BioCyc & Pathway Tools: Summary

The BioCyc Database Collection™ is a collection of 20,049 Pathway/Genome Databases™ (PGDBs). A PGDB contains the annotated genome of an organism and its computationally inferred reactome, metabolic pathways, and operons; some curated PGDBs contain regulatory networks. Databases are available as downloadable flatfiles or combined with the Pathway Tools software. There are three tiers of databases with varying amounts of curation:

Tier 1
Highly curated, with more than one person-year of manual, literature-based curation. Includes databases for human (HumanCyc), *E. coli* K-12 (EcoCyc), and *Saccharomyces cerevisiae* (YeastCyc); MetaCyc is a database of metabolic pathways and enzymes from all domains of life curated from more than 76,000 publications.

Tier 2
Moderately curated, with no more than one person-year of manual, literature-based curation. Includes databases for *Bacillus subtilis, Pseudomonas putida, Mycobacterium tuberculosis, Salmonella enterica, Corynebacterium glutamicum*, and sixty other organisms.

Tier 3
Not curated. Tier 3 databases exist for more than 19,000 organisms.

Pathway Tools™ is a PGDB-management software system consisting of four independent modules:

Pathway/Genome Navigator™
Software tools for visualizing, analyzing, and searching PGDBs, including tools for omics data analysis, metabolic route searching, comparative analysis, and sequence analysis. In webserver mode, Navigator makes custom PGDBs available via a web portal.

PathoLogic™
Computationally infer an organism's reactome, metabolic pathways, protein complexes, missing enzymes, and operons from an annotated genome.

Pathway/Genome Editors™
Interactively update a PGDB by adding a new metabolite or metabolic pathway, changing a gene function, etc.

MetaFlux™
Construct a quantitative metabolic flux model from a PGDB; allows temporal simulations and simulation of spatial interactions.

Tools within the Pathway/Genome Navigator:

Genome Browser
Display genomic regions—genes, operons, promoters, terminators, etc.—using ultra fast zooming.

Cellular Omics Viewer
Paint omics data onto a metabolic-map diagram; supports animation of time course or comparative omics datasets.

Omics Dashboard
Survey transcriptomic and metabolomic responses of cellular systems—e.g., biosynthesis, degradation, regulation—to stimuli.

Regulatory Overview
Display and navigate the genetic regulatory network of an organism.

Cellular Overview
Display and navigate the metabolic and transporter networks of an organism.

Information pages
Find detailed information on biochemical pathways and reactions, chemical compounds, genes, proteins, RNAs, operons, and regulatory interactions.

RouteSearch
Find minimum-cost paths between existing metabolites in a metabolic network and design pathways to new metabolites.

SmartTables
Store and analyze groups of genes, metabolites, pathways, etc.

Comparative Tools
Compare genomes, proteomes, reactomes, metabolic pathways, etc.

Structured Advanced Query Form
Interactively formulate advanced SQL-like queries to PGDBs.

Export Tool
Export PGDBs in multiple formats including BioPAX, SBML, GFF, and Genbank.

APIs/Web services
PGDBs can be queried through Python, Perl, Java, and Lisp APIs and via web services.

Hardware requirements

There are no specific hardware requirements for working with individual PGDB flatfiles.

Pathway/Genome Navigator with up to 500 PGDBs: Linux, Windows (webserver mode not available), MacOS with at least a 2.4 GHz processor, 16 GB RAM, 5 GB disk.

Pathway/Genome Navigator plus all PGDBs: Linux with at least a 2.4 GHz processor, 128 GB RAM, 3 TB disk.