Genedata Analyst

Technical Data Sheet

Genedata Analyst™ is the premier software solution for the integration and interpretation of experimental data in life science R&D. Built on a scalable client-server architecture, Genedata Analyst supports billions of data points and hundreds of users. It puts rigorous statistical algorithms, interactive data analysis tools, and intuitive visualizations into the hands of scientists and data analysts.

Applications

Genedata Analyst enables researchers to analyze data from different experimental platforms. With applications ranging from biomarker discovery and patient stratification to trait development, Genedata Analyst is a powerful asset in any R&D software portfolio.

Biomarker discovery:

- Identification of disease markers and prognostic biomarkers
- Correlation between phenotype information and molecular profiling data

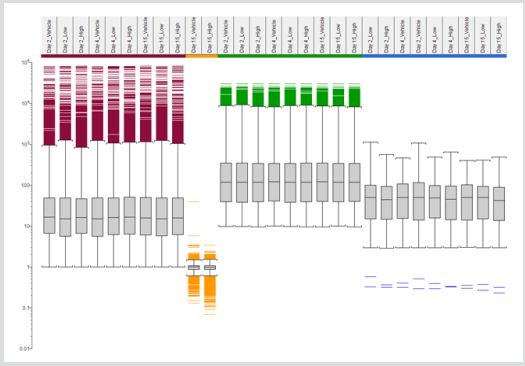
Modeling and prediction:

- Prediction of compound toxicity (toxicogenomics)
- Mode-of-action classification
- Patient stratification and personalized medicine
- Molecular diagnostics

Mapping:

- Integration of diverse data through proprietary mapping technology
- Incorporation of public domain data with proprietary experiments
- Combination of omics profiling and phenotype data

Data Integration



Data Integration

Distribution of metabolomics, proteomics, transcriptomics, and clinical chemistry data. The box plot highlights the need for data normalization to enable cross-technology comparisons.

Statistical Tools

Genedata Analyst provides a range of powerful data analysis tools designed to support statistical analysis of even the most complex life science data.

Data normalization:

- Linear normalization to standardize experiments and remove artifacts
- Nonlinear normalization methods like LOWESS, Quantile Normalization and Median Polish

General statistics:

- Data overview and QC using tools like Principal Components Analysis (PCA)
- Unsupervised clustering and network analyses help facilitate generation of new hypotheses

Statistical tests:

- Parametric and non-parametric tests
- Mixed linear model, ANOVA, ANCOVA supporting complex experimental designs
- Multiple testing corrections controlling false positives from high-throughput experiments
- Trend identification and time series analyses

Machine learning:

- Modeling and prediction using tools like Decision Trees, PLS, and Linear Discriminant Analysis
- Feature selection methods like ANOVA and Recursive Feature Elimination
- Model validation using leave-one-out and Monte Carlo cross validation

Genedata
solutions in silico

Data Visualization

Fully interactive and scalable data visualization enables researches to understand and interpret complex statistical results.

Interactive data visualization:
- Heat maps, bar charts, line graphs, and pie charts
- 2D and 3D scatter plots, log-log plots, and MA plots
- Histograms and volcano plots
- Clustering results with dendograms, heat maps, and tree maps

Data exploration:
- Shared selections across visualizers and tables support data exploration and data mining
- Linked data tables with full statistical details

Publication ready:
- Export of results to image and text files, and as Excel and PDF documents
- Integration with system clipboard for easy pasting into MS Office applications

Integration APIs

Genedata Analyst is an open and flexible platform built on a plugin framework with supported APIs. The integrated SDK enables Java developers to expand Analyst with integrated plugins.

Data import:
- Parser API supports custom file import options
- Default parsers for all major microarray and RT-PCR vendors
- Data import API to load data and annotation from any external source (incl. web services)
- Standard integration with Gene Expression Omnibus (GEO)
- Out-of-the-box integration with all Genedata products

Data export:
- Export as text files and Excel workbooks
- Export to pathway analysis tools
- Simple configuration of URL-based exports

Data analysis:
- Java analysis API enables integration with proprietary and third-party tools
- Support for R analysis scripts, utilizing built-in visualizations

Architecture

Built on a client-server architecture, Genedata Analyst ensures highest scalability, ease of installation and maintenance. Server solution enables end-to-end traceability for GxP compliance.

Client-server enterprise solution:
- Light-weight client with fully interactive visualizations
- Automated deployment of patches and updates

Highest scalability and performance:
- Proprietary disk caching for the processing of unlimited data sets
- Parallel processing scales linearly with available system resources

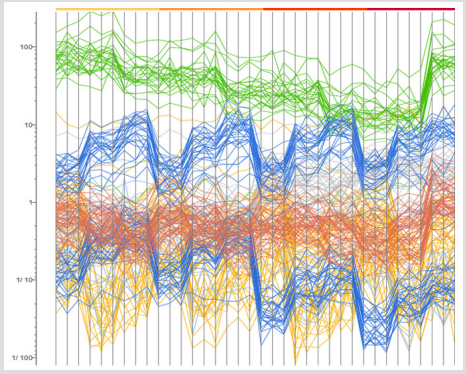
System Requirements

Genedata Analyst works on standard server hardware without the need for clusters or grids. This allows seamless and cost-efficient integration into existing IT environments.

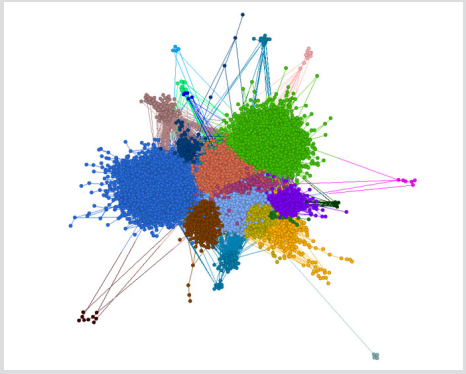
Server recommendation:
- SUSE Linux Enterprise on x64
- 4 GB of free memory
- 100 GB of free disk space

Client recommendation:
- Windows 7 on Intel Core2 or newer
- 2 GB RAM
- 10 GB of free disk space
- 1280 x 1024 color display

Network configuration:
- LAN with 100 MB/sec
- Fixed server IP address
- Forward and reverse DNS lookups
- No firewall between client and server



Linear Model
Marker genes from a two-factor model considering time, dose, and interaction effects.



Network Analysis
Human protein-protein-interaction network showing >13,000 proteins with ~250,000 interactions.

Genedata Analyst is part of the Genedata portfolio of advanced software solutions that serve the evolving needs of drug discovery, industrial biotechnology, and other life sciences.

Basel | Boston | Munich | San Francisco | Tokyo
www.genedata.com | analyst@genedata.com

Genedata
solutions in silico