MetaPathways

Elizabeth Suter MAR 550 November 2015

Metapathways is an open source pipeline for analyzing environmental sequences, annotating them, and visualizing them based on metabolic pathways or reactions. It can also build phylogenetic trees.

5 General Steps

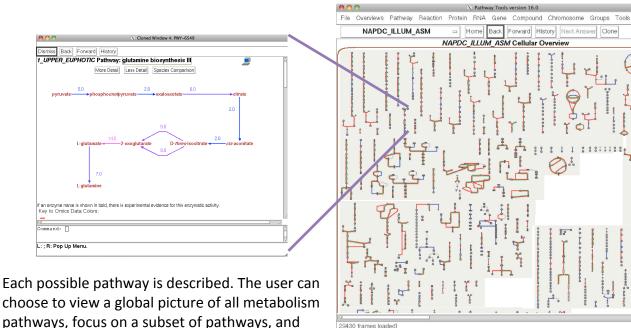
- 1- Quality control by trimming sequences and removing duplicates. Then protein-coding regions are predicted.
- 2- Functional and taxonomic annotation using BLAST.
- 3- Further taxonomic annotation using a Lowest Common Ancestor algorithm (MEGAN).
- 4- Environmental Pathway Genome Database (ePGDB) creation. The sequence, gene, and pathway data are integrated to predict metabolic pathways based on the MetaCyc pathways database.
- 5- Pathway export. The MetaCyc pathways can be exported for analysis by the user.

even on individual reactions and environmental sequences from which the pathways were based.

Benefits of Using Metapathways:

Metapathways lets the user annotate their sequences using multiple reference protein databases. It can handle both metagenomic and metatranscriptomic data. It integrates all steps of analyses that are necessary in metagenomics, from quality control to construction of environmental pathways, making the process more user-friendly. And recently, they developed a GUI version.

Example of Output



References:

Konwar KM, Hanson NW, Pagé AP, and SJ Hallam (2013) MetaPathways: a modular pipeline for constructing pathway/genome databases from environmental sequence information. *BMC Bioinformatics* 14:202

Hanson NW, Konwar KM, Wu S, and SJ Hallam (2014) MetaPathways v2. 0: A master-worker model for environmental Pathway/Genome Database construction on grids and clouds. *Computational Intelligence in Bioinformatics and Computational Biology, 2014 IEEE Conference* IEEE: 1-7 Hanson N. (2015) Metapathways v2.5 Wiki https://github.com/hallamlab/metapathways2>