



**HAL**  
open science

## Microbial diversity in four Mediterranean irciniid sponges

Katerina Skaraki, Christina Pavludi, Thanos Dailianis, Jacques Lagnel, Adriani Pantazidou, Antonios Magoulas, Georgios Kotoulas

► **To cite this version:**

Katerina Skaraki, Christina Pavludi, Thanos Dailianis, Jacques Lagnel, Adriani Pantazidou, et al.. Microbial diversity in four Mediterranean irciniid sponges. *Biodiversity Data Journal*, 2024, 12, pp.e114809. 10.3897/bdj.12.e114809. hal-04575995

**HAL Id: hal-04575995**

**<https://hal.inrae.fr/hal-04575995>**

Submitted on 15 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# Microbial diversity in four Mediterranean irciniid sponges

Katerina Skaraki<sup>‡,§</sup>, Christina Pavloudi<sup>‡</sup>, Thanos Dailianis<sup>‡</sup>, Jacques Lagnel<sup>¶</sup>, Adriani Pantazidou<sup>§</sup>, Antonios Magoulas<sup>‡</sup>, Georgios Kotoulas<sup>‡</sup>

<sup>‡</sup> Institute of Marine Biology, Biotechnology & Aquaculture, Hellenic Centre for Marine Research, Heraklion, Crete, Greece

<sup>§</sup> Department of Ecology & Systematics, Faculty of Biology, National & Kapodistrian University of Athens, Athens, Greece  
| European Marine Biological Resource Centre (EMBRC-ERIC), Paris, France

<sup>¶</sup> INRAE, UR1052, Génétique et Amélioration des Fruits et Légumes (GAFL), Centre de Recherche PACA, Montfavet, France

Corresponding author: Katerina Skaraki ([kskaraki@gmail.com](mailto:kskaraki@gmail.com)),

Christina Pavloudi ([christina.pavloudi@embrc.eu](mailto:christina.pavloudi@embrc.eu))

Academic editor: Lyubomir Penev

Received: 27 Oct 2023 | Accepted: 14 Jan 2024 | Published: 19 Jan 2024

Citation: Skaraki K, Pavloudi C, Dailianis T, Lagnel J, Pantazidou A, Magoulas A, Kotoulas G (2024) Microbial diversity in four Mediterranean irciniid sponges. Biodiversity Data Journal 12: e114809.

<https://doi.org/10.3897/BDJ.12.e114809>

## Abstract

This paper describes a dataset of microbial communities from four different sponge species: *Ircinia oros* (Schmidt, 1864), *Ircinia variabilis* (Schmidt, 1862), *Sarcotragus spinosulus* Schmidt, 1862 and *Sarcotragus fasciculatus* (Pallas, 1766). The examined sponges all belong to Demospongiae (Class); Keratosa (Subclass); Dictyoceratida (Order); Irciniidae (Family). Samples were collected by scuba diving at depths between 6-14 m from two sampling sites of rocky formations at the northern coast of Crete (Cretan Sea, eastern Mediterranean) and were subjected to metabarcoding for the V5-V6 region of the 16S rRNA gene.

## Keywords

sponge metagenome, marine metagenome, amplicon sequencing, 454 GS FLX Titanium, pyrosequencing, eastern Mediterranean

## Introduction

Porifera (sponges) is one of the oldest metazoan Phyla (Müller 2003), comprising species very important for the shaping and functioning of benthic ecosystems. Sponges maintain symbiotic microorganisms that, in many cases, can comprise up to 40% of their biomass (Vacelet and Donadey 1977, Hentschel et al. 2003). Microbes associated with sponges belong to phylogenetically diverse phyla of Prokaryotes and Eukaryotes (Taylor et al. 2007). The main objective of this paper was to study the microbial diversity associated with a group of Porifera species that exhibit certain characteristics of interest, such as inhabiting an oligotrophic environment, i.e. the South Aegean Sea (east Mediterranean) and being difficult to identify taxonomically while, at the same time, considered as model taxa in host-associated microbiological studies (Hardoim and Costa 2014, Pita et al. 2013), like species of the family Irciniidae. In addition, by collecting samples from various species and tissues, we were able to identify if there is host and/or tissue specificity in the sponge microbiome.

