A custom ARB database of SSU rRNA gene sequences from corals, as well representative cultivated and environmental sequences from public sources (Coral Microbial Relationships project)

Website: https://www.bco-dmo.org/dataset/724355
Data Type: experimental
Version: 2
Version Date: 2021-06-10

Project
» Fundamental Coral-Microbial Associations (Coral Microbial Relationships)

Contributors	Affiliation	Role
Apprill, Amy	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator, Contact
Huggett, Megan	University of Newcastle (UON)	Scientist
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Table of Contents

Dataset Description

A custom ARB database of SSU rRNA gene sequences from corals, as well representative cultivated and environmental sequences from public sources. These data are described in Huggett and Apprill (in press).

Metadata for the database is available by clicking the "Get Data" button on this project page. The database is available for download as the following ARB file: coralmicrobiome_database.arb (37 MB)

ARB software (version arb-6.06) is available from http://www.arb-home.de/downloads.html

Methods & Sampling

The coral-microbial database was built over time using the ARB software (Ludwig et al., 2004). Initially, the All-Species Living Tree Project (LTP) s95 database containing SSU rRNA gene sequences for type strains of bacteria, archaea and eukarya was used as the database backbone (Yarza et al., 2008). In 2009, we used the search criteria 'coral + bacteria' and 'coral + archaea' within BLAST (Altschul et al., 1990) to obtain coral-associated SSU rRNA gene sequences from studies that applied cultivation-dependent or independent approaches from the GenBank database (Benson et al., 2008). These sequences were aligned using SINA (Pruesse et al., 2007) and imported into ARB. This database was then revised by first searching for all sequences coded as 'coral*' in any field. From these, sequences that matched 'eukaryot*' in the field tax_slv were removed, and all others were marked. An initial search of these marked sequences for those that had *coral* in the field isolation_source was done and these were manually checked. For those that were isolated from a coral (soft or hard, tropical or deep sea, etc.) were assigned 'coral' in the remark field. For those that were not isolated from a coral all were assigned 'checked' in the remark field. At this stage, there were 1333 sequences remaining with the term *coral* in any field. These were manually checked for their description and isolation details and marked either 'coral' or 'checked' in the remark field. From these, a database was created

that contained just the sequences marked 'coral' from our in-house (2009) database. This small curated database was then merged with the SSU_Ref111_SILVA_NR (Pruesse et al., 2007) database (released 19 July 2012) which created new names for the living tree sequences. All sequences that were brought in from the small living tree database were marked with an identifier and the SSU_Ref database was searched as above to locate sequences in the SSU_Ref database that were bacteria or archaea sequences derived from corals. Information was added to these sequences in the location, author, journal and host species fields and all were given the term 'coral' in the remark field.

Next, the ISI web of science electronic database was searched for any publications that contained the search term 'coral*' and 'bacteria*' or 'coral*' and 'archaea* in 'topic' from 2010 to present. From these, the publications were manually checked and any manuscript that mentioned sequence data or appeared likely to contain Sanger sequence data was obtained. If sequence data was associated with a manuscript, the corresponding sequences were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) in fasta format, aligned using the online SINA aligner (Pruesse et al., 2007) and imported into arb. If replicates were located they were removed from arb. Newly imported sequences were manually curated to include as much metadata as possible. These included the location (in country field), host (in host field), author names (in author field) and any other accessible data (e.g. Journal).

Data Processing Description

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 in this dataset are displayed as "nd" for "no data." nd is the default missing data identifier in the BCO-DMO system.
* converted non-delimiter commas to semicolons to support export as csv
* removed duplicate column of lengths
* data version 2: 2018-08-06 is an update of data version 1: 2018-05-11 with the following change. In the file coralmicrobiome_database.arb, phylogenetic trees were updated to include those available in Huggett and Apprill (in press) which describes this dataset.

Dataset version 2 (2021-06-10) replaces version 1(2018-08-06):
* Converted file to UTF-8

[table of contents | back to top]

Data Files

File
coral_microb_suppl.csv(Comma Separated Values (.csv), 5.11 MB) MD5:2bb47e331812054a27d2a09f53dca824
Primary data file for dataset ID 724355

[table of contents | back to top]

Related Publications

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403–410. doi:10.1016/s0022-2836(05)80360-2
https://doi.org/10.1016/S0022-2836(05)80360-2
Methods

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2007). GenBank. Nucleic Acids Research, 36(Database), D25–D30. doi:10.1093/nar/gkm929
Methods

Huggett, M. J., & Apprill, A. (2018). Coral microbiome database: Integration of sequences reveals high diversity and relatedness of coral-associated microbes. Environmental Microbiology Reports, 11(3), 372–385.

doi:
Methods

Ludwig, W. (2004). ARB: a software environment for sequence data. Nucleic Acids Research, 32(4), 1363–1371. doi:10.1093/nar/gkh293
Methods

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glockner, F. O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research, 35(21), 7188–7196. doi:10.1093/nar/gkm864
Methods

Yarza, P., Richter, M., Peplies, J., Euzeby, J., Amann, R., Schleifer, K.-H., … Rosselló-Móra, R. (2008). The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. Systematic and Applied Microbiology, 31(4), 241–250. doi:10.1016/j.syapm.2008.07.001
Methods

[table of contents | back to top]

Parameters

Parameter	Description	Units
Sequence_Identifier	Sequence identifier	unitless
gi_number	gi number; A series of digits that are assigned consecutively to each sequence record processed by NCBI.	unitless
Accession_source	Source of genetic accession; Database containing accession	unitless
Accession_number	Genetic accession number for the database supplied in Accession_source	unitless
Accession_link	Link to the genetic accession in the database specified in Accession_source	unitless
Length_bp	Length of base pairs (bp) in the genetic accession	count
Taxonomy	Taxonomic heirarchy of the sampled organism	unitless
Organism_details	Description of sequence and organism source	unitless
Host	Scientific name of organism sampled or description of organism	unitless
Location	Latitude and longitude of sampled organism	various
Accession_Number	Accession number (either relative to NCBI or ARB)	unitless
Isolate_Clone	Indication of whether clone" or "isolate"	unitless

[table of contents | back to top]

Project Information

Fundamental Coral-Microbial Associations (Coral Microbial Relationships)

Coverage: Florida Keys, Federated States of Micronesia, Red Sea, & Bermuda

Description from NSF award abstract:
Reef-building corals are in decline worldwide due in part to climate change and other human activities, and it is becoming increasingly important to understand what aspects of coral biology are degraded by environmental stress which then leads to coral mortality. It is now widely known that corals harbor communities of bacteria and archaea that are believed to play important roles in maintaining the health of their hosts, but we lack any appreciable understanding about the identity of the microbial associates regularly residing within healthy, reef-building corals. This project asks the central question: do reef-building corals harbor fundamental or persistent microbial associates that are symbiotic within their tissues? In order to address this hypothesis, the investigator will assess the identity of the bacterial and archaeal microbes using a variety of molecular and

microscopy approaches that includes the identification and localization of a widespread group of coral bacterial associates belonging to the genus Endozoicomonas. The results of this study will then be used to develop additional questions about the role of these microbial associates in nutrient cycling and how they contribute to the health and survival of corals.

Funding