Universität Bielefeld

BioLayout Express$_{3D}$

Abel Nikoghosyan
09.07.2014, Bielefeld

Inhaltsverzeichnis

- BioLayout Express3D
  - Charakteristics
  - Functionality
  - Algorithms & Tools
  - Imput Formats
- Implementation
- Demonstration
- Results

BioLayout Express3D Charakteristics

- …..is an application Designed for integration, visualisation and analyses of large network graphs derived fromm biological data
- …..is based on an earlier program, Bio Layout Java, written in Java and utilizing Java 2D for graphics
- …. is an independant platform and runs on Windows, Apple Mac or Unix Systems. It has been designed to work on all standard computing platforms (Windows 32/64bit, MacOSX, Linux 32/64bit) and to make best use of the available hardware
- …. can render graphs of up to 30,000 nodes and 2-3 mil. edges using standard workstation hardware with 3D accelerated graphics card
- ….is suited for user with interaction with larger and complicated data set. Allows the display to pack information more effectively in the viewing area

BioLayout Express 3D Functionality

- Hardware-accelerated Open GL framework is used, combined an optimized layout algorithm and graph clustering (which allows the user to effectively explore graphs and data sets beyond the reach of other tools with standard consumer computer)
- It also supports the input of data in a number of standard graph formats and appropriate data from any source can be visualized
- Transforming gene expression data into a network graph based on correlation network
- Able to display and cluster large graphs in two- or three-dimensional space

  ⇨ Visualization of networks in 3-dimensions provides a fast and powerful interface to understand the often complex relationships between the entities represented in the graph

BioLayout Express 3D Algorithms & Tools

- Markov clustering algorithm (heavily optimizied C-based MCL, Graph clustering)
- Java 1.6 language
- Java 1.6 Generics and Iterator compliant code
- OOP for graph modeling and processing
- Modified Fruchterman-Rheingold layout algorithm for 2D/3D graph positioning and display
- JOGL Sun Api (For fast native OpenGL 2D/3D rendering engenes)
- Apache Xerces SAX XML parser (GraphML parsing and file creation)
- NSIS library (istaller support for Windows platform, JRE included)

BioLayout Express 3D Algorithms – MCL

- MCL - Is state-of-the-art general purpose clustering algorithm
- Represents a powerful approach to dividing graphs nonsubjectively into discrete chunks of genes (These genes share similarites in their expression)
- Uses the general properties of clusters
- - lot of internal edges
- - less of external edges
- Based on simulation of stochastic flow in graphs
- Fast and scalable
- Used for identification of ortholog groups, protein compleses, molecular pathway discovery

BioLayout Express 3D
Input Formats - Conventions

- Usually text files representing columns of data
- Data points separated by tabs
- Each node has a unique identifier
- Text entries are enclosed by quotations (e.g., „Protein Kinase Alpha“)
- Comments placed in file by proceding them with „//“
- Advanced options placed at the end of the input file (e.g., // NODESIZE)

BioLayout Express 3D
Input Formats

- Simple multicolumn format (.layout, .txt)
- Reactome OWL format (.owl)
- Cytoscape SIF format (.sif)
- Matrix files (.matrix)
- Expression data input format (.expression)
- Graphml (.graphml)
- Creation of classes

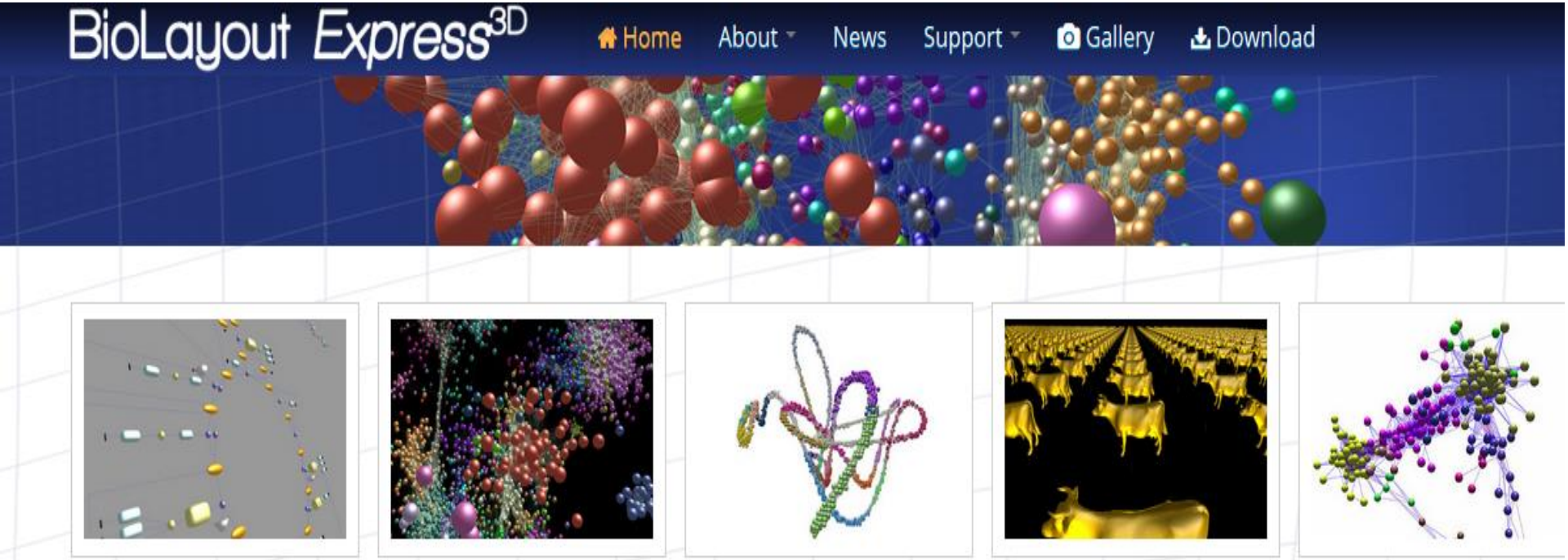
Expression data input format

Unique;ProbeID	Description	Annot.1	Annot.2	Data1	Data2	Data3
Tub;gnf1m00002 _f_at	tubulin, alpha 7	Term1	Term1	245.6	278.9	364.6
Il16;gnf1m00009_s_at	interleukin 16	Term2	Term1	125	203	235.2
Cul7;gnf1m00122_a_at	c cullin 7	Term3	Term2	302	288	134.7

