

# Sharing Data Through The Biological and Chemical Oceanography Data Management Office



---

Gordon and Betty Moore Foundation  
Marine Microbiology Initiative  
January 15, 2020

# Outline

1. Who we are; what we do
  - a. History and mission
  - b. Why share your data
  - c. Supporting the data lifecycle
2. Sharing data through BCO-DMO
  - a. Prepare and submit data files & metadata
  - b. What happens next...
3. Discovery & access of data
4. Tips for successful data sharing
  - a. Data type consideration
  - b. Data formatting tips that speed processing!

4	A	B	C	D	E	F
1	Table . Pore water chemical concentration data and location from push cores th					
2	Some Mn and Fe values are below detection and are indicated by the less than					
3						
4	Core	Depth	pH	alkalinity	Nitrate	Chlori
5		cm		mmol/kg	umol/kg	mmol
6	METHOD		electrode	titration	colorimetric	titrat
7						
8	2014 bottom water		7.92	2.32	21.1	544
9						
10	J2-733-PC 1 - Did not hit bottom (18:24) and was positioned next to PC 2. 22					
11		2	7.68	2.08	22.3	546
12		6	7.69	2.11	23.8	546
13		10	7.71	2.20	25.1	545
14		13	7.7	2.22	25.8	547
15		16	7.69	2.22	24.7	544
16		18	7.71	2.22	24.6	546
17						
18	J2-733-PC 2 - Did not hit bottom (18:27) and was positioned next to PC 1. 22					
19		2	7.69	2.13	21.9	548
20		7	7.70			
21		11	7.73			
22		15	7.70			
23		18	7.72			
24		20	7.70			
25		22	7.72			
26						
27	J2-733-PC 4 - Hit bottom (18:01) and					
28		3	7.67			
29		6	7.69			
30		8	7.71			
31		10	7.75			
32		12	7.71			
33		14	7.73			
34						
35	J2-733-PC 6 - Did not hit bottom (19:					
36		4	7.71			
37		7	7.71			



	A	B
1	Site Code	Site Code
2	1	Dittlif Point
3		
4	2	Cocoloba Cay
5		
6		
7	3	Joel's Shoal
8		
9	4	White Point
10		
11	5	Europa Bay
12		
13		
14	6	
15		
16		
17	7	
18		
19		
20	8	
21		
22		
23	9	
24		
25		
26	S	
27		
28		

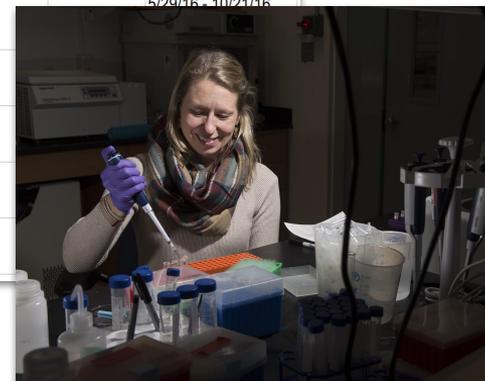


Photo credits: WHOI image bank

# The Biological & Chemical Oceanography - Data Management Office

- Established in 2006, through the merging of the data management offices for the NSF-funded U.S. GLOBEC and JGOFS programs.
- BCO-DMO is primarily supported by NSF to provide data management services at no cost to the investigators funded through the same programs that fund BCO-DMO.
- We now manage data from thousands of projects from researchers (PIs) across the U.S. studying a wide variety of oceanographic subdisciplines.



# Who are we?



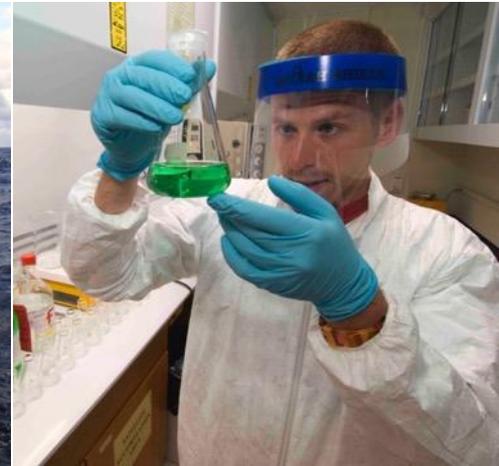
**Peter Wiebe, Mak Saito, Amber York, Karen Soenen, Nancy Copley**



**Danie Kinkade, Adam Shepherd, Shannon Rauch, Mathew Biddle, Tina Haskins**

# Mission

To work closely with principal investigators to curate, publish, and serve data and information from federally and privately funded research projects within the ocean sciences.



# Why Share Your Data? Benefit to...



## YOU:

- Credit for your hard work!
- Can increase citations
- Can boost collaborations
- Increases exposure
- Satisfies funder requirements



## Research Community:

- Builds a resource
- Enables new discoveries
- Sparks collaborations
- Allows for transparency and reproducibility of research results



## Society at-large:

- Transparency boosts public confidence in scientific process
- Can contribute to management and policy
- Availability to audiences outside of research (education, general public)

# Making Data ...

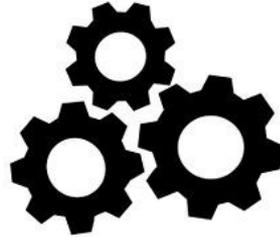
F  
indable



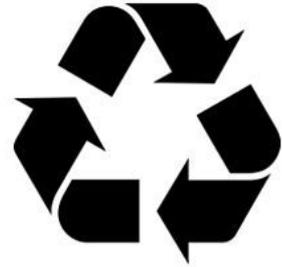
A  
ccessible



I  
nteroperable



R  
eusable



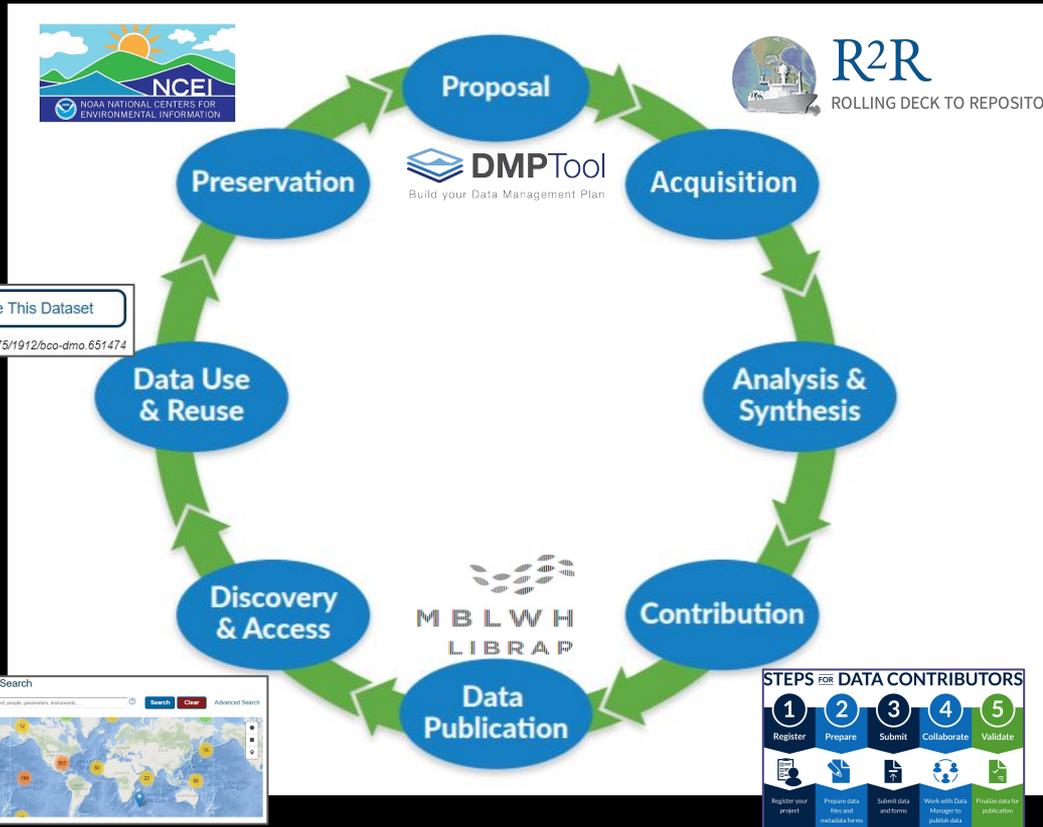
**F**indable: Data are linked to descriptive persistent metadata.

