MGnify Workshop Program

GSC23 Bangkok, 11 August 2023

Genomic Standards for Precision Medicine, Agriculture, Comparative Genomics, and Metabolomics
Location: 1st Floor, SiMR Building, Siriraj Hospital Mahidol University, 2 Thanon Wang Lang,
Siriraj, Bangkok Noi, Bangkok 10700, Thailand

MGnify is a freely available resource, provided by EMBL-EBI, for the analysis and archiving of microbiome-derived sequencing data. Users can request assembly of raw-read data, and taxonomic and functional analysis of publicly available (or pre-publication) datasets available in ENA. MGnify also provides microbiome-derived genome catalogues for multiple biomes, and a non-redundant protein database of over 3 billion sequences.

Two half-day workshops (a mixture of hands-on practicals and lectures) are available on Friday 11th August, provided by the MGnify resource. In the morning there will be an introductory session for users new to the field and/or the resource, and in the afternoon a more advanced session. More details are provided below about what will be covered in each session. Registration will be for both sessions, but delegates should specify which sessions they will attend - either one or both of the sessions, as they prefer based on their experience.

Please note that attendees will be expected to bring along their own laptop computer for use within the practical sessions. All practical aspects will be browser-based and so the only software required for the course will be access to a recent version of either Chrome, Firefox, or Safari.

Introductory session (9:00-12:00)

Topics covered:
- What is MGnify?
- Analysis and results of metabarcoding and metagenomic/metatranscriptic datasets
- Assembly of metagenomic data - the advantages and disadvantages
- Submitting data to ENA for analysis by MGnify
- MGnify quality control and data cleaning steps
- Why metadata is important
- Searching MGnify and accessing results

Advanced session (14:00-17:00)

Topics covered:
- Generating MAGs from metagenomic data
- Biome-specific genome catalogues
- Comparing new MAGs against genome catalogues
- Programmatic access to MGnify
- Downstream analysis in Jupyter notebooks, e.g. comparing across datasets