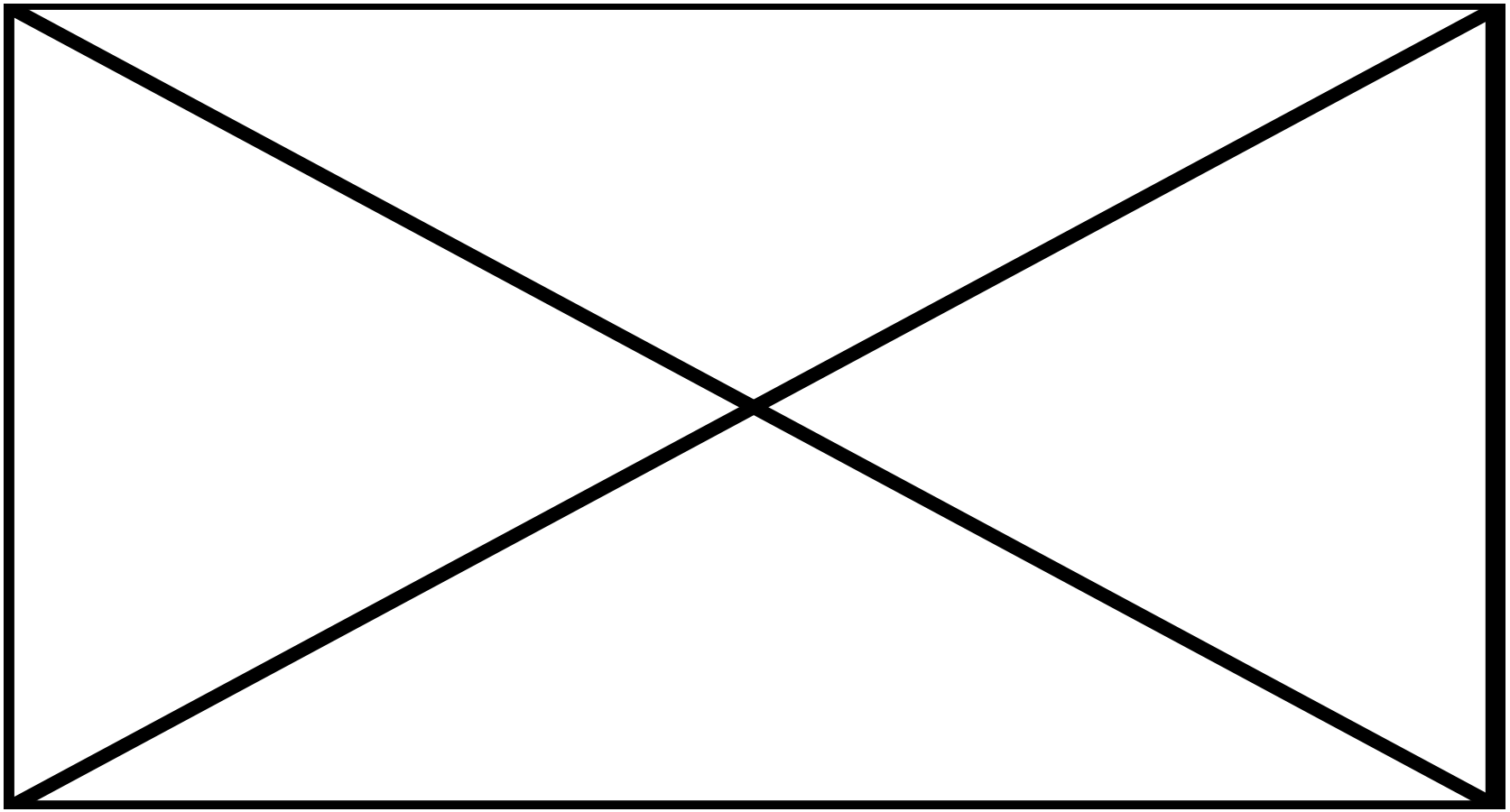
Aim Of The Work

- Network graphs represent coexpression relationships within a data set (above the threshold level)
- Breaking them down into modules/unites of coexpressed genes by clustering the graphs
- Aims to help biologists to find the degree of differential expression between experimental conditions or biological degree
- Make genes, that share the same function and are tightly coexpressed (Ribosomes), form cliques within a network
- By dropping the correlation threshold the graph is getting larger → basal Genes, weakly coexpressed, with little differential expression account for more of the graph structure
- True exploration of the data
- Provide an interface for exploration of large portions of data and integration of other information

