GelCompar II®

GelCompar II 6.5 consists of Basic Software and 4 analysis modules: Tree and Network Inference, Identification & Library Manager, Dimensioning techniques, and Database Sharing Tools.

The Basic Software can be combined with any or all modules. The descriptions below list most, but not all functions and possibilities of the GelCompar II modules. Please call for details and prices.

▪ *Basic software functionality*

Database. Object-oriented relational multi-user database interface. Display of up to 150 information fields per database entry. Store information fields as text, number or date. Lock information fields or limit editing options by using pick lists. Easy drag & drop linkage of multiple experiments to database entries. List up to 250000 entries in one view. Powerful search engine for combined database searches on information fields and experiment presence and/or contents (character values, sequences, ranges, bands, etc.). Creation of object queries to retrieve data from any database object. Storage and management of database queries and external attachments. Multi-database system: each database can contain any combination of different experiment types. Creation of levels and relations for a richer hierarchical data structure. Creation of XML files from any selection of entries and techniques in the database for export. XML files can be imported as fully editable database entries. XML exchange is the preferred way of exchanging database entries in a peer-to-peer network.

Image processing and normalization. Input of any bitmap images, densitograms, and chromatograms of unlimited file size. Image pre-editing and cleaning. 3D representation of bitmaps. Automatic lane finding for all types of gels. Gelstrip borders and tracking splines adjustable for individual lanes. Automated and manual alignment by pattern recognition using external reference patterns and/or internal reference bands. On-screen normalization of bitmap images with indication of reliability and possible misalignments. Direct processing of sequencer chromatogram files and fragment analysis files with inline reference tracks. Adjustable background subtraction and curve smoothing. Spot removal. Display of any combination of normalized 2D-bitmap strips, densitograms or reconstructed patterns. Direct comparison of patterns normalized with different reference systems.

Quantification. Band-search algorithms with adjustable sensitivity for shoulder and double-band finding. Possibility to find and mark uncertain bands/peaks. Quantification of molecular sizes or any other metric unit using linear, logarithmic, combined logarithmic-third power regression, cubic spline or pole functions. Accurate expression of protein or nucleic acid quantities or concentrations based on cubic spline regression using known calibration peaks. Comparative quantification of bands/peaks between groups of patterns. Generation of tables and reports for unlimited numbers of patterns, indicating molecular weight, fragment length, absence/presence or absolute amounts of protein or DNA per band/peak. Search for discriminative bands/peaks between selected groups of patterns; search for unique and common bands/peaks. Binary and quantitative band matching tables of multiple combined fingerprints. Possibility to define named band classes based upon size and position (e.g. for DGGE/TGGE analysis). Add/edit bands directly in the comparison window in band matching mode.

Error values. Import data from high-throughput molecular fingerprint techniques (such as MALDI, SELDI,...) with error values (e.g. standard deviations). Show errors in the curve window or compare them with an error-weighted correlation coefficient.

▪ *Tree and Network Inference*

Methods. Comparisons of up to 20,000 database entries, various similarity/distance coefficients for different data types: Pearson product-moment correlation, cosine correlation, Dice or Nei and Li, Jaccard, Jeffrey's X, Ochiai,... Fuzzy logic and area sensitivity for banding patterns. Similarity-based clustering: Unweighted pair-grouping (UPGMA), complete linkage (furthest neighbor), single linkage (nearest neighbor), Ward or Neighbor Joining clustering. Adjustable trace-to-trace optimization and tolerance settings for banding patterns. Statistical determination of most justified tolerance settings for banding patterns. Interactive wizard-driven input of parameters, options and choices make the advanced clustering window more intuitive for users with little statistical background.

Phylogenetic inference methods. Generalized Parsimony. Population modeling: Analysis of categorical data such as using Minimum Spanning Trees to reconstruct evolutionary models. Advanced presentation and editing tools, including faithful tree representation ('rendered trees').

