mock community DNA extracts that they can then process using their laboratory's normal extraction, amplification, and sequencing protocols. The data and associated metadata are then uploaded to a community portal where it becomes publicly available to the community along with all metadata. Importantly, these metadata include information on experimental protocols at a high level of granularity to facilitate identification of sample processing variables that influence the final data product. The project has invested in a significant web presence to facilitate community data sharing and communication (https://platform.mosaicbiome.com/challenges/8).

4.3. Intercomparison efforts in the environmental molecular biology and microbiology communities

Table 1 provides a summary of previous intercomparison efforts of various na'omics protocols for environment sequencing. Although this list is not exhaustive, it does provide the most relevant

studies for the proposed S&I activities to date.

Table 1. Selected prior intercomparisons of marine microbiome sampling and analysis methods

Variables tested	Target community	Variables impacting community composition	Environment	Reference
Filtration volume, filter pore size	Bacteria	Sample volume, volume x filter pore size	Marine water column	Padilla et al. (2015)
DNA extraction	Microbial eukaryotes	Extraction method	Marine sediments	Lekang et al. (2015)
DNA extraction	Bacterial diversity	Extraction method	Marine sediments	Luna et al. (2006)
Sample volume, carrier DNA	Prokaryotic qPCR	Sample volume, carrier DNA	Marine water column	Boström et al. (2004)
Filter type, DNA extraction	Multiple trophic levels (bacteria, eukaryotes, vertebrates)	Extraction method	Marine water column	Djurhuus et al. (2017)
In situ vs. bottle collection, filtration methods, filter type	Bacterial and eukaryotic diversity	Filter size, in situ vs. bottle collection	Oxygen minimum zone water column	Torres-Beltrán et al. (2019)

5 Internal and external standards

Processing and analysis of na'omics data is prone to several challenges that limit usefulness of the data. The most common of these include: 1) lack of quantitative data (i.e., the results are proportional in nature) 2) biases from the library construction and sequencing that can skew the results and 3) biases in bioinformatic quality control steps and read processing that can also influence results and interpretation. These challenges can be overcome by the use of both internal and external standards that are composed of material from a single strain or from more complex mock communities. There is a need to standardize the types of materials used and the protocols for these additions to samples along with their downstream analysis.

5.1. Internal standards

Most na'omics data is analyzed in terms of relative abundances; that is, read counts are normalized as a proportion of either the total community pool, or sometimes in the case of transcriptomics, as a proportion of transcripts binning to a particular taxon. The addition of DNA or RNA of known sequence and quantity into a sample prior to nucleic acid extraction enables enumeration of reads originating from the standard after sequencing and then the calculation of a recovery ratio (reads recovered over standard molecules added). The recovery ratio enables one to back calculate the total reads in the sample, as well as the absolute abundances of individual genes or transcripts. For aquatic samples, by dividing this value by the volume of water filtered, one can then calculate the volumetric abundances of a gene or transcript in each sample. This technique thus enables 'omics data to be placed on a quantitative scale that can be compared with other oceanographic rate and standing stock measurements, as well as integrated into ocean models.

Workshop participants recommended that internal standards should be added to every nucleic acid sample, prior to cell lysis. The internal standards can be alien nucleic acid material such as extremophile genomic DNA or synthetically designed sequences. The aim is that internal standards should have little to no homology over the entire length of the sequence with native organisms from the sampled environment. This gives the standard a unique signature to distinguish it from the native reads in a sample. Multiple different internal standards should be added independently to the sample. This helps control for human and pipetting error. Further, by adding the standards at different concentrations one can see how sequencing depth affects read recovery (linearity of recovery) as well as identify any processing or sequencing biases resulting from difference in sequence composition between the standards. Relatedly, nucleic acid processing and library preparation often introduce fragment length biases, and the participants therefore suggested adding internal standards of different fragment lengths to control for how these biases affect sequencing results and quantification.

Similar to purified nucleic acids, whole cells can also be added to the sample before sample storage to assess sample preservation and lysis efficiency. These "live" internal standards should be absent in the native community under analysis (e.g., for the purposes of marine systems, organisms from terrestrial systems are options). These cell-based internal standards can be added as a single strain or as a set of different strains with variable susceptibility to lysis during filtration and subsequent DNA

extraction. Trade-offs exist between internal standards based on purified nucleic acids versus whole cells. The former enable more precise quantification of recovery ratios but do not account for sample storage conditions or differential lysis efficiencies. The latter requires assumptions to be made about the number of chromosomes per cell in order to estimate recovery ratios, but can still be useful as an internal standard to assess performance of filtration, storage, and extraction protocols.

Some hesitancy persists within the community surrounding the addition of internal standards to samples, partly due to the concern that the standards will dominate the downstream sequence library. However, this has rarely been observed in prokaryotic metagenomes, metatranscriptomes, and amplicon datasets. Current sequencing depths allow internal standards added at <1% of total nucleic acid mass to be recovered efficiently for accurate quantification. The participants suggested that protocols be developed for determining the optimal amount of internal standards added to eukaryotic samples that would enable accounting for substantial changes in eukaryotic cell abundances (e.g., phytoplankton blooms).

5.2. External Standards

It is commonly assumed that current sequencing technologies will produce similar results across different sequencing runs and that any potential sequencing biases are known and controlled for. However, several studies and discussions during the S&I workshop highlighted that different sequencing runs can produce variant community compositions and spurious artifacts that would not be detected without a controlling factor (Yeh et al. 2018). External standards (i.e., mock communities) are one approach to validate that sequencing runs are comparable and biases are controlled for. Cell based external standards are artificial assemblages of different organisms combined in a known proportion, such that the relative proportion determined from the sequence dataset can be compared to the actual community composition to identify any potential DNA extraction, library preparation, or sequencing artifacts that might skew the results. External standards are often designed to mimic natural assemblages in order to best replicate biases that might arise in the sequencing of the target sample. Given their close sequence similarity to the targeted natural community, external standards are processed separately from the environmental samples. In other words, external standard DNA is not added directly to the environmental sample (as is done for the internal standards above), instead it is processed and barcoded independently before pooling with the environmental sample library preps or even sequenced on its own independent lane.

The workshop participants recommended that every sequencing run contain an external standard. The ideal external standard would contain 20 to 30 different organisms that are either mixed in an even proportion or staggered in proportions that mimic the organisms' relative abundance in the marine environment. The participants noted that external standards can be made of different types of material, including custom or commercially synthesized plasmids (both circular or linear), amplicons, genomic DNA, and possibly whole cells; there was discussion on the pros and cons of each. Metagenomic external standards should be made from genomic DNA, while amplicon studies have more leeway in the type of material used for external standard construction. There was general agreement that the organisms in external standards should be as similar as possible to marine community organisms and that separate external standards should be made for both prokaryotic

and eukaryotic assemblages. Validated external standards should be made available to the greater marine 'omics community for optimization of procedures and intra-lab calibration. Cell based external standards could be stored and distributed through institutions such as the National Center for Marine Algae and Microbiota (NCMA). The participants recommended that rigorous testing of external standards be conducted before their implementation. This included sequencing of even and staggered communities to determine if relative abundances affect interpretation and testing for genome copy number artifacts that might arise in the genomic external standards.

6 Common and approach-specific protocols for standardization and intercalibration

Variables identified as important to assess in the context of standardization and intercalibration for each of the na'omics approaches were separated into the following categories: a) seawater collection, sample filtration and storage, b) internal and external standards, c) DNA/RNA extraction and sequencing and d) sequencing analysis. Some variables are common to all three approaches whereas others are specific to each approach. An overview of variables examined is provided in Fig. 3.

	Amplicon Sequencing	Metagenome Sequencing	Metatranscriptome Sequencing
Seawater collection	• Sample collection methods – Niskin bottle, PTFE diaphragm pump, autonomous underwater vehicle • Length of time between seawater collection and sample processing – impact on homogeneity of seawater and DNA/RNA content • Volume considerations to detect rare taxa		
			• Time of water collection; i.e., diel cycles • Seawater dwell time before filtration
Sample filtration & storage	• Filter materials and pore sizes – polyethersulfone versus polycarbonate and the impact of size fractionation • Filter discs versus filter capsules (e.g. Sterivex) • Flash freezing or slow freezing ; impact of storage temperature ; long-term sample stability		
	• Impact of RNA preservatives on DNA yields		• RNA preservation chemicals • Sample processing time
Internal & external standards	• Quality control of standard production, concentration, and storage • Optimal point of workflow to add internal standards (e.g., prior to sample filtration in the field or before extraction in the lab) • Selection, storage, and distribution of community accessible reference material		
	• DNA or cell-based internal standards • Primer site diversity of internal standards • Plasmid-based, genomic, or whole cell external (mock community) standards.	• Genomic DNA or cell-based internal standards • Design and performance of external (mock community) standards	• Type of internal standard • Diversity of internal standard sequences • Feasibility of external (mock community) standards
DNA/RNA extraction & sequencing	• Effect of DNA/RNA extraction methodology on yields and observed sequence diversity – e.g., chemical, mechanical, or enzymatic cell lysis ; extraction kits versus phenol/chloroform ; and concentration/purification of DNA/RNA extracts • Selection of library preparation technology and inter-comparability of sequencing platforms		
	• Taxonomically informed primer selection • Selection of minimal sequencing depth	• Impact of biases associated with mechanical versus transposon-based fragmentation • PCR-dependent versus PCR-independent approaches	• Impact of poly-A selection or rRNA subtraction on observed sequence diversity
Sequence analysis	• Sequence quality control • Metrics to assess precision and accuracy of DNA/RNA omics data ; metrics for documentation of intercalibrated data • Selection of appropriate reference databases for taxonomy and function		
	• Assessment of missing taxa • Detection limits for rare taxa	• Assembly, mapping, and annotation • Metagenome assembled metagenomes • Concentration of taxonomically resolved genome equivalents	• Gene expression analysis

Figure 3. Variables for considerations for standardization and intercalibration of amplicon sequencing, metagenomics, and metatranscriptomics.

