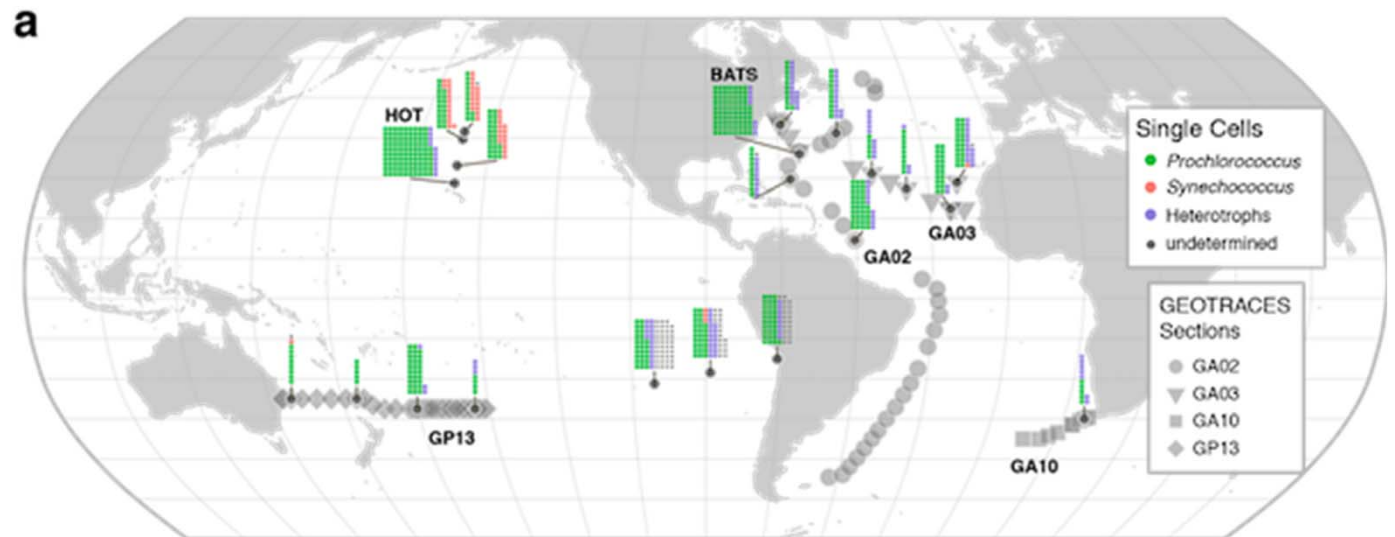


EXPLORATION: LINKING TEI & GENOMICS DATA IN IDP2021

We wish to strengthen and intensify the collaboration between the marine geochemical community and the biological / omics' community.

GEOTRACES
Sections
with omics data

GA02
GA03
GA10
GP13



Berube et al. 2018
<http://doi.org/10.1038/sdata.2018.154>



www.geotraces.org

#GEOTRACESDataProduct

#OceanData #OceanScience



EXPLORATION: LINKING TEI & GENOMICS DATA IN IDP2021

A major challenge, as the omics data is dynamic (changes depending on ever-growing genome reference databases) while the TEI data is static.

Acknowledgements to:

Reiner Schiltzer (AWI)

Paul Berube (MIT)

Jesse McNichol (U. Southern California)

And many others who made this possible, specially members of Penny Chisholm lab

4 Types of Sample Descriptors for omics data linkages in the IDP2021:

to NCBI website

480 Metagenomes
14 Single-Cell genomes
273 16S-18S-rRNA gene analyses

to EMBL-EBI website

480 16S/18S amplicons (microbial community taxonomic & functional structure and abundance)



www.geotraces.org

#GEOTRACESDataProduct

#OceanData #OceanScience



EXPLORATION: LINKING TEI & GENOMICS DATA IN IDP2021

OMICs linkages

4 Sample Descriptors in the Metadata variables:

- Accession numbers at NCBI or EMBL-EBI
- Or links to functional and taxonomic analyses webpages

OMICs linkages in Metadata VARIABLES #:

