

## **BION meta pre-release**

BION meta is an open-source package for 16S rRNA based bacterial community analysis. Like Qiime and Mothur it is open source, but has a growing number of advantages. The package has so far been developed mostly by the undersigned, one of the original RDP authors. It is as yet unpublished, but is being used "in production" by Danish institutions that also partly sponsored its development.

### **Features and advantages**

- Works on Linux, Unix, Mac OSX.
- Installs as regular user from source with one short command.
- Automatic installation of Greengenes, RDP and Silva databases.
- Works with non-amplicon data.
- Slices and de-replicates sub-databases to match amplicons.
- Accepts all standard sequence formats and quality encodings.
- Flexible de-multiplexing, cleaning, de-replication and clustering.
- Similarity matcher detects very low similarities and can ignore low quality.
- Includes a new fast chimera checker.
- Taxonomic mapping to species and sequence with multiple methods.
- Can merge and reorder profiles.
- Uses complete workflow "recipes" that are easy to understand and edit.
- Good performance and low resource usage (see below).
- Outputs a web-site with results, detailed statistics and file access.
- Results include a copy of the source code used to process the data.

Important missing features include signature based taxonomic mapping, combination of cluster linkage analysis with taxonomic mapping and good dynamic web visualizations and statistics. BION meta is command line based, but recipes are trivial to run; a first class web-interface for operation could be made, but has not yet been highest priority. The roadmap includes expansion to other molecules and whole transcriptomes.

### **Performance and outputs**

Taxonomic overviews of 250,000 454-reads takes 30-90 minutes on a four-core machine, 700,000 about 3 hours. The run-time for 200 million Illumina reads is 2-6 hours, depending on redundancy. Maximum RAM usage is usually 1-2.5 Gb. Output files go to a single flat directory, where files are consistently named with suffixes. This directory is also a static web site: point a browser to Results/index.html, and all statistics and data are available from there. The site can be copied to a USB stick, viewed on any system with a browser, or published onto the internet as is.

### **Support and contact**

BION meta comes with full source code and optional paid support, including running customer data remotely. If BION meta is offered as a paid service, the author must be contacted for arrangements.

Niels Larsen, CEO, PhD  
Danish Genome Institute  
Aarhus, Denmark

E-mail: niels@genomics.dk  
Mobile: +45-3091-5426 (GMT+1)  
Skype: niels\_larsen\_denmark