Mini-workshop: Metagenomics using DIAMOND+MEGAN
15 September 2023

Abstract

Interactive microbiome analysis using DIAMOND+MEGAN

Daniel Huson and Banu Cetinkaya, Institute for Bioinformatics and Medical Informatics, University of Tübingen

Computational analysis of microbiome samples is usually a two-step process. First, metagenomic sequencing datasets are subjected to computational analysis on a server. This usually involves alignment of metagenomic data against a protein reference database such as NCBI-nr, which is a computational intensive computation and requires specialized software such as DIAMOND. The resulting alignments will then be processed to perform taxonomic and functional binning, using additional command-line tools such MEGANIZER. Once server-based analysis has been completed, the second step is that investigators then explore and compare the taxonomic and functional content of the samples on a personal computer, using an interactive tool such as MEGAN.

In the first half of the tutorial, participants will learn how to setup and run the main computational steps of metagenomic analysis on a server. In the second half of the tutorial, the focus will be on how to interactively explore and compare the taxonomic and functional content of metagenomic datasets.

Learning Objectives:

The overall aim of this tutorial is to enable participants to setup and perform metagenomic analysis using the DIAMOND+MEGAN pipeline. At the end of the tutorial, participants will be familiar with the main steps of pipeline and will be able to:

- setup and run DIAMOND alignment of short read datasets against a protein reference database such as NCBI-nr, on a server,
- adjust the parameters of DIAMOND to align assembled contigs and long reads,
- use alternative databases such as AnnoTree or UniRef,
- perform basic taxonomic and functional computational analysis using MEGAN tools on a server,
- open and interactively explore the taxonomic and functional content of metagenomic datasets using MEGAN, on a personal computer,
- interactively compare multiple datasets,
- import and work with sample metadata,
- export assignments, reads and alignments, and
- run gene-centric assembly for genes of interest.

In preparation of the tutorial, please visit the tutorial website and install the software and files discussed there: https://github.com/husonlab/tutorials/wiki/Tutorial