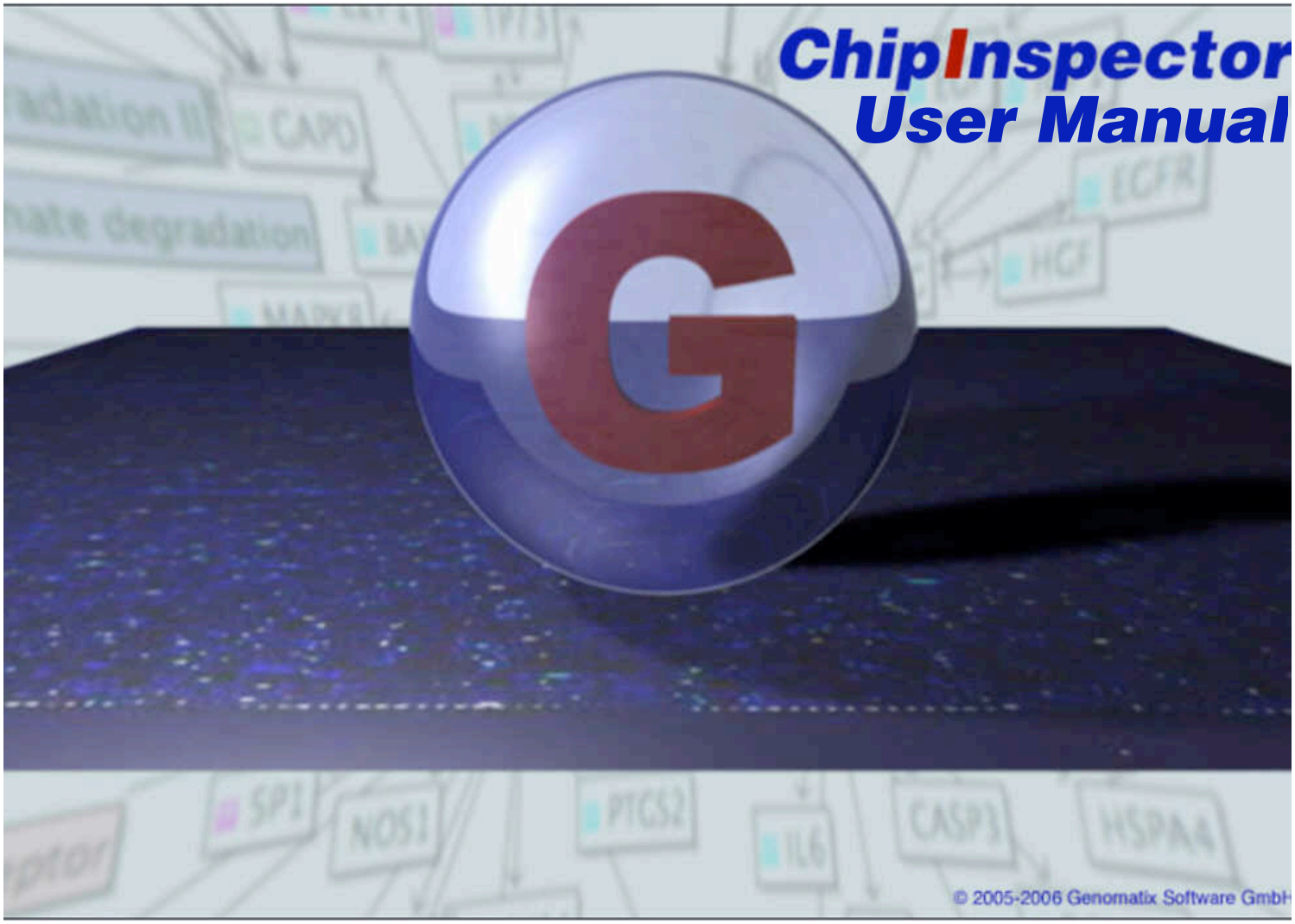
Genomatix
understanding gene regulation



For more information please contact:

Genomatix Software GmbH
Bayerstr. 85a
80335 Munich
Germany

Phone: **+49 89 599766 0**
Fax: **+49 89 599766 55**
Email: **info@genomatix.de**
WWW: **http://www.genomatix.de**

Table of Contents

Introduction to ChipInspector

What is Genomatix ChipInspector?

ChipInspector extracts significant information from the expression level of single probes of Affymetrix GeneChip© microarrays. Sophisticated analysis techniques and state-of-the-art genomic knowledge increase the number of significant features while simultaneously reducing false positive rates by an order of magnitude.
ChipInspector uses the world's largest database of alternative transcripts and promoters to achieve superior signal-to-noise ratios in microarray analysis. It is unique in removing statistical and gene calling errors at the single probe level. This technology provides the basis for unmatched accuracy in significance analysis of microarray data. The resulting lists of significantly regulated genes from the experiment are directly usable as input for Genomatix BiblioSphere PathwayEdition.