6.1. General Considerations

6.1.1. Biomass filtrations

Sample collections of seawater commonly involve deployment of Niskin bottles on a CTD rosette and emptying water into a collection vessel or filtering directly off the rosette using gravity or positive pressure. Surface seawater piped through ship underway systems is often sampled, particularly for research questions related to the spatial distribution of organisms in the surface ocean. Underway systems may vary in how often they are cleaned (i.e., there may be biofilm growth in the plumbing) and how well the cell integrities of diverse organisms tolerate shear forces in different seawater distribution systems. Trace metal clean Towfish instruments and diaphragm pumps can also be used for large collections of uncontaminated surface seawater. Newer methods such as in situ McLane pumps or filtering AUVs like Clio (Breier et al. 2020) are also being used. Shipboard filtration of collected water with peristaltic pumps is one of the most widely deployed filtration methods amongst workshop participants, particularly as they can be set up with connected in-line filters used for size fractionation. Other positive pressure methods are employed as well. Investigators interested in concentrated biomass from larger plankton size classes also use net tow samples with mesh sizes ranging from 150-330 µm.

Pumps: Although peristaltic pumps are one of the most commonly used filtration methods in the na'omics community, vacuum pumps are often the most accessible and used for common, routine filter-based sampling methods (e.g., chlorophyll analysis). Vacuum pumps may be the easiest for others in the oceanography community to use when collecting na'omics samples opportunistically in conjunction with other biogeochemical sampling. One way vacuum pumps could be used to accommodate in-line fractionation with capsule filters (e.g., Sterivex filters) is to use vacuum manifolds (e.g., see here). Several members of our community use systems with headspace and positive pressure as these can be relatively inexpensive to set up with an aquarium pump as opposed to specialized equipment, but they require bottle modifications and familiarity with the systems.

Volume: Sampling volumes required depend on the organisms of interest. For instance, there may be planktonic organisms that are rare but important in an ecosystem. It would be advantageous for our community to be able to sample more water than typical constraints imposed by Niskin bottles in order to provide sufficient biomass for the detection of these rare organisms. Employing pumps or autonomous underwater vehicle (AUV)-based systems are a possible alternative. A clear advantage to an AUV system is that it liberates wire time on a ship but it also requires berths for engineers to deploy it.

6.1.2. Filtration Matrices and Size Fractionation

The most commonly used methods amongst workshop participants for collecting biomass from whole seawater is direct collection onto 0.22 µm Sterivex capsule filters or 0.2 µm PES (Supor or Durapore) filter discs. Other filter matrices used for the 0.2 µm fraction include cellulose nitrate and polycarbonate. Some groups pre-filter water through a larger mesh (80-153 µm) before biomass collection to remove larger organisms (mostly macrozooplankton). Often size-fractionation methods are used with an upstream filter that varies in pore size (e.g., 10, 5, 3, 1.1, 1 µm) and filter matrix (polyester, polycarbonate, polyethersulfone (PES), glass fiber filters). Investigators studying viral dynamics collect the filtrate from the 0.2 µm fraction onto a 0.02 µm Anodisc (aluminum oxide, 25 mm diameter) connected in series, others collect the filtrate from 0.2 µm fraction and use a FeCl3

flocculation method and collect the flocculate onto a 0.45 µm filter.

Global ocean sampling programs like TARA routinely deploy size fractionation in sample collections to separate microbial groups based on cell size. The TARA Oceans Global sampling program has used the following size fractions i) enriched in prokaryotic viruses: <0.22 µm; ii) enriched in virus and prokaryotes: 0.1-0.22 µm, 0.22-0.45 µm, 0.45-0.8; iii) enriched in prokaryotes: 0.22-1.6 µm or 0.22-3 µm; iv) enriched in protists and metazoans: 0.8–5 µm (or 0.8–3 µm), 0.8–2000 µm, 5–20 µm (or 3–20 µm), 20–180 µm and 180–2000 µm. Many in our community in ocean sampling use two size fractions of a 3 or 5 µm prefilter to capture eukaryotic microbes and larger organisms, and prokaryotes onto a 0.2 µm filter.

Filter type: Some investigators report that PES filters may not be optimal for RNA extraction, whereas it is a preferred filter matrix for many working with DNA. This raises the question of whether there is an optimal filter type for all variables. Still, there was community consensus that PES filters can be a good place to start intercomparison of downstream variables.

Utilization of Size Fractionation: The main advantage of size fractionation is that this approach expands the research questions that can be addressed, for example, separating particle-attached bacteria. It also may be a way to collect more biomass of organisms within a specific size range that would rapidly clog a 0.2 µm filter. For prokaryotic metagenomics, removing the large size fractions also removes eukaryotic organisms with huge genomes that would divert sequencing effort away from the organisms of interest. Workshop participants discussed methods of size fractionation where filters are directly in-line vs. parallel size fractionation where water is directly collected onto a specific larger filter size and filtrate subsequently collected on a smaller pore size. Different filtration setups (peristaltic, vacuum, or positive pressure) favor one size-fractionation approach over another. How do these methods compare with filtration time and organismal integrity? Are they worth the increased analytical burden? What is missed in size fractionation? Alternatively, are there advantages for the detection of some organisms by being able to pass more volume through a larger pore size?

6.1.3. Sample preservation methods

Our community's most commonly used storage method for biomass collected onto filters is rapid flash-freezing right after collection into liquid nitrogen and transfer to -80 °C storage. Filters are also added directly to nucleic acid lysis buffers, flash-frozen in liquid nitrogen, and transferred to -80 °C storage. Other preservatives that help samples be shelf-stable at room temperature or in storage at cool temperatures (4˚C, -20˚C) are also used (RNAlater, DNA/RNA Shield (Zymo Research)). Samples for single-cell genomes are collected in 10% glycerol-TE buffer and flash-frozen in liquid nitrogen.

Our community should assess the impact of sample preservation methods on the sequence data recovered, but the literature to date shows that nucleic acid extraction methods play a much greater role in impacting downstream results (Djurhuus et al., 2017). Several studies have been conducted comparing sample preservation methods, and some found biases in community composition, whereas others did not. One study used known mock microbial communities to compare storage on chemically treated filter cards with liquid preservatives (e.g., RNA Later) and found that the liquid methods outperformed the card-based methods. Still, all had their own biases (Gray et al., 2013). Another study that used mock communities comparing storage materials such as no buffer, kit lysis buffer, or RNA later found no significant influence on taxa richness or evenness (Hallmaier-Wacker et

al. 2018). Another study used human fecal samples and found that RNA-later moderately impacted bacterial and fungal community structures (Angebault et al., 2018). Taken together, these results are encouraging in that using a sample preservative may be protective and probably won't greatly skew community composition. Newer sample preservation methods such as DNA/RNA Shield (https://www.zymoresearch.com/collections/dna-rna-shield) that are more compatible with downstream methods than RNA later are being used by community members and should be included in intercomparisons of sample storage method.

Workshop participants ranked priorities for intercomparison. Regardless of the na'omics method, all participants prioritized size-fractionation and prefiltration first, followed by preservation methods and storage conditions. Filtration method and filter type received a lower prioritization for assessment. I.e., they may be less likely to influence downstream sequence results, and the community has already converged on a small set of practices, thus facilitating harmonization of filtration methods.

6.2. Amplicon Sequencing

Almost all workshop participants had experience with amplicon sequencing for profiling either eukaryotic communities using the 18S rRNA genes or prokaryotes using the 16S rRNA genes. Tables 2 and 3 summarize the most commonly used primers for these loci. Most workshop participants agreed that internal standards for amplicon sequencing are a good idea, and there is a wide range of internal standards being used; general approaches are summarized below.

Table 2. Commonly used universal prokaryotic primers by workshop participants

Prokaryotic universal amplicon primers (16S)	Region	Reference and Notes
515Y (5'GTGYCAGCMGCCGCGGTAA'3); 926R (5'CCGYCAATTYMTTTRAGTTT'3)	V4-V5	Parada et al. (2016); most frequently used in community
V4: 515F-Y above and 806RB 5'GGACTACNVGGGTWTCTAAT-3'	V4	Parada et al. (2016), Apprill et al. (2015)
515F-Y and 806RB (modified 515F/806R)	V4	Wear et al. (2018)
Bacteria-specific B969F = ACGCGHNRAACCTTACC BA1406R = ACGGGCRGTGWGTRCAAC	V6-V8	Comeau et al. (2017)

Table 3. Commonly used universal eukaryotic primers by workshop participants (continues on next page)

Eukaryotic universal amplicon primers (18S)	Region	Notes
1389F 5'- TTGTACACACCGCCC -3' and 1510R 5'-CCTTCYGCAGGTTCACCTAC -3'	V9	Amaral-Zettler et al. (2009); most frequently used in community
1380F 5'-CCCTGCCHTTTGTACACAC-3') and 1510R	V9	Amaral-Zettler et al. (2009)

Eukaryotic universal amplicon primers (18S)	Region	Notes
Euk_1391f (5'-GTACACACCGCCCGTC-3')/EukBr (5'-GTACACACCGCCCGTC-3')	V9	Stoeck et al. (2010)
V8f (5'ATAACAGGTCTGTGATGCCCT-3') and 1520r (5'CCTTCYGCAGGTTCACCTAC-3')	V8-V9	Bradley et al. (2016)
TAReuk454FWD1 (5'-CCAGCASCYGCGGTAATTCC-3',V4 EukV4RB ACTTTCGTTCTTGATYRR	V4	Stoeck et al. (2010), Balzano et al. (2015)
TAReuk454FWD1, V4r (5'ACTTTCGTTCTTGAT-3')	V4	Stoeck et al. (2010), Bradley et al. (2016)
Diatoms: D512F (5' -ATTCCAGCTCCAATAGCG-3') D978R (5'-GACTACGATGGTATCTAATC-3')	Diatom V4	Zimmermann et al. (2011)
HCO2198R TAIACYTCIGGRTGICCRAARAAYCA	metazoan COI	Leray et al. (2013)

6.2.1. Sampling and Extraction.

The na'omics pipelines focusing on DNA all share common issues related to sampling and extraction. Please see the section 6.c.i and 6.c.ii for details related to sampling and extraction in the context of both amplicon sequencing and metagenomics.

