THE
UNIVERSITY
OF RHODE ISLAND
COLLEGE OF
THE ENVIRONMENT
AND LIFE SCIENCES

Dr. Bethany D. Jenkins
Professor
Department of Cell and Molecular Biology and
Graduate School of Oceanography
279 CBLS, 120 Flagg Rd.
Kingston, RI 02881
401-874-7551 ◆ bjenkins@uri.edu

Dec. 2, 2018

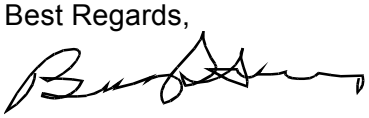
Dear OCB Steering Committee Members,
I along with my steering committee members are submitting a Scoping Workshop proposal that was deemed to address an urgent need for developing a concept for a program focused on intercalibration of nucleic acids 'omics methodology to study prokaryotic and eukaryotic marine microbes. This area is one of OCB relevance as stated in the proposal is relevant to sampling at ocean time-series locales, the EXport Processes in the Ocean from Remote Sensing (EXPORTS) program, and ongoing (BioGEOTRACES) and emerging Biogeoscapes programs coupling inventories of microbial abundance and activities with intercalibrated measures of ocean chemistry.

The urgency for this Scoping Workshop emerged from a recent meeting in Nov 2018 to discuss the development of a new program focused on chemical and microbial sampling of the ocean (Biogeoscapes) as part of the legacy of the GEOTRACES program. It was felt that the time is right to develop a series of plans by the marine microbiology community to intercalibrate nucleic acids 'omics measures so they could be integrated with other intercalibrated measurements. As microbes are the engines of major biogeochemical cycles in the ocean we feel that such a workshop would be of service to the OCB community as we hope to develop a means by which nucleic acids 'omics are better standardized and intercalibrated.

Our proposed workshop will be held in the fall of 2019 at the University of North Carolina Chapel Hill, a venue selected to be low cost in terms of hotels, venue pricing and the fact that two of our steering committee members have faculty appointments there.

Thank you for considering our proposal for an **Ocean nucleic acids 'omics intercalibration and standardization workshop**

Best Regards,



Bethany D. Jenkins, PhD and Steering Committee (Andrew Allen, Paul Berube, Scott Gifford, Adrian Marchetti and Alyson Santoro)

Ocean nucleic acids 'omics intercalibration and standardization workshop proposal

Steering Committee:
Bethany Jenkins (University of Rhode Island), Andrew Allen (University of California, San Diego and UCSD and the J. Craig Venter Institute), Paul Berube (Massachusetts Institute of Technology), Scott Gifford (University of North Carolina, Chapel Hill), Adrian Marchetti (University of North Carolina, Chapel Hill), and Alyson Santoro (University of California, Santa Barbara)

This OCB scoping workshop proposal is aimed at developing a focused marine microbial nucleic acid (na) 'omics intercomparison and intercalibration effort. Increasingly, field programs of relevance to the OCB community have major components that use high throughput molecular barcoding, metagenomics and transcriptomics (nucleic acid 'omics or na'omics herein) to understand the functioning of prokaryotic and eukaryotic microbes in the ocean. Examples of these programs include integration of microbial sampling at ocean time-series sites, the NASA-NSF funded EXPORTS program, and the international BioGEOTRACES program that couples these measurements in conjunction with GEOTRACES cruise sampling. A pressing challenge for both our community and the broader microbiome research community is the poor degree of standardization and intercalibration that would facilitate comparison between na'omics data (Nayfach et al. Cell. 2016 Aug 25; 166(5): 1103–1116; Stulberg et al. Nature Microbiology 2016. Article number: 15015). These issues were highlighted as a critical need at a recent OCB-supported workshop held in Nov. 2018 to discuss the next wave of potential international field programs (Biogeoscapes) that would tightly couple ocean chemistry sampling with na'omics measurements. To ensure the success of this and other such programs, we seek to bring together experts and stakeholders in order to identify a path towards na'omics standardization and intercalibration solutions.