**A**ccessible: Data and metadata are open, free, and machine accessible.

**I**nteroperable: Data and metadata are standardized and use vocabularies.  
Data points to related metadata.

**R**eusable: Metadata are rich, and employ usage licenses, provenance, and community standards.

# BCO-DMO supports data stewardship throughout the data lifecycle



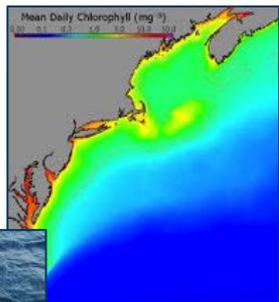
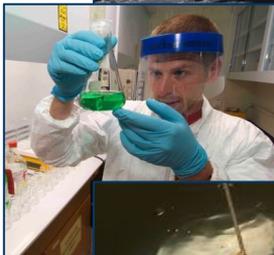
- Provide guidance on data formatting and standards;
- Apply gross QC (e.g lat/lon checking);
- Capture and record metadata;
- Make data and metadata available online (restricting access as appropriate);
- Publish with digital object identifiers (DOIs) and recommended citations.
- Ensure final archive in appropriate national data center

# BCO-DMO Data Catalog

DATASETS = >9000  
 CONTRIBUTORS = >2500  
 PROJECTS = >1000

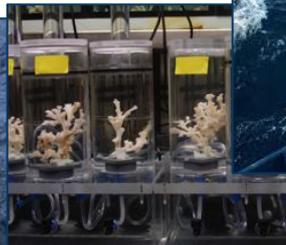
- Biological
- Chemical
- Biogeochemical
- Physical
- Geophysical

- ASCII Text (tabular)
- Binary (e.g., NetCDF)
- Images
- Acoustics
- Application (e.g., Matlab)
- Links to other data (e.g., NCBI accessions)



- Molecular to Megafaunal
- Local to Global
- Discrete to synoptic

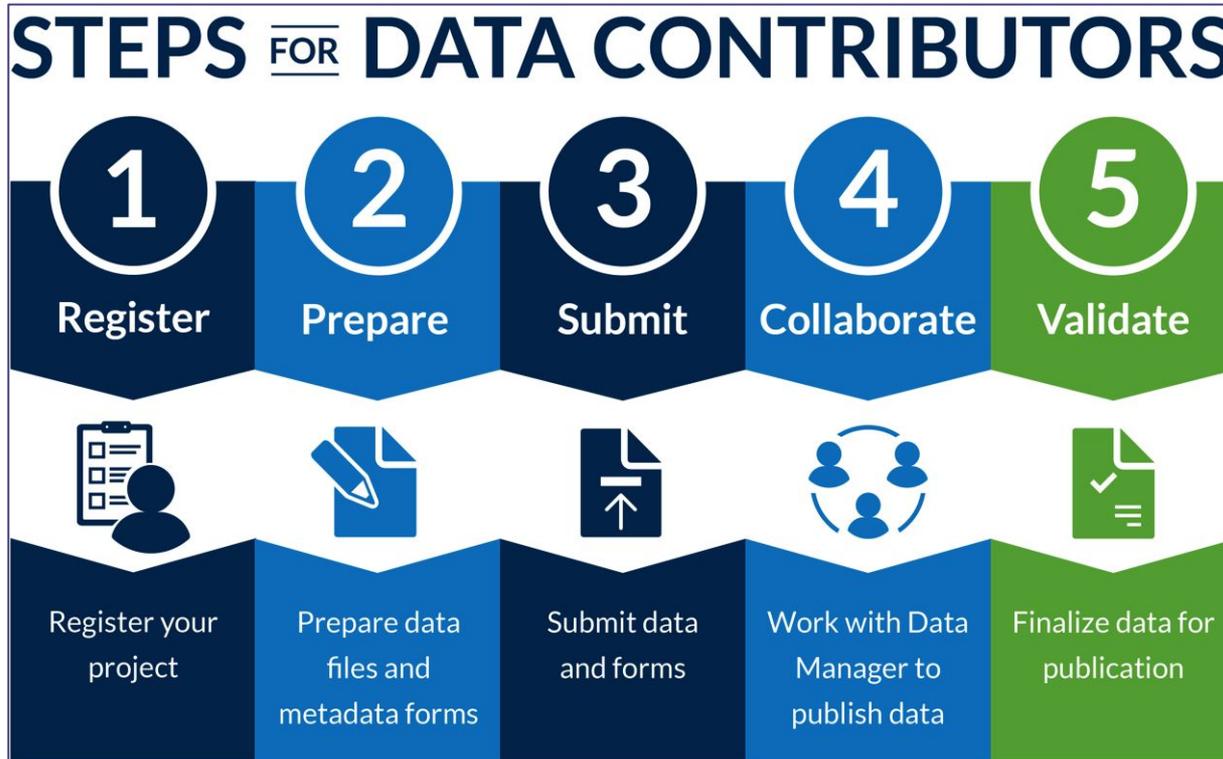
1	Site	Date	Plot	Species	Weight	Adult									
2	DeepWell	2/13/2010		1 DIPO	12.1j										
3	Deep Well	Feb-10		2 Pero	13.22j										
4	rioSalado	2/13/2010	1a	pero	16. N										
5	riuSladu	*	1*	CleGap	18.92	gut away									
6				Mean1	15.06j										
7															
8															
9															
10															
11															
12	Rodent Trapping			MJK & ALN	10-Apr-10										
13	Site	Plot	Adult	Species	grams	Comments									
14	deep well		1 y	woodrat	13										
15	riosalado		2 y	PERO	24.5										
16	riosalado		3 y	Clegap	91										
17															
18															
19															
20															



- *In situ*
- Laboratory
- Remotely sensed
- Synthetic/derived

- Parameters = > 1400
- From Abundance to Zenith Sun Angle

# Sharing Your Data Through BCO-DMO





# Prepare to Share: Metadata

Metadata: describes your data (who, what, where, when and how it was collected. Includes funding sources, instrumentation, etc.).

Metadata templates are available on our Resources page help guide you in the types of metadata needed to describe your project output.

The image shows a screenshot of the BCO-DMO website. At the top, the navigation menu includes 'DATA', 'RESOURCES' (circled in red), and 'ABOUT US'. Below the menu is a 'DATABASE' table with the following data:

Category	Count
Programs	44
Projects	1,116
Deployments	2,924
Platforms	603
Datasets	9,490
Instruments	490
Parameters	1,420
People	2,784
Affiliations	601
Funding	94
Awards	2,091

Below the table is a 'GEOSPATIAL ACCESS' section with a world map. Underneath is a 'CONTRIBUTE DATA' section with a 'Getting started' subsection containing links for 'How-to Guide', 'FAQs', and 'Quick Start Guide (pdf)'. A 'Metadata Forms (.rtf files)' subsection lists 'Program Metadata Form', 'Project Metadata Form', and 'Deployment Metadata Form'. The 'How to Register a Project' link is circled in red, with an arrow pointing to the 'Dataset Metadata Submission Form' on the right.

The 'Dataset Metadata Submission Form' is titled 'BCO-DMO Dataset Metadata Submission Form'. It includes instructions: 'All data should be reported in table form. For full instructions, see our "See How to Get Started" page: <http://www.bco-dmo.org/how-get-started>. Please send your completed form or questions to [info@bco-dmo.org](mailto:info@bco-dmo.org).' The form has fields for 'Dataset Name', 'Dataset Description', 'Originating PI name and contact information' (Name, Email, Phone, Mailing Address), 'ORCID', 'Affiliation/Institution during data acquisition', 'Co-PI name(s) and contact information' (Name, Email, Phone, Mailing Address), and another 'ORCID' field.

# Prepare to Share: Metadata

The contents of your metadata form are directly used to populate the public dataset landing page.

Allows your data to be discovered, understood, and re-used by others.