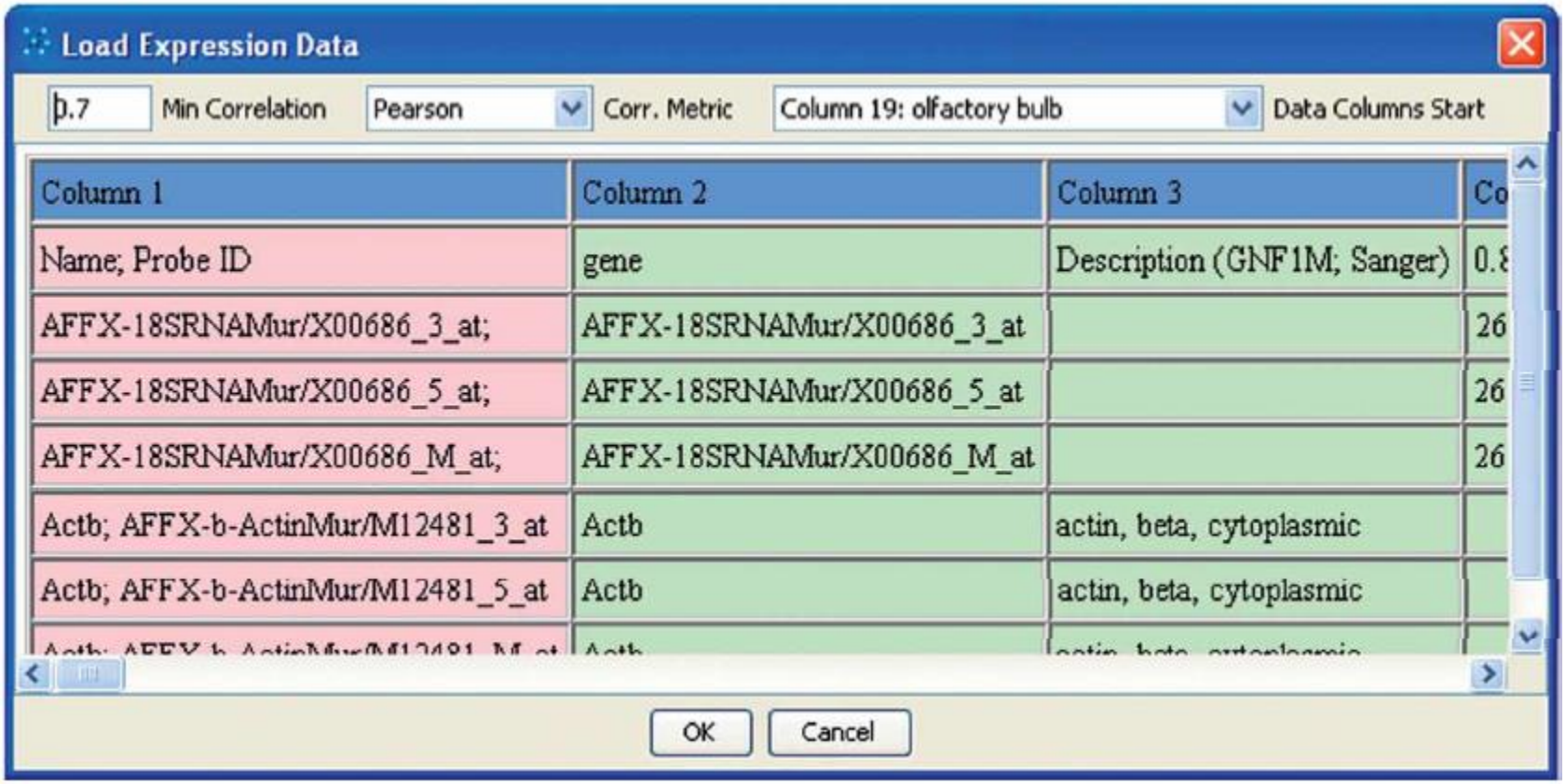
BioLayout Express 3D



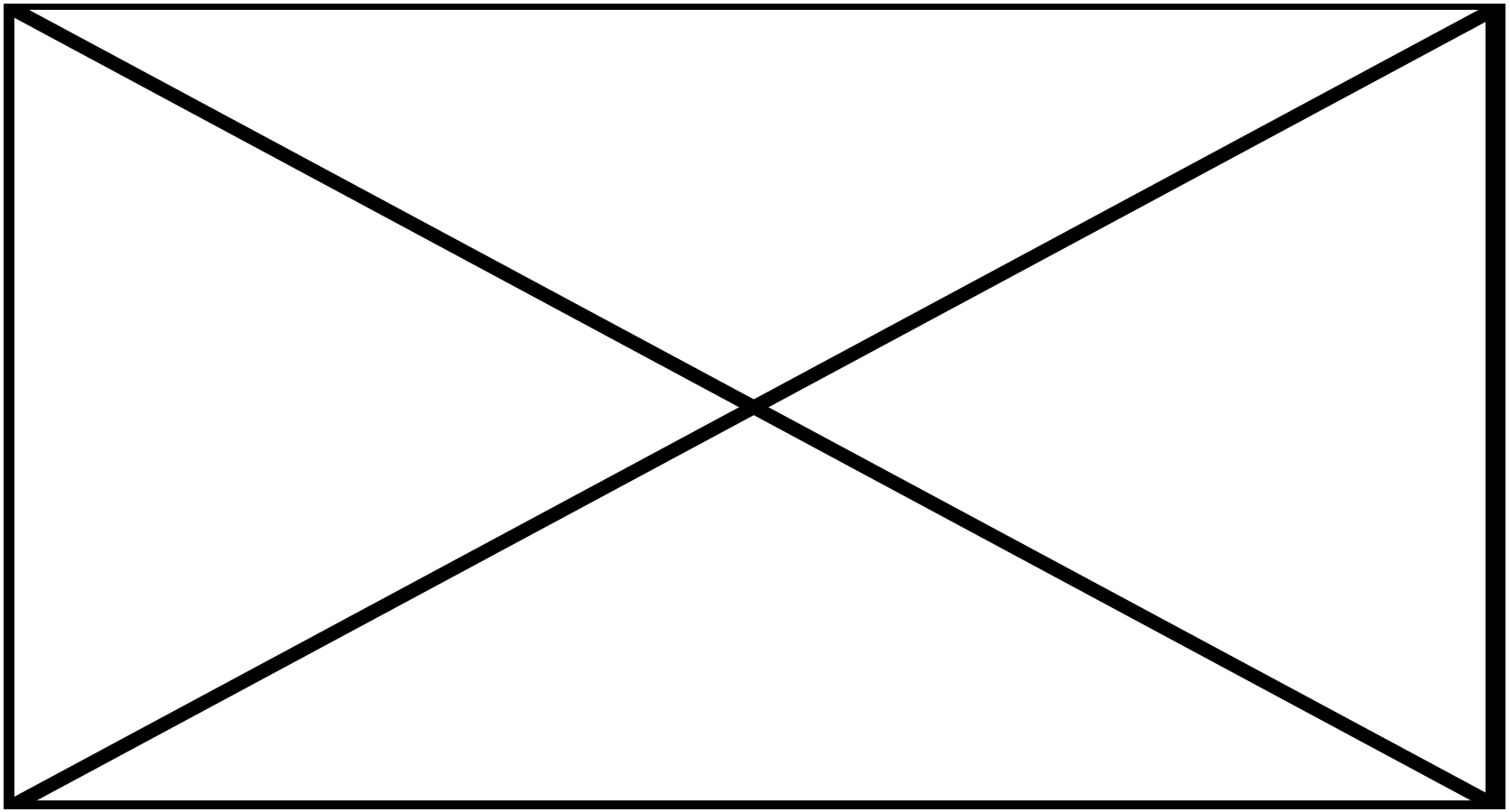
Computer Setup



