Exploring Online Resources and Repositories (II)

Miguel Semedo

2024/09/05

Funding

Sponsorship

Support

Online Resources and Repositories

MG-RAST

- Upload **raw and assembled reads** and perform taxonomic and functional annotation

- **Download stored data** (512,000 metagenomes)

- Explore stored data (find certain genes and/or functions)

- Export analysis results

- *Meyer et al., 2008, BMC Bioinformatics*

- Website link: https://www.mg-rast.org

IMG/MER

- Upload **assembled reads** and perform taxonomic and functional annotation

- **Download stored data** (24,000 metagenomes + 25,000 genomes, etc.)

- Explore stored data (find certain genes and/or functions)

- Export analysis results

- Chen et al., 2019, Nucleic Acids Res

- Website link: https://img.jgi.doe.gov/cgi-bin/mer/main.cgi

- Available workshops (next: April 28th – May 2nd, 2025)

KBase

- Upload **any metagenome input (reads, contigs, bins)** and perform taxonomic and functional annotation

- **Highly customized pipeline.** Most current tools available

- **Share "narratives"**

- Explore stored data (find certain genes and/or functions)

- Export analysis results

- *Arkin et al., 2018, Nat Biotechnol*

- Website link: https://www.kbase.us

Online Resources and Repositories

MGnify

- Upload **raw and assembled reads (DNA, RNA, proteins)** and perform taxonomic and functional annotation

- Metabarcoding and metagenomics data

- All bioinformatic tools available, independently

- Download stored data – <u>linked to ENA</u> as well

- Explore stored data (find certain genes and/or functions)

- Export analysis results

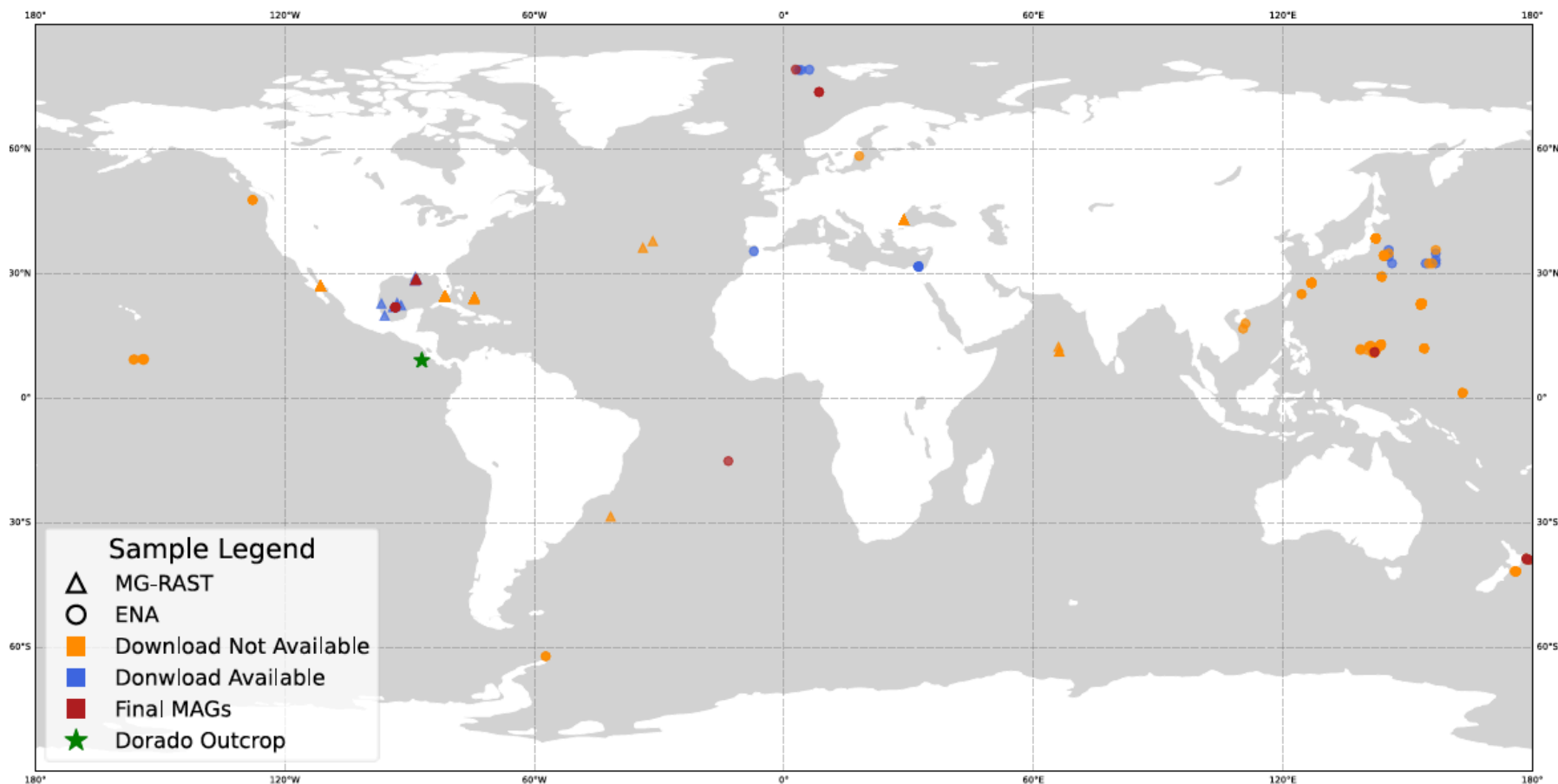- *Richardson et al., 2023, Nucleic Acids Res*

- Website link: https://www.ebi.ac.uk/metagenomics

ENA
European Nucleotide Archive

- Upload and make public **raw, assembled, or functionally annotated reads (DNA, RNA). Not used to perform analysis.**

- Metabarcoding and metagenomics data

- Explore and download stored data **using their metadata**.

- Website link: https://www.ebi.ac.uk/ena/

**All of these are free, but require registration (sometimes with different access levels).
Check webpage documentation.**

Online Resources and Repositories

Sample metadata is essential, but can be very difficult to acess..

- Querying public metagenome databases for **raw shotgun sequencing data** from **deep-sea marine sediments**, 277 metagenome entries (samples) were obtained: 79 from **MG-RAST**, 198 from **ENA**, and 0 from **Mgnify**

- **Available raw reads for download:** 20 MG-RAST entries and 46 ENA entries (66 metagenomes in total)

- **After our metagenomics full workflow**: 81 MQ or HQ MAGs found – **still extremely worth it** the hard work

Sample Legend
△ MG-RAST
○ ENA
■ Download Not Available
■ Donwload Available
■ Final MAGs
★ Dorado Outcrop

Online Resources and Repositories

Brief tour on IMG/JGI and KBase

IMG/MER

Better for exploring data (lots of genomes, annotations, and relatively good metadata).

https://img.jgi.doe.gov/cgi-bin/mer/main.cgi

- Show around
- Specific task:
    - Find the nosZ gene (K00376) in all available bacterial genomes?
    - Isolated from deep-sea marine sediments?
    - What are the gene neighborhoods in selected genes?
- Anything else you would like to do? Thousands of genes from thousands of genes and metagenomes from different environments (with metadata)

KBase

Better for analyzing your own data (or from collaborators). Many current tools and highly flexible and modular narratives. However, computational speed is limited, since this is a public server.

https://www.kbase.us

Questions?