## Value of the dataset

This dataset includes microbial taxa inhabiting the demosponges *Ircinia oros*, *I. variabilis*, *Sarcotragus spinosulus* and *S. fasciculatus* and contributes to the ongoing efforts of the Ocean Biogeographic Information System ([OBIS](#)) which aims at filling the gaps in our current knowledge of the world's oceans. The dataset has been also published in [GBIF](#) (Skaraki et al. 2023).

## Methods

### Sampling

Sponge samples were collected in 2009-2010 by scuba diving from two sampling sites of rocky formations at the northern coast of Crete (Cretan Sea, eastern Mediterranean). Specimens of *Sarcotragus spinosulus*, *Ircinia oros* and *I. variabilis* were collected from Alkes along with a seawater sample (2 litres), whereas samples of *Sarcotragus fasciculatus* were collected from Elounda. Samples were transferred to the laboratory in a cooler with ice packs. Upon arrival, sponge samples were preserved in 97% ethanol and stored at -20°C prior to DNA extraction; the seawater sample was filtered through a Whatman 0.45 µm mixed cellulose ester membranes filter, which was stored at -80°C prior to DNA extraction. Sponge samples were identified both morphologically, as well as by molecular markers (18S rRNA and COI).

### Geographic range

The dataset's geographical range includes two locations in Crete (Greece) (Table 1).

Table 1.

Locality, geographical coordinates and depth (m) of the sampling stations.

Locality	Coordinates	Depth (m)
Alikes	35.41568N, 24.9881E	6-10
Elounda	35.2521N, 25.7592E	12-14

## Sample processing

For each sponge specimen, mesohyl and ectosomal tissue samples (10-15 mg each) were cut and proceeded separately. DNA from those tissues was extracted using DNeasy Blood & Tissue™ kit (Qiagen) following the manufacturer's guidelines. DNA from the filter was extracted according to a freeze–thaw–boiling protocol of Riemann and Middelboe (2002). DNA extracts were stored at -20°C prior to amplification.

## Amplification & Sequencing of 16S rRNA

The V5-V6 hypervariable region (~ 280 bp) of the 16S ribosomal RNA (rRNA) gene was amplified using the degenerate universal primers 802F: 5'-GGATTAGATACCCBNGTA-3' (originally designed as reverse primer by Claesson et al. (2009)) and 1027R: 5'-CGACRRCCATGCANCACT-3' (Claesson et al. 2009) that target Bacteria and Archaea. In order to pool samples prior to sequencing, the primers contained a 10 bp length of multiplex identifiers (MIDs) between the 454 adaptor A (forward primer) and B (reverse primer), according to Roche. PCR reactions were performed in a total volume of 25 µl containing 0.3 µM of each primer, 0.3 mM dNTPs, 0.5 U KAPA high-fidelity HotStart DNA polymerase (Kapa Biosystems), 2.0 mM MgCl<sub>2</sub> and approximately 20 ng of template DNA. PCR reactions were carried out in a thermal cycler (Bio-Rad, MyCycler) under the following conditions: initial denaturation at 95°C for 5 min; followed by 25 cycles of denaturation at 98°C for 20 s, annealing at 55°C for 15 s and extension at 72°C for 60 s, followed by a final extension step at 72°C for 5 min. PCR amplicons were checked by electrophoresis in 1% agarose gel. Amplicons were purified by Agencourt AMPure XP purification system (Beckman Coulter Genomics) and quantified with Quantit PicoGreen dsDNA (Invitrogen) in a Quantiflour fluorometer (Promega). Equal concentrations of amplicons were pooled prior to sequencing.

## Technologies used

Amplicon libraries' preparation and sequencing were performed at the [DNA Sequencing platform](#) of the Institute of Marine Biology, Biotechnology and Aquaculture ([IMBBC](#), [HCMR](#)) using 454 GS FLX Titanium technology (Roche, 454 Life Sciences), according to the manufacturer's recommendations.