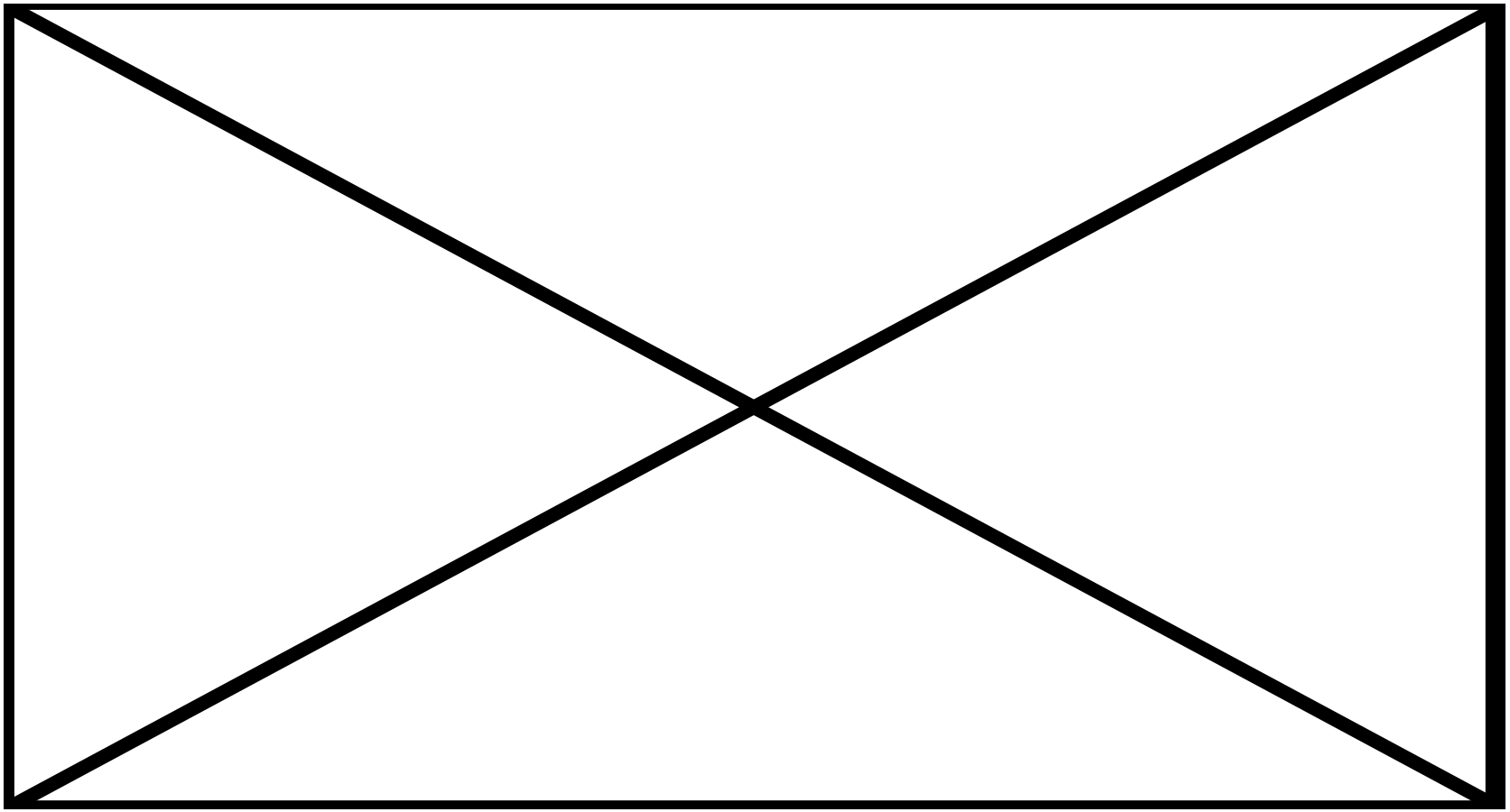
The Load Expression Data Dialog



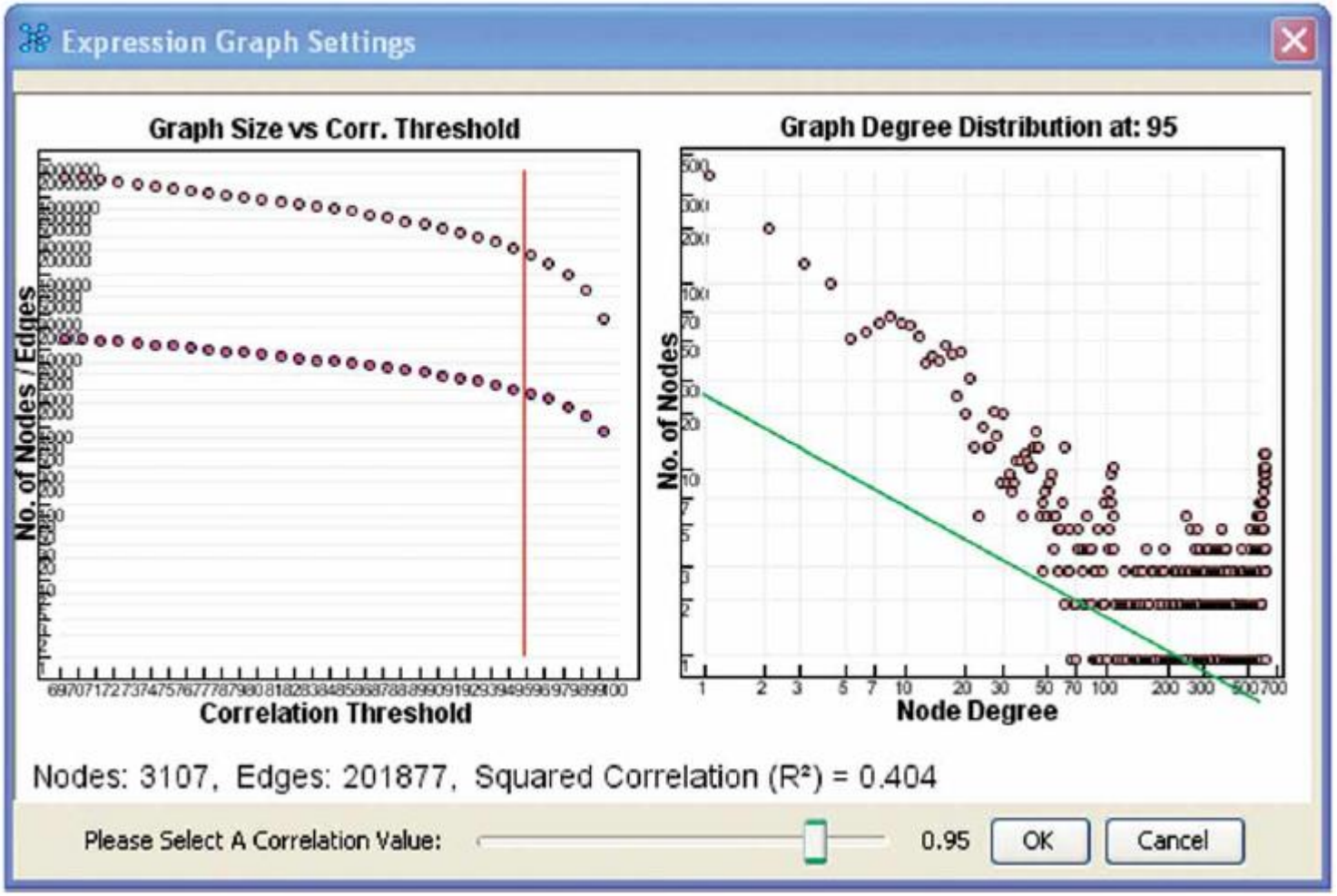
Navigation in 2D and 3D