6.2.2. Primer Design

Primer design and testing are essential for amplicon sequencing-based approaches. Where possible, workshop participants recommended sets of standard primers for both 16S and 18S rRNA genes. As Tables 1 and 2, show there are variations on primers of choice, but that some are more commonly used. If standard primers exist for a given group (e.g., diatoms, Table 3) they should be recommended. Where possible, information should be added about the taxa each primer set would miss if it is known. Primers are available in different purity levels, and as amplicon primers with Illumina adapters can be quite long, it is worthwhile to order them with HPLC purification. Storage (e.g., freeze-thaw cycles) can impact primer integrity and should be taken into consideration.

6.2.3. Amplicon Library Preparations

The community considers a one-step amplicon library preparation preferable to a 2-step library preparation with addition of adaptors in a second set of PCR as performed by some commercial sequencing facilities. That being said, many researchers use 2-step library preparation as primer costs to an individual laboratory are lower as sequencing facilities have multiplexing barcodes on Illumina indices that are added in a second amplification. The choice of Taq polymerase and its published error rate used for library construction is to be considered, as a high-fidelity Taq will incorporate lower erroneous bases than a regular Taq enzyme.

6.2.4. Sequencing Platform

Choice of sequencing platform is another consideration, comparisons have shown that the Illumina

family of sequencers (MiSeq, HiSeq, NovaSeq) give similar results, but they can vary in coverage of amplicon length. Newer sequencers such as Nanopore MinION are being used for amplicon sequencing, but comparisons between Illumina and Nanopore data are emerging. Studies show platforms largely agree with some showing Nanopore biased against certain taxa in a nasal microbiome (Heikema et al. 2020) and others showing greater taxonomic resolution in a dust microbiome from the MinION compared to the Illumina MiSeq (Nygaard et al. 2020).

Sequencing depth is a consideration for amplicon libraries and typically ~10,000 reads (after QA/QC filtering) is enough to saturate rarefaction. The optimal depth will depend on the environment and types of organisms (e.g., protists may be under-sampled in deep ocean samples), but this is also highly dependent on initial volumes filtered. Workshop participants have found that amplicon libraries can be multiplexed with metagenomic libraries on Illumina platforms. The level of multiplexing and how it impacts read depth needs to be considered.

6.2.5. Internal Standards and Mock Communities for amplicon sequencing

The use of internal standards spiked into samples and external mock community standards that are included as separate libraries on sequencing runs are encouraged for evaluating extraction efficiency (Table 4), amplification biases and sequencing run variation. Researchers in our community have generated external standards from PCR amplicons of cultivated organisms relevant to the environment of interest and combined them in different ratios for 16S rRNA gene amplicon sequencing. Yeh et al., 2018 did this for an Illumina run, spiking in a ~27-member mock community for 16S rRNA gene amplicon sequencing and saw that 1 taxon disappeared completely from both the external standard and field samples and had to repeat the sequencing run. Thus, external standards used as input to PCR amplification and subsequent sequencing are essential to check sequencing run performance (Yeh et al. 2018). Other researchers purchase external mock community standards from commercial sources. Other approaches to lab-designed mock communities are enumerating cells, combining them in different rations, extracting DNA from the different ratio combination samples, amplifying them with the same primers used for samples and inclusion on a sequencing run.

A readily obtainable source of internal standards for normalizing extraction, amplification and sequencing are plasmids – these constructs have target primer sequences flanking an insert of completely synthetic DNA that is designed to not to match anything in a current database. These plasmid inserts can be tuned such that GC content and length is representative of the sample and can be constructed so that they can match multiple primer sets that might be used on a given sample. One challenge is determining the concentration of plasmid to add to the sample so that it does not swamp the signal or is too low. Tests over a range of concentrations are recommended for this approach (Tkacz et al. 2018). Although it can be challenging to construct mock communities for use as external standards, they are recommended at least as external samples to be amplified and included on every sequencing run. Ideally, consistently used internal standards would also be included.

Table 4. Characteristics of Internal and External Standards

Internal standards possibilities/ approaches	Utility	Considerations
Cells (exogenous material)	Capture biases with extraction/ lysis Quantification	Added to sample lysis Need to ensure no-related taxa in sample Labor intensive, might be throughput limited to generate enough material for long-term use
Nucleic acid-gDNA (alien material)	Normalize extraction efficiency Quantification	Can be added to sample lysis Need to ensure no related taxa are in sample, might be material limited or challenging to generate more
Nucleic acid-(synthetic plasmid)	Normalize extraction efficiency quantification	Can be added to sample lysis Primers flank synthetic, artificial sequences with no known matches No limit in supply as plasmids can be commercially resynthesized
External standard possibilities/ approaches	**Utility**	**Considerations**
gDNA extracted from mock communities in different ratios	Test amplification bias Normalization between sequencing runs	Designed to mimic natural assemblage and relevant environment can be generated from known cell numbers 20-30 organisms recommended where possible
gDNA extracted from reference seawater sample(s)	Normalization between sequencing runs	True (complex) natural assemblage Limited availability of sample material which would need to be collected, homogenized, and stored for long-term use and distribution Facilitate run-to-run variability of complex samples and assess precision between laboratories and sequencing centers

6.2.6. Recommendations for an intercalibration experiment

Workshop participants with amplicon experience recommended the following variables as priorities for intercalibration: size fractionation, extraction method, and PCR methodologies. Use of a standard reference sample for intercomparison is the best approach since everyone cannot use the same facility or ensure the same conditions. A recommendation is to use a range of reference samples with

different levels and types of diversity (e.g., coastal vs. open ocean vs. deep sea vs. mock community samples). Samples could have spiked in quantities of taxa (cells, genes, plasmids) that are not marine in origin. A subset of samples for the spike in could be evaluated by qPCR for absolute quantification.

6.3. Metagenomics

Over the past 15 years, metagenomics – the sequencing of an entire microbial community, typically through culture-independent shotgun approaches – has emerged as one of the most widely used measurement techniques in marine microbial ecology. Metagenomics offers a powerful lens to observe both the structure and potential function of marine microbial communities. Although metagenomics methods are now considered robust, realizing their full potential has been hampered by insufficient adoption of standardization and intercalibration protocols by the broader research community. Most marine microbial metagenomics datasets only provide relative abundance information (i.e., no internal standards to facilitate quantification on a volumetric basis) and lack accompanying external mock community controls to assess variation between sequencing runs. The further absence of frameworks for intercalibration of data produced by different labs (using potentially different data production pipelines) has also limited the ability of researchers to compare and combine multiple data sets.

The general lack of standardization and intercalibration protocols during the inception of widespread metagenomics-based analyses was to be expected – initially, high costs associated with sequencing precluded broad adoption of external mock community standard sequencing, replicate sequencing runs, and cross-laboratory intercalibration studies. But, at present, the workshop participants agreed that the biological oceanography community is well positioned to tackle this problem. High-throughput sequencing is now relatively inexpensive (in comparison to 10-15 years ago) and access to sequencing centers (academic, governmental, or commercial) is widely available. Among the workshop participants, metagenomics targeting the Bacteria and Archaea was identified as a measurement pipeline well suited for the development and implementation of standardization and intercalibration protocols. Considerations specific to viruses and the Eukarya, which have additional hurdles, are further detailed below.

Workshop participants recognized that sampling approaches are dependent on the questions, target organisms, and sampling situation (e.g., resources, water budget, and space). Higher order concerns include prioritization of community goals for scientific discovery, e.g., characterization of community diversity versus targeting specific organisms. This information will influence where we sample, what volume of water is collected, the depth of sequencing, and the extraction protocols employed.

6.3.1. Sampling protocols

Workshop participants identified field sampling and storage protocols as among those deserving the highest priority for assessment by the broader community with regard to measurement biases. Given that these upstream procedures ultimately influence all other downstream procedures (e.g., nucleic acid extraction, library preparation, sequencing, and final analyses), sampling and storage was generally viewed as having the greatest influence on the precision and accuracy of nucleic acids 'omics measurements.

General sampling protocols with a focus on Bacteria and Archaea:

1. Filtration methods – peristaltic pump, pressurized headspace, or under vacuum.
2. How do the volumes of seawater filtered affect reproducibility and overall biases?
3. Surface area of filter
4. What effect does pre-filtration and filter capacity have? Are smaller organisms inadvertently captured in the larger size fractions and does the frequency of this increase with higher biomass loads on the filter?
5. Filter types – e.g., Sterivex and flat filters in particular.
6. Filter materials – e.g., polycarbonate or polyethersulfone
7. How long can samples be stored at various temperatures?
8. How long can samples be stored in various preservatives – e.g., RNAlater or sucrose lysis buffer?

Eukarya sampling protocols:

1. Sample volume is an issue; larger plankton are patchier so you will naturally lose elements of the community.

Virus sampling protocols:

1. How do different sampling methods compare? Examples include iron chloride flocculation, tangential flow filtration, and filtration through 0.01 µm pore size filters (mixed cellulose ester for large diameter filters and Anodisc for small diameter filters).
2. Are there biases associated with prefiltration through different pore sizes – e.g., <0.2 µm or <0.1 µm (depending on microbial decision) followed by iron chloride flocculation or filtration through a 0.01 µm filter?

