

Biostatistics

Use of various software in microbiology

Deepa Anwar

Metaxa

A software package for automated identification and classification of small subunit (SSU) rRNA sequences from archaea, bacteria, eukaryotes, chloroplasts and mitochondria in metagenomic and environmental sequencing data sets.

Megraft –

A software package to graft ribosomal small subunit (16S/18S) fragments onto full-length sequences for accurate species richness and sequencing depth analysis in pyrosequencing-length metagenomes.

ITSx –

A software tool for detection and extraction of ITS₁ and ITS₂ from ribosomal ITS sequences of fungi and other eukaryotes for use in environmental sequencing

PETKit –

A set of tools to ease the use of sequences from large-scale sequencing projects, generating paired-end reads.

metaorf –

A very simple gene finder adapted for the fragmentary nature of metagenomic data. There is no advanced probabilistic model involved in this piece of software, it just looks for open reading frames, allowing (as specified by the user) sequences to start and end outside of the nucleotide reads.

bloutminer -

A script for mining data from BLAST reports. Allows you to specify e.g. an E-value cutoff, a length cutoff and a percent identity cutoff, and extract a list of the hits satisfying these cutoffs. normally used an E-value cutoff of 10 for BLAST searches, and then extracts hits with bloutminer, allowing to change the cutoffs at a later stage without redoing the whole BLAST search. Is also able to “pool” sequences into groups, based on their sequence tags. In addition, it can also output the N number of best matches to a query.

Scripts collection –

Most of the things found here are just small handy perl or shell scripts.