18s DNA sequence of potential prey of siphonophores

Website: https://www.bco-dmo.org/dataset/935469
Data Type: Cruise Results, experimental, Other Field Results
Version: 1
Version Date: 2024-08-13

Project
» Collaborative research: The effects of predator traits on the structure of oceanic food webs (SiphWeb)

Contributors	Affiliation	Role
Dunn, Casey W.	Yale University	Principal Investigator
Damian-Serrano, Alejandro	Yale University	Scientist
Merchant, Lynne M.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This dataset contains 18s DNA sequences of planktonic animals that are potential prey of siphonophores. This dataset expands the ability to study the feeding ecology of siphonophores and the structure of the open-ocean food web by facilitating the molecular identification of siphonophore gut contents. The new data presented here are for species that were underrepresented in existing sequence datasets.

Table of Contents

Coverage

Location: Water off California
Spatial Extent: **N**:41.06 **E**:-64.58 **S**:19.58 **W**:-160.37
Temporal Extent: 2016-06-11 - 2020-10-04

Methods & Sampling

The methods below are adapted from Damian-Serrano, Hetherington et al. (2022).

We built an 18S gene barcoding database of potential siphonophore prey items to expand on the available reference sequences in public databases. To do this, we collected 60 specimens of 30 species of zooplankton and micronekton from the California Current using a Tucker trawl. We targeted plausible prey species from motile open-ocean taxa that cohabitate with siphonophores and are underrepresented in SILVA databases (high quality ribosomal RNA databases, arb-silva.de), including fishes, crustaceans, jellyfishes, urochordates, chaetognaths, polychaetes, and mollusks. Specimens were photographed alive, then tissue was sampled and frozen, and finally the rest of the animal was fixed in formalin as a voucher to be identified and preserved at the Yale Peabody Museum of Natural History.

DNA extraction, quality control, PCR, and amplicon cleanup was carried out in a similar fashion as the metabarcoding protocol described above (and detailed in Damian Serrano (2022)), using the PCR program with an annealing temperature of 54°C, and a single pair of primers (166F and 134R), spanning the full extent of the sequence containing all barcode regions used in the gut content metabarcoding (from V3 to V9). Purified amplicons were sent in plates with the forward and reverse primer separately for Sanger sequencing from both ends at the Yale DNA Analysis Facility. A total of 89 newly-submitted sequences were then assembled and trimmed at a 95% quality cutoff in Geneious and concatenated with the latest SILVA database (SILVA_138_SSURef_NR99 downloaded on February 23, 2021) pruned to remove non-eukaryotic sequences.

Taxonomic identities of the reads in siphonophore gut contents were assigned using the assignment software METAXA2 (Bengtsson-Palme 2015) with a 70% reliability cutoff, comparing the sequences against our custom-built library built using SILVA138 that includes the new prey sequences.

Siphonophore collection

In order to sample a representative set of taxa across the siphonophore phylogeny, we targeted a set of 41 species (aiming for 10 specimens per species) including cystonects, apolemiids, pyrostephids, euphysonects, and calycophorans from shallow and deep waters. Most species were sampled from the Offshore California Current Ecosystem (OCCE) except for the Portuguese man-o-war *P. physalis*, which was collected off Bermuda in the Sargasso Sea; *Sulculeolaria chuni* and some *Nanomia* spp. (labeled as "Atlantic") which were collected off Rhode Island in the Block Island sound; *Forskalia* sp. M123-SS8 and shallow *Nanomia* sp. KiloMoana2018-BW7-4 which were collected off the coast of Hawaii. While all the *Nanomia* populations sampled in this study have been referred to as *Nanomia bijuga*, we suspect that there may be undescribed cryptic *Nanomia* species among the specimens sampled based on the disparate tentillum morphologies we observed. Therefore, we decided to have them labeled at the genus level. One *Nanomia* specimen (KiloMoana2018−BW7−4) was collected off the coast of Kona, Hawaii. The pleustonic (surface floating) *Physalia physalis* samples were collected manually using a bucket from a small boat. Species found between the 0-20m deep were collected using blue water diving techniques following the guidelines in Haddock & Heine (2005). Species from 200-4000m were collected using ROVs. All animals were collected live and brought back to the ship (or field station in Bermuda for *P. Physalis*) for dissection. Live colonies were photographed (sometimes recorded on video), and zooids of diagnostic value (nectophores, bracts, tentacles) were dissected, when possible, fixed in 4% formalin, and stored as vouchers at the Yale Peabody Museum of Natural History (voucher catalog numbers provided in specimen metadata S15 Table of Damian-Serrano et al., 2015).