6.3.2. Extraction protocols

Given that there is the opportunity to collect replicate samples and that nucleic acid extraction protocols are downstream of initial sampling, the consideration of variables related to DNA/RNA extraction received a medium-high priority for assessment. On balance, a greater degree of work has already been done by the broader research community in the context of evaluating the biases associated with different extraction techniques (e.g., in the human microbiome research community), and we should leverage this work. Regardless, some issues that are specific to marine organisms likely persist.

General extraction protocols with a focus on Bacteria and Archaea:

1. What is the reproducibility of different extraction methods – e.g. phenol chloroform extraction versus various commercial kits? Can we count on commercial platforms being around for the long-term or will we be forced to continually reevaluate new platforms and make them back compatible with prior platforms?
2. There is a need to evaluate methods for cell lysis, including mechanical, chemical, and/or enzymatic. The diversity of cell structure in these organisms suggests that a single method

might not be suitable for all, but this needs to be rigorously evaluated.
3. How does lysis efficiency vary for different types of cells of Bacteria, Archaea, Eukarya, and viruses?

Eukarya extraction protocols:

1. A number of Eukarya are known to be difficult to lyse. Current methods will need to be evaluated to assess the variability in lysis and extraction efficiency for various eukaryotic taxa.

Virus extraction protocols:

1. Assess the use of protease to remove proteins bound to nucleic acids and the use of chloroform to remove cells.
2. There is a need to determine whether certain extraction protocols are virus-specific; will they enrich for viruses over cellular nucleic acids?

6.3.3. Sequencing and bioinformatics

Workshop participants generally agreed that sequencing (library construction, sequencing platforms, preliminary data processing, etc.) and bioinformatics pipelines for post sequencing analysis were not a bottleneck from a methods perspective. Upstream methods for sampling and extraction were identified as likely to be the source of the greatest methods-specific and laboratory-specific biases. Costs for sequencing and computing resources/time, however, need to be taken into consideration to ensure that adoption of standardization and intercalibration protocols is tractable. It's generally unknown how much run-to-run variability there is for different sequencing platforms, but this can be assessed through the analysis of external mock community standards on each sequencing run at a relatively low cost (e.g., if 20 metagenomes are multiplexed, adding a single mock community metagenome would only modestly increase the cost per bp of the experimental metagenomes).

Eukarya sequencing and bioinformatics:

1. Library construction and sequencing can generally follow the methods developed for marine Bacteria and Archaea as well as those developed for eukaryotic organisms examined in other research fields.
2. Metagenomic annotation pipelines for Eukarya will need further development. In part, this will depend on expanding reference genomic databases.

6.3.4. Internal and external standards

The development of standards for metagenomics has advanced rapidly over the past decade. See Section 5 -- Internal and external standards.

6.3.5. General considerations related to Eukarya metagenomics

Workshop participants with expertise on the Eukarya considered the current and future potential of metagenomics for these taxonomic groups. In general, they observed that there were a number of potential benefits related to metagenomics-based inquiry of the Eukarya. In particular, metatranscriptomics – which is more broadly used for Eukarya – is known to miss a number of

important eukaryotic features. These genomic features include binding sites for RNA polymerase and other transcriptional regulators, transposons, chromosomal heterogeneity, strain level diversity, intron sequences, and DNA methylation patterns. Overall, metagenomics could facilitate gene enrichment studies in abundant taxa, support the improvement of reference databases, and further enhance the interpretation of eukaryotic metatranscriptomics data.

While there are a number of potential benefits that might accrue from eukaryotic metagenomics, there were still questions regarding the viability of this method (i.e., is it worth the effort when function can be gleaned from metatranscriptomics and the abundance and distribution of taxa can be obtained for rRNA gene amplicon sequencing?). The Eukarya have large and complicated genomes in comparison to Bacteria, Archaea, and viruses – thus, the assembly of eukaryotic metagenomic sequencing reads and the depth of sequencing required to observe statistically significant enrichment of gene families present a challenge. Long-read sequencing technologies and the sequencing of single amplified genomes (SAGs) from the same samples may ameliorate some of the hurdles regarding metagenomic assembly. Further, limited numbers of closed genomic reference sequences from cultured organisms inhibits the ability to identify eukaryotic metagenomic sequences through read mapping. Many of these challenges may need to be addressed first before investing the funds and resources on standardization and intercalibration.

6.4. Metatranscriptomics

Recent complete-genome and transcriptome sequencing of marine isolates represent exciting developments that foster the possibility of using molecular and genomic approaches to study biochemical pathways and associated genes in ecologically important marine microorganisms. Metatranscriptomics enables identification of expressed genes of microbes within their natural environment. Because microorganisms have fairly distinct genetic codes, sequences can be used to categorize expressed genes into specific taxonomic groups based on their similarity to existing reference sequences. This can lead to valuable insights into the diversity and ecology of metabolically active organisms in nature and may also be a useful tool in evaluating how these organisms are affected by changes in their environment through examination of targeted expressed genes, gene clusters or entire transcriptome profiling.

There are numerous variations and analyses of RNA-Seq data that have been established. This may result in a source of confusion for new users when becoming familiar with all the steps necessary to conduct RNA-Seq studies properly. A relatively recent review by Conesa et al. (2016) provided a survey of best practices for RNA-Seq data analysis, although this is by no means specific to marine environments, which may have additional challenges due to issues such as sampling limitations and high diversity. At present, there does not seem to be one optimal pipeline, as each may be catered to the specific analysis strategies that are dependent on the organisms being studied and the research goals. Generally speaking, if genome sequences or transcriptomes are available, these can be used to map RNA-Seq reads. Alternatively, for organisms without sequenced genomes, which is often the case, the reads must be first assembled de novo, BLASTed to reference databases for both taxonomy and functional identification, and the gene expression is determined. Conesa et al. (2016) also highlight that every RNA-Seq experimental scenario could have different optimal methods for transcript quantification, normalization and gene expression. They also recommend quality control checks be applied at different stages of the analysis to ensure both reproducibility and reliability of

results.

6.4.1. Sampling protocols

Workshop participants identified filtration, storage, extraction, and rRNA subtraction protocols as key areas for need of standardization and bias determination. Many of these areas directly overlap with the priorities identified for metagenomics (Section 6.c.i). However, the short half life of mRNAs and dynamic regulation of cellular transcriptomes in response to cell collection were identified as high priority areas specific to metatranscriptomes.

General sampling protocols with a focus on Bacteria and Archaea:

1. Filtration methods -- peristaltic pump, pressurized headspace, or under vacuum.
2. How do the volumes of seawater filtered affect reproducibility and overall biases?
3. How does filtration time alter transcriptomes in comparison to when cells are in the environment?
4. Storage time -- What effect does storage time have on RNA quality? Do storage times at -80°C for months or years result in significant degradation? If so, is this degradation severe enough to be detrimental to sequencing success?
5. Preservatives -- e.g. RNAlater or flash freezing in liquid nitrogen?

Eukarya sampling protocols:

1. Due to larger filtration volumes typically required for Eukarya metatranscriptomes, important sources of variability include the larger filter volumes and longer filtration time.

6.4.2. Extraction protocols

General extraction protocols with a focus on Bacteria and Archaea

1. Simultaneous collection -- Can RNA be extracted with DNA without introducing substantial bias or highly reduced RNA yields?
2. Internal standards for absolute quantification and checking for downstream biases.

Eukarya extraction protocols:

1. A number of Eukarya are known to be difficult to lyse. Current methods will need to be evaluated to assess the variability in lysis and extraction efficiency for various eukaryotic taxa.

6.4.3. Sequencing and bioinformatics

1. For bacteria and archaea, rRNA subtraction is considered a necessary step to increase mRNA sequence yields (subtraction typically leads to ten to 50-fold increase in the proportion of mRNA reads in a library). However, there are a variety of rRNA subtraction methodologies (probe based, enzymatic, etc.). Which of these approaches leads to the best mRNA enrichment with the smallest introduction of bias?
2. Normalization - What is the appropriate biological level to compare differences in transcript abundances between samples (i.e., strain, species, genus, or community)? How do different normalization methods, such as DESeq or absolute abundance alter perceived changes in transcription between samples?

Eukarya sequencing and bioinformatics:

1. Either a polyA selection or rRNA depletion is typically performed to increase yields of mRNA. Preferences tend to be lab-specific although which approach leads to the highest yields of mRNA enrichment has not been determined.
2. Normalization – There are similar issues for normalization of eukaryotic sequences as those described for Bacteria and Archaea.

6.4.4. Internal and external standards

Many of the internal and external standard protocols and standardization needs for metatranscriptomics overlap with those for metagenomes (See Section 6.c.iii). However, there were several areas workshop participants identified specific to working with RNA.

1. How should RNA standards be constructed? There are several approaches to doing this (plasmid versus amplicon templates, in vitro transcription, commercial synthesis), and each has potential benefits and downsides.
2. How long and in what aliquots should RNA standards be stored?
3. Does adding naked DNA to a filter and lysis buffer result in over- or under- representation in the total RNA pool?

7 Sequencing and Bioinformatics

7.1. Data management

Data management for the na'omics approaches continues to be a challenge for our community. In addition to the proposed S&I activities, sequence management development will be a priority. The long-term vision motivating our project is to develop nucleic acid extraction protocols for the marine sciences and a comprehensive data resource for each of the na'omics approaches. Specific deliverables include community-accessible: (1) nucleic acid extraction protocols in protocols.io or other open access public repository, (2) raw omics data in Genbank SRA, (3) harmonized metadata using the Open Biological and Biomedical Ontology (OBO) Foundry standards for marine 'omics and environmental data, (4) code in Github for all data analysis, including quality assessment, data visualization, and statistical analyses and (5) final computational workflows that are encapsulated in containers, to enable the community to replicate the code, dependencies, and computational environment and encourage re-use. Finally, all data (including raw and post-processed data) will be consolidated and accessible to the team and community in the CyVerse data store and linked to Planet Microbe to make data searchable and accessible to the community. Furthermore, users will be able to run computational workflows (defined by the team and advisory committee) on any data (including their own) in Planet Microbe using freely-available XSEDE supercomputing resources. These efforts will provide FAIR software and data practices for the marine science community by promoting standardized nucleic acid extraction protocols to create interoperable and comparable datasets that can be analyzed in a consistent manner.

