Transcriptional profile of marine bacterium Ruegeria pomeroyi in a three-member co-culture study

Website: https://www.bco-dmo.org/dataset/719970
Data Type: experimental
Version: 1
Version Date: 2017-11-29

Project
» Bacterial Taxa that Control Sulfur Flux from the Ocean to the Atmosphere (OceanSulfurFluxBact)

Program
» Dimensions of Biodiversity (Dimensions of Biodiversity)

Contributors	Affiliation	Role
Moran, Mary Ann	University of Georgia (UGA)	Principal Investigator
Rauch, Shannon	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Transcriptional profile of marine bacterium Ruegeria pomeroyi in a three-member co-culture study. This dataset contains the processed, QC'ed, normalized sequence data. The full raw data file is deposited in the NCBI BioProject database under accession PRJNA381627.

Table of Contents

- Dataset Description
  - Methods & Sampling
  - Data Processing Description
- Data Files
- Related Publications
- Parameters
- Instruments
- Project Information
- Program Information
- Funding

Dataset Description

Transcriptional profile of marine bacterium *Ruegeria pomeroyi* in a three-member co-culture study. This dataset contains the processed, QC'ed, normalized sequence data. The full raw data file is deposited in the NCBI BioProject database under accession PRJNA381627.

These data are published in table 2 of Landa et al., 2017, ISME Journal, in press.

Methods & Sampling

Published methodology in Landa et al., 2017, ISME Journal, in press.

One liter samples were filtered through 2-um polycarbonate (PC) membranes to collect bacterial cells and flash-frozen in liquid nitrogen and stored at -80C. Filters were extracted for RNA by the acid phenol:chloroform:isoamylalcohol method. Potential traces of DNA were removed using the Turbo DNA-free kit (Invitrogen, Waltham, MA, USA). Samples were tested for residual DNA by a 40-cycle PCR targeting the 16S rRNA gene of *R. pomeroyi*. Samples were depleted of rRNA using custom probes for small and large subunit rRNA genes from all three microbes (Stewart et al., 2010). Libraries were prepared for two replicate cubitainers at 8 time points (16 samples) using the KAPA Stranded mRNA-Seq Kit (Kapa Biosystems, Wilmington, MA, USA) at the Georgia Genomics Facility (University of Georgia) and sequenced on a HiSeq Illumina 2500 at the Hudson

Alpha Institute for Biotechnology (AL, USA).

Reference cited:
Stewart FJ, Ottesen EA, DeLong EF. (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. ISME J 4: 896–907. doi:10.1038/ismej.2010.18

Data Processing Description

Sequence quality control was performed using the FASTX toolkit, imposing a minimum quality score of 20 over 80% of read length. Reads aligning to an in-house rRNA database were removed (blastn, score cutoff greater than or equal to 50). Remaining reads were mapped to the *R. pomeroyi* genome (Bowtie 2; Langmead and Salzberg, 2012) and counted (HTSeq; Anders et al., 2015). Counts were converted to transcripts per million (TPM, Table S2) and data were also deposited in the NCBI BioProject database under accession PRJNA381627. Genes with differential expression between selected time points were determined with DESeq2 (Love et al., 2014), and linear relationships of gene expression with phytoplankton abundance with Pearson correlations.

References cited:
Langmead B, Salzberg SL. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359. doi:10.1038/nmeth.1923

Anders S, Pyl PT, Huber W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169. doi:10.1093/bioinformatics/btu638

Love MI, Huber W, Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550. doi:10.1186/s13059-014-0550-8

BCO-DMO Processing:
- modified parameter names to conform with BCO-DMO naming conventions;
- replaced spaces with underscores in Description column.

[table of contents | back to top]

Data Files

File
Switch_RNAseq.csv (Comma Separated Values (.csv), 641.69 KB) MD5:710349f0860d73cd73c7318474aa4db7 Primary data file for dataset ID 719970

[table of contents | back to top]

Related Publications

Anders, S., Pyl, P. T., & Huber, W. (2014). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics, 31(2), 166–169. doi:10.1093/bioinformatics/btu638
Methods

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357–359. doi:10.1038/nmeth.1923
Software

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15(12). doi:10.1186/s13059-014-0550-8
Methods

Stewart, F. J., Ottesen, E. A., & DeLong, E. F. (2010). Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. The ISME Journal, 4(7), 896–907. doi:10.1038/ismej.2010.18

Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
Locus_Tag	gene identifier	unitless
Description	gene function, if known	unitless
Day7_R1	Transcript count by day and replicate: day 7, replicate 1	TPM (transcripts per million)
Day7_R2	Transcript count by day and replicate: day 7, replicate 2	TPM (transcripts per million)
Day9_R1	Transcript count by day and replicate: day 9, replicate 1	TPM (transcripts per million)
Day9_R2	Transcript count by day and replicate: day 9, replicate 2	TPM (transcripts per million)
Day12_R1	Transcript count by day and replicate: day 12, replicate 1	TPM (transcripts per million)
Day12_R2	Transcript count by day and replicate: day 12, replicate 2	TPM (transcripts per million)
Day15_R1	Transcript count by day and replicate: day 15, replicate 1	TPM (transcripts per million)
Day15_R2	Transcript count by day and replicate: day 15, replicate 2	TPM (transcripts per million)
Day18_R1	Transcript count by day and replicate: day 18, replicate 1	TPM (transcripts per million)
Day18_R2	Transcript count by day and replicate: day 18, replicate 2	TPM (transcripts per million)
Day23_R1	Transcript count by day and replicate: day 23, replicate 1	TPM (transcripts per million)
Day23_R2	Transcript count by day and replicate: day 23, replicate 2	TPM (transcripts per million)
Day30_R1	Transcript count by day and replicate: day 30, replicate 1	TPM (transcripts per million)
Day30_R2	Transcript count by day and replicate: day 30, replicate 2	TPM (transcripts per million)
Day37_R1	Transcript count by day and replicate: day 37, replicate 1	TPM (transcripts per million)
Day37_R2	Transcript count by day and replicate: day 37, replicate 2	TPM (transcripts per million)

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	HiSeq Illumina 2500
Generic Instrument Name	Automated DNA Sequencer
Dataset-specific Description	Transcripts were sequenced on a HiSeq Illumina 2500 at the Hudson Alpha Institute for Biotechnology (AL, USA).
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.

[table of contents | back to top]

Project Information

Bacterial Taxa that Control Sulfur Flux from the Ocean to the Atmosphere (OceanSulfurFluxBact)

Surface ocean bacterioplankton preside over a divergence point in the marine sulfur cycle where the fate of dimethylsulfoniopropionate (DMSP) is determined. While it is well recognized that this juncture influences the fate of sulfur in the ocean and atmosphere, its regulation by bacterioplankton is not yet understood. Based on recent findings in biogeochemistry, bacterial physiology, bacterial genetics, and ocean instrumentation, the microbial oceanography community is poised to make major advances in knowledge of this control point. This research project is ascertaining how the major taxa of bacterial DMSP degraders in seawater regulate DMSP transformations, and addresses the implications of bacterial functional, genetic, and taxonomic diversity for global sulfur cycling.

The project is founded on the globally important function of bacterial transformation of the ubiquitous organic sulfur compound DMSP in ocean surface waters. Recent genetic discoveries have identified key genes in the two major DMSP degradation pathways, and the stage is now set to identify the factors that regulate gene expression to favor one or the other pathway during DMSP processing. The taxonomy of the bacteria mediating DMSP cycling has been deduced from genomic and metagenomic sequencing surveys to include four major groups of surface ocean bacterioplankton. How regulation of DMSP degradation differs among these groups and maps to phylogeny in co-occurring members is key information for understanding the marine sulfur cycle and predicting its function in a changing ocean. Using model organism studies, microcosm experiments (at Dauphin Island Sea Lab, AL), and time-series field studies with an autonomous sample collection instrument (at Monterey Bay, CA), this project is taking a taxon-specific approach to decipher the regulation of bacterial DMSP degradation.

This research addresses fundamental questions of how the diversity of microbial life influences the geochemical environment of the oceans and atmosphere, linking the genetic basis of metabolic potential to taxonomic diversity. The project is training graduate students and post-doctoral scholars in microbial biodiversity and providing research opportunities and mentoring for undergraduate students. An outreach program is enhance understanding of the role and diversity of marine microorganisms in global elemental cycles among high school students. Advanced Placement Biology students are participating in marine microbial research that covers key learning goals in the AP Biology curriculum. Two high school students are selected each year for summer research internships in PI laboratories.

[table of contents | back to top]

Program Information

Dimensions of Biodiversity (Dimensions of Biodiversity)

Website: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503446

Coverage: global

(adapted from the NSF Synopsis of Program)
Dimensions of Biodiversity is a program solicitation from the NSF Directorate for Biological Sciences. FY 2010 was year one of the program. [MORE from NSF]

The NSF Dimensions of Biodiversity program seeks to characterize biodiversity on Earth by using integrative, innovative approaches to fill rapidly the most substantial gaps in our understanding. The program will take a broad view of biodiversity, and in its initial phase will focus on the integration of genetic, taxonomic, and functional dimensions of biodiversity. Project investigators are encouraged to integrate these three dimensions to understand the interactions and feedbacks among them. While this focus complements several core NSF

programs, it differs by requiring that multiple dimensions of biodiversity be addressed simultaneously, to understand the roles of biodiversity in critical ecological and evolutionary processes.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1342694

[table of contents | back to top]