One of the reasons for the success of the international GEOTRACES program is that measurements made between research groups are intercomparable and this is due to a series of intercalibration efforts and standard reference materials developed by the chemical oceanography community. It is clear and urgent that a parallel effort needs to be undertaken by the microbial oceanography community to intercalibrate nucleic acids 'omics measurements with a goal of developing community benchmarks. It is our goal that we work together to consider issues and develop a plan and framework for conducting a na 'omics intercalibration effort.

Many of us have participated in OCB and Earth Cubed sponsored workshops where issues of intercomparabilty of 'omics methodologies have been discussed and we feel that the time is right and our community is committed to advancing these efforts. As we describe our recommendations below, it may appear that the methodological variables are extremely high but we feel that from our combined expertise and converging agreement on methodological considerations we can develop recommendations for an intercalibration framework. We would like to highlight that our goal is not to develop one prescribed set of methods but rather a best practices framework that could be tested in a formal intercalibration effort.

We are envisioning the development of a series of recommendations around the following issues to be tested in future intercalibration activities.

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

Typically researchers use filtration methods to capture biogenic particles from seawater. Filtration can be done with water collected from niskin casts and onboard pumps. Alternatively, the use of in situ pumping systems (e.g. McLane pumps) are increasing and it is unclear how results from different pump systems compare. There are different filter matrices, construction (e.g. flat vs capsule) and pore sizes used with variations in filtration time. After particle collection, filters are flash frozen in liquid N_2 or chemically preserved. As gene expression (transcriptomics) can be rapidly altered as a function of time from in situ sampling, filtration times are a particular issue for these methods. Additionally, biomass in nutrient rich and nutrient poor areas of the ocean can vary widely which also influences filtration times and yields. Filtration volume has been shown to impact the community detected in seawater samples (Padilla et al. 2015 Front. Micro. Jun. 02). We will suggest tests that can be done to compare pump methods and we will develop a series of recommended protocols for filtration methods and times that are appropriate for different ocean environments and depths.

II. Extraction protocols for nucleic acids

There are variations in methods for nucleic acid extraction that range from protocols with reagents made up in the laboratory to commercially available kits that may utilize proprietary reagents. Optimal extraction protocols would be ones that maximize extraction efficiency across organisms. Additionally, cost and throughput of the methods are worthy of consideration. Some methods use mechanical shearing to disrupt cells that is a consideration for the sequencing technology utilized. 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 include shearing steps may impact quality of material for some sequencing platforms. Some researchers have been directly involved in comparisons of nucleic acid isolation methods and we feel that we can develop a confined set of protocols to test in a future intercalibration effort.

III. 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 (e.g. those developed by Mike Lomas and Dave Emerson at Bigelow (https://www.bigelow.org/news/articles/2016-07-26-3.html) or synthetically engineered pieces of DNA or RNA that do not correspond to any living organism (Tourlousse et al., 2017, Nucl. Acids. Res. 45 (4)). 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 and from the National Institute of Standards and Technology's (NIST) Microbiome Community Measurement group that have been converging on the use of exogenous reference materials to sample extraction. We feel that this is an area in which the microbial oceanography community would benefit from discussion with members of a different microbiome communities in such a workshop.

IV. Isolation methods unique to RNA

Total RNA isolated from a biomass sample is a matrix of highly abundant structural RNAs (rRNA) and molecules derived from protein coding genes (mRNA) present in

lesser abundance but that contain metabolic signal. There are methods that enrich for mRNA via rRNA removal steps and other methods that take advantage of the presence of rRNA for taxonomic measures. Additionally, RNA molecules from prokaryotes and eukaryotes have sequence differences that are utilized in isolation protocols so it is 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. They are also subject to the whims of manufacturers for kit utilized for sequence library construction (see below). A goal for this workshop will be to determine what is the state of the art for RNA based methods and a series of recommendations that are feasible across laboratories.

V. Sequence library construction

Once nucleic acids are isolated they are subject to downstream protocols for preparation for sequencing on a high throughput platform. Generally these library preparation methods are performed outside an individual user's laboratory and are contracted through typically 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. As this will require financial resources 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. We will include workshop participants from different sequencing centers to help us develop these recommendations.