**BCO-DMO**  
Biological & Chemical Oceanography Data Management Office

DATA RESOURCES ABOUT US Enter search terms

**DATABASE**

- Programs 39
- Projects 923
- Deployments 2,740
- Platforms 502
- Datasets 9,105
- Instruments 489
- Parameters 1,414
- People 2,506
- Affiliations 561
- Funding 87
- Awards 1,753

**Dataset: Wave Glider - Fluorescence - C3**

Get Data Map It Cite This Dataset

Spatial Extent: N 29.47969 E -144.82868 S 19.99672 W  
Temporal Extent: 2015-05-20 - 2015-11-02

Project: Long Duration AUVs as tools to explore scale feature aggregate interactions (MAGI) (MAGI)  
Principal Investigator: Dr. Tracy A. Villareal (University of Texas at Austin, UT Austin)

Co-Principal Investigator: Dr. Cara Wilson (National Oceanic and Atmospheric Administration - Southwest Fisheries Science Center, NOAA SWSFC, ERD)

Contact: Dr. Tracy A. Villareal (University of Texas at Austin, UT Austin)

BCO-DMO Data Manager: Amber York (Woods Hole Oceanographic Institution, Woods Hole, MA)

Version Date: 2017-07-07

Restricted: No

Validated: Yes

Current State: Final no updates expected

Data URL: <https://www.bco-dmo.org/dataset/653>

Fluorescence data (C3) collected by the AUV Honey Badger (Wave Glider) (MAGI project)

Expand/Collapse All

- Description**

This dataset includes chlorophyll, phycoerythrin, and CDOM data from the AUV Honey Badger (Wave Glider model V2) during a 2015 deployment in the North Pacific Ocean.

For more information on project MAGI and a description of Honey Badger, see: <http://oceanview.pfeg.noaa.gov/MAGI/>

Additional support was provided by the National Science Foundation (NSF).
- Acquisition Description**

Data were collected at the surface by the the AUV Honey Badger (Wave Glider) (MAGI project). This deployment in the North Pacific Ocean is part of the MAGI project. For more information on project MAGI and a description of Honey Badger, see: <http://oceanview.pfeg.noaa.gov/MAGI/>

Chlorophyll, phycoerythrin, and CDOM data were acquired from two float-mounted Turner Designs' C3™ Submersible Fluorometers. Biofouling is a big issue for any long deployment vehicle, and having two sensors provided the redundancy needed for data quality assurance. No calibration was deemed useful due to the duration of the mission and nature of questions asked.
- Processing Description**

No calibration was deemed useful due to the duration of the mission and nature of questions asked. Sensors returned only RFU.

BCO-DMO Data Manager Processing Notes:

  - \* added a conventional header with dataset name, PI name, version date
  - \* modified parameter names to conform with BCO-DMO naming conventions
  - \* blank values replaced with no data value 'nd'
  - \* values of "NaN" and "nan" changed to "nd"

## Dataset Metadata Form

**BCO-DMO**  
Dataset Metadata Submission Form

All data should be reported in table form. For full instructions, see our "See How to Get Started" page: <http://www.bco-dmo.org/how-get-started>. Please send your completed form or questions to [info@bco-dmo.org](mailto:info@bco-dmo.org).

Dataset Name [ Preferred short name for the dataset (preferably 30 characters or less) ]

Dataset Description: [ Brief sentence describing these data (preferably 60 characters or less) ]

Originating PI name and contact information:  
Name:

If you have an ORCID number, you can register for one at <https://orcid.org/register>. ORCID's are unique persistent identifiers assigned to individuals to enable easy identification and ensure proper credit for work. It is free to register. ]

Acquiring data acquisition:  
Contact information: [ Contact information for Co-PIs. Repeat as necessary ]

Email:  
Phone:  
Mailing Address:

ORCID: [ If you don't currently have an ORCID number, you can register for one at <https://orcid.org/register>. ORCID's are unique persistent identifiers assigned to individuals to enable easy identification and ensure proper credit for work. It is free to register. ]

Affiliation/Institution during data acquisition:

**Dataset Name**

**Dataset Description**

**Methodology**

**Sampling and analytical procedures**

# Prepare to Share: Metadata Considerations

- How were the data generated?
  - How were samples collected?
  - How were they processed and analyzed?
    - Methods - include references and citations
    - Instruments - include manufacturer and model; as well as calibration information! (for both lab instruments and shipboard instruments)
    - Be specific; e.g. describe all the sensors on the CTD (not just "CTD")
- Describe the quality control
  - Replicate samples, inter-comparisons, reference materials used
- What software or scripts did you use?
  - Are they publicly accessible?
  - Modeling project? Describe the inputs, parameters, conditions, etc.

# Submit Data and Metadata

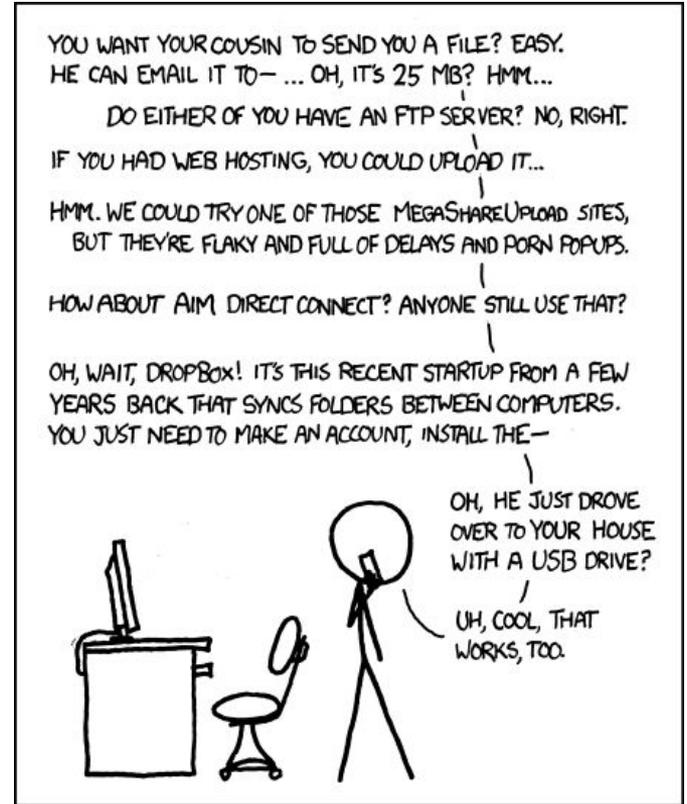
Email metadata forms and data files

[info@bco-dmo.org](mailto:info@bco-dmo.org)

Identify your MMI Project



Talk to us if data are too large for email



I LIKE HOW WE'VE HAD THE INTERNET FOR DECADES, YET "SENDING FILES" IS SOMETHING EARLY ADOPTERS ARE STILL FIGURING OUT HOW TO DO.

# Collaborate and Validate

**/BCO-DMO/MedusanFeeding**

Directory Documentation Downloads

Level 0 Next Level Flat Listing

# Medusae Predator-Prey Interactions  
# S. Colin  
# Version 1 March 2017

Cnidaria species  
Cassiopea  
food\_type  
Artemia  
individual  
7

video\_id  
Clip006  
time evaded ingested tentacle  
0:14:09 XXX nd nd  
0:31:18 XXX nd nd

video\_id  
Clip009  
time evaded ingested tentacle  
0:13:15 nd nd nd  
0:14:26 X nd nd  
0:20:19 XX nd nd  
0:21:14 X nd nd  
0:22:00 nd nd nd  
0:22:29 X nd nd

AN	AO	AP	AQ	AR	AS	A1	notes
Time (minute)	E	D	C.T.	C.O.	C.E.		
2:13	X						
3:23	X		X				
4:18	X						
4:41	XX						
8:06	X						
9:18	X						
12:16			X				
13:11	X						
15:05	X						moved out
15:24			X				
19:1	X						
21:05	XX						
23:08			X				
23:08	X						
25:13	X						moved out
24:1				XX			
36:26	XXX						
40:15	X						moved out
41:12	X						
41:12			X				
46:12	XXX						
49:13	X						
53:19			X				
55:13			X				

**BCO-DMO**  
Biological & Chemical Oceanography Data Management Office

DATA RESOURCES ABOUT US Enter search terms

**Dataset: Capture efficiency of scyphomedusae**

View Edit Request a DOI Delete Revisions GCMD DIF Record Add new dataset Copy Contact Status Node Queues

Linked Data URI: <http://dx.doi.org/10.1575/1912/bco-dmo.683750> [Cite This Dataset](#)

Get Data DOI:10.1575/1912/bco-dmo.683750.1 Temporal Extent: 2015 - 2015

VIEW TABLE SUBSET DATA CSV TSV MATLAB netCDF

**Project:** RUI: Collaborative Research: What's their impact?: Quantification of medusan feeding mechanics as a tool for predicting medusan predation (Medusan Feeding Mechanics)