Gut content metabarcoding

Shortly after collection of the live specimens, we dissected and pooled several gastrozooids from each colony, making sure that those with visible gut contents are included in addition to several other without conspicuous prey, and also including visible egested food pellets at the bottom of the sampling container.

To extract DNA, we digested the samples with proteinase K at 56°C for 1-2h, and used the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) eluting twice at 56°C for 10min into a final volume of 100µl. For barcode amplification, we used a set of six primer pairs that amplify six barcode regions within the 18S gene ('V3', 'V5-V7S', 'V5-V7L', 'V7', 'V7p+V8', and 'V9'). The primers were designed using Geneious 11.1.5 (Kearse 2012), constraining the search to short (>300 bp) amplicon products with a high chance of remaining uncleaved after digestion in the gastrozooid, flanked by priming sites conserved (to a maximum mismatch of 3bp) across metazoans. The search for conserved priming sites was conducted on an alignment of 18S genes from 975 species across all metazoan phyla downloaded from GenBank (available in github.com/dunnlab/siphweb_metabarcoding/Primer_design). The primer search was optimized to only retrieve non-degenerate primer pairs with compatible annealing temperatures and without problematic dimerization and hairpin temperatures. Primer sequences are shown in Table 1 (Damian Serrano 2022), and their properties can be found in Table T1 in the protocol (Damian Serrano 2022).

Prey reference database

In order to enhance the accuracy of the taxonomic assignments of reads, we also built an 18S gene barcoding database of potential prey items to expand on the available reference sequences in public databases. To do this, we collected 60 specimens of 30 species of zooplankton and micronekton from the OCCE using a Tucker trawl. We targeted plausible prey species from motile open-ocean taxa that cohabitate with siphonophores and are underrepresented in SILVA databases, including fishes, crustaceans, jellyfishes, urochordates, chaetognaths, polychaetes, and mollusks. Specimens were photographed alive, then tissue was sampled and frozen, and finally the rest of the animal was fixed in formalin as a voucher to be identified and preserved at the Yale Peabody Museum of Natural History. DNA extraction, quality control, PCR, and amplicon cleanup was carried out in a similar fashion as the metabarcoding protocol described above (and detailed in Damian Serrano 2022), using the PCR program with an annealing temperature of 54°C, and a single pair of primers (166F and 134R), spanning the full extent of the sequence containing all barcode regions used in the gut content metabarcoding (from V3 to V9). Purified amplicons were sent in plates with the forward and reverse primer separately for Sanger sequencing from both ends at the Yale DNA Analysis Facility. A total of 89 newly-submitted sequences were then assembled and trimmed at a 95% quality cutoff in Geneious and concatenated with the latest SILVA database (SILVA_138_SSURef_NR99 downloaded on February 23, 2021) pruned to remove non-eukaryotic sequences.