Data Analysis

Analysis Background

Genomatix calculates a proprietary annotation for the database ElDorado. ChipInspector data is based on this. On 86 of the 87 currently supported chips, more than 85% of the perfect match probes are used to calculate the statistics. The following tables show the data for each chip.

Affymetrix exon arrays	Number of columns / rows	Perfect match probes (Genomatix optimized)	Transcripts (Genomatix annotated)	Species
Human Exon 1.0 ST	2560	4983374	60194	*H.sapiens*
Human Gene 1.0 ST	1050	737465	52642	*H.sapiens*
Mouse Exon 1.0 ST	2560	4401613	143709	*M.musculus*
Rat Exon 1.0 ST	2560	3729669	21779	*R.norvegicus*

Affymetrix tiling arrays	Number of columns / rows	Perfect match probes (Genomatix optimized)	Analysis is annotation independent	Species
Human Promoter 1.0 R	2166	3967233		*H.sapiens*
Human Tiling 1.0 R Chip 1	2560	t.b.d.		*H.sapiens*
Human Tiling 1.0 R Chip 2	2560	2972683		*H.sapiens*
Human Tiling 1.0 R Chip 3	2560	3067917		*H.sapiens*
Human Tiling 1.0 R Chip 4	2560	3046879		*H.sapiens*
Human Tiling 1.0 R Chip 5	2560	3020832		*H.sapiens*
Human Tiling 1.0 R Chip 6	2560	3036710		*H.sapiens*
Human Tiling 1.0 R Chip 7	2560	t.b.d.		*H.sapiens*
Human Tiling 1.0 R Chip 8	2560	2900753		*H.sapiens*
Human Tiling 1.0 R Chip 9	2560	3018060		*H.sapiens*
Human Tiling 1.0 R Chip 10	2560	3045500		*H.sapiens*
Human Tiling 1.0 R Chip 11	2560	2942269		*H.sapiens*
Human Tiling 1.0 R Chip 12	2560	2942359		*H.sapiens*
Human Tiling 1.0 R Chip 13	2560	2951350		*H.sapiens*
Human Tiling 1.0 R Chip 14	2560	1254212		*H.sapiens*
Human Tiling 2.0 R Chip 1 ver 1	2560	t.b.d.		*H.sapiens*
Human Tiling 2.0 R Chip 2 ver 1	2560	t.b.d.		*H.sapiens*
Human Tiling 2.0 R Chip 3 ver 1	2560	t.b.d.		*H.sapiens*
Human Tiling 2.0 R Chip 4 ver 1	2560	5887585		*H.sapiens*
Human Tiling 2.0 R Chip 5 ver 1	2560	5829518		*H.sapiens*
Affymetrix tiling arrays (cont.)	Number of columns /	Perfect match probes (Genomatix	Analysis is annotation	Species

	rows	optimized)	independent	
Human Tiling 2.0 R Chip 6 ver 1	2560	5860214		*H.sapiens*
Human Tiling 2.0 R Chip 7 ver 1	2560	5746726		*H.sapiens*
Chromosome 21/22 1.0F Chip A	914	312159		*H.sapiens*
Chromosome 21/22 1.0R Chip A	914	313581		*H.sapiens*
Chromosome 21/22 1.0F Chip B	914	295946		*H.sapiens*
Chromosome 21/22 1.0R Chip B	914	296329		*H.sapiens*
Chromosome 21/22 1.0F Chip C	914	320535		*H.sapiens*
Chromosome 21/22 1.0R Chip C	914	320231		*H.sapiens*
Chromosome 21/22 2.0R	2166	2058473		*H.sapiens*
ENCODE01-Forward_4x	1280	2082877		*H.sapiens*
ENCODE01-Reverse_4x	1280	2082877		*H.sapiens*
ENCODE 2.0R	1280	760199		*H.sapiens*
Arabidopsis Tiling 1.0R	2560	2888551		*A.thaliana*
Arabidopsis Tiling 1.0F	2560	2888550		*A.thaliana*
Drosophila Tiling 1.0R	2166	3004387		*M.musculus*
Mouse Promoter 1.0 R	2166	3943515		*M.musculus*
Mouse Tiling 2.0R Chip 1 ver2	2560	5618536		*M.musculus*
Mouse Tiling 2.0R Chip 2 ver2	2560	4952181		*M.musculus*
Mouse Tiling 2.0R Chip 3 ver2	2560	t.b.d.		*M.musculus*
Mouse Tiling 2.0R Chip 4 ver2	2560	5539462		*M.musculus*
Mouse Tiling 2.0R Chip 5 ver2	2560	5481706		*M.musculus*
Mouse Tiling 2.0R Chip 6 ver2	2560	5607960		*M.musculus*
Mouse Tiling 2.0R Chip 7 ver2	2560	5373354		*M.musculus*