**Principal Investigator:** Dr Sean Colin (Roger Williams University, RWU)

**BCO-DMO Data Manager:** Hannah Ake (Woods Hole Oceanographic Institution, WHOI BCO-DMO)

**Version:** 1

**Version Date:** 2017-03-01

**Restricted:** No

**Validated:** Yes

**Current Status:** Final no updates expected

**BCO-DMO Issue Tracking ID:** <http://redmine.bco-dmo.org/issues/1927>

**Data URL:**  
[http://dmoserv3.bco-dmo.org/g/serve/BCO-DMO/Medusan\\_Feeding\\_Mechanics/medusae\\_interactions.html%0%7Bdir=dmoserv3.whoi.edu/g/dir/BCO-DMO/Medusan\\_Feeding\\_Mechanics/info=dmoserv3.bco-dmo.org/g/info/BCO-DMO/Medusan\\_Feeding\\_Mechanics/medusae\\_interactions%7D](http://dmoserv3.bco-dmo.org/g/serve/BCO-DMO/Medusan_Feeding_Mechanics/medusae_interactions.html%0%7Bdir=dmoserv3.whoi.edu/g/dir/BCO-DMO/Medusan_Feeding_Mechanics/info=dmoserv3.bco-dmo.org/g/info/BCO-DMO/Medusan_Feeding_Mechanics/medusae_interactions%7D)

**Dataset Title:** Raw capture efficiency data of scyphomedusae from video analysis collected in Woods Hole, MA beginning in 2015.

**Brief Description:** Raw capture efficiency data of scyphomedusae from video analysis.

**Abstract:**  
Raw capture efficiency data of scyphomedusae from video analysis collected in Woods Hole, MA beginning in 2015.

**Account:** dkinkade

**Logout**

**GEOSPATIAL ACCESS**

# Data Publication

Once data & metadata are validated by the submitter, a Digital Object Identifier (DOI) is assigned to the dataset.

This makes it easier to cite the dataset and discover it (e.g. from a publication to the repository).

The screenshot shows the BCO-DMO (Biological & Chemical Oceanography Data Management Office) website. The header includes the BCO-DMO logo and navigation buttons for DATA, RESOURCES, and ABOUT US, along with a search bar. The main content area is titled 'Dataset: Carbon flux' and features a 'Cite This Dataset' button highlighted with a red arrow. Below the button is the DOI: 10.1575/1912/bco-dmo.685952. A map shows the spatial extent of the dataset in the Caribbean region, with coordinates: N:24.94972 E:-80.45361 S:24.94972 W:-80.45361 and a temporal extent of 2013-05-01. The dataset details include: Project: Testing the sponge-loop hypothesis for Caribbean coral reefs (sponge-loop); Principal Investigator: Christopher Finelli (University of North Carolina - Wilmington, UNC-Wilmington); Co-Principal Investigator: Steven McMurray (University of North Carolina - Wilmington, UNC-Wilmington) and Dr Joseph Pawlik (University of North Carolina - Wilmington, UNC-Wilmington); Contact: Steven McMurray (University of North Carolina - Wilmington, UNC-Wilmington); BCO-DMO Data Manager: Nancy Copley (Woods Hole Oceanographic Institution, WHOI BCO-DMO); Version Date: 2017-03-27; Restricted: No; Validated: Yes; Current State: Final no updates expected; Data URL: <https://www.bco-dmo.org/dataset/685783/data>. The dataset is titled 'Carbon flux for the Caribbean giant barrel sponge Xestospongia muta (Sponge-loop)'. The left sidebar shows a 'DATABASE' table with counts for various categories: Programs (39), Projects (917), Deployments (2,736), Platforms (578), Datasets (9,168), Instruments (469), Parameters (1,414), People (2,493), Affiliations (557), Funding (87), and Awards (1,736). Below this is a 'GEOSPATIAL ACCESS' map and a 'CONTRIBUTE DATA' section with links for 'Getting started', 'How-to Guide', 'FAQs', and 'Metadata Forms'.

DATABASE	
Programs	39
Projects	917
Deployments	2,736
Platforms	578
Datasets	9,168
Instruments	469
Parameters	1,414
People	2,493
Affiliations	557
Funding	87
Awards	1,736

GEOSPATIAL ACCESS	

CONTRIBUTE DATA	
<b>Getting started</b>	
» How-to Guide	
» FAQs	
<b>Metadata Forms</b> (.rtf files)	
» Program Metadata Form	
» Project Metadata Form	

# Data Publication

Suggested citation format  
and license information is  
provided, facilitating  
attribution

The screenshot shows the BCO-DMO (Biological and Chemical Oceanography Data Management Office) website. The top navigation bar includes 'DATA', 'RESOURCES', and 'ABOUT US', along with a search bar. The main content area displays 'Dataset: Carbon flux' with a 'Get Data' button and a 'Cite This Dataset' button. A sidebar on the left lists database statistics: Programs (39), Projects (917), Deployments (2,736), Platforms (578), Datasets (9,168), and Instruments (469). A map of the Caribbean region is visible in the background. A white popup window is overlaid on the map, containing the following text:

**Data Citation:**

Finelli, C., Pawlik, J., McMurray, S. (2017) Carbon flux for the Caribbean giant barrel sponge *Xestospongia muta* (Sponge-loop). Biological and Chemical Oceanography Data Management Office (BCO-DMO). Dataset version 2017-03-27 [if applicable, indicate subset used]. doi:10.1575/1912/bco-dmo.685952 [access date]

**Terms of Use**

All data sets are licensed under a [Creative Commons Attribution 4.0 International License \(CC BY 4\)](#). Per the CC BY 4 license it is understood that any use of the data set will properly acknowledge the individual(s) listed above using the suggested data citation. If you wish to use this data set, it is highly recommended that you contact the original principal investigator(s) (PI). Should the relevant PI be unavailable, please contact BCO-DMO ([info@bco-dmo.org](mailto:info@bco-dmo.org)) for additional guidance. For general guidance please see the BCO-DMO [Terms of Use](#) document.

**Version Date:** 2017-03-27

**Restricted:** No

**Validated:** Yes

**Current State:** Final no updates expected

**Data URL:** <https://www.bco-dmo.org/dataset/685783/data>

**Carbon flux for the Caribbean giant barrel sponge *Xestospongia muta* (Sponge-loop)**

Expand/Collapse All

Archival Copy

At the bottom left, there is a 'CONTRIBUTE DATA' section with links for 'Getting started', 'How-to Guide', 'FAQs', and 'Metadata Forms (.rtf files)' including 'Program Metadata Form' and 'Project Metadata Form'.

# Data Discovery & Access

- BCO-DMO data holdings are freely accessible to the public
- No login or account creation needed
- Discoverable via text and geospatial search interfaces

**BCO-DMO**  
Biological & Chemical Oceanography Data Management Office

DATA RESOURCES ABOUT US Enter search terms

**DATABASE**

Programs	39
Projects	916
Deployments	2,737
Platforms	578
<b>Datasets</b>	<b>9,168</b>
Instruments	469
Parameters	1,414
People	2,497
Affiliations	558
Funding	87
Awards	1,737

