Metabarcoding samples collected from surface and chlorophyll maximum depths from R/V Pt. Sur PS 18-09 Legs 01 and 03, Hurricane Harvey RAPID Response cruise (western Gulf of Mexico) September-October 2017.

Website: https://www.bco-dmo.org/dataset/824599
Data Type: Cruise Results
Version: 1
Version Date: 2020-09-11

Project
» RAPID: Hurricane Impact on Phytoplankton Community Dynamics and Metabolic Response (HRR)

Contributors	Affiliation	Role
Campbell, Lisa	Texas A&M University (TAMU)	Principal Investigator
Henrichs, Darren W.	Texas A&M University (TAMU)	Co-Principal Investigator
Gaonkar, Chetan	Texas A&M University (TAMU)	Contact
Copley, Nancy	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Metabarcode samples collected from surface and chlorophyll maximum depths from R/V Pt. Sur PS 18-09 Legs 01 and 03, Hurricane Harvey RAPID Response cruise (western Gulf of Mexico) September-October 2017.

Table of Contents

Coverage

Spatial Extent: **N**:29.0649 **E**:-94.9 **S**:27.2286 **W**:-97.268
Temporal Extent: 2017-09-23 - 2017-10-01

Dataset Description

Metabarcode samples collected from surface and chlorophyll maximum depths from R/V Pt. Sur PS 18-09 Legs 01 and 03, Hurricane Harvey RAPID Response cruise (western Gulf of Mexico) September-October 2017.

Methods & Sampling

On each of 2 cruise legs 01 and 03, samples were collected at 7 stations (S01, S06, S11, S16, S21, SS and GI) from 2 depths [surface and chlorophyll maximum depth when possible; see HRR-bottle data]) and triplicate 500-1000 ml samples were filtered and immediately fixed in RNALater. Triplicate samples from each

station/depth were extracted with AllPrep DNA/RNA MiniKit (Qiagen, USA) following the manufacturer's instructions. DNA concentration and quality were evaluated using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc, USA). All the samples extracted for DNA were normalized to 5ng/µl concentration for the amplicon library construction.

The V4 regions of the 18S rRNA genes were amplified using customized V4 primers (Bradley et al. 2016; Kozich et al. 2013). Library construction and amplicon sequencing was performed at Texas A&M University Agrilife's Genomics and Bioinformatics Services (https://www.txgen.tamu.edu) using custom designed primers (Bradley et al. 2016; Kozich et al. 2013). Output of MiSeq results as fasta files were deposited in GenBank under the project number PRJNA592369.

Leg 01 Station 01 sample was not collected.

Sampling locations:

Sample ID	**Station**	**Leg**	**Location Lat ^{o}N/Long ^{o}W**
na		1	27.2286 -97.2686
L3_S01	S01	3	
L1_S06	S06	1	27.8358 -96.9874
L3_S06	S06	3	
L1_S11	S11	1	28.2614 -96.4129
L3_S11	S11	3	
L1_S16	S16	1	28.5366 -95.8656
L3_S16	S16	3	
L1_S21	S21	1	28.7644 -95.2978
L3_S21	S21	3	
L1_SS	SS	1	28.9600 -95.0946
L3_SS	SS	3	
L1_GI	GI	1	29.0649 -94.9000
L3_GI	GI	3	

Data Processing Description

Fastq files for all 13 station samples were quality checked using FastQC. Illumina paired end reads (2x300 bp) were processed in mothur v1.39.0 (Schloss et al. 2009). Contigs were assembled and pre-cleaned processed for homopolymers and ambiguities. These sequences were then screened for chimera using UCHIME in denovo mode (Edgar et al. 2011). Sequences were de-noised by pre-clustering at 1 bp per 100 bp and generate unique sequence for taxonomic annotation and characterization. Sequences less than three were eluted out in our study and rest OTUs were characterized using BLAST search using PR2 database v4.12.0. The BLAST analysis used a assignment approach with similarity was ≥ 90% and query coverage was ≥ 70% against the reference sequence. Any OTU that did not compile with this criterion was not used in this study. The following thresholds for identity with BLAST results were used for taxonomic assignment clustering: species (97%), genus (94%), family (93%), class (92%) and order (90%).

HTS-metabarcoding procedures were followed as mentioned in Gaonkar et al., (2020). Only HTS metabarcodes haplotypes (OTUs) with reads more than 3 allocated to Chaetocerotaceae (Chaetoceros and Bacteriastrum) were used with a selection criterion of ≥ 90% similarity and ≥ 70% similarity after BLAST analysis using the protistan PR2 dataset. A total of 206 (n=82881) OTUs annotated as Chaetocerotacean haplotypes were obtained from the HRR dataset. See figures 1 and 2 in the Supplemental Files section.