## Data processing

The 454 GS FLX Titanium metabarcoding data were processed in the following steps: a) raw data were denoised using the software package AmpliconNoise v.1.29 (Quince et al. 2011) by removing pyrosequencing noise and correcting for putative PCR errors; b) the Perseus algorithm, included in AmpliconNoise, was used to check for chimeras and the denoised reads were clustered to Operational Taxonomic Units (OTUs), with a complete linkage algorithm on a 97% sequence identity level (Quince et al. 2011); c) taxonomy assignment was performed using GTDB r207 (Parks et al. 2022) using the IDTAXA Classifier (Murali et al. 2018) and a confidence threshold  $\geq 60\%$  (cautious). AmpliconNoise and Perseus were run with the default parameters. IDTAXA Classifier was run using the [IDTAXA web tool](#).

The phyloseq package (version 1.42.0) (McMurdie and Holmes 2013) was used to generate bar charts and calculate beta diversity, using Bray-Curtis distances visualised with non-metric multidimensional scaling (nMDS).

## Technologies used

The analysis was carried out through [Zorba](#), the High Performance Computing (HPC) system of IMBBC (Zafeiropoulos et al. 2021).

## Biodiversity scope

The majority of the reads were unclassified (~ 26%). The most abundant phyla were Proteobacteria (~ 19%), Chloroflexota (~ 9%), Cyanobacteria (~ 8%) and Poribacteria (~ 6%). The processed sequences, along with their corresponding taxonomic information and metadata, can be found in the occurrence dataset available from [GBIF](#). In addition, processed sequences using the MGnify pipeline (4.1) and their respective taxonomic information is available with the MGnify study id MGYS00004687 (available at <https://www.ebi.ac.uk/metagenomics/studies/MGYS00004687>) (Richardson et al. 2022).

## Target

The target of the dataset was to amplify prokaryotic taxa associated with the studied sponges.

## Taxonomic range

Archaea, Bacteria

## Data Resources

Details for the samples can be found in Suppl. material 1 and Suppl. material 2. All the raw sequence files of this study were submitted to the European Nucleotide Archive (ENA) (Burgin et al. 2022) with the study accession number PRJEB23640 (available at <http://www.ebi.ac.uk/ena/data/view/PRJEB23640>).

### Resource 1

#### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR220/001/ERR2204961/ERR2204961.fastq.gz>

#### Resource identifier

ERR2204961

#### Data format

FASTQ

### Resource 2

#### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/008/ERR2486828/ERR2486828.fastq.gz>

#### Resource identifier

ERR2486828

#### Data format

FASTQ

### Resource 3

#### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/009/ERR2486829/ERR2486829.fastq.gz>

#### Resource identifier

ERR2486829

**Data format**

FASTQ

**Resource 4****Download URL**<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/001/ERR2486831/ERR2486831.fastq.gz>**Resource identifier**

ERR2486831

**Data format**

FASTQ

**Resource 5****Download URL**<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/003/ERR2486833/ERR2486833.fastq.gz>**Resource identifier**

ERR2486833

**Data format**

FASTQ

**Resource 6****Download URL**<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/005/ERR2486835/ERR2486835.fastq.gz>**Resource identifier**

ERR2486835

**Data format**

FASTQ

## Resource 7

### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/006/ERR2486826/ERR2486826.fastq.gz>

### Resource identifier

ERR2486826

### Data format

FASTQ

## Resource 8

### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/007/ERR2486827/ERR2486827.fastq.gz>

### Resource identifier

ERR2486827

### Data format

FASTQ

## Resource 9

### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/004/ERR2486834/ERR2486834.fastq.gz>

### Resource identifier

ERR2486834

### Data format

FASTQ

## Resource 10

### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/007/ERR2486837/ERR2486837.fastq.gz>



**Resource identifier**

ERR2486837

**Data format**

FASTQ

**Resource 11****Download URL**<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/009/ERR2486839/ERR2486839.fastq.gz>**Resource identifier**

ERR2486839

**Data format**

FASTQ

**Resource 12****Download URL**<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/002/ERR2486842/ERR2486842.fastq.gz>**Resource identifier**

ERR2486842

**Data format**

FASTQ

**Resource 13****Download URL**<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/003/ERR2486843/ERR2486843.fastq.gz>**Resource identifier**