**DATASET SEARCH**

carbon flux + Pawlik ⓘ Hide Advanced Search

**Collection Date** Start: 1995 End: 2018

**Status** Select a status **Validated** Yes

**Embargoed** No

**GEOSPATIAL ACCESS**

**CONTRIBUTE DATA**

**Getting started**

- » How-to Guide
- » FAQs

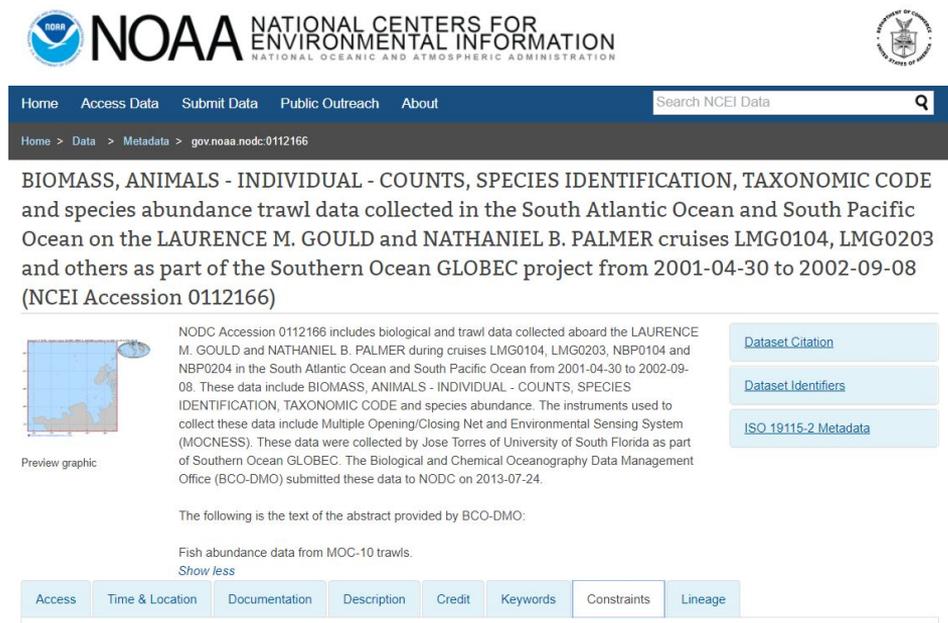
**Total: 137 (1 of 10)**

# of Results 10 Sort by Relevance

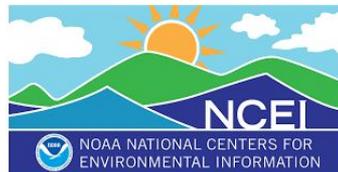
[+] Dataset	People	Award
[+] Carbon flux for the Caribbean giant barrel sponge <i>Xestospongia muta</i>	Principal Investigator	OCE-1558580

# Preservation

Once a project's data and metadata are published online at BCO-DMO, they are then submitted to an appropriate national data center for long-term preservation, e.g., the National Centers for Environmental Information (NCEI).



The screenshot displays the NOAA National Centers for Environmental Information (NCEI) website. At the top, there are navigation links: Home, Access Data, Submit Data, Public Outreach, and About. A search bar is located on the right. Below the navigation is a breadcrumb trail: Home > Data > Metadata > gov.noaa.nodc:0112166. The main heading reads: BIOMASS, ANIMALS - INDIVIDUAL - COUNTS, SPECIES IDENTIFICATION, TAXONOMIC CODE and species abundance trawl data collected in the South Atlantic Ocean and South Pacific Ocean on the LAURENCE M. GOULD and NATHANIEL B. PALMER cruises LMG0104, LMG0203 and others as part of the Southern Ocean GLOBEC project from 2001-04-30 to 2002-09-08 (NCEI Accession 0112166). To the left of the text is a small map of the Southern Ocean with a red box indicating the study area. Below the map is a 'Preview graphic' link. To the right of the text are three buttons: 'Dataset Citation', 'Dataset Identifiers', and 'ISO 19115-2 Metadata'. Below the main text, there is a section for the abstract provided by BCO-DMO, which includes the text: 'Fish abundance data from MOC-10 trawls. Show less'. At the bottom, there is a navigation bar with tabs: Access, Time & Location, Documentation, Description, Credit, Keywords, Constraints, and Lineage.



# STEPS FOR DATA CONTRIBUTORS



# Data Type Considerations

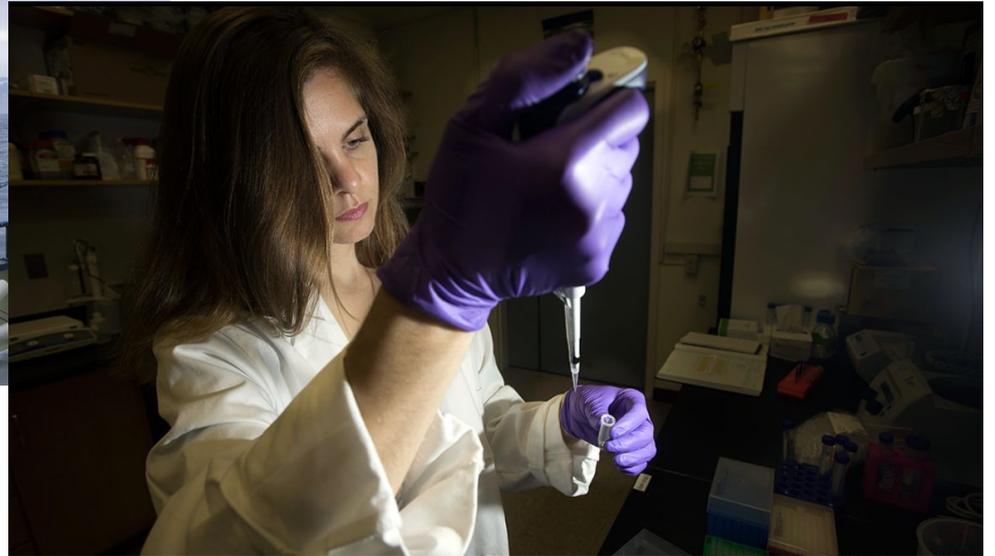
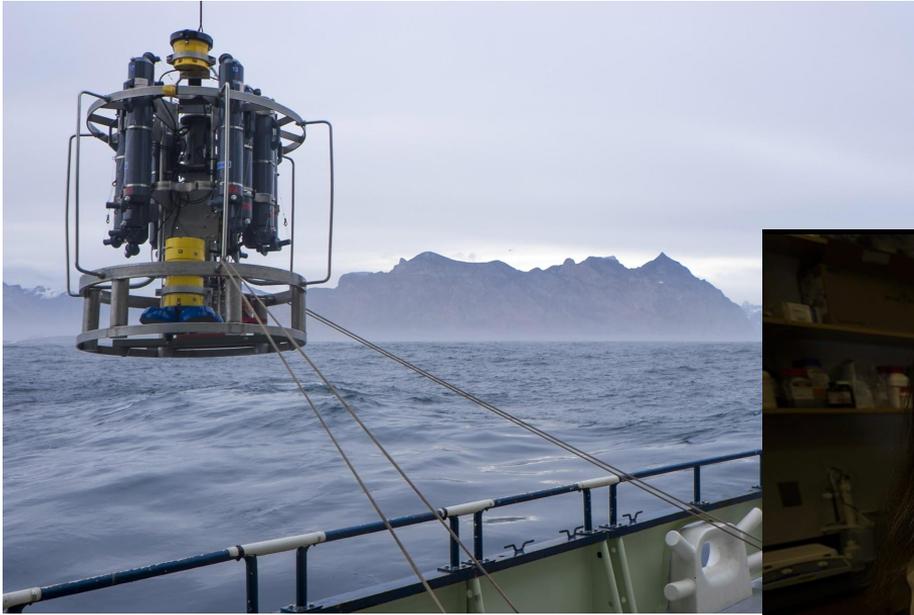