11. NCBI_Metagenome_BioSample accession #:

13. NCBI_16S-18S-rRNA-gene_BioSample accession #:

@ <https://www.ncbi.nlm.nih.gov/biosample/>

12. NCBI_Single_cell_genome_BioProject accession #:

@ <https://www.ncbi.nlm.nih.gov/bioproject/>

14. EMBL_EBI_Metagenome_MGNIFY_Analysis_Accession #:

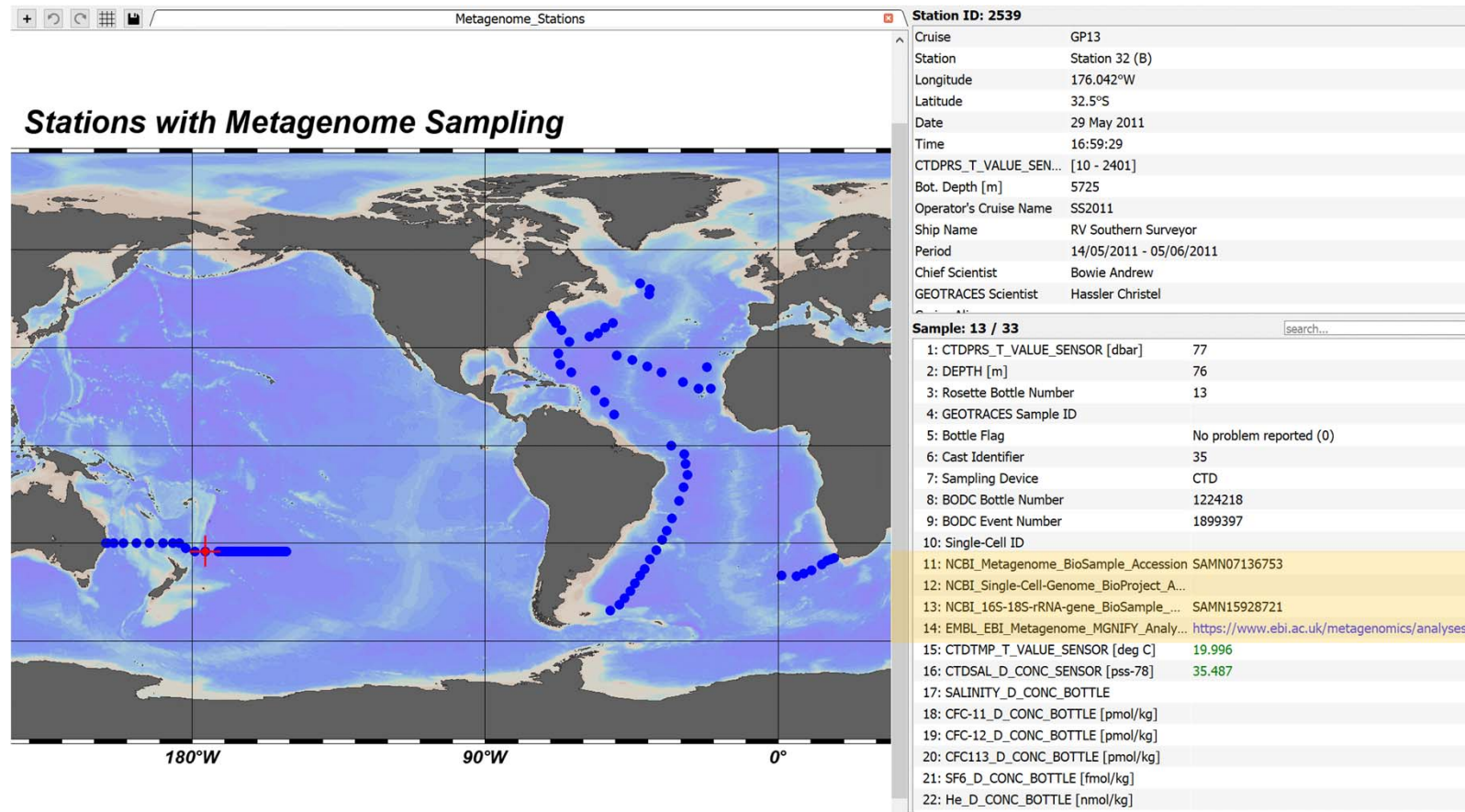
@<https://www.ebi.ac.uk/metagenomics/analyses/MGYA>

EXPLORATION: LINKING TEI & GENOMICS DATA IN IDP2021

OMICs linkages

4 Sample Descriptors in the Metadata variables:

- Accession numbers at NCBI or EMBL-EBI
- Or links to functional and taxonomic analyses webpages



METAGENOME BIOSAMPLE ACCESSION NUMBERS AT NCBI

Example 1.
Metagenome sample
accession numbers at
NCBI

www.ncbi.nlm.nih.gov

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

All Databases SAMN07136753 Search

COVID-19 Information
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

UNITE
A new NIH initiative to end structural racism and achieve racial equity in the biomedical research enterprise.
[LEARN MORE](#)

Ending Structural Racism
nih.gov/ending-structural-racism

NCBI Home Welcome to NCBI Popular Resources

METAGENOME BIOSAMPLE ACCESSION NUMBERS AT NCBI

Example 1.
Metagenome sample
accession numbers at
NCBI

www.ncbi.nlm.nih.gov



NCBI Resources How To Sign in to NCBI

BioSample BioSample Advanced Search

COVID-19 Information
[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

Full Send to

MIMS Environmental/Metagenome sample from marine metagenome

Identifiers **BioSample: SAMN07136753** Sample name: S0286; SRA: SRS2329440

Organism [marine metagenome](#)
unclassified entries; unclassified sequences; metagenomes; ecological metagenomes

Package [MIMS: metagenome/environmental_water_version_5.0](#)

Attributes

collection date	2011-05-29
depth	75m
broad-scale environmental context	ocean_biome
local-scale environmental context	ocean
environmental medium	water
geographic location	Pacific Ocean
latitude and longitude	32.5 S 176 W
geotraces_section	GP13
cruise_id	SS2011
cruise_station	32
bottle_id	SS2011Station32Cast35Niskin13

Description Keywords: GSC:MixS;MIMS:5.0

BioProject [PRJNA385854](#) marine metagenome
Retrieve [all samples](#) from this project

Related information
BioProject
SRA
Taxonomy

Recent activity [Turn Off](#) [On](#)

- MIMS Environmental/Metagenome sample from marine metagenome
- PRJNA445865 (1)
- Single cells of cyanobacteria and sympatric microorganisms
- SAMN07136495 (1)
- GP13-S0286

See more

FUNCTIONAL & TAXONOMIC ANALYSES EMBL-EBI

Example 2.