ERR2486843

**Data format**

FASTQ

**Resource 14**

**Download URL**

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/000/ERR2486830/ERR2486830.fastq.gz>

**Resource identifier**

ERR2486830

**Data format**

FASTQ

**Resource 15**

**Download URL**

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/002/ERR2486832/ERR2486832.fastq.gz>

**Resource identifier**

ERR2486832

**Data format**

FASTQ

**Resource 16**

**Download URL**

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/006/ERR2486836/ERR2486836.fastq.gz>

**Resource identifier**

ERR2486836

**Data format**

FASTQ

**Resource 17****Download URL**

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/008/ERR2486838/ERR2486838.fastq.gz>

**Resource identifier**

ERR2486838

**Data format**

FASTQ

**Resource 18****Download URL**

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/000/ERR2486840/ERR2486840.fastq.gz>

**Resource identifier**

ERR2486840

**Data format**

FASTQ

**Resource 19****Download URL**

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/001/ERR2486841/ERR2486841.fastq.gz>

**Resource identifier**

ERR2486841

**Data format**

FASTQ

## Resource 20

### Download URL

<ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR248/004/ERR2486844/ERR2486844.fastq.gz>

### Resource identifier

ERR2486844

### Data format

FASTQ

## Results

The most abundant phyla across all the samples are Proteobacteria (~ 19%), Chloroflexota (~ 9%), Cyanobacteria (~ 8%) and Poribacteria (~ 6%) (Fig. 1). It seems that the microbial community profiles are clustering, based on the host species (PERMANOVA:  $R^2 = 0.56$ ,  $p = 0.001$ ), but not based on the sampled tissue (PERMANOVA:  $R^2 = 0.073$ ,  $p = 0.23$ ) (Fig. 2).

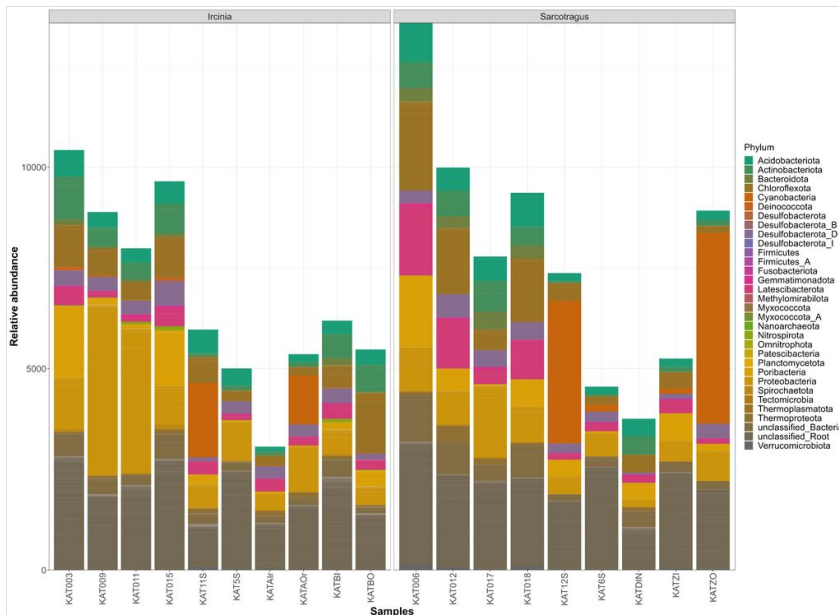


Figure 1. [doi](#)

Bar chart showing the abundances of the main microbial taxa, at the phylum level, at each of the sponge samples.