7.2. Sequence bioinformatics

Although examining different bioinformatic approaches and pipelines for sequence analyses will be encouraged, standardization and intercomparisons of bioinformatics pipelines will not be a specific objective of this project. Standardized sequence bioinformatics pipelines used for the processing of our S&I activities will be recommended through the formation of a bioinformatics advising committee composed of project team members as well as others from the broader scientific community.

8 Design of Standardization and Intercalibration Studies

One major goal of this workshop was to leverage the expertise of participants to circumscribe potential research activities aimed at addressing the lack of standardization and intercalibration in marine na'omics measurements. As described in the preceding sections of this report, workshop participants identified a multitude of 'omics data generation variables that are potentially prone to biases – e.g., lack of precision or lack of accuracy/trueness. Following the identification of these concerns, workshop participants worked in teams to design tractable ways of resolving questions related to measurement biases and further develop best practices for intercalibration among marine na'omics researchers. These planning exercises have the potential to drive community-wide standardization and intercalibrations studies that will improve scientific inquiry, enhance the intercomparison and reuse of diverse data sets, and support the planning and development of the next generation of field campaigns.

Workshop participants were randomly assigned to several teams and charged with developing an outline for a na'omics standardization and intercalibration study. The themes emerging from these activities are outlined below beginning with themes common to all teams as well as more specific details on the activities each team proposed. These ideas provide a foundation for the design of studies aimed at assessing bias in marine na'omics measurements and identifying best practices for the generation of intercalibrated and harmonized data products.

General considerations from all team-based discussions:

- The main purpose of any S&I effort should be to understand variability
- As a broader community, we need to decide whether we are searching for the truth or searching for a consensus solution. Are we trying to calibrate measurements (get as close to the true value as possible) or harmonize measurements (ensure that measurements from different labs are intercomparable)? Regardless, we need to decide on baseline standards for acceptable results
- While cell based mock community standards for metagenomics and amplicon sequencing seem relatively straightforward, mock communities for RNA-based experiments are expected to be extremely difficult to develop and may require the use of chemostats
- Reference material for both intercalibration and long-term use would need to be obtained using different seawater samples with contrasting microbial diversity (e.g., coastal, open ocean, surface, epipelagic, mesopelagic, etc.)
- Given the number of variables, hundreds/thousands of distinct filters and nucleic acid samples would be required for any methods development and intercalibration exercise. All teams noted the challenge of designing a study that could be feasibly executed. These limitations necessitated prioritization of pipelines and variables under consideration

- Data needs to be accessible and shared freely among all participants. This will facilitate extended analysis and reuse of data produced during any intercalibration study. These values for data sharing will likely help to build community-wide consensus about the need for the adoption of standardization and intercalibration protocols

- Prior to any S&I activity, consultations with biostatisticians should occur in order to determine what data are needed to evaluate and correct biases in nucleic acid 'omics data generation pipelines. Bias is probably introduced at all steps in a pipeline, but is mechanistic (McLaren et al. 2019). Overall, it will be important to clearly define the statistics used in evaluating variation among samples and come to some consensus on meaningful effect sizes

- Encouragement of international participation would be ideal. There is no obvious barrier to sending filters. It would be helpful to consider identifying international partner collaborations that might be pursuing different aspects of DNA/RNA 'omics methods development

- Data repositories (e.g., BCO-DMO) need to be identified for collecting and serving up metadata. Considerations need to be made for the diversity of data types expected to be produced in any S&I study (e.g., protocol metadata, environmental variables during sample collection, as well as raw and processed sequence data). Further development of cyberinfrastructure should be considered to better deal with these diverse data types specific to marine na'omics and also facilitate community-wide engagement

- Sample archives for filters and extracted nucleic acids will need to be maintained to facilitate the repetition of downstream processing and data analysis steps (e.g., labs should maintain backups for a minimum of 10 years before destroying). These archives can be used to test alternative protocols and assess the impact of long-term storage of filters and nucleic acid materials

- Lab-based seawater systems (e.g., at coastal field stations) will help constrain biological variability for all tests. This will require large volume containers that would enable the homogenization of seawater. Preliminary work would need to be done to identify coastal research sites with the right infrastructure and expertise

Team 1 proposed an iterative, multi-phased approach. Phase 1 would entail the collection of a large number of replicate filters using the same sampling methodology. Filters would be preserved and stored using up to 3 different methods (e.g. RNAlater, sucrose lysis buffer, or frozen on liquid nitrogen) and then sent to a large number of volunteer labs. These labs would extract DNA or RNA from the filters using a common protocol provided by the lead organizer (e.g., phenol-chloroform extraction) and optionally, any lab specific protocol for extraction of nucleic acids. Extracted nucleic acids would be sent to a core facility that would handle further processing and sequencing for amplicons, metagenomics, and metatranscriptomics. The goal of Phase 1 focuses on limiting variability at the sample collection step in order to assess user precision (through use of a common provided extraction protocol) as well as identify alternative lab specific extraction methods that could be leveraged. Phase 2 would entail more specific testing of sample collection variables (e.g., pressurized versus vacuum filtration, filter type and pore sizes, the effect of pre-filtration and fractionation, sample storage buffers, etc.). Finally, Phase 3 would focus on a quantitative

intercalibration exercise that leverages internal and external standards to assess both precision and accuracy. The team recognized that there would need to be significant support for the development of external and internal standards and that this work could proceed in parallel with Phases 1 and 2. This quantitative intercalibration phase would proceed similarly to Phase 1 where external mock community standards, internal standards, and replicate filters from natural communities are distributed to volunteer labs for extraction. Ancillary measurements, such as qPCR of distinct taxa or genes as well as flow cytometry-based measurements of picocyanobacteria, could be further leveraged to help assess the trueness of the 'omics measurements.

Team 2 modeled part of a marine nucleic acids 'omics S&I activity on the U.S. National Institute of Standards and Technology's (NIST) MOSAIC Standards Challenge (Jackson 2019). During year 1, volunteer labs would receive a set of diverse reference samples (e.g., surface seawater, deep seawater, coastal eutrophic seawater, relatively oligotrophic seawater, high biomass filters, and low biomass filters). Samples would be prepared on the same filter type, flash frozen in liquid nitrogen, and stored at ultra-cold temperatures. Volunteer labs would extract nucleic acids from these samples (DNA or RNA) using their own lab specific protocols and providing highly granular metadata on their method. Extracted nucleic acids would be sequenced at a centralized sequencing center to minimize differences related to library preparation and sequencing. Sequencing data would then be compared to identify laboratory protocols that influence variability. During year 2, sampling and storage methods (e.g., vacuum filtration, pressurized filtration, filter types, filter pore sizes, fractionation, storage buffers, and storage temperatures, etc.) would be tested by smaller teams of scientists. After a better understanding of the variability related to extraction and sampling methods has been developed, year 3 would focus on a research expedition to test ship-based operations and their influence on 'omics measurements variability. Some examples of these operations include filtration from pressurized Niskin bottles, filtration using McLane pumps, and the use of automated samplers like CLIO (WHOI) or the ESP (MBARI).

Team 3 had four tiers of testing and standardizing/harmonizing:

1. Conduct a community survey and literature meta-analysis to determine what treatments to test
2. Design of external mock communities and NA spike-ins
    a. Mock Marine Cell Community – e.g., to be used independently (not added to sample) to evaluate the effect of protocols on different taxa
    b. Mock Non-marine Cell Community – e.g., to be used by adding to a marine sample to evaluate methods, such as recovery
    c. Nucleic Acids Standards – for spiking in after extraction
3. Sampling and storage testing
    a. 3-4 sampling platforms: e.g., surface, deep, coastal, NELHA (SW pumped from surface and deep)
    b. >1000 L filtered
    c. Pump type, pre-filtration, filter size classes, filter type (flat vs. sterivex)
    d. Storage method, including buffers
    e. Will need to decide how each method will be evaluated.

4. Extraction to sequence generation (using shared Reference Material)
    a. "Round robin" – Small number of groups generate multiple reference materials and distribute broadly to large number of groups.
    b. Large volume samples to generate reference material. Ship filters.

Team 4 focused on a single at sea research expedition to obtain reference material for an intercalibration exercise and for testing various cruise specific sampling procedures.

1. Preliminary work to evaluate nucleic acids extraction procedures and downstream biases in each na'omics pipeline should occur before any research expedition. Collaborations with other microbiome research communities (e.g., NIST) should be formed to help ensure that currently established best practices in sampling, storage, and extraction are considered -- i.e. while there will be marine specific issues that need to be addressed, researchers studying the marine microbiome should not have to "reinvent the wheel" for more generally applicable methodologies.
2. A research cruise should be conducted in area with a lot of variability in small geographic area, i.e. Santa Barbara Basin or Monterey Bay.
    a. Sample the surface water during the day and the deep water at night to avoid vertical migrators. During each sampling, obtain a very large homogenized sample -- this will require a holding tank suitable for deployment at sea.
    b. 50-80 µm pre-filter onto lots of smaller filters (Sterivex 0.2 µm) – save prefilters too.
    c. Evaluate commonly used storage techniques (e.g., RNAlater, sucrose lysis buffer, storage temperatures).
3. One lab or small group of labs will be tasked with collecting and storing filters for later distribution to various volunteer labs who will use their own preferred extraction method.
4. Distribution of internal standards and mock communities should be considered but it is recognized that there is potentially a need to develop new internal/external standard resources.
5. The extracted DNA/RNA should then be distributed to other volunteer labs with expertise in downstream parts of na'omics pipelines (e.g., amplicon sequencing, metagenomics, metatranscriptomics) for sample processing and submission of materials for sequencing.
6. Informatics – everyone starts with the same mock communities and then go to natural communities. Protocols/workflows would be made available to the scientific community.

