

Ex



# Genedata Expressionist

for Mass Spectrometry

## Technical Data Sheet

### Applications

With applications ranging from bio-marker discovery and patient stratification to biopharma, Genedata Expressionist® is a powerful asset in any R&D software portfolio.

#### Proteomics:

- ▶ Label-free LC-MS
- ▶ Fragment analysis with LC-MS/MS
- ▶ 4D data from LCxLC-MS experiments
- ▶ Labeled experiments using SILAC, iTRAQ, and TMT
- ▶ MRM for targeted analysis

#### Metabolomics:

- ▶ Label-free LC-MS and GC-MS
- ▶ Fragmentation analysis via LC-MS/MS and GC-MS/MS
- ▶ ESI-MS and Infusion MS
- ▶ Adduct clustering, formula detection
- ▶ Compound identification with small molecule libraries

#### Lipidomics:

- ▶ ESI and Infusion MS
- ▶ Integration with lipid libraries and databases

### Data Import

Genedata Expressionist is an open and flexible platform that imports raw data from all major vendors, instruments and technologies.

#### Vendor platforms:

- ▶ AB SCIEX

- ▶ Agilent
- ▶ Bruker
- ▶ LECO
- ▶ Shimadzu
- ▶ Thermo
- ▶ Waters

#### Public formats:

- ▶ mzXML
- ▶ NetCDF

#### Search engines and libraries:

- ▶ Mascot and SEQUEST
- ▶ X!Tandem and TPP
- ▶ ChemSpider
- ▶ HMDB and Metabocard files
- ▶ MSP libraries
- ▶ Glycome DB

### Data Processing

Genedata Expressionist provides powerful tools supporting pre-processing and cleansing of MS data irrespectively of technology, vendor, and application. All tools provide highest scalability and can be used to process 1000s of samples in parallel.

#### Calibration:

- ▶ Lock masses ensure standardization across samples
- ▶ Retention indices enable comparison across runs and columns

#### Noise removal:

- ▶ Smoothing and baseline subtraction clean even the noisiest MS data

- ▶ Proprietary noise subtraction eliminates most sources of background signal
- ▶ Process control with subtraction of blank control samples

#### Alignment corrections:

- ▶ Nonlinear retention time drift correction using raw intensity signals
- ▶ RT correction of samples with or without reference samples
- ▶ Automatic QC based on RT shift analysis

#### Peak detection:

- ▶ Proprietary peak picking capable of handling large number of samples
- ▶ Reference peaks support long-running clinical studies
- ▶ Isotope clustering to determine charge state and masses
- ▶ Mass deconvolution for intact proteins

### Visualization

Genedata Expressionist visualization tools display raw and processed data down to individual measurement values. This interactive visualization allows for real-time quality control and result validation.

#### Chromatogram plane:

- ▶ Fully interactive viewer showing RT, m/z, and intensity
- ▶ Integrated visualization of MS/MS fragment scans and peaks
- ▶ Easy navigation across many samples using bookmarks
- ▶ Unlimited scalability

#### Montage view:

- ▶ Montage view to inspect features across 1000s of samples
- ▶ Effortless validation of biomarker peaks across large experiments

#### Spectrum viewer:

- ▶ Display of individual scans, transitions, and single ion chromatograms
- ▶ Integrated visualization of peaks, peak clusters, and compounds

#### 3D viewer:

- ▶ 3D view for ultimate data inspection and validation

#### Reporting:

- ▶ 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

## Statistical Tools

Genedata Expressionist 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, or 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 experiments
- ▶ 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

## Integration APIs

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

#### Data import:

- ▶ Data import API to load data and annotation from any external source
- ▶ Standard integration with Gene Expression Omnibus (GEO)

#### Data export:

- ▶ Export as text files and Excel
- ▶ 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

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

#### Client-server enterprise solution:

- ▶ Light-weight client with fully interactive visualization
- ▶ Automated server-side deployment of patches and updates
- ▶ Data management based on proven Oracle technology

#### Highest scalability and performance:

- ▶ Proprietary swapping system for the processing of unlimited data volumes
- ▶ Parallel processing scales linearly with available system resources

## System Requirements

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

#### Server recommendation:

- ▶ SUSE Linux Enterprise on x64
- ▶ 8 GB of free memory
- ▶ 100 GB of free disk space

#### Client recommendation:

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

#### Network configuration:

- ▶ 100 Mb/s LAN without firewalls
- ▶ Fixed server IP address



Genedata Expressionist 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 | expressionist@genedata.com

© 2013 Genedata AG. All rights reserved. Genedata Expressionist is a registered trademark of Genedata AG. All other product and service names mentioned are the trademarks of their respective companies. 05E13