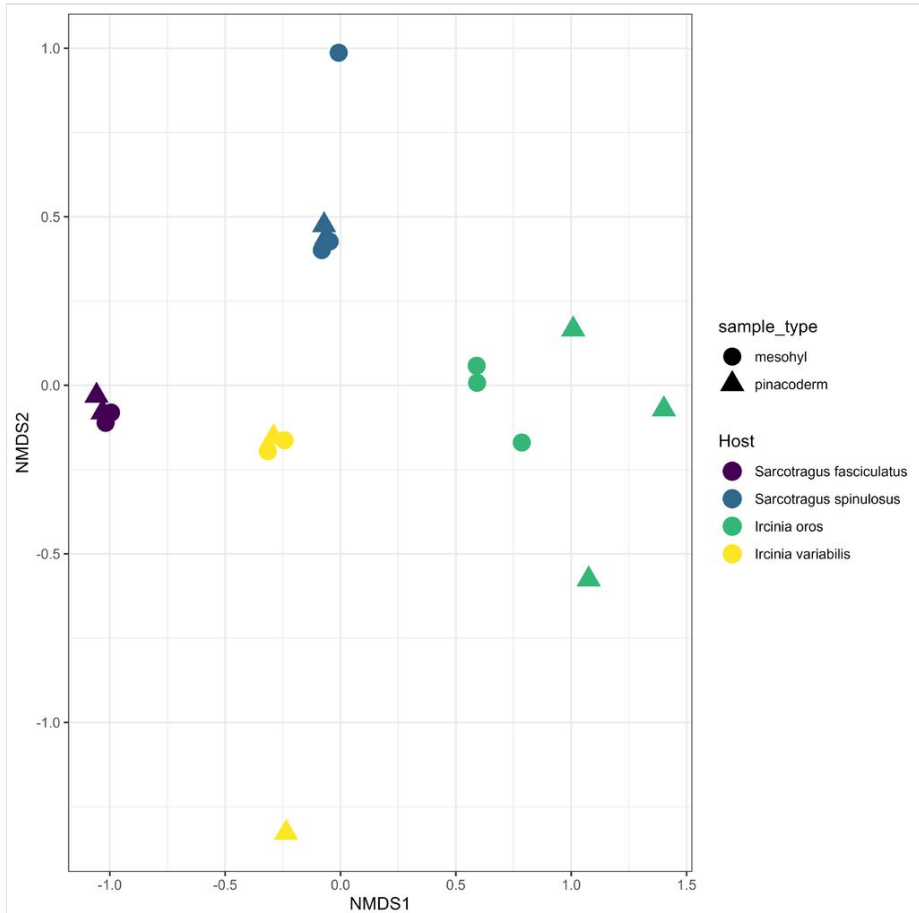


Figure 2. [doi](#)

nMDS plot of the Bray-Curtis similarity matrix of the sponge samples, based on the microbial OTUs abundances.

## Discussion

Since the advent of pyrosequencing, research papers describing and discussing findings, based on eDNA and DNA metabarcoding data, have been a major part of the scientific literature. However, submission of such data into repositories was not possible until very recently, with the development of the Darwin Core (DwC) [DNA derived data extension](#) (Finstad et al. 2023). This extension is based on the [MlxS checklists](#) (Yilmaz et al. 2011) which have been developed by the Genomic Standards Consortium (GSC) and are the golden standard for describing sample metadata information in a standardised way when submitting raw sequence data to the relevant repositories (e.g. ENA, NCBI etc.). To ensure the interoperability between the different repositories and platforms, as well as between the different metadata standard formats, Biodiversity Information Standards (TDWG) and GSC

have agreed on the creation of a continuous model to synchronise their standards; undoubtedly, such efforts on enhancement of data-sharing and utilisation across domains will definitely have a major impact on future scientific discoveries (Meyer et al. 2023).

It should be noted, however, that certain metabarcoding limitations should be taken into account when preparing DwC DNA occurrence data for submission to GBIF and/or OBIS. Notably, the uncertainty of organism occurrences, along with the less precise taxonomic identifications inferred from sequencing data, should be clearly mentioned in the dataset descriptions. Moreover, MlxS checklists in host-associated datasets, while capturing the host species information in the sample metadata, fail to do so in the run, i.e. sequence, metadata; an example would be this particular dataset, where the taxonomy associated with the run accession numbers in ENA is "[sponge metagenome](#)", with no further details shown, although the host species is mentioned in the sample metadata, i.e. it can be retrieved when searching ENA with the sample accession number. In addition, standards should be continuously updated to incorporate changes in sequencing technologies and protocols, in order to aid researchers document such data in a more FAIR way. If discrepancies between repositories and platforms are minimised and even eliminated in the future, the already high intrinsic value of DNA derived datasets will increase even further.