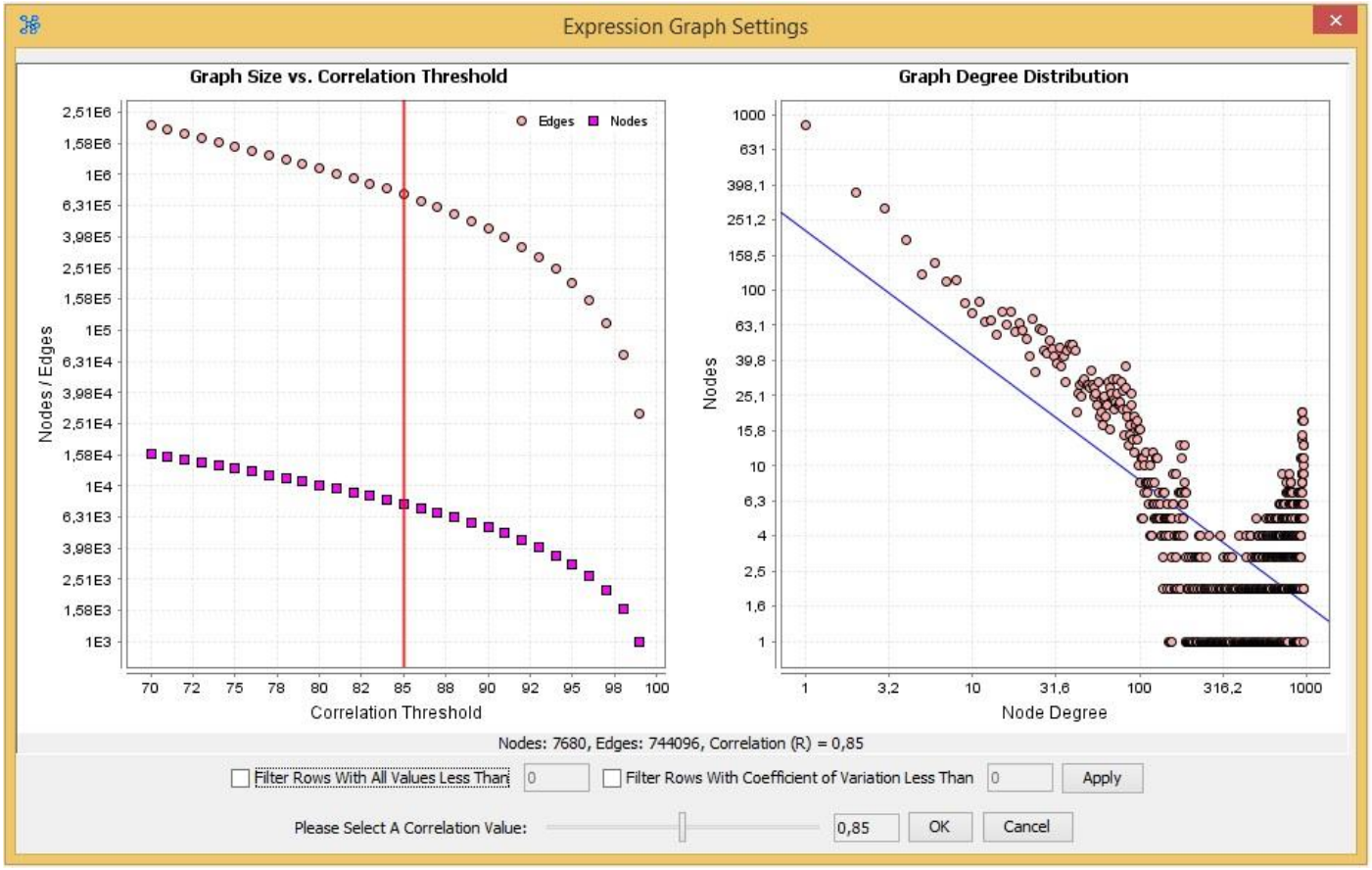
Visual Preferences for Nodes and Edges



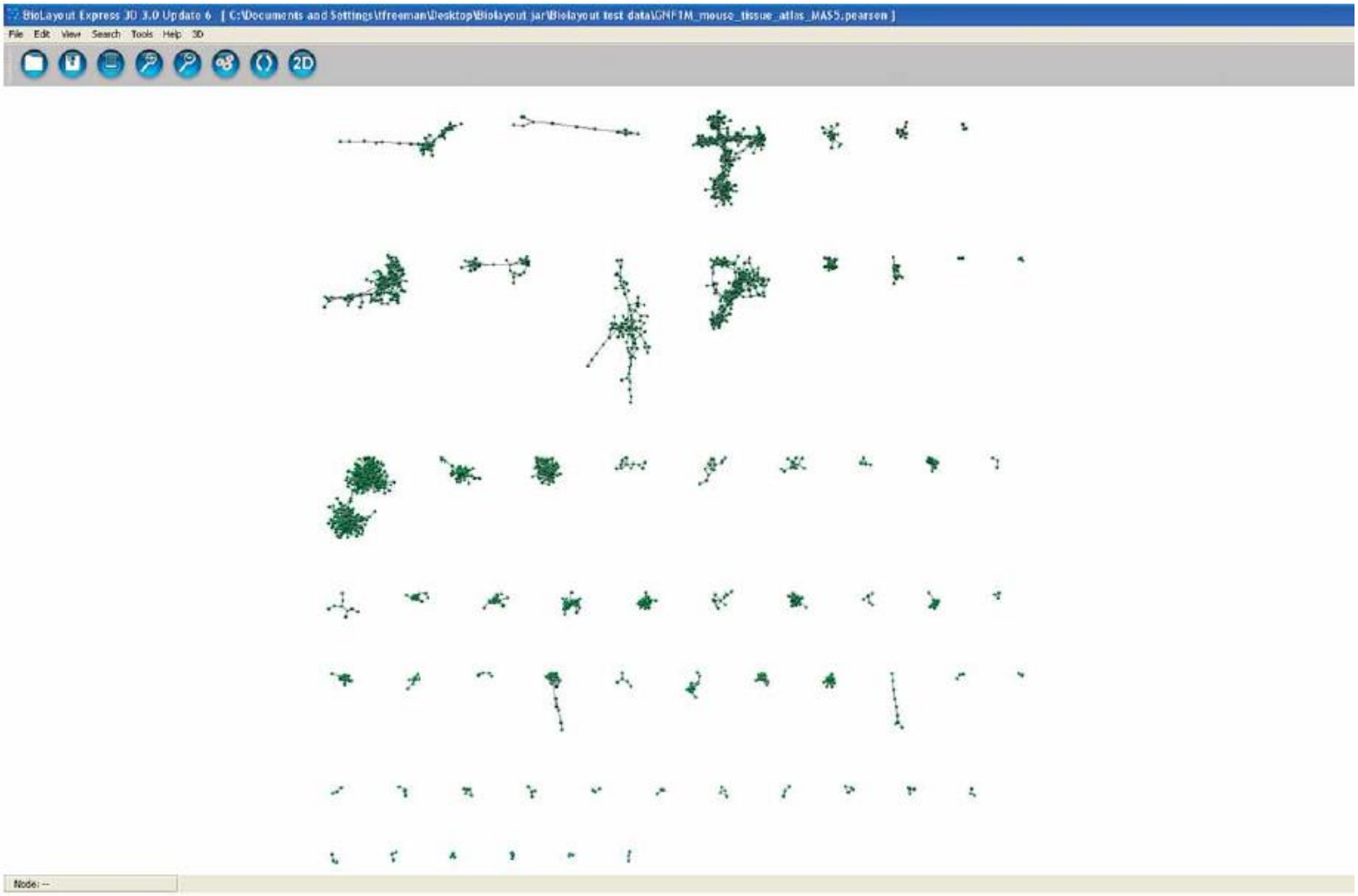
EGS Dialog – CV 0,95