Affymetrix expression arrays	Number of columns / rows	Perfect match probes (Genomatix optimized)	Transcripts (Genomatix annotated)	Species
Arabidopsis Genome	534	114960	10635	*A.thaliana*
Arabidopsis ATH1 Genome	712	220039	29840	*A.thaliana*
Bovine Genome	732	199713	16861	*B.taurus*
C.elegans Genome	712	213496	20501	*C.elegans*
Canine Genome Ver 2	732	383133	39164	*C.familiaris*
Chicken Genome	984	315499	15996	*G.gallus*
Drosophila Genome	640	192332	17686	*D.melanogaster*
Drosophila Genome 2.0	732	243002	19174	*D.melanogaster*
Human Genome Focus	448	82235	23040	*H.sapiens*
Human Genome U133 Plus 2.0	1164	525438	61158	*H.sapiens*
Human Genome U133A	712	207689	39876	*H.sapiens*
Human Genome U133A 2.0	732	207689	39876	*H.sapiens*
Human Genome U133B	712	222339	22693	*H.sapiens*
Human Genome U95Av2	640	169901	24755	*H.sapiens*
Human Genome FL (6800)	536	103884	15267	*H.sapiens*
Human X3P	1164	582006	59331	*H.sapiens*
500K_Sty	2560	1610660	12329	*H.sapiens*
500K_Nsp	2560	1612024	10303	*H.sapiens*
Mouse Expression Set 430 A	712	207750	62161	*M.musculus*
Mouse Expression Set 430 B	712	220386	34676	*M.musculus*
Mouse Genome 430 2.0	1002	427307	89895	*M.musculus*
Mouse Genome 430A 2.0	732	207750	62161	*M.musculus*
Murine Genome U74v2 A	640	141087	37949	*M.musculus*
Murine Genome U74v2 B	640	159254	28735	*M.musculus*
Murine Genome U74v2 C	640	101224	14067	*M.musculus*
Rat Expression Set 230 A	602	144141	18676	*R.norvegicus*
Rat Expression Set 230 B	602	141572	10871	*R.norvegicus*
Affymetrix expression arrays (cont.)	Number of columns / rows	Perfect match probes (Genomatix optimized)	Transcripts (Genomatix annotated)	Species
Rat Genome 230 2.0	834	284875	26353	*R.norvegicus*

Rat Genome U34 A	534	100954	8056	*R.norvegicus*
Rat Genome U34 B	534	106552	7533	*R.norvegicus*
Rat Genome U34 C	534	106005	8160	*R.norvegicus*
Rhesus Macaque Genome	1164	590073	40644	*M.mulatta*
Rice Genome	1164	513875	55024	*O.sativa*
Zebrafish Genome	712	152434	8716	*D.rerio*

Other Array providers	Number of identifiers	Perfect match probes (Genomatix optimized)	Transcripts (Genomatix annotated)	Species
Illumina Human Expression BeadChip Version 2	48701	t.b.d.	t.b.d.	*H.sapiens*
Illumina Mouse Expression BeadChip Version 1.1	46643	t.b.d.	t.b.d.	*M.musculus*
Agilent Human Genome	43931	t.b.d.	t.b.d.	*H. sapiens*
Agilent Mouse Genome	41174	t.b.d.	t.b.d.	*M.musculus*
Agilent Human Promoter	476024	t.b.d.		*H. sapiens*
Agilent Mouse Promoter	474380	t.b.d.		*M.musculus*

Each of the Genomatix optimized perfect match probes is analyzed separately for its behavior under the experimental conditions. Statistical analysis is carried out after calculating the base 2 logarithm of the fold-changes between experiment and control.
The user chooses the statistical stringency of the observed signals over the background. The measure of stringency is expressed as False Discovery Rate (FDR). The probes are then mapped onto the relevant transcripts using Genomatix' proprietary genome annotation. The resulting Coverage value shows the number of significantly regulated probes for each transcript.
More than one transcript can be annotated at a locus, therefore many (if not most) probes are mapped to multiple transcripts.
If a time-course/titration experiment was performed, it is possible to cluster the significantly regulated probes according to their profiles over the experimental point. This is done by hierarchical tree clustering (average linking) of the Euclidean distances of the profiles

File Requirements

ChipInspector has a number of requirements for the data files. The files as they are produced in the experiment usually meet all of them, but if the files cannot be analyzed, it might be advisable to check the following list:

1. The data files should not have names with spaces in them (this is possible on Windows systems).
2. The chip type given in the data file must be compliant with the (currently) 87 chips supported (cf. the list of accepted chip types).
3. The files should be stored locally or on a mounted drive. Please be aware that, depending on the file format and your network protocol, remote storage could cause increased time demand.
4. File extension: ChipInspector analyzes files with the .cel or .CEL extensions in case of Affymetrix microarrays. For other chip providers, tab-delimited files are expected and a data import interface is shown.