Team 5 focused their design on the central goal of simplifying na'omic sampling and understanding the variability among different labs. The development of mock communities and internal standards was deemed essential to addressing these issues. External standards (mock communities) provide samples of known composition to assess accuracy/trueness and internal standards allow for the direct cross-lab comparison of absolute measurements. The following key objectives were proposed:

1. Develop a mock community (cells) and a set of internal standards (DNA) to facilitate method validation and intercomparison.
    a. Design based on consensus of research community opinions of what is important -- this

is critical for building community support, especially among field program participants.

b. Multiple mock communities are recommended:
    i. Mesopelagic and bathypelagic mock communities will depend on what is available in culture collections.
    ii . Multiple strains of the same organisms could help evaluate how relatedness/ similarity influences results.
c. Consider the use of synthetic strains in addition to available cultures.
d. High versus low complexity mock communities; use average nucleotide differences among taxa as a metric for mock community specifications.
    i. Even versus staggered mixture of organisms.

2. Test key variables in the sample collection and extraction process that contribute to between lab variability. Mock communities on filters and reference seawater material (from research cruises or coastal sites) should be distributed to multiple volunteer labs for parallel nucleic acids extraction.
3. Establish a data portal to make all data available to the scientific community in a timely manner.

8.1. Community-wide intercalibration

The organizers synthesized the recommendations of workshop participants by designing a potential S&I effort that we are tentatively naming 'OMICS' (Ocean Molecular InterCalibration and Standardization) to be proposed to NSF for funding. One part will cover the testing of different sampling and extraction methodologies, while another part of the proposal will focus on homogenized reference material collection at land-based sites and at sea for community wide intercalibration.

Internal and external standards: We will develop a set of internal molecular standards, external mock community standards, and in silico spike-ins, as well as protocols for their implementation, in order to quantify the precision and accuracy of the OMICS molecular datasets. Internal DNA and RNA molecular standards will be constructed and added to all genomic samples, enabling estimates of volumetric abundances of taxa, genes, and transcripts as well as determination of sequence error rates. External mock communities composed of organisms common to oligotrophic and coastal systems will be processed and sequenced in parallel to the native na'omics samples to identify biases introduced during sample processing and bioinformatics analyses. Internal standards and mock communities will also be distributed with the reference materials to all US-based labs participating in the intercalibration effort. A set of in silico standards to control for biases introduced in the bioinformatic steps will be generated and made publicly available. Together, we expect the development of these standards, protocols, and best practices to increase their adoption by the na'omics community, enhancing the reproducibility and quantitative accuracy of marine omics datasets.

8.2. Field work

8.2.1. Coastal activities

Work at field stations with access to plumbed seawater can facilitate the development of best practices towards sample processing, including filtration and preservation protocols/techniques for each of the three na'omics approaches (i.e., amplicon sequencing, metagenomics and metatranscriptomics), specifically for bacteria, archaea and protists. Field station locations should likely include those with contrasting characteristics in order to better assess the role that diversity has in the introduction of sampling biases. Ideally, a high productivity coastal site and a relatively low productivity should be identified. Sample processing variables to be examined and compared include filtration methods and duration, filter sizes and types, sample volumes and flow rates and filter preservation methods. One critical need identified by workshop participants was the ability to homogenize large volumes of seawater to reduce variability and enable determination of bias across all levels of any na'omics pipeline. Further, large volumes of homogenized seawater can be used to prepare replicate filters for community wide intercalibration activities during which the extent of lab-specific processing errors can be identified. Field stations with infrastructure that enables the homogenization of large volumes of seawater (e.g., holding tanks or mesocosm facilities) should likely be prioritized. Alternatively, investments can be made in the acquisition and/or design and production of portable systems for homogenizing seawater.

8.2.2. Cruise-based activities

A S&I OMICS research cruise will focus on the development of best practices towards sea-going sample collection techniques. The 10-day cruise will be conducted along an oceanographic transect where steep environmental gradients exist within short spatial scales providing contrasts in microbial biomass, productivity, and community composition. Samples will be collected within and below the euphotic zone. Selected protocols for filtration and preservation will be based on the outcomes of the land-based activities. Field sample collection intercomparisons for each of the three na'omics approaches include the ship's underway seawater system, pumping systems (both on-deck and in situ) and Niskin bottles mounted on a Rosette sampler.

8.3. Centralized versus community-based activities

Workshop participants thought it would be most efficient to break up the work into two tracks. The first track would consist of several labs with expertise (i.e., the Ninjas) enabling them to target specific questions related to sampling, storage, nucleic acids extraction, and bioinformatics. This work would be iterative and make particular use of the field work conducted at coastal stations. During this work, the methods specific biases would be assessed and improved practices for marine na'omics developed – with the primary goal of identifying the key sources of bias and cost-effective ways of mitigating them. The second track would consist of a broad swath of research community members (i.e., the Armies) already doing studies using marine na'omics. The goal of the second track is to identify lab-specific bias and better understand how well the results of different labs can be compared. The results of such an intercalibration effort will be used to develop metrics for

"compliance" and reevaluate methods for the production of marine 'omics data at scale.

8.4. Nucleic acid extraction intercalibration activities

Homogenized samples for the community-wide intercalibration of nucleic acid extraction protocols will be obtained from the marine stations and on the research cruise and distributed to labs willing to participate in this intercalibration effort. Filters of homogenized samples, mock communities, and internal standards will be distributed to participating labs. A website will be created to track information on the nucleic acid extraction protocols and requests for samples. Following extraction, the samples will be sent to a single core sequencing facility for sequence library construction and sequencing. Contributing labs will be invited to participate in a workshop to compare results and development of best practices for nucleic acid extraction protocols for each of the na'omics approaches.

8.5. Data management

See section 7.1. for details on a proposed data management structure.

9 Towards building international programs that include na'omics activities

Increasingly, field programs of relevance to the oceanographic community have major components that use high throughput molecular amplicon, metagenomics and metatranscriptomics to understand the functioning of microbes in the ocean. When intercalibrated and combined with intercalibrated biogeochemical measurements, 'omics measures will be even more powerful for inferring microbial function. Additionally, sequencing approaches have converged, providing a better understanding of sample requirements for sequencing efforts. With this growing interest and gained perspective, a na'omics S&I program is critically needed to move the field from a largely qualitative to a quantitative discipline. Embarking on an S&I program for na'omics is timely, given rapid expansion of use of 'omics methods in biological oceanography. The outcomes of the OMICS Project will not only benefit US-based scientists but are recognized as imperative by the international community. Best practices will be compiled in a "cookbook" and made freely available through international data sharing portals such as Ocean Best Practices. Such activities have been identified as a critical first step towards the development of large omics-based initiatives such as BioGeoSCAPES (Fig. 4), similar to those conducted by the trace metal chemistry community prior to the implementation of GEOTRACES. Undergraduate, graduate students and post-docs across numerous institutions will be involved with these efforts, providing technical training and professional development through these important community-building efforts. Together, this project will build the human resources and knowledge infrastructure on which to build a global program.