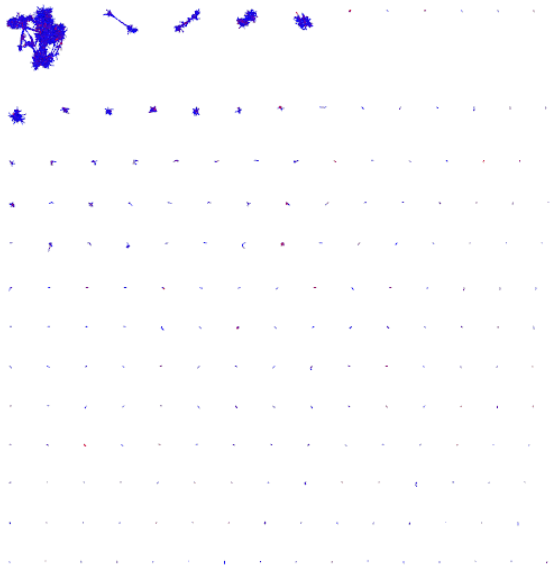
EGS Dialog – CV 0,85



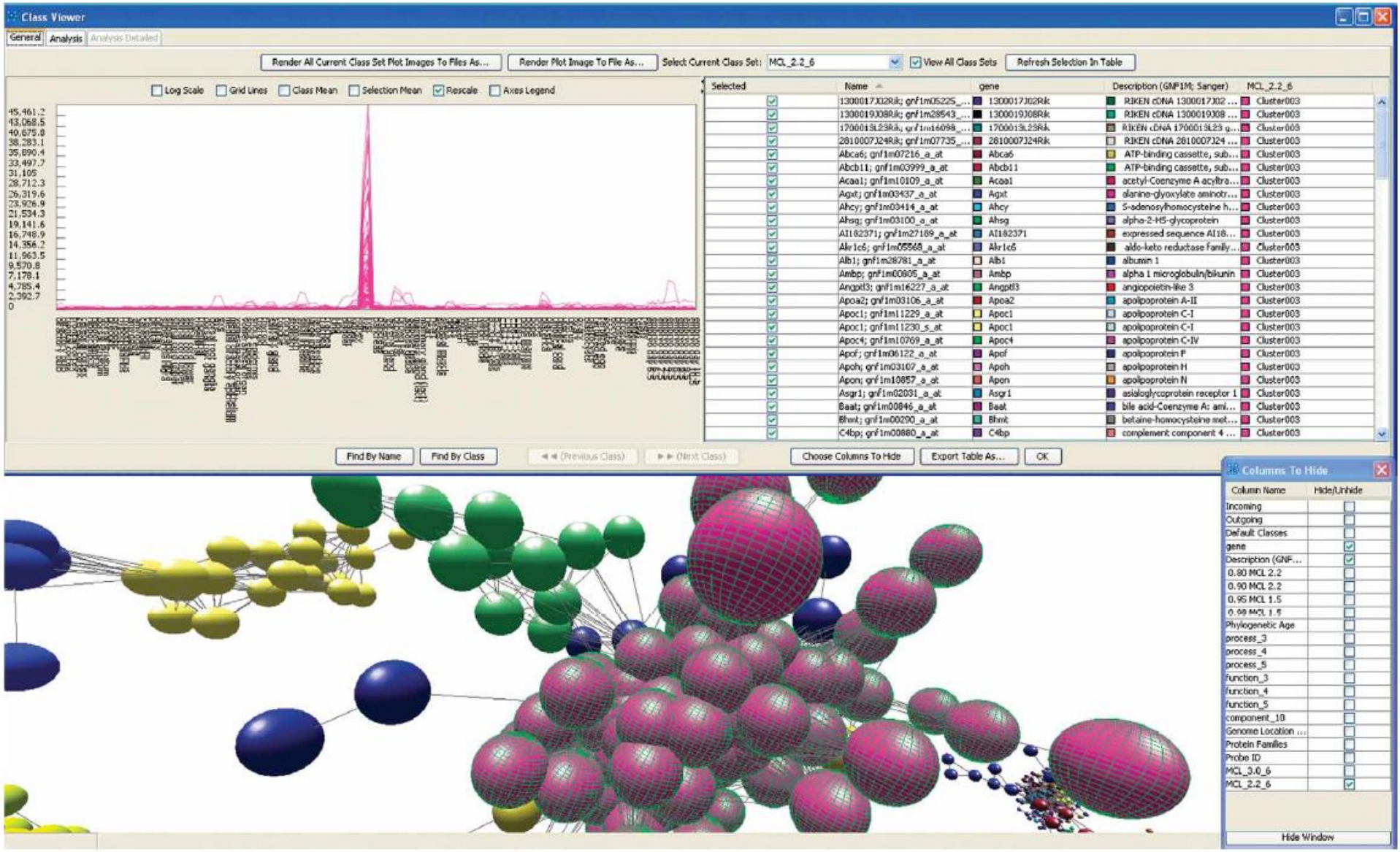
The Main BioLayout *Expressi3D* graph window