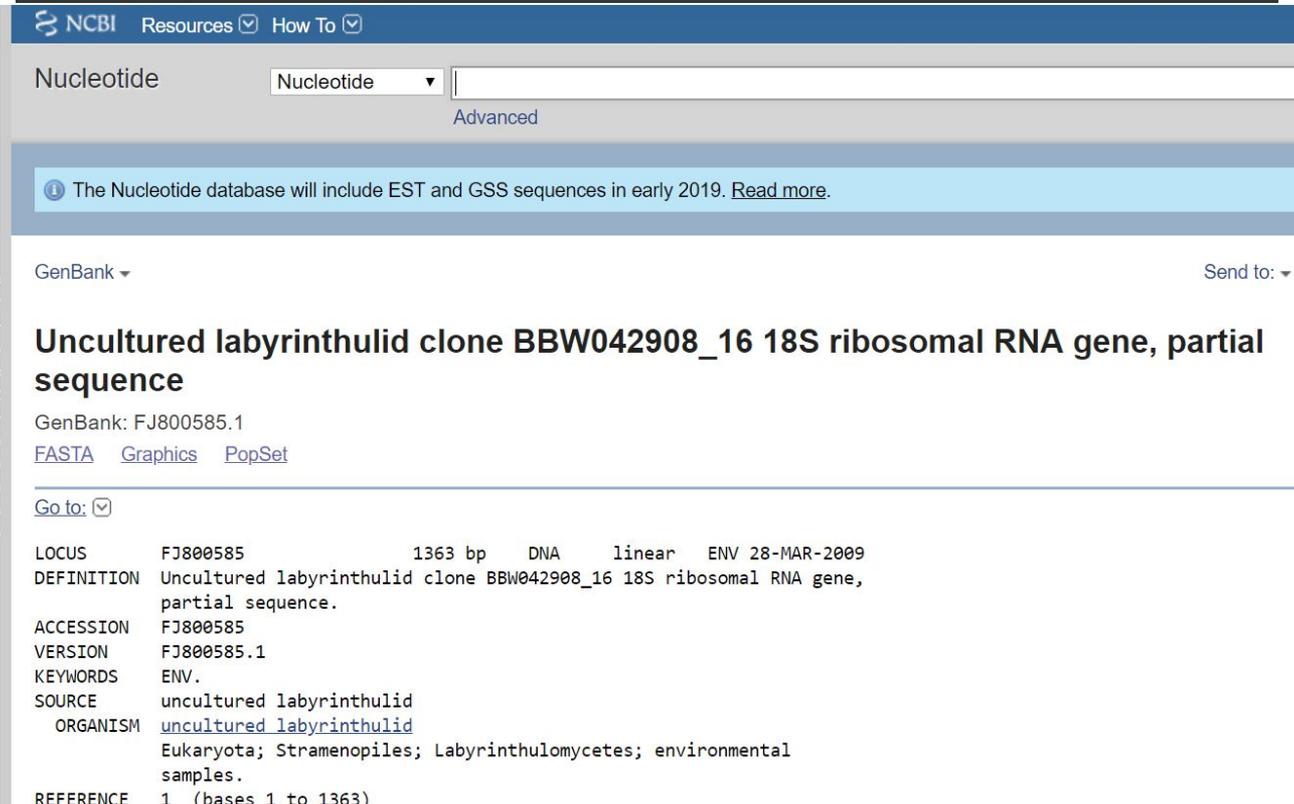
Photo credits: WHOI image bank

# Contributing Sequence Data

Sequences themselves are best served by specialized repositories, like NCBI's GenBank.

BCO-DMO serves the related environmental data and links out to NCBI, etc.

This allows all your project data to be discoverable from one place, but served by the domain repositories best suited to do so.



The screenshot shows the NCBI Nucleotide database search results for the entry FJ800585.1. The search criteria are set to 'Nucleotide' in the 'Advanced' search mode. A notification banner indicates that the database will include EST and GSS sequences in early 2019. The search results show the entry 'Uncultured labyrinthulid clone BBW042908\_16 18S ribosomal RNA gene, partial sequence' with GenBank ID FJ800585.1. Links for FASTA, Graphics, and PopSet are provided. The entry details include the locus (FJ800585, 1363 bp, DNA, linear, ENV 28-MAR-2009), definition (Uncultured labyrinthulid clone BBW042908\_16 18S ribosomal RNA gene, partial sequence), accession (FJ800585), version (FJ800585.1), keywords (ENV), source (uncultured labyrinthulid), organism (uncultured labyrinthulid), and reference (1 (bases 1 to 1363)).

NCBI Resources How To

Nucleotide Nucleotide Advanced

The Nucleotide database will include EST and GSS sequences in early 2019. [Read more.](#)

GenBank Send to:

**Uncultured labyrinthulid clone BBW042908\_16 18S ribosomal RNA gene, partial sequence**

GenBank: FJ800585.1

[FASTA](#) [Graphics](#) [PopSet](#)

Go to:

LOCUS FJ800585 1363 bp DNA linear ENV 28-MAR-2009

DEFINITION Uncultured labyrinthulid clone BBW042908\_16 18S ribosomal RNA gene, partial sequence.

ACCESSION FJ800585

VERSION FJ800585.1

KEYWORDS ENV.

SOURCE uncultured labyrinthulid

ORGANISM [uncultured labyrinthulid](#)

Eukaryota; Stramenopiles; Labyrinthulomycetes; environmental samples.

REFERENCE 1 (bases 1 to 1363)

# Contributing Seq

**BC-DMO**  
Biological & Chemical Oceanography Data Management Office

**DATABASE**

Dataset: 16S rRNA sequence data at East Pacific Rise /BCO-DMO/Microbe\_Vent

Directory Documentation Down

Level 0 Next Level Flat Listing

# 16S rRNA sequence data at East Pacific Rise  
# Alvin dives from RV/Atlantis cruises, 2004  
# PI: S. Sievert (WHOI)  
# version: 2017-04-20

=====  
sample\_descrip  
-----  
16S\_tags  
=====  
cruise\_id date\_deployed date\_recovered  
-----  
AT11-07 2004-02-07 2004-02-16  
AT11-07 2004-02-07 2004-02-16  
AT11-07 2004-02-03 2004-02-16  
AT11-07 2004-02-03 2004-02-16  
AT11-07/AT11-10 2004-02-03 2004-04-19  
AT11-07/AT11-20 2004-02-03 2004-11-12  
AT11-07/AT11-20 2004-02-03 2004-11-22  
AT15-06 nd 2006-06-27

=====  
sample\_descrip  
-----  
16S\_clones  
=====  
cruise\_id date\_deployed date\_recovered  
-----  
AT11-07/AT11-20 2004-02-03 2004-11-12  
AT11-07/AT11-21 2004-02-03 2004-11-12

**Dataset Title:** GenBank accession numbers for microbial DSV/Alvin during R/V Atlantis research cruises (Jan 2004 - June 2006) (Microbial Communities at Deep-Sea Vents project)

**BC-DMO**  
Biological & Chemical Oceanography Data Management Office

DATA RESOURCES ABOUT US Enter search terms

**DATABASE**

New Entry

Programs 45

Projects 1,116

Deployments 2,924

Platforms 603

Datasets 9,491

Related Resources

Instruments 490

Parameters 1,420

People 2,786

Affiliations 601

Funding 94

Awards 2,092

Contact Status 270

Archive 1,226

Account: dkinkade

Logout

**Project: An Integrated Study of Energy Metabolism, Carbon Fixation, and Colonization Mechanisms in Chemosynthetic Microbial Communities at Deep-Sea Vents**

View Edit Delete Add new project Node Queues

**Acronym/Short Name:** Microbial Communities at Deep-Sea Vents  
**Start Date:** 2011-10  
**End Date:** 2014-09  
**Datasets:** 14  
**Collections:** 7  
**Deployments:** 9  
**Cruises:** 7  
**MannedSubmarine:** 2

**Programs:**  
Dimensions of Biodiversity [Dimensions of Biodiversity]  
Center for Dark Energy Biosphere Investigations [C-DEBI]

Expand/Collapse All

- Description
- More Information
- Funding
- Dataset Collections

Dataset Short Name	Full Dataset Title
16S rRNA sequence and collection data	GenBank accession numbers for microbial 16S rRNA sequences collected at the East Pacific Rise by DSV/Alvin during R/V Atlantis research cruises (Jan. 2004 - June 2006) (Microbial Communities at Deep-Sea Vents project)
Chemosynthetic biofilm diversity	Samples sequenced from chemosynthetic biofilm communities in deep-sea hydrothermal vents collected on the R/V Atlantis AT26-10 in the East Pacific Rise, Pacific Ocean from 2013 - 2014 (Microbial Communities at Deep-Sea Vents project)
Incubation in	Results from shiboard high-pressure incubations of diffuse flow vent fluids collected from the

**GEOSPATIAL ACCESS**

# Contributing Experimental Data

**BCO-DMO**  
Biological & Chemical Oceanography Data Management Office

DATA RESOURCES

**DATABASE**

- New Entry
- Programs 45
- Projects 1,116
- Deployments 2,924
- Platforms 603
- Datasets 9,491**
- Related Resources
- Instruments 490
- Parameters 1,420
- People 2,786
- Affiliations 601
- Funding 94
- Awards 2,092
- Contact Status 270
- Archive 1,226

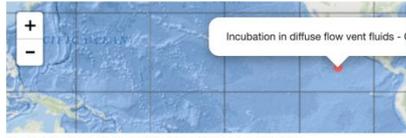
Account: dkinkade  
Logout

**GEOSPATIAL ACCESS**

## Dataset: Incubation in diffuse flow vent fluids

View Edit Request a DOI Delete Revisions GCMD DIF Record  
 Linked Data URI: <http://lod.bco-dmo.org/id/dataset/828993>  
 Get Data

VIEW TABLE SUBSET DATA CSV TSV MATLAB netCDF



Spatial Extent: N:9.83992 E:-104.2915 S:9.83992 W:-104.2915

**Project:** An Integrated Study of Energy Metabolism, Carbon Fixation, and Chemosynthetic Microbial Communities at Deep-Sea Vents (M)