## Usage Rights

[CC BY 4.0](#)

## Acknowledgements

This research was supported in part through computational resources provided by IMBBC (Institute of Marine Biology, Biotechnology and Aquaculture) of the HCMR (Hellenic Centre of Marine Research). Funding for establishing the IMBBC HPC has been received by the MARBIGEN (EU Regpot) project, LifeWatchGreece RI and the CMBR (Center for the study and sustainable exploitation of Marine Biological Resources) RI. In addition, the authors would like to acknowledge the [BiCIKL project](#) (Grant No 101007492).

## Conflicts of interest

The authors have declared that no competing interests exist.

## References

- Burgin J, Ahamed A, Cummins C, Devraj R, Gueye K, Gupta D, Gupta V, Haseeb M, Ihsan M, Ivanov E, Jayathilaka S, Balavenkataraman Kadhivelu V, Kumar M, Lathi A, Leinonen R, Mansurova M, McKinnon J, O’Cathail C, Paupério J, Pesant S, Rahman N, Rinck G, Selvakumar S, Suman S, Vijayaraja S, Waheed Z, Woollard P, Yuan D, Zyoud

- A, Burdett T, Cochrane G (2022) The European Nucleotide Archive in 2022. *Nucleic Acids Research* 51 <https://doi.org/10.1093/nar/gkac1051>
- Claesson M, O'Sullivan O, Wang Q, Nikkilä J, Marchesi J, Smidt H, de Vos W, Ross RP, O'Toole P (2009) Comparative Analysis of Pyrosequencing and a Phylogenetic Microarray for Exploring Microbial Community Structures in the Human Distal Intestine. *PLOS One* 4 (8). <https://doi.org/10.1371/journal.pone.0006669>
  - Finstad AG, Andersson A, Bissett A, Fossøy F, Grosjean M, Hope M, Køljalg U, Lundin D, Nilsson H, Prager M, Jeppesen TS, Svenningsen C, Schigel D, Abarenkov K, Provoost P, Suominen S, Frøslev TG (2023) Publishing DNA-derived data through biodiversity data platforms. GBIF Secretariat <https://doi.org/10.35035/doc-vf1a-nr22>
  - Hardoim C, Costa R (2014) Microbial communities and bioactive compounds in marine sponges of the family Irciniidae—A review. *Marine Drugs* 12 (10): 5089-5122. <https://doi.org/10.3390/md12105089>
  - Hentschel U, Fieseler L, Wehrl M, Gernert C, Steinert M, Hacker J, Horn M (2003) Microbial diversity of marine sponges. In: Müller WE (Ed.) *Sponges (Porifera). Progress in molecular and subcellular biology*. Vol. 37. Springer, Berlin, Heidelberg, 59-88 pp. [https://doi.org/10.1007/978-3-642-55519-0\\_3](https://doi.org/10.1007/978-3-642-55519-0_3)
  - McMurdie P, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS One* 8 (4). <https://doi.org/10.1371/journal.pone.0061217>
  - Meyer R, Appeltans W, Duncan W, Dimitrova M, Gan Y, Stjernegaard Jeppesen T, Mungall C, Paul D, Provoost P, Robertson T, Schriml L, Suominen S, Walls R, Sweetlove M, Ung V, Van de Putte A, Wallis E, Wiczorek J, Buttigieg P (2023) Aligning Standards Communities for Omics Biodiversity Data: Sustainable Darwin Core-MIxS Interoperability. *Biodiversity Data Journal* 11 <https://doi.org/10.3897/bdj.11.e112420>
  - Müller WG (2003) The origin of metazoan complexity: Porifera as integrated animals. *Integrative and Comparative Biology* 43 (1): 3-10. <https://doi.org/10.1093/icb/43.1.3>
  - Murali A, Bhargava A, Wright ES (2018) IDTAXA: a novel approach for accurate taxonomic classification of microbiome sequences. *Microbiome* 6 (1): 140. <https://doi.org/10.1186/s40168-018-0521-5>
  - Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P, Hugenholtz P (2022) GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic acids research* 50 (D1): D785-D794. <https://doi.org/10.1093/nar/gkab776>
  - Pita L, Turon X, López-Legentil S, Erwin P (2013) Host rules: spatial stability of bacterial communities associated with marine sponges (*Ircinia* spp.) in the Western Mediterranean Sea. *FEMS Microbiology Ecology* 86 (2): 268-276. <https://doi.org/10.1111/1574-6941.12159>
  - Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ (2011) Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* 12 (38). <https://doi.org/10.1186/1471-2105-12-38>
  - Richardson L, Allen B, Baldi G, Beracochea M, Bileschi M, Burdett T, Burgin J, Caballero-Pérez J, Cochrane G, Colwell L, Curtis T, Escobar-Zepeda A, Gurbich T, Kale V, Korobeynikov A, Raj S, Rogers A, Sakharova E, Sanchez S, Wilkinson D, Finn R (2022) MGnify: the microbiome sequence data analysis resource in 2023. *Nucleic Acids Research* 51 <https://doi.org/10.1093/nar/gkac1080>