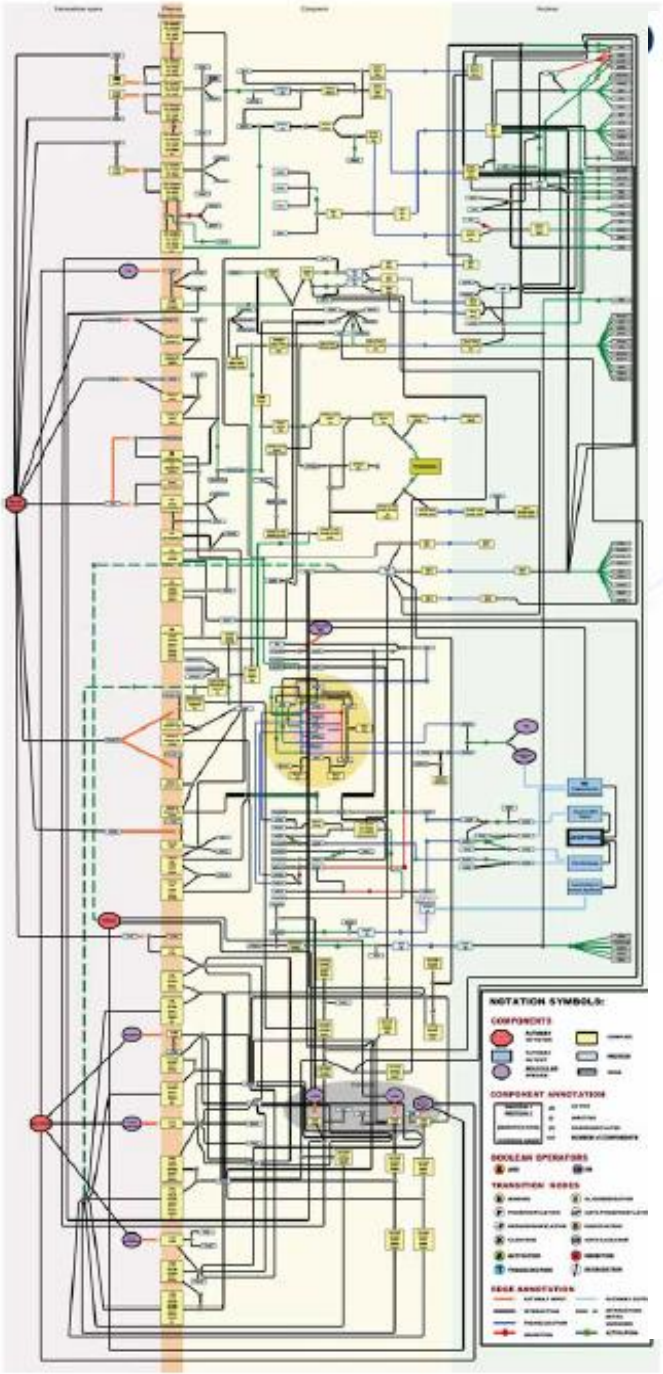


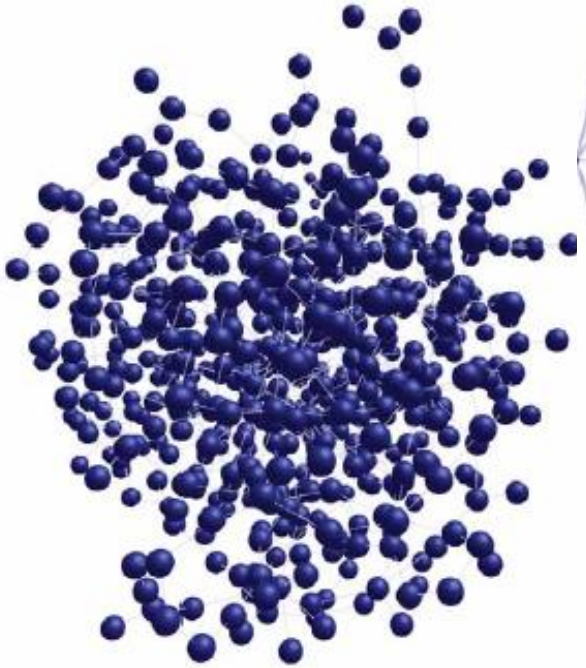
The Main BioLayout *Expressi3D* graph window CV 0,85