Data Processing Description

Bioinformatic pipeline

Amplicon libraries were demultiplexed by primer sequence using custom bash code. Primer sequences were removed using *cutadapt* (Martin 2011). The forward and reverse reads were matched and repaired using *bbtools* (Bushnell 2017), then denoised and de-replicated using the DADA2 (Callahan 2016) plugin in QIIME2 (Bolyen 2019) with a truncation quality threshold of 28. We *de novo* clustered the unique features into operational taxonomic units (OTUs) using the VSEARCH (Rognes 2016) plugin in QIIME2 with a similarity threshold of 95%. To reduce computational load, only the top 100 most abundant features among the clustered OTUs were selected for taxonomic assignment. Taxonomic identities were assigned using the assignment software METAXA2 (Bengtsson-Palme 2015) with a 70% reliability cutoff, comparing the sequences against the SILVA123.1 reference library (Quast 2012), and against our custom-built library built using SILVA138 as a foundation. The SILVA123.1 database contains 61383 eukaryotic reference sequences, while our custom database (built off SILVA138.1) contains 79044. Animals in the SILVA123.1 taxonomy are annotated to the ranks of superphylum, phylum, subphylum, class, subclass, order, family, genus, and species. However, the SILVA138.1 animal taxonomy was annotated at the levels of clade (e.g. Bilateria, Protostomia, Deuterostomia, Ecdysozoa, Lophotrochozoa), phylum, class, subclass, order, suborder, and species. All bioinformatics analyses were carried out in the Yale High Performance Computing Cluster. The taxonomic assignments and read count data were merged, then parsed to match the sample of origin and the DNA sequence they derived from. Sequence post-processing scripts can be found in the GitHub repository (https://github.com/dunnlab/siphweb_metabarcoding/Scripts).

[table of contents | back to top]

Related Publications

Bengtsson-Palme, J., Hartmann, M., Eriksson, K. M., Pal, C., Thorell, K., Larsson, D. G. J., & Nilsson, R. H. (2015). metaxa2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. Molecular Ecology Resources, 15(6), 1403–1414. Portico. https://doi.org/10.1111/1755-0998.12399
Software
,
Software

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... Asnicar, F. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology, 37(8), 852–857. doi:10.1038/s41587-019-0209-9
Software

Bushnell, B., Rood, J., & Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. PLOS ONE, 12(10), e0185056. https://doi.org/10.1371/journal.pone.0185056
Software

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13(7), 581–583. doi:10.1038/nmeth.3869
Software

Damian Serrano, A. (2022). DNA metabarcoding protocol for siphonophore gut contents v2. https://doi.org/10.17504/protocols.io.5qpvo57o7l4o/v2
Methods

Damian-Serrano, A., Hetherington, E. D., Choy, C. A., Haddock, S. H. D., Lapides, A., & Dunn, C. W. (2022). Characterizing the secret diets of siphonophores (Cnidaria: Hydrozoa) using DNA metabarcoding. PLOS ONE, 17(5), e0267761. https://doi.org/10.1371/journal.pone.0267761
Results

Haddock, S. H. D., Heine, J. N., United States National Oceanic and Atmospheric Administration, California Sea Grant College Program, & National Sea Grant College Program (U.S.). (2005). Scientific blue-water diving. California Sea Grant College Program. https://isbnsearch.org/isbn/9781888691139
Methods

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647–1649. doi:10.1093/bioinformatics/bts199
Software

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, 17(1), 10. doi:10.14806/ej.17.1.200
Software

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(D1), D590–D596. doi:10.1093/nar/gks1219
Methods

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ, 4, e2584. Portico. https://doi.org/10.7717/peerj.2584
Software

[table of contents | back to top]

Parameters

Parameter	Description	Units
BioProject	The NCBI identifier for the BioProject associated with the data	unitless
SRA Study	The NCBI identifier for the SRA study associated with the data	unitless
Experiment	The NCBI identifier for the sequencing experiment	unitless
Run	The unique NCBI identifier for each sequencing run	unitless
BioSample	The NCBI identifier for the BioSample associated with the data	unitless
geo_loc_name	The specific geographical location where the sample was collected	unitless
lat	latitude, South is negative	decimal degrees

lon	longitude, West is negative	decimal degrees
Collection_Date	The date when the sample was collected	unitless
BioSampleModel	The model describing the BioSample (e.g., human, environmental)	unitless
Organism	The scientific name of the organism from which the sample was taken	unitless
TaxID	NCBI taxon identifier	unitless
isolation_source	The source from which the sample was isolated (e.g., soil, blood)	unitless
Host	The host organism from which the sample was obtained	unitless
Sample Name	The name of the sample	unitless
source_material_id	The NCBI identifier for the source material	unitless
spots	The spot model is Illumina GA centric. The flowcells have the locations where the adapters have stuck them to the glass of the lane. There are X and Y coordinates that identify these 'spots'. As the camera reads the fluorescent flashes during sequencing, the coordinates indicate which spot the new base is added to. All of the bases for a single location constitute the spot.	unitless
Bases	The total number of bases sequenced	unitless
spots_with_mates	spots with mates	unitless
AvgSpotLen	The average length of the spots (reads) in the run	Base pairs
size_MB	The total size of the sequencing data files	MB
Extraction	The identifier for the extracted DNA	unitless
Index	Index sequence used in the library preparation	unitless
samp_collect_device	The device used to collect the sample	unitless
Platform	The sequencing platform used (e.g., Illumina, PacBio)	unitless

