

Metabolomics applications of the MetaCyc family of metabolic databases

Peter D Karp

SRI International, USA

The MetaCyc family of pathway databases consists of more than 2,500 Pathway/Genome Databases (PGDBs). MetaCyc itself is a multi-organism PGDB that contains 2,000 experimentally elucidated metabolic pathways found in 2,400 organisms. The MetaCyc data were curated from 35,000 publications. Most PGDBs in the MetaCyc family were created for organisms with sequenced genomes. Many are highly curated, including PGDBs for humans, *E. coli*, yeast, mouse, and *Arabidopsis*.

Common to MetaCyc-family PGDBs is the Pathway Tools software, which contains a large suite of algorithms for manipulating biological networks and genome data. Given the annotated genome of an organism, Pathway Tools will infer the metabolic pathways of the organism, and will generate a genome-scale quantitative metabolic flux model from the predicted metabolic network. The software includes extensive visualization tools for individual metabolic pathways, for complete metabolic networks, and for complete regulatory networks. These visualization tools can be used for analysis of metabolomics datasets, such as painting metabolomics data onto a zoomable metabolic map diagram. The software also finds metabolic pathways that are statistically over represented in metabolomics datasets.

pkarp@ai.sri.com