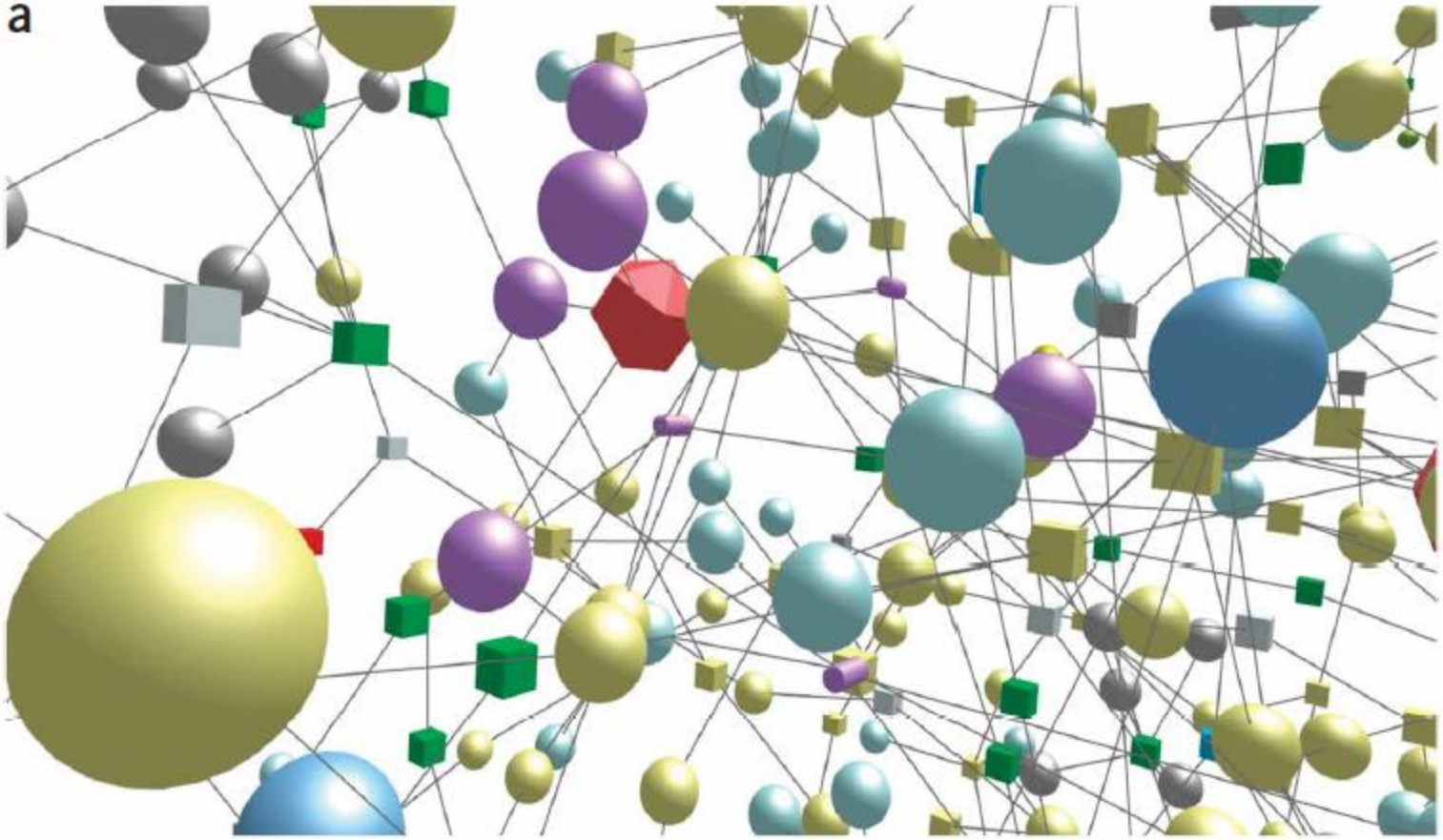


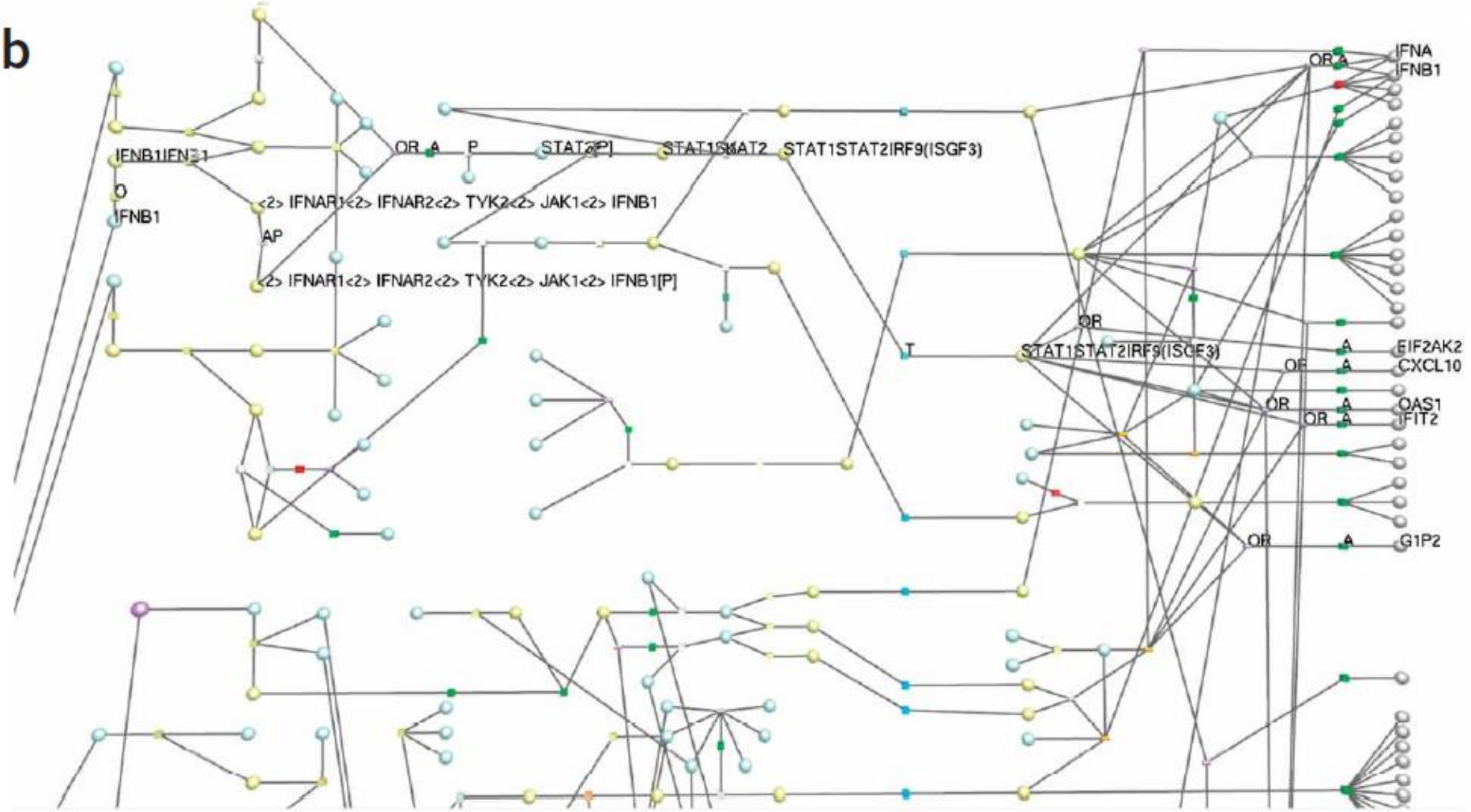
The Class Viewer on the top with main BioLayout *Express3d* graph window below




NOTATION SYMBOLS:
COMPONENTS
COMPONENT ANNOTATION
BOOLEAN OPERATORS
TRANSITION NODES
EDGE ANNOTATION




a


b
IFNA
IFNB1
OR
A
P
STAT2[P]
STAT1STAT2
STAT1STAT2IRF9(ISGF3)
IFNB1IFNB1
O
IFNB1
<2> IFNAR1<2> IFNAR2<2> TYK2<2> JAK1<2> IFNB1
AP
<2> IFNAR1<2> IFNAR2<2> TYK2<2> JAK1<2> IFNB1[P]
T
STAT1STAT2IRF9(ISGF3)
EIF2AK2
CXCL10
OAS1
IFIT2
G1P2

Results

- Pro: Visualisation of huge datasets
- Contra: Could not run the GraphML Input
- Network Navigation: Mostly navigible
- 3-D Visualisation: Useful, if works
- Installation: Very easy on Windows platform
- User: Generally scientists

Results – Platforms & Installation

- Latest Version - BioLayout *Express*^{3D} 3.2
- Windows
  - JAVA SE Runtime Envirenment 7
  - 32 Bit / 64 Bit
- Mac OS X
  - Installer
- Linux
  - JAVA SE Runtime Envirenment 7
  - Installer
  - 32 Bit / 64 Bit
- Android
  - ----------

Literature

- Theocharidis, A., Van Dongen, S., Enright, A.J, Freeman, T.: Network visualisation ans analysis of gene expression data using BioLayout Express 3D, 2009
- http://www.biolayout.org/

Thank you for your Attention