**Principal Investigator:** Dr Dionysis Foustoukos (Carnegie Institution for Science)

**BCO-DMO Data Manager:** Nancy Copley (Woods Hole Oceanographic Institution)

**Version:** 2

**Version Date:** 2017-02-07

**Restricted:** No

**Validated:** Yes

**Current State:** Final no updates expected

**BCO-DMO Issue Tracking ID:** <http://redmine.bco-dmo.org/issues/1021>

**Data URL:**  
[http://dmosev3.bco-dmo.org/g/serv/BCO-DMO/Dim\\_Biodiversity/crab\\_spa\\_incubation\\_v2.html%7Bdir=dmoserv3.bco-dmo.org/g/info/BCO-DMO/Dim\\_Biodiversity/info-dmoserv3.bco-dmo.org/g/info/BCO-DMO/Dim\\_Biodiversity/crab\\_spa\\_incubation%20%7D](http://dmosev3.bco-dmo.org/g/serv/BCO-DMO/Dim_Biodiversity/crab_spa_incubation_v2.html%7Bdir=dmoserv3.bco-dmo.org/g/info/BCO-DMO/Dim_Biodiversity/info-dmoserv3.bco-dmo.org/g/info/BCO-DMO/Dim_Biodiversity/crab_spa_incubation%20%7D)

## /BCO-DMO/Dim\_Biodiversity/crab\_spa\_incubation\_v2 ---- Level 3

Directory Documentation Download & Other Operations

Level 0 Next Level Flat Listing

```

# Microbial incubations in diffuse flow vent fluids
# D. Foustoukos
# version: 2017-02-07 (added cell concentration and d15N data)
# replaces version: 2015-12-17
#
=====
description                date_start  date_end    flow_rate
-----
Crab_spa_diffuse_flow_fluids  2014-01-03  2014-01-20  0.042
=====
temp
-----
30
=====
press
-----
25
=====
time_elapsed  cell_concentration  NO3_uM  NH4_uM  H2_uM  H2S_uM  CH4_uM  pH  d15N_NO3_ppt  d15N_Biomass_ppt
-----
0             nd                12.8    9.9     nd      109     5.1     5.8    nd             nd
19            nd                11.6    7.4     0.13   nd       5.1     5.8    nd             nd
27            nd                13.6    10.8    nd      0.7     nd      6.4    nd             nd
44            nd                6.7     nd      nd      nd      nd      nd     2.1           nd
47            1420000          12.9    nd      nd      28.1    nd      nd     -2.3          nd
53            2720000          419     12.0    nd      nd      3.8     6.1    nd             nd
57            3805000          417     nd      nd      nd      nd      nd     nd            nd
68            6635000          1568    1.7     0.12   nd      nd      nd     5.15          nd
77            8825000          1289    1.4     nd      nd      nd      nd     5.10          nd
90            7270000          592     nd      nd      nd      nd      nd     6.4           nd
99            6980000          141     5.5     4.60   nd      nd      nd     6.5           6.2
106           5810000          105     23.2    0.10   nd      nd      nd     nd            nd
114           7380000          80      5.1     nd      nd      nd      5.5    8.1           -0.5
=====
press
-----
5
=====
time_elapsed  cell_concentration  NO3_uM  NH4_uM  H2_uM  H2S_uM  CH4_uM  pH  d15N_NO3_ppt  d15N_Biomass_ppt
-----
123           2550000          12.3    5.2     nd      nd      nd      5.8    nd            nd
133           4350000          5.3     12.3    nd      nd      nd      6      nd            nd
  
```

# Taxonomy - WoRMs

The World Register of Marine Species (WoRMs) provides an authoritative and comprehensive list of names of marine organisms, including information on synonymy. **/BCO-DMO/Southern\_Ocean\_Pinnipeds/fossil\_seal\_aa\_isotopes ---- Level 0**

[Directory](#) [Documentation](#) [Download & Other Operations](#)

[Level 0](#) [Next Level](#) [Flat Listing](#)

```
# Fossil Seal Amino Acid Isotopes
# PIs: Paul L. Koch (UC Santa Cruz) & Brenda Hall (UMaine)
# Co-PIs: Daniel P. Costa (UC Santa Cruz) & A. Rus Hoelzel (Durham University)
# Version: 26 March 2018
=====
Common_name      Scientific_name      WoRMS_LSID           AphiaID
-----
Crabeater_seal   Lobodon_carcinophaga  urn:lsid:marinespecies.org:taxname:231416  231416
Southern_elephant_seal  Mirounga_leonina     urn:lsid:marinespecies.org:taxname:231413  231413
Weddell_seal     Leptonychotes_weddellii  urn:lsid:marinespecies.org:taxname:195932  195932
```

We recommend checking species names against WoRMs and including identifiers in your data when possible.

# Taxonomy - NCBI for Microbial taxa and other IDs

The NCBI Taxonomy Database provides a source for names of marine microorganisms. Other NCBI databases can be linked to information on molecule models, protein IDs, etc.

## /BCO-DMO/En-Gen\_DMSP\_Cycling/DMSP\_d

[Directory](#)
[Documentation](#)
[Download & Other Operations](#)

[Level 0](#)
[Next Level](#)
[Flat Listing](#)

# Links to published DMSP-dependent protein structures in *P. ubique*  
 # at NCBI Molecular Modeling Database (MMDB)  
 # PI: Mary Ann Moran (UGA)  
 # Co-PIs: Ronald Kiene (DISL) & William Whitman (UGA)  
 # Notes: MMDB\_ID links open in new window/tab. PDB\_ID = Protein Data Bank ID.  
 # Version: 19 Nov 2012

protein_name	taxon	strain	PDB_ID	MMDB_ID
Demethylase_With_Cofactor_THF	Pelagibacter_ubique	HTCC1062	3TFJ	96080
Demethylase_With_Substrate_Dmsp	Pelagibacter_ubique	HTCC1062	3TFI	96079
Demethylase_APO	Pelagibacter_ubique	HTCC1062	3TFH	96078

Search for: **Candidatus Pelagibacter ubique**  
 Taxonomy ID: 198252 (for references in articles please use NCBI:taxid:198252)

**Entrez records**

Database name	Subtree links	Direct links	Links from type
Nucleotide	3,305	3,037	14
Protein	28,020	23,254	-
Structure	11	-	-
Genome	1	1	-
Popset	13	13	-
GEO Datasets	157	13	-
PubMed Central	55	55	-
Gene	1,392	1	-
SRA Experiments	26	4	-
Protein Clusters	1,321	1,321	-
Identical Protein Groups	23,865	23,240	-
Bio Project	22	4	-
Bio Sample	59	28	13
Bio Systems	156	-	-
Assembly	34	23	1
Taxonomy	14	1	-

**Notes:**  
 1) A provisional name for well characterised but as-yet uncultured organisms.

**Comments and References:**  
[Murray, R.G. & Schleifer, K.H. \(1994\)](#)  
 Murray, R.G., and Schleifer, K.H. "Taxonomic notes: a proposal for recording the properties of putative taxa of prokaryotes." *Int. J. Syst. Bacteriol.* (1994) 44:174-176.

[Murray, R.G. & Stackebrandt, E. \(1995\)](#)  
 Murray, R.G., and Stackebrandt, E. "Taxonomic note: implementation of the provisional status Candidatus for incompletely described prokaryotes." *Int. J. Syst. Bacteriol.* (1995) 45:186-187.

# International Geo Sample Numbers (IGSNs)

The IGSN is a persistent unique identifier for physical samples and specimens that eliminates the problems associated with the ambiguous naming of samples. In the U.S., you can obtain IGSNs using the System for Earth Sample Registration (SESAR) at IEDA Data Facility.

See <https://igsn.github.io/overview/> and <http://www.geosamples.org/>



Search IGSN

HOME ABOUT THE IGSN SERVICES SAMPLES NEWS HELP ABOUT US LOG IN TO MYSESAR



## get your igsn

Register your samples with SESAR to obtain IGSNs for unique sample identification.

## search the catalog

Search the SESAR catalog to find registered samples and their current location.

## sample curation

Learn about the DESC initiative to build a Digital Environment for Sample Curation.

## interoperability

Access IGSN metadata profiles and register samples via web services.

## new user?

Get a MySESAR account to register your samples.

# Event Logs

A chronological record of all scientific sampling events that happened during a cruise, wherein each sampling event is assigned a unique identifier.