- Riemann L, Middelboe M (2002) Stability of bacterial and viral community compositions in Danish coastal waters as depicted by DNA fingerprinting techniques. *Aquatic Microbial Ecology* 27: 219-232. <https://doi.org/10.3354/ame027219>
- Skaraki K, Dailianis T, Pavlouidi C (2023) Microbial diversity in four Mediterranean irciniid sponges. Occurrence dataset. Hellenic Centre for Marine Research. Release date: 2023-9-20. URL: <https://doi.org/10.15468/3uxuba>
- Taylor M, Radax R, Steger D, Wagner M (2007) Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiology and Molecular Biology Reviews* 71 (2): 295-347. <https://doi.org/10.1128/mnbr.00040-06>
- Vacelet J, Donadey C (1977) Electron microscope study of the association between some sponges and bacteria. *Journal of Experimental Marine Biology and Ecology* 30 (3): 301-314. [https://doi.org/10.1016/0022-0981\(77\)90038-7](https://doi.org/10.1016/0022-0981(77)90038-7)
- Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PSG, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, Kuczynski J, Kyrpides N, Larsen R, Lauber CL, Legg T, Ley RE, Lozupone CA, Ludwig W, Lyons D, Maguire E, Methé BA, Meyer F, Muegge B, Nakielny S, Nelson KE, Nemergut D, Neufeld JD, Newbold LK, Oliver AE, Pace NR, Palanisamy G, Peplies J, Petrosino J, Proctor L, Pruesse E, Quast C, Raes J, Ratnasingham S, Ravel J, Relman DA, Assunta-Sansone S, Schloss PD, Schriml L, Sinha R, Smith MI, Sodergren E, Spo A, Stombaugh J, Tiedje JM, Ward DV, Weinstock GM, Wendel D, White O, Whiteley A, Wilke A, Wortman JR, Yatsunenko T, Glöckner FO (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nature biotechnology* 29 (5): 415-20. <https://doi.org/10.1038/nbt.1823>
- Zafeiropoulos H, Gioti A, Ninidakis S, Potirakis A, Paragkamian S, Angelova N, Antoniou A, Danis T, Kaitetzidou E, Kasapidis P, Kristoffersen JB, Papadogiannis V, Pavlouidi C, Ha QV, Lagnel J, Pattakos N, Perantinos G, Sidirokastritis D, Vavilis P, Kotoulas G, Manousaki T, Sarropoulou E, Tsigenopoulos CS, Arvanitidis C, Magoulas A, Pafilis E (2021) 0s and 1s in marine molecular research: a regional HPC perspective. *GigaScience* 10 (8). <https://doi.org/10.1093/gigascience/giab053>

## Supplementary materials

### Suppl. material 1: BioSamples MixS checklists [doi](#)

**Authors:** HELLENIC CENTRE FOR MARINE RESEARCH

**Data type:** sample metadata

**Filename:** biosamples\_mixs\_checklist.csv - [Download file](#) (24.76 kb)



**Suppl. material 2: Sample metadata** [doi](#)

**Authors:** Katerina Skaraki, Christina Pavloudi

**Data type:** metadata

[Download file](#) (4.20 kb)