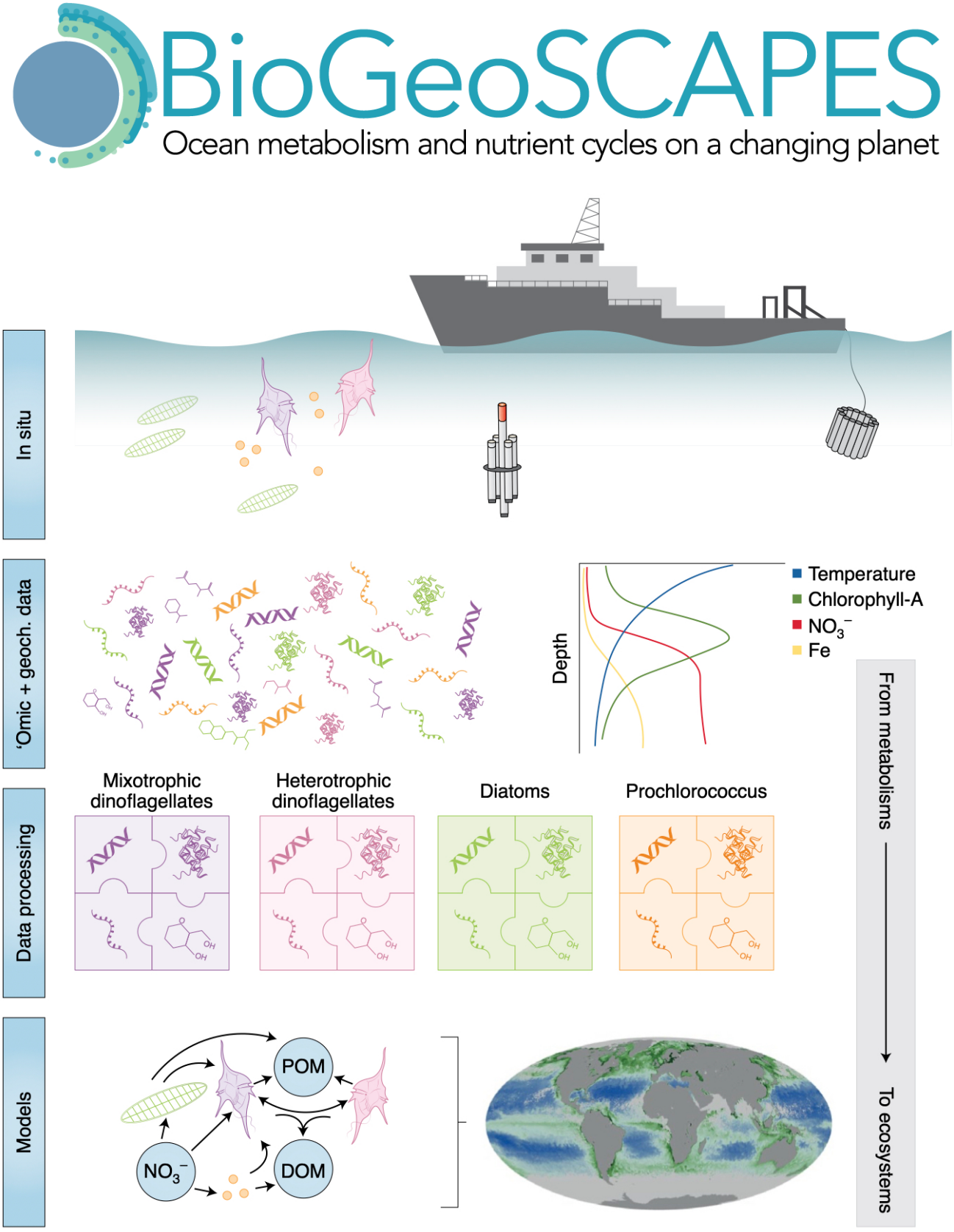


Figure 4. (Top) BioGeoSCAPES is a microbial biogeochemistry program being developed by the international scientific community. (Bottom) The nascent program will integrate in situ 'omic and geochemical observations with data science and models to transform our understanding of ocean microbial biogeochemistry. From Levine and Leles 2021.

Citation: Levine, NM and SG Leles (2021), Marine plankton metabolisms revealed, Nature Microbiology, doi:10.1038/s41564-020-00856-x

10 References

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W. & Huse, S. M. 2009. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLOS ONE* 4:e6372.

Anderson, R. F., Fleisher, M. Q., Robinson, L. F., Edwards, R. L., Hoff, J. A., Moran, S. B., Van der Loeff, M. R., Thomas, A. L., Roy-Barman, M. & Francois, R. 2012. GEOTRACES intercalibration of 230Th, 232Th, 231Pa, and prospects for 10Be. *Limnology and Oceanography: Methods* 10:179-213.

Angebault, C., Ghozlane, A., Volant, S., Botterel, F., d'Enfert, C. & Bougnoux, M.-E. 2018. Combined bacterial and fungal intestinal microbiota analyses: Impact of storage conditions and DNA extraction protocols. *PLOS ONE* 13:e0201174.

Apprill, A., McNally, S., Parsons, R. & Weber, L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial Ecology 75:129-37.
Balzano, S., Abs, E. & Leterme, S. C. 2015. Protist diversity along a salinity gradient in a coastal lagoon. *Aquatic Microbial Ecology* 74:263-77.

Becker, S., Aoyama, M., Woodward, E. M. S., Bakker, K., Coverly, S., Mahaffey, C. & Tanhua, T. 2020. GO-SHIP repeat hydrography nutrient manual: the precise and accurate determination of dissolved inorganic nutrients in seawater, using continuous flow analysis methods. Frontiers in Marine Science 7:908.

Boström, K. H., Simu, K., Hagström, Å. & Riemann, L. 2004. Optimization of DNA extraction for quantitative marine bacterioplankton community analysis. *Limnology and Oceanography: Methods* 2:365-73.

Boyle, E. A., John, S., Abouchami, W., Adkins, J. F., Echegoyen-Sanz, Y., Ellwood, M., Flegal, A. R., Fornace, K., Gallon, C. & Galer, S. 2012. GEOTRACES IC1 (BATS) contamination-prone trace element isotopes Cd, Fe, Pb, Zn, Cu, and Mo intercalibration. *Limnology and Oceanography: Methods* 10:653-65.

Bradley, I. M., Pinto, A. J. & Guest, J. S. 2016. Design and Evaluation of Illumina MiSeq-Compatible, 18S rRNA Gene-Specific Primers for Improved Characterization of Mixed Phototrophic Communities. *Applied and Environmental Microbiology* 82:5878-91.

Breier, J. A., Jakuba, M. V., Saito, M. A., Dick, G. J., Grim, S. L., Chan, E. W., McIlvin, M. R., Moran, D. M., Alanis, B. A., Allen, A. E., Dupont, C. L. & Johnson, R. 2020. Revealing ocean-scale biochemical structure with a deep-diving vertical profiling autonomous vehicle. *Science Robotics* 5:7104.

Comeau, A. M., Douglas, G. M., Langille, M. G. I. & Eisen, J. 2017. Microbiome Helper: a Custom and Streamlined Workflow for Microbiome Research. *mSystems* 2:e00127-16.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X. & Mortazavi, A. 2016. A survey of best practices for RNA-seq data analysis. *Genome Biology* 17:13.

Cutter, G. A. 2013. Intercalibration in chemical oceanography—getting the right number. *Limnology and Oceanography: Methods* 11:418-24.

Dickson, A. G., Afghan, J. & Anderson, G. 2003. Reference materials for oceanic CO2 analysis: a method for the certification of total alkalinity. *Marine Chemistry* 80:185-97.

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., Goldsmith, D. B., Michisaki, R., Breitbart, M. & Boehm, A. B. 2017. Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Frontiers in Marine Science* 4:314.

Grasse, P., Brzezinski, M. A., Cardinal, D., De Souza, G. F., Andersson, P., Closset, I., Cao, Z., Dai, M., Ehlert, C. & Estrade, N. 2017. GEOTRACES inter-calibration of the stable silicon isotope composition of dissolved silicic acid in seawater. *Journal of Analytical Atomic Spectrometry* 32:562-78.

Gray, M. A., Pratte, Z. A. & Kellogg, C. A. 2013. Comparison of DNA preservation methods for environmental bacterial community samples. FEMS Microbiol Ecol 83:468-77.
Hallmaier-Wacker, L. K., Lueert, S., Roos, C. & Knauf, S. 2018. The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. *Scientific Reports* 8:6292.

Heikema, A. P., Horst-Kreft, D., Boers, S. A., Jansen, R., Hiltemann, S. D., de Koning, W., Kraaij, R., de Ridder, M. A. J., van Houten, C. B., Bont, L. J., Stubbs, A. P. & Hays, J. P. 2020. Comparison of Illumina versus Nanopore 16S rRNA Gene Sequencing of the Human Nasal Microbiota. *Genes* 11:1105.

Jackson, S. 2019. The MOSAIC Standards Challenge: Capturing the Extent and Sources of Variability in NGS-Based Microbiome Profiling Protocols and Analyses. *Journal of Biomolecular Techniques : JBT* 30:S63-S63.

Lekang, K., Thompson, E. & Troedsson, C. 2015. A comparison of DNA extraction methods for biodiversity studies of eukaryotes in marine sediments. *Aquatic Microbial Ecology* 75:15-25.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T. & Machida, R. J. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* 10:34.

Levine, N. M. & Leles, S. G. 2021. Marine plankton metabolisms revealed. Nat Microbiol 6:147-48.
Luna, G. M., Dell'Anno, A. & Danovaro, R. 2006. DNA extraction procedure: a critical issue for bacterial diversity assessment in marine sediments. *Environmental Microbiology* 8:308-20.

McLaren, M. R., Willis, A. D. & Callahan, B. J. 2019. Consistent and correctable bias in metagenomic sequencing experiments. *eLife* 8:e46923.

Nygaard, A. B., Tunsjø, H. S., Meisal, R. & Charnock, C. 2020. A preliminary study on the potential of Nanopore MinION and Illumina MiSeq 16S rRNA gene sequencing to characterize building-dust microbiomes. *Scientific Reports* 10:3209.

Padilla, C. C., Ganesh, S., Gantt, S., Huhman, A., Parris, D. J., Sarode, N. & Stewart, F. J. 2015. Standard filtration practices may significantly distort planktonic microbial diversity estimates. *Frontiers in Microbiology* 6:547.

Parada, A. E., Needham, D. M. & Fuhrman, J. A. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* 18:1403-14.

Schlitzer, R., Anderson, R. F., Dodas, E. M., Lohan, M., Geibert, W., Tagliabue, A., Bowie, A., Jeandel, C., Maldonado, M. T., Landing, W. M., Cockwell, D., Abadie, C., Abouchami, W., Achterberg, E. P., Agather, A., Aguliar-Islas, A., van Aken, H. M., Andersen, M., Archer, C., Auro, M., de Baar, H. J., Baars, O., Baker, A. R., Bakker, K., Basak, C., Baskaran, M., Bates, N. R., Bauch, D., van Beek, P., Behrens, M. K., Black, E., Bluhm, K., Bopp, L., Bouman, H., Bowman, K., Bown, J., Boyd, P., Boye, M., Boyle, E. A., Branellec, P., Bridgestock, L., Brissebrat, G., Browning, T., Bruland, K. W., Brumsack, H.-J., Brzezinski, M., Buck, C. S., Buck, K. N., Buesseler, K., Bull, A., Butler, E., Cai, P., Mor, P. C., Cardinal, D., Carlson, C., Carrasco, G., Casacuberta, N., Casciotti, K. L., Castrillejo, M., Chamizo, E., Chance, R., Charette, M. A., Chaves, J. E., Cheng, H., Chever, F., Christl, M., Church, T. M., Closset, I., Colman, A., Conway, T. M., Cossa, D., Croot, P., Cullen, J. T., Cutter, G. A., Daniels, C., Dehairs, F., Deng, F., Dieu, H. T., Duggan, B., Dulaquais, G., Dumousseaud, C., Echegoyen-Sanz, Y., Edwards, R. L., Ellwood, M., Fahrbach, E., Fitzsimmons, J. N., Russell Flegal, A., Fleisher, M. Q., van de Flierdt, T., Frank, M., Friedrich, J., Fripiat, F., Fröllje, H., Galer, S. J. G., Gamo, T., Ganeshram, R. S., Garcia-Orellana, J., Garcia-Solsona, E., Gault-Ringold, M., George, E., et al. 2018. The GEOTRACES Intermediate Data Product 2017. Chemical Geology 493:210-23.
Sharp, J. H., Carlson, C. A., Peltzer, E. T., Castle-Ward, D. M., Savidge, K. B. & Rinker, K. R. 2002. Final dissolved organic carbon broad community intercalibration and preliminary use of DOC reference materials. *Marine Chemistry* 77:239-53.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D., Breiner, H. W. & Richards, T. A. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* 19 Suppl 1:21-31.