Functional & taxonomic analyses of metagenomes at EMBL-EBI

- <https://www.ebi.ac.uk/metagenomics/analyses/MGYA00452481#overview>



The screenshot shows the MGnify web interface. At the top, there is a navigation bar with links for EMBL-EBI, Services, Research, Training, and About us. The MGnify logo is prominently displayed, along with the tagline "Submit, analyse, discover and compare microbiome data". A search bar is located on the right, with examples like "MGYS00000410", "Tara Oceans", and "Human Gut". Below the navigation bar, there are tabs for "Overview", "Submit data", "Text search", "Sequence search", "Browse data", "API", "About", "Help", and "Login". The main content area shows the analysis ID "MGYA00452481" and a breadcrumb trail: "HOME / STUDY MGYS00005294 / SAMPLE SRS2329440 / RUN SRR5788301 / ANALYSIS MGYA00452481". Underneath, there are tabs for "Other analyses" including "Overview", "Quality control", "Taxonomic analysis", "Functional analysis", and "Download". The "Description [-]" section contains a table with the following data:

Study:	MGYS00005294
Sample:	SRS2329440
Run:	SRR5788301
Pipeline version:	4.1

The "Experiment details [-]" section contains a table with the following data:

Experiment type:	metagenomic
Instrument model:	NextSeq 550
Instrument platform:	ILLUMINA

FUNCTIONAL & TAXONOMIC ANALYSES EMBL-EBI

Analysis MGYA00452481

Other analyses

Overview Quality control **Taxonomic analysis** Functional analysis Download

rRNA

small subunit rRNA

large subunit rRNA

These are the results from the taxonomic analysis steps of our pipeline. You can switch between different views of the data using the menu of icons below (pie, bar, stacked and interactive krona charts). The data used to build these charts can be found under the "Download" tab.

Search: x

all Total: 24605

Max depth: 8 +

Font size: 11 +

Chart size: +

Collapse

Snapshot

Link

?

Example 2.

Functional & taxonomic analyses of metagenomes at EMBL-EBI

- <https://www.ebi.ac.uk/metagenomics/analyses/MGYA00452481#overview>

FUNCTIONAL & TAXONOMIC ANALYSES EMBL-EBI

HOME / STUDY MG900005294 / SAMPLE SRR2329440 / RUN SRR5788301 / ANALYSIS MG9A00452481

Analysis MG9A00452481

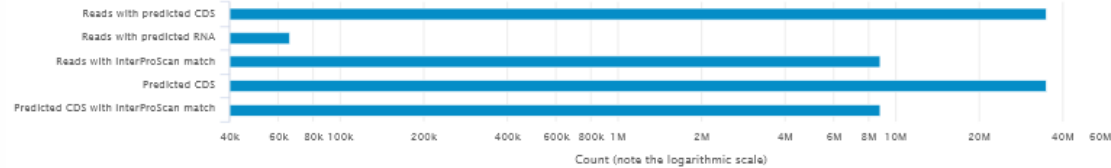
Other analyses

Overview Quality control Taxonomic analysis **Functional analysis** Download

These charts present the functional analysis of our pipeline, which focuses on matches to the [InterPro database](#) and [GO terms](#). These summarise the functional content of the sequences in the sample. The full set of results files may be found under the "Download" tab.

InterPro GO Terms

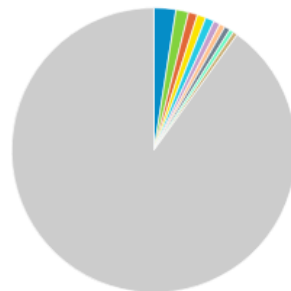
Sequence feature summary



InterPro match summary

InterPro match summary

Total: 8639785 InterPro matches



Show: 25 results from a total of 10112 (page 1 of 405).

Download results

Entry name	ID	pCDS matched	%
Aldolase-type TM barrel	IPR013785	215649	2.50
Rossmann-like alpha/beta/alpha sandwich fold	IPR014729	123660	1.43
Pyridoxal phosphate-dependent transferase, subdomain	IPR015422	89750	1.04
FAD/NAD(P)-binding domain	IPR023753	85430	0.99
Pyridoxal phosphate-dependent transferase, major	IPR015421	84283	0.98

Example 2.
Functional & Taxonomic analyses of metagenomes at EMBL-EBI

- <https://www.ebi.ac.uk/metagenomics/analyses/MG9A00452481#overview>