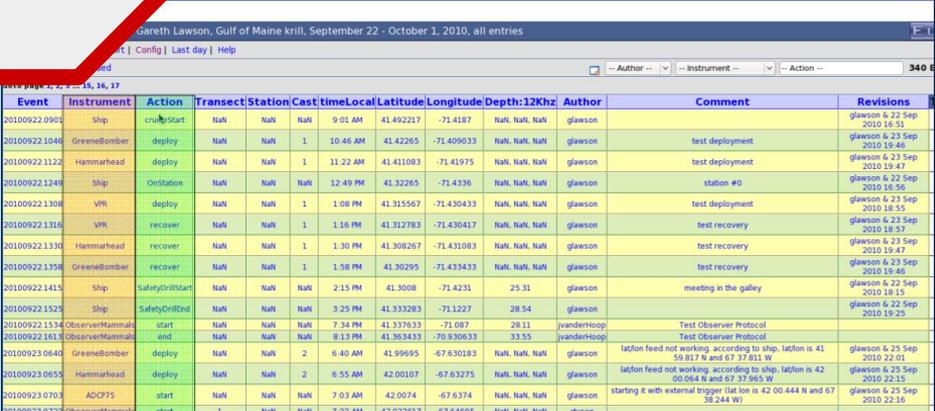
- Event #: unique to the cruise
- Instrument type/name/model
- Station #
- Cast # (if applicable)
- Date and time (UTC or local + time zone)
- Latitude and longitude
- Sampling depth
- Depth of the water
- Investigator
- Notes/comments

The event log allows investigators to integrate data from different sampling devices used during a cruise.

## Scientific Sampling Event Log

The R2R Event Logger is installed on many R/Vs and other SLS vessels.

The R2R Eventlogger is a program that creates a log of the scientific sampling events that occurred during a cruise. Final log can be downloaded as in csv format.

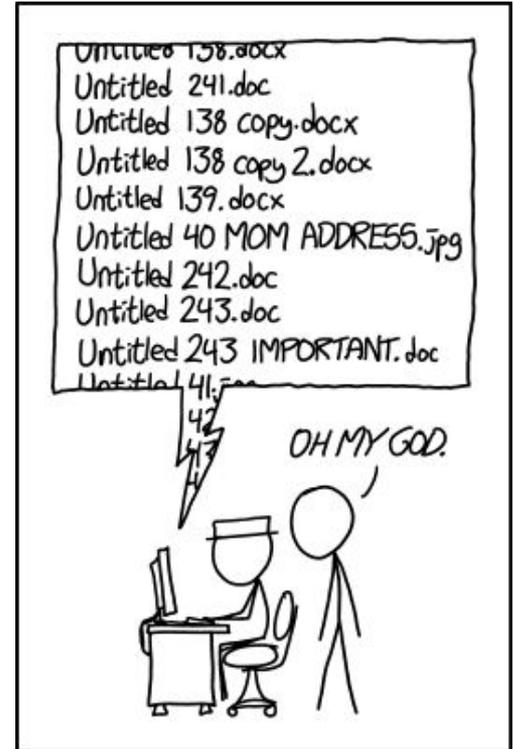


Event	Instrument	Action	Transect	Station	Cast	timeLocal	Latitude	Longitude	Depth	12Khz	Author	Comment	Revisions	
20100922.0903	Ship	crucStart	NaN	NaN	NaN	9:01 AM	41.492217	-71.4187	NaN	NaN	NaN	glawson	glawson & 22 Sep 2010 16:51	
20100922.1048	GreeneBomber	deploy	NaN	NaN	1	10:46 AM	41.42265	-71.409033	NaN	NaN	NaN	glawson	glawson & 23 Sep 2010 19:46	
20100922.1122	Hammarhead	deploy	NaN	NaN	1	11:22 AM	41.411083	-71.41975	NaN	NaN	NaN	glawson	glawson & 23 Sep 2010 19:47	
20100922.1249	Ship	OnStation	NaN	NaN	NaN	12:49 PM	41.32265	-71.4336	NaN	NaN	NaN	glawson	glawson & 22 Sep 2010 16:56	
20100922.1308	VR	deploy	NaN	NaN	1	1:08 PM	41.315567	-71.430433	NaN	NaN	NaN	glawson	glawson & 23 Sep 2010 18:55	
20100922.1316	VR	recover	NaN	NaN	1	1:16 PM	41.312783	-71.430417	NaN	NaN	NaN	glawson	glawson & 23 Sep 2010 18:57	
20100922.1330	Hammarhead	recover	NaN	NaN	1	1:30 PM	41.308267	-71.431083	NaN	NaN	NaN	glawson	glawson & 23 Sep 2010 19:47	
20100922.1358	GreeneBomber	recover	NaN	NaN	1	1:58 PM	41.30295	-71.433433	NaN	NaN	NaN	glawson	glawson & 23 Sep 2010 19:46	
20100922.1415	Ship	SafetyDriftStart	NaN	NaN	NaN	2:15 PM	41.3008	-71.4231	25.31		glawson	meeting in the gallery	glawson & 23 Sep 2010 18:15	
20100922.1525	Ship	SafetyDriftEnd	NaN	NaN	NaN	3:25 PM	41.333283	-71.1227	28.54		glawson		glawson & 22 Sep 2010 19:25	
20100922.1534	ObserverMammals	start	NaN	NaN	NaN	7:34 PM	41.337633	-71.087	28.11		jvanderstoep	Test Observer Protocol		
20100922.1613	ObserverMammals	end	NaN	NaN	NaN	8:13 PM	41.363433	-70.930633	33.55		jvanderstoep	Test Observer Protocol		
20100923.0640	GreeneBomber	deploy	NaN	NaN	2	6:40 AM	41.99695	-67.630183	NaN	NaN	NaN	glawson	latlon feed not working according to ship. latlon is 41 59.817 N and 67.37811 W	glawson & 23 Sep 2010 22:01
20100923.0655	Hammarhead	deploy	NaN	NaN	2	6:55 AM	42.00107	-67.63275	NaN	NaN	NaN	glawson	latlon feed not working according to ship. latlon is 42 00.064 N and 67.37865 W	glawson & 23 Sep 2010 22:15
20100923.0703	ACCF75	start	NaN	NaN	NaN	7:03 AM	42.0074	-67.6374	NaN	NaN	NaN	glawson	starting it with external trigger (lat lon is 42 00.444 N and 67 38.244 W)	glawson & 23 Sep 2010 22:16
20100923.0723	ObserverMammals	start	NaN	NaN	NaN	7:23 AM	42.026617	-67.64685	NaN	NaN	NaN	glawson		



# Data Formatting Best Practices

- Create descriptive column names without spaces or special characters. Use underscores instead of symbols.
  - Avoid using numbers at the beginning of a column name (some programs have trouble with this)
  - E.g., Temp 30 meters → **Temp\_30\_m**
  - E.g., Species Code → **species\_code**
- Use descriptive file names to provide useful information about the data.
  - Consistent naming allows for sorting and organizing files (esp. images and video).
    - E.g., **PIV\_E\_gracilis\_20180524.csv** (type = Particle Image Velocimetry, species = E. gracilis, date= May 24, 2018).
    - E.g., **KM1104\_12\_1.csv** (CTD files containing cruiseID\_station\_cast.csv)



PRO TIP: NEVER LOOK IN SOMEONE ELSE'S DOCUMENTS FOLDER.

# Data Formatting Best Practices

- Missing Data:
  - Again, be consistent!
  - Blank cells have no meaning...use "nd" or "NaN" (-999)
  - Bear in mind that "0" has meaning (0 = measured and not found as opposed to "not measured")
- Round data to appropriate number of decimal places for given property
- Document all codes and quality flag definitions/conventions in the metadata
- Don't rely on Excel formatting to convey meaning (e.g. colored cells)



\*Any/all data: include lat/lon, date, time, and depth whenever possible.

# Additional Resources

- BCO-DMO: <https://www.bco-dmo.org/resources> (Metadata Forms, DMP Template, BCO-DMO Quick Guide)
- R2R: <https://www.rvdata.us/>
- NCBI Taxonomy Database: <https://www.ncbi.nlm.nih.gov/taxonomy>
- World Register of Marine Species: <http://www.marinespecies.org/index.php>
- IGSN: <http://www.geosamples.org/>
- Data Management Short Course: <http://commons.esipfed.org/datamanagementshortcourse>

# Questions?

[info@bco-dmo.org](mailto:info@bco-dmo.org)  
[www.bcodmo.org](http://www.bcodmo.org)



(Lawson, 2002)