Stulberg, E., Fravel, D., Proctor, L. M., Murray, D. M., LoTempio, J., Chrisey, L., Garland, J., Goodwin, K., Graber, J., Harris, M. C., Jackson, S., Mishkind, M., Porterfield, D. M. & Records, A. 2016. An assessment of US microbiome research. *Nature Microbiology* 1:15015.

Tkacz, A., Hortala, M. & Poole, P. S. 2018. Absolute quantitation of microbiota abundance in environmental samples. *Microbiome* 6:110.

Torres-Beltrán, M., Mueller, A., Scofield, M., Pachiadaki, M. G., Taylor, C., Tyshchenko, K., Michiels, C., Lam, P., Ulloa, O. & Jürgens, K. 2019. Sampling and processing methods impact microbial community structure and potential activity in a seasonally anoxic fjord: Saanich Inlet, British Columbia. *Frontiers in Marine Science* 6:132.

Wear, E. K., Wilbanks, E. G., Nelson, C. E. & Carlson, C. A. 2018. Primer selection impacts specific population abundances but not community dynamics in a monthly time-series 16S rRNA gene amplicon analysis of coastal marine bacterioplankton. Environmental Microbiology 20:2709-26.
Yeh, Y.-C., Needham, D. M., Sieradzki, E. T., Fuhrman, J. A. & Caporaso, J. G. 2018. Taxon Disappearance

from Microbiome Analysis Reinforces the Value of Mock Communities as a Standard in Every Sequencing Run. *mSystems* 3:e00023-18.

Zimmermann, J., Jahn, R. & Gemeinholzer, B. 2011. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms Diversity & Evolution* 11:173.

Appendix A Participant List

Harriet Alexander
WHOI
halexander@whoi.edu

Heather Benway
Woods Hole Oceanographic Institution
hbenway@whoi.edu

Paul Berube
MIT
pmberube@mit.edu

Ann Bucklin
University of Connecticut
ann.bucklin@uconn.edu

Craig Carlson
University of California Santa Barbara
carlson@lifesci.ucsb.edu

Dreux Chappell
Old Dominion University
pdchappe@odu.edu

Alex Crocker
UNC Chapel Hil
dinosaur@live.unc.edu

David Emerson
Bigelow Laboratory for Ocean Sciences
demerson@bigelow.org

Jed Fuhrman
University of Southern California
fuhrman@usc.edu

Scott Gifford
UNC Chapel Hill
sgifford@email.unc.edu

Stephen Giovannoni
Oregon State University
steve.giovannoni@oregonstate.edu

Bethany Jenkins
University of Rhode Island
bjenkins@uri.edu

Rob Lampe
Scripps Institution of Oceanography
rlampe@ucsd.edu

Julie LaRoche
Dalhousie University
julie.laroche@dal.ca

Adrian Marchetti
UNC Chapel Hill
amarchetti@unc.edu

Adam Martiny
UC Irvine
amartiny@uci.edu

Jesse McNichol
University of Southern California
mcnichol@alum.mit.edu

Maria Pachiadaki
Woods Hole Oceanographic Institution
mpachiadaki@whoi.edu

Ryan Paerl
NCSU
rpaerl@ncsu.edu

Juan Pierella Karlusich
École normale supérieure (ENS) de Paris
pierella@biologie.ens.fr

Tatiana Rynearson
University of Rhode Island
rynearson@uri.edu

Mak Saito
WHOI
msaito@whoi.edu

Alyson Santoro
UCSB
asantoro@ucsb.edu

Michael E Sieracki
NSF OCE/BO
MSIERACK@nsf.gov

Sachia Traving
University of British Columbia
straving@mail.ubc.ca

Ben Tully
University of Southern California
tully.bj@gmail.com

Emma Wear
University of Montana
emma.wear@flbs.umt.edu

Appendix B Workshop Agenda

Thursday, January 9: Plenary Talks and Brainstorming

9:00 Welcome and introductions (Adrian Marchetti and Jaye Cable, UNC Chapel Hill)
Workshop goals and outcomes (Bethany Jenkins, URI)
Participant demographics (Adrian Marchetti, UNC Chapel Hill)
Individual participant introductions

Plenary 1. Lessons learned from previous S&I efforts in the oceanographic community (Moderator: Bethany Jenkins, URI)

9:25 History and overview of GEOTRACES efforts to standardize and intercalibrate trace metal measurements (Alyson Santoro, UCSB)
9:40 History and overview of BioGEOTRACES (Paul Berube, MIT)
9:55 NIST efforts on S&I for microbiome research (Scott Jackson, NIST)
10:20 Q&A and discussion
10:30 Break

Plenary 2. Quantitative measurements in molecular data (Moderator: Adrian Marchetti, UNC Chapel Hill)

10:45 Amplicon sequencing S&I (Tatiana Rynearson, URI)
11:00 Prokaryotic Metagenomics & Metatranscriptomics S&I (Scott Gifford, UNC Chapel Hill)
11:15 Eukaryotic Metagenomics & Metatranscriptomics S&I (Andrew Allen, UCSD)
11:30 Development of internal/external standards for 'omics intercalibration (David Emerson, Bigelow)
11:45 Comparison of field sample collection methods and relationship to data output (Craig Carlson, UCSB)
12:00 Q&A and discussion

Plenary 3. The relationship of bioinformatics to S&I

12:10 Bioinformatics for assessing sample intercomparison (Harriet Alexander, WHOI)
12:30 Q&A and discussion
12:40-14:30 Lunch on your own

14:30 Group brainstorming exercise - Musical chairs for discussing critical issues for S&I (Moderator: Paul Berube, MIT)

14:30 Introduction to brainstorming exercise (Paul Berube, MIT)
14:40 10-minute brainstorming discussions (participants rotate through 5 topics in mixed groups)
- Sampling and Extraction (Facilitator: Alyson Santoro, UCSB)
- Amplicon Sequencing (Facilitator: Bethany Jenkins, URI)
- Standards and Mock Communities (Facilitator: Scott Gifford, UNC Chapel Hill)
- Community Intercalibration Issues, Cross-comparison, and Assessment of S&I (Facilitator: Paul Berube, MIT)

- Freeform Brainstorming (Facilitator: Adrian Marchetti, UNC Chapel Hill)

15:30 Break

15:45 Facilitators report out on discussions

16:15 Group discussion on critical S&I issues from pre-workshop solicited information and the musical chairs activity to identify:
- Key topics and methods that need to be addressed in the workshop
- Priority list of topics

17:30–18:30 Welcome Reception (TRU Deli and Wine Bar - 114 Henderson St., Chapel Hill)

18:00–19:00 Organizing committee meeting

Friday, January 10: Small Group Discussions

9:00 Review of previous day and introduction to Day 2 (Organizing committee)

9:10 Presentation on BioGEOSCAPES, Metaproteomic Intercomparison, and CLIO (Mak Saito, WHOI)

9:30-11:30 Morning breakout sessions (DNA theme)
- Sampling and extraction for DNA
- Eukaryotic metagenomics
- Viral/Archaeal/Bacterial metagenomics
- Amplicon sequencing

11:30–12:30 Write-ups for morning breakout sessions
12:30–14:00 Lunch
14:00–16:00 Afternoon breakout sessions (RNA theme)
- Sampling and extraction for RNA
- Eukaryotic metatranscriptomics
- Archaeal/Bacterial metatranscriptomics
- Viral metatranscriptomics
- Biases associated with library construction and sequencing platforms
- Bioinformatics

16:00–17:00 Write-ups for afternoon breakout sessions
17:00–18:00 Overview of key outcomes by facilitator from each breakout session
18:30–20:30 Workshop dinner (Tandem, 200 N Greensboro St #1a, walking directions)
20:30–21:30 Organizing committee meeting

Saturday, January 11: Planning a Nucleic Acids 'Omics S&I Activity

9:00 Review of previous day and introduction to Day 3 (Organizing committee)</strong
9:15 Small group sticky note prioritization activity (for variables to include in intercalibration exercise)
9:45 Report back results

10:00–12:00 Small group discussions (5 groups) to design an intercalibration activity
Discussion topics
- Participants - how to engage participation and feedback
- Scope of activity
- Study sites/sample sources
- Variables to focus on
- Sequencing needs
- Potential funding sources
- Publicizing effort

12:00 Lunch
13:00 Groups present their intercalibration activities (~5 mins. each) and discussion
14:30-16:30 Small group discussions to flesh out different aspects of an intercalibration exercise
- Mock communities
- Sampling methods
- Variables to consider

16:30 Small groups report back
17:00 Group discussion
17:30 Adjourn
19:00 Working dinner (organizing committee only)

View presentation slides on the workshop website:
https://www.us-ocb.org/ocean-nucleic-acids-omics-workshop/

For more information visit:
https://www.us-ocb.org/ocean-nucleic-acids-omics-workshop/

Ocean Carbon & Biogeochemistry Project Office
Woods Hole Oceanographic Institution
266 Woods Hole Road MS#25
Woods Hole, MA 02543
hbenway@whoi.edu | 508.289.2838
www.us-ocb.org | Twitter @us_ocb

Ocean Carbon & Biogeochemistry

OCB acknowledges support from these US agencies:

This report was developed with federal support of NSF (OCE-1558412) and NASA (NNX17AB17G). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsoring agencies.