

EXPERIMENT READERS

Affymetrix® and Agilent experiments can be read and processed using several BioConductor packages (affy, affyPLM, plier, limma). Processing includes background correction, normalization, and summarization. Excel experiment readers can import data directly from existing spreadsheets. GEO and SOFT readers enable processing of data from NCBI's Gene Expression Omnibus.

EXPERIMENT ANALYSIS

Fold change and pairwise differential expression are quickly calculated for full experiments using Student's t-test or Wilcoxon tests with optional control for multiple comparisons. Gene subsets can be defined based on calculations or annotations—all without copying or cloning data. Subsets can also be defined using generic components like Outlier Filter or Top N Filter. Set operations (e.g., union, intersection, and subtraction) can be performed on existing subsets. Clustering components, including hierarchical and k-means, are also provided.

DATA MANIPULATORS

Gene expression experiments can be quite large. Once a subset of interesting genes has been identified, the irrelevant portions of the experiment can be removed from the experiment record, greatly speeding up downstream processing. Additionally, individual genes can be extracted, leading to easy integration with the Sequence Analysis Collection.

RESULTS ANNOTATORS

Annotations (e.g., descriptions, pathway identifiers, Gene Ontology terms) can be added using vendor-provided data via BioConductor's annaffy package. In addition, annotations can be imported from existing flat files.

VIEWERS AND REPORTING TOOLS

Use BioConductor graphics or create your own graphics and reports using Pipeline Pilot's Reporting Collection. While BioConductor provides sophisticated plots and charts, they can be difficult to configure and use. The Gene Expression Collection makes it easy to use these standard BioConductor reporting capabilities, while also giving you the option to implement additional data views using the Reporting Collection. This gives you more control over display features, including hyperlinks and tooltips. Supplied components include heat maps, annotated tables, and parallel coordinate plots.

INTEGRATION WITH OTHER PIPELINE PILOT COLLECTIONS

The components in the Gene Expression Collection are designed to work seamlessly with components from existing collections, including the Sequence Analysis, R Statistics, Data Modeling, and Chemistry (via pathways) Collections.

To learn more about Pipeline Pilot, go to accelrys.com/pipeline-pilot