BCO-DMO Processing Notes:
- added conventional header with dataset name, PI name, version date
- modified parameter names to conform with BCO-DMO naming conventions

[table of contents | back to top]

Data Files

File
metabarcode.csv (Comma Separated Values (.csv), 1.80 MB) MD5:ffd076a7f8f909d5d3eb308581830ac0 Primary data file for dataset ID 824599

[table of contents | back to top]

Supplemental Files

File	
Chaetocerotacean diversity filename: HRR_Chaetocerotacean_phylogeny.pdf	(Portable Document Format (.pdf), 73.72 KB) MD5:9510b43f8eefaaa59cac95a51fe7ba71
Fig. 1. Chaetocerotacean diversity: Assessing the Chaetocerotacean diversity using the V4-hypervaiable region of 18S rDNA using HTS-metabarcoding. Maximum likelihood tree generated using raxmlGUI v2.0 (Edler et al. 2019) for the Chaetocerotacean species from the sample collected in the Gulf of Mexico, Texas coast. A total of 206 validated Chaetocerotacean metabarcode haplotypes along with 150 taxonomically validated Chaetocerotacean reference sequences and 26 outgroup sequences were used to generate the phylogenetic trees to assess species diversity. The OTU ID consists of the dataset name, unique identification number and number of reads at the end. Numbers on the internodes indicate bootstrap values if ≥50 (1000 replicates).	
Karenia phylogeny filename: HRR_Karenia_phylogeny.pdf	(Portable Document Format (.pdf), 1.22 MB) MD5:8dff398a441a5cdcebe6c8220edac9a9
Fig 2. Karenia phylogeny: Selection of appropriate metabarcoding marker for species detection and delineation. Maximum likelihood tree generated using raxmlGUI v2.0 (Edler et al. 2019) for the Karenia species. OTUs based on (a) V4-region and (b) V8-V9 regions of 18S rDNA gene yield different results. Numbers on the internodes indicate the bootstrap values if ≥50 (1000 replicates). Only those reference sequences which had full length 18S rDNA sequences were selected, and Karenia reference sequences are indicated with a colored dot. Only OTUs those with read abundance more than 100 were selected.	
Station locations filename: locations_metatrans.csv	(Comma Separated Values (.csv), 445 bytes) MD5:335177c22a60a318da8e433442c6a71a
Table with locations for each of the station-leg collection sites.	
Station map filename: station_locations.png	(Portable Network Graphics (.png), 461.29 KB) MD5:7c598067f3efca47dec504ee13dc5121
Sampling locations and identification of the station abbreviations	

[table of contents | back to top]

Related Publications

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(19), 5878–5891. doi:10.1128/aem.01630-16 https://doi.org/10.1128/AEM.01630-16
Methods

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics, 27(16), 2194–2200. doi:10.1093/bioinformatics/btr381
Methods

Edler, D., Klein, J., Antonelli, A., & Silvestro, D. (2019). raxmlGUI 2.0: a graphical interface and toolkit for phylogenetic analyses using RAxML. doi:10.1101/800912
Software

Gaonkar, C. C., Piredda, R., Sarno, D., Zingone, A., Montresor, M., & Kooistra, W. H. C. F. (2020). Species detection and delineation in the marine planktonic diatoms Chaetoceros and Bacteriastrum through metabarcoding: making biological sense of haplotype diversity. Environmental Microbiology, 22(5), 1917–1929. doi:10.1111/1462-2920.14984
Results

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., ... & Christen, R. (2012). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic acids research, 41(D1), D597-D604. https://doi.org/10.1093/nar/gks1160
General

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. Applied and Environmental Microbiology, 79(17), 5112–5120. doi:10.1128/aem.01043-13
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
Sequence_ID	OTU sequence generated from the HRR data	Unitless
Taxonomy	Taxonomic annotation of the OTU	Unitless
Similarity	percentage of identical matches	Unitless
Length	alignment length (sequence overlap)	Unitless
Mismatches	number of mismatches	Unitless
Gaps	number of gap openings	Unitless
Q_start	start of alignment in query	Unitless
Q_end	end of alignment in query	Unitless
R_start	start of alignment in reference sequence	Unitless
R_end	end of alignment in reference sequence	Unitless
e_value	number of expected hits of similar quality	Unitless
Score	Bit-score	Unitless
OTU	representative OTUs	Unitless
L1_S06	number of copies of the OTU at Leg 1 Station 06	Unitless
L1_S11	number of copies of the OTU at Leg 1 Station 11	Unitless
L1_S16	number of copies of the OTU at Leg 1 Station 16	Unitless
L1_S21	number of copies of the OTU at Leg 1 Station 21	Unitless
L1_SS	number of copies of the OTU at Leg 1 Surfside station SS	Unitless
L1_GI	number of copies of the OTU at Leg 1 Galveston Island station GI	Unitless
L3_S01	number of copies of the OTU at Leg 3 Station 01	Unitless
L3_S06	number of copies of the OTU at Leg 3 Station 06	Unitless
L3_S11	number of copies of the OTU at Leg 3 Station 11	Unitless
L3_S16	number of copies of the OTU at Leg 3 Station 16	Unitless
L3_S21	number of copies of the OTU at Leg 3 Station 21	Unitless
L3_SS	number of copies of the OTU at Leg 3 Surfside station SS	Unitless
L3_GI	number of copies of the OTU at Leg 3 Galveston Island station GI	Unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Illumina MiSeq
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Used to obtain DNA sequences. See https://www.illumina.com/systems/sequencing-platforms/miseq.html
Generic Instrument Description	General term for a laboratory instrument used for deciphering the order of bases in a strand of DNA. Sanger sequencers detect fluorescence from different dyes that are used to identify the A, C, G, and T extension reactions. Contemporary or Pyrosequencer methods are based on detecting the activity of DNA polymerase (a DNA synthesizing enzyme) with another chemoluminescent enzyme. Essentially, the method allows sequencing of a single strand of DNA by synthesizing the complementary strand along it, one base pair at a time, and detecting which base was actually added at each step.