Interpretation. Combined display of normalized pattern images from different gel types, with similarity matrices and sorted according to dendrogram(s). Indication of statistical error at all linkage levels and calculation of co-phenetic correlation. "Seaweed" and pseudo-rooted representation for unrooted trees. Bootstrap analysis for single or composite datasets. Display of sorted similarity matrices, shaded or with numerical similarity values. Impressive edit and publishing functions. Enhanced presentation and printing facilities, in a WYSIWYG environment. Direct interaction between database and dendrogram. Incremental and decremental clustering: new entries can be added to or deleted from existing cluster analyses, without having to recalculate the complete analysis. All features of a comparison can be stored to disk.

Trusted Clusters. Patented method allowing the reliability of clusters to be calculated and visualized for any clustering algorithm and data set. The method enumerates the reliability of dendrograms or networks in function of degeneracies as well as poorly resolved clusters and can calculate consensus trees or networks that impose a minimum reliability threshold on each resolved cluster.

Composite cluster analysis. Different data sets can be combined into one consensus clustering. Calculation of global similarity by merging characters or by averaging experiment-related similarities. Optional weighting based on number of characters or defined by the user.

Plots and graphs. Creation of 2-D and 3-D bar graphs, contingency tables, 2-D and 3-D scatterplots from database fields and characters. Professional presentation, printing and exporting tools.

Identification

Database screening. Fast identification of batches of entries with entire databases or selections from databases, using all available coefficients.

Libraries. Creation of highly characteristic identification libraries using the open multi-library system. Specific similarity measures and settings can be defined for specific experiment types. Comprehensive identification reports showing results for each available experiment. Many different viewing options and statistical tools to facilitate interpretation.

Neural Networks. Neural Networks can be trained for each experiment type and used for quick and accurate identification of complex groupings.

Dimensioning techniques

Principal Component Analysis. Non-hierarchic grouping by PCA. Spatial representation of clouds of entries in user-definable X-Y-Z coordinate systems. Indication of total discrimination of axes. Real-time rotation of coordinate system to enhance perception of 3-D structures. Advanced Open-GL presentation and layout for publication. Delineation of populations using colors and/or codes. Plotting of dendrogram branches on PCA for advanced grouping comparisons and methodological validations.

Multi-Dimensioning Scaling. Non-hierarchic grouping by MDS. Iterative optimization of distances according to similarity matrix. Same presentation features as PCA.

Statistics. A number of parametric and non-parametric statistical tests can be performed in an easy and intuitive environment (Chi-square test, T-test, Wilcoxon signed-ranks test, Kruskal-Wallis test, ANOVA, Pearson correlation test, Spearmann rank-order test. Automatic display of available tests for each input data type. Kolgomorov-Smirnov test for normality. Clear significance reporting.

Database Sharing Tools

ODBC connectivity. Possibility to link GelCompar II to ODBC compatible high performance database management systems such as Oracle, SQL Server, MySQL.

User management. Comprehensive set of User and Security tools, including creation of Users with logins and passwords and User Groups defining specific privileges. Control password timeout and strength, user activity logging and data input consistency. Possibility to define access privileges for each individual database object. Create, Modify, Delete, Sign, Restore, Lock and Unlock privileges can be granted to specific users.

Bundles. Any selected information and experiment data for selections of entries from the database can be condensed into a Bundle. Bundles can be exchanged over internet and opened in a recipient database. Automatic remapping makes full comparison between different fingerprint systems possible.

Client-Server setup. Client functions come with the Database Sharing Tools. Functions include querying and downloading entries from a central BioNumerics Server database; upload of data to the Server for identification; receipt of detailed identification report from Server. Call for prices of BioNumerics Server package.

Geographical coding. Perform geocoding and plot database entries on an interactive geographical map based on locations present in the database.

GelCompar II Network Licensing Solutions

Powerful network solution. HTTP-based licensing protocol, ensuring compatibility with most firewall and LAN restrictions. Full information and control of license usage at any time by any user and optional logging of critical events or all actions. Compatible with Windows XP, Windows Vista, Windows 7, Windows Server 2003, Windows Server 2008.

License limits. Network versions are available from 2 up to any number of users. Contact us for details and pricing.