Instrument	The sequencing instrument used (e.g., Illumina MiSeq)	unitless
Assay_Type	The type of sequencing assay performed (e.g., RNA-Seq, WGS)	unitless
Library_Name	The name of the sequencing library	unitless
LibrarySource	The source material for the library (e.g., genomic, transcriptomic)	unitless
LibraryLayout	The layout of the sequencing library (e.g., paired-end, single-end)	unitless
LibrarySelection	The method used to select the nucleic acid library (e.g., PCR, random)	unitless
Consent	Information on the consent for data usage	unitless
create_date	The date when the submission was created	unitless
ReleaseDate	The date when the data was released to the public	unitless
version	The version of the NCBI submission	unitless
Center_Name	The name of the sequencing center	unitless
DATASTORE_filetype	The file type stored in the NCBI DataStore (e.g., FASTQ, BAM)	unitless
DATASTORE_region	The geographical region of the DataStore	unitless
DATASTORE_provider	The provider of the DataStore where files are kept	unitless

[table of contents | back to top]

Project Information

Collaborative research: The effects of predator traits on the structure of oceanic food webs (SiphWeb)

Coverage: North Pacific

Food webs describe who eats whom, tracing the flow of energy from plants up to large animals. While many connections in food webs on land are quite familiar (lions eat antelope and antelope eat grass, for example), there are large gaps in our understanding of ocean food webs. Closing these gaps is critical to understanding how nutrients and energy move through ocean ecosystems, how organisms interact in the ocean, and how best to manage ocean resources. This project will study ocean food web structure with a focus on siphonophores, an abundant group of predators in the open ocean that range in length from less than an inch to more than one hundred feet. Siphonophores are closely related to corals and many jellyfish. They are known to be important predators within ocean food webs, but they are difficult to study because they live across great ocean depths and are gelatinous and fragile. The details of what they eat, as well as many other features of their biology, remain poorly known. This project will combine direct observations of feeding, genetic analysis of siphonophore gut contents, and stable isotope analyses to identify what different species of siphonophores eat. The team will also examine why they eat what they do. This will provide a new understanding of how the structure of food webs arise, aiding in our ability to predict future changes to food webs as the global climate shifts. Siphonophores feed in a very unique manner--they have highly specialized tentacles that are used solely for capturing prey--thus, the prey captured is determined largely by the anatomy and function of these tentacles. The project will describe these tentacles, reconstruct their evolutionary history, and investigate how evolutionary shifts in tentacle structure have led to changes in diet. This project will train one PhD student, one Master's student, a postdoc, and undergraduate students, including individuals of underrepresented groups. This project will support the production of scientifically rigorous yet engaging videos, foster the expansion of a citizen-science program, and create K-12 teaching modules.

This project will advance three scientific aims: First, it will identify the diet of a diverse range of siphonophores using DNA metabarcoding of gut contents and prey field, remotely operated vehicle (ROV) video of prey encounters, and stable isotope analysis. These approaches are highly complementary and allow for extensive cross validation. Second, the project will characterize the selectivity of siphonophore diets by comparing them to the relative prey abundances in the habitats of each of these species. Third, the project will characterize the structure of the siphonophore prey capture apparatus across species through detailed morphological analysis of their tentacles and nematocysts. These data will be integrated in an ecological and evolutionary framework to identify predator features associated with prey specialization. In a larger context, addressing these questions will advance our understanding of oceanic predation by revealing how evolutionary changes in predator selectivity correspond to evolutionary changes in habitat and feeding apparatus and how these changes shape current food web structure in the open ocean. We will test and refine an integrated approach to describing the structure and origin of food web topology, and evaluate the potential for phylogenetic relationships to explain prey selectivity.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1829835

[table of contents | back to top]