Dataset-specific Instrument Name	
Generic Instrument Name	Niskin bottle
Dataset-specific Description	Used to collect samples
Generic Instrument Description	A Niskin bottle (a next generation water sampler based on the Nansen bottle) is a cylindrical, non-metallic water collection device with stoppers at both ends. The bottles can be attached individually on a hydrowire or deployed in 12, 24, or 36 bottle Rosette systems mounted on a frame and combined with a CTD. Niskin bottles are used to collect discrete water samples for a range of measurements including pigments, nutrients, plankton, etc.

Dataset-specific Instrument Name	Nanodrop spectrophotometer (Thermo Fisher Scientific Inc, USA)
Generic Instrument Name	Spectrophotometer
Dataset-specific Description	Used for DNA concentration and quality evaluation.
Generic Instrument Description	An instrument used to measure the relative absorption of electromagnetic radiation of different wavelengths in the near infra-red, visible and ultraviolet wavebands by samples.

[table of contents | back to top]

Deployments

PS1809

Website	https://www.bco-dmo.org/deployment/784313
Platform	R/V Point Sur
Start Date	2017-09-22
End Date	2017-10-03
Description	HRR study with three legs. Chief Scientists: Steve DiMarco (Leg 1); Kristen Thyng (Leg 2); Lisa Campbell (Leg 3). R2R Cruise Page: https://www.rvdata.us/search/cruise/PS1809

[table of contents | back to top]

Project Information

RAPID: Hurricane Impact on Phytoplankton Community Dynamics and Metabolic Response (HRR)

Coverage: Texas coast

NSF Award Abstract:
Hurricane Harvey is the strongest hurricane to hit the Texas coast in decades and the resulting tidal surges, flooding and terrestrial runoff have had a severe impact on the coastal ocean. The effects on the phytoplankton, the first link in the food chain, may be unprecedented. To determine how the phytoplankton community will respond to such drastic changes in salinity, nutrient inputs, and potential toxins, immediate and continuous sampling is the only way to fully capture the effects and to identify when conditions return to "normal". An automated, continuous phytoplankton imaging instrument that is deployed on the Texas coast records images of the phytoplankton and permits calculation of the abundance of different species. Together with molecular information on the genes that have been "turned on", or expressed, outcomes of this project will help determine the responses of individual types of phytoplankton. Extreme storms are expected to increase in frequency with future climate change, so the responses identified now will be valuable in predicting how such events will affect these primary producers, which in turn support most of the food webs in marine ecosystems, in the future.

High temporal resolution observations from the Imaging FlowCytobot (IFCB) have revealed that hurricanes in the Gulf of Mexico cause drastic changes in the phytoplankton community structure. The objectives of this RAPID project are: 1) to characterize the dynamics of the phytoplankton species in relation to the environmental variables along the Texas coast; 2) to assess the short and long-term changes in the phytoplankton community; and 3) to identify the strategies of the phytoplankton community for resource acquisition. To accomplish these objectives, this project will utilize IFCB time series to follow phytoplankton community structure during the recovery period from Hurricane Harvey. In addition, two RAPID response cruises (in late September and early October) to sample at 5 sites along a transect from Galveston to Port Aransas, TX. At each station, CTD profiles and water samples from surface and the chlorophyll maximum will be collected for nutrients, carbonate chemistry, and RNA sequencing for metatranscriptomic analysis. Metatranscriptomics can provide an indication of the metabolic strategies employed and functional relationships within the plankton community in response to changes in the environment. The advantage of a metatranscriptomic approach is that the entire molecular response to the environment is captured. So, while the response of phytoplankton to increased nutrient inputs from floodwater runoff is targeted, the responses to other environmental stresses (toxics, hypoxia, acidification) are also captured. Analyses of this time series using multivariate statistical techniques, such as principal component analysis (PCA), and network analysis, a powerful technique for identifying potential interactions among taxa, will provide insights on the environmental factors and metabolic responses structuring the community during the aftermath of the hurricane.

Related data from the The Texas Observatory for Algal Succession Time-Series (TOAST) can be found at the following: https://toast.tamu.edu/timeline?dataset=HRR_Cruise

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1760620

[table of contents | back to top]