VI. Considerations unique to barcoding methods

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

VII. Integration of bioinformatic considerations

This workshop is an opportunity to have researchers with bioinformatics expertise inform discussion of na ‘omics standardization and intercalibration. Members of the marine na 'omics bioinformatic community work with data generated by many different methods and can make recommendations for standardized pipelines by which to verify intercalibration. In particular, bioinformaticists would offer a perspective on challenges they face when working with data that lack of standardization and intercalibration. While this workshop would not focus on specific methodological aspects of individual software pieces incorporated into bioinformatic pipelines as these rapidly change, we would want to include discussion of overarching bioinformatic pipeline architecture tailored to each type of ‘omics as this will have bearing on the best methods for standardization.

Conceptualizing what an 'omics intercalibration program would entail

Ultimately, from this workshop we would like to develop concrete plans for an intercalibration exercise, with the goal of submitting a formal proposal for an intercalibration cruise and subsequent laboratory work. We are inspired by the SAFe cruises conducted prior to the GEOTRACES program and intercalibration efforts conducted by the GEOTRACES community. We envision developing plans and a written document guiding a concerted 'omics intercalibration effort that combines optimization and comparison of field and laboratory sampling protocols.

Structure of a 'omics scoping workshop
We would like to bring members of our community together that have experience in different ocean environments, different taxa, different target molecules (e.g. DNA, RNA, barcodes), and bioinformatics. We will also include members of the NIST Microbiome Community Measurement group as well as members of the human microbiome community that have expertise in standardization methods to help us develop a concerted set of plans for intercalibration. Additionally, we will include participants from sequencing centers that are involved in sequence library construction and knowledge of sequencing technologies. We plan to hold the workshop over 3 days at the University of North Carolina, Chapel Hill in the fall of 2019. Two of our steering committee members are faculty members at UNC, which also has a sequencing center used by members of our community, and we feel that this is a low cost location for our workshop. Ahead of the workshop we will circulate a draft of a document to participants to help them prepare what we develop with a detailed list of issues we think are important topics to discuss under the headings I-VII above.

We propose the following structure for this 3 day workshop
Day 1
Morning: short presentations highlighting the state of the art in I-VII above
Afternoon: working group breakouts to identify areas of best practices under I-VII above

Day 2
Morning: working groups identifying issues that need further resolution under I-VII above
Afternoon: working group written summaries focusing on best practices and most important issues that need additional discussion/resolution prior to intercalibration

Day 3
Conceptualization of an intercalibration program that would then be introduced to the community at a town hall at the 2020 ALSO Ocean Sciences meeting, the 2020 International Society for Microbial Ecology Meeting, and the 2020 OCB summer workshop. We anticipate that we will need to meet virtually after the fall 2019 workshop to finalize written outcomes from the workshop and articulation of a concept for an intercalibration program to present to the community.

Workshop Participants:
We would like this workshop to include 20-25 participants. We chose to focus on a smaller participant number than would typically be included in a larger scoping workshop to favor focused discussion and production of workshop products that can then be shared with a larger stakeholder group. Workshop participants will include the 6 steering committee members and a combination of participants from solicited invitations and an open application process. We will seek additional funds to include international participants. Communities we would like represented include participants with expertise in ocean ‘omics (suggested participants: Ginger Armbrust, Craig Carlson, Ed DeLong, Sonya Dyhrman, Jed Fuhrman, Adam Martiny), BioGEOTRACES(suggested participants Dreux Chappell, Maite Maldonado, NIST (participants TBD), human microbiome (participants TBD), bioinformatics (suggested participant Harriet Alexander) TARA Oceans: (suggested participant Chris Bowler), and members of sequencing centers (participants TBD from MIT and UNC).

Proposed Workshop Budget

We are requesting $35,810 from OCB for this workshop and show a budgetary breakdown below. The requested costs reflect the lower costs (e.g. no flights or hotels) for UNC participants. We also anticipate that it might be possible to save money on catered breakfast and lunch over individual per diems but we do not have the cost estimates for catering at this time.

	Budget item	duration	per day costs	total
venue	UNC, Hyde Hall, 3 rooms	3 days	900	2700
hotel	10 participants	5 days	150	7500
	8 participants	4 days	150	4800
meal per diem	10 participants	4 days	69	2760
	10 participants	5 days	69	3450
flights	domestic east (8)		600	4800
	domestic west (10)		800	8000
misc. taxi, parking	18 participants		100	1800
			Total Request	**$35,810**