We recommend a minimum of three replicates per experimental point. It is possible to work with two replicates, but it is not recommended. It is not possible to have less than two replicates per experimental point, because this makes statistics non-utilizable.

Steps in the Statistical Analysis of Microarrays in ChipInspector

Design Correction

Previous annotations of the single oligonucleotide probes are disregarded together with the grouping of the probes in probe sets. Mismatch probes are disregarded. The sequence of each single probe is mapped against the current genome of its target organism and against ElDorado, Genomatix' database of transcripts. Only probes that meet quality criteria such as uniqueness in the genome, mismatch proof and other criteria are used for the analysis. Generally, more than 500.000 single probes (depending on the chip type) fulfill these quality criteria.
As knowledge on the genomic sequences grows and consolidates, these mappings are repeated and ChipInspector is automatically updated with this information.

Normalization

A linear total intensity normalization algorithm is used.

Statistical Analysis

A significance test is performed at the single probe level. This is done basically via a standard permutational T-Test, similar to SAM (Tusher et al., 2001). The exact method depends on the type of the experiment:

One class analysis (Experiment versus Control)

A single sided permutation T-test analysis is performed.

Multiclass analysis

For a timeline analysis or an analysis including multiple stages a multiclass permutational T-test analysis is performed.

Presence/Absence calling

It is also possible to measure expression values relative to the average expression on the chip, e.g. for gene expression values in one specific tissue. In this case a permutational T-test analysis detecting probes which are significant above the experiment average is performed.

Hierarchical Clustering

For the time series/dose response option of analysis, ChipInspector offers a Hierarchical Tree Clustering of the significantly regulated single probes. It is based on Euclidean distance matrix calculations. It doubles as a quality check, when those single probes that describe the same transcript are also found to cluster together.

Mapping the Significantly Regulated Probes to the Transcripts

The probes determined to be significantly regulated in the experiment are subsequently matched with the transcripts that they describe. For each transcript, coverage of regulated probes is thus calculated. Previous experiments have shown that coverage of 3 or more probes per transcript provides sufficiently stringent evidence of the transcript being regulated in the experiment.

Exporting the Resulting Transcript Lists

The list of regulated transcripts is saved in MS Excel format and can be directly uploaded into Genomatix' BiblioSphere Pathway Edition, where they can be displayed as gene networks together with their signal values.

Technical Requirements

Memory Requirements

The following chapter explains the technical requirements to install the ChipInspector client application on your computer.

The table shows the maximum possible number of CEL-files in a control/treatment setup (exhaustive combinations) for one single analysis run in relation to different computer configurations. This table is intended to give the user a perception of possible setups with the current version of ChipInspector, depending on the available main memory (RAM in Gigybyte GB).

Chip Type	Gene Chip	Promoter Tiling		Exon Tiling	
Chip Size	712 x 712**	2166 x 2166		2560 x 2560	
Mapping Type	Annotation	Annotation	Position	Annotation	Position
1 GB RAM 32bit O/S	10 x 10	3 x 3*	2 x 2*	3 x 3*	Not possible***
2 GB RAM 32bit O/S	10 x 10	4 x 4*	3 x 3	4 x 4*	Not possible***
4 GB RAM 64bit O/S	30 x 30	6 x 6	4 x 4*	5 x 5	4 x 4*

*) The memory that is allocated to the program may need to be increased from the default settings. It is best to first test how much memory to allocate: Go to ~GenomatixApplications/apps/chipinspector/conf and edit the file chipinspector.bat. Change the parameter –Xmx895m to e.g. –Xmx1400m.
Then save the file and start ChipInspector by double-clicking on the file chipinspector.bat. If it works, then the program starts with an additional console window, but otherwise works normally. If the parameter is false, then the program does not start at all. In this way, the allocated memory can be maximized by trial and error. Theoretical limitations with a 32bit O/S are around 1600m depening on the individual configuration of the computer.

**) different GeneChip types have different sizes

***) this problem will be addressed in future versions

Operating Systems

The application runs on the following operating systems:

Windows systems:

- Windows 98, SE, 2000, ME, XP
- 5 GB hard disc space
- Minimum of 1 GB RAM required (*)
- 1 GHz processor speed

Macintosh systems:

- At least MacOS X 10.3
- 5 GB hard disc space
- Minimum of 1 GB RAM required (*)

- 1 GHz processor speed

Linux/Unix systems:

- SuSE Linux 8.0 or higher, or equivalent version of other distributors
- 5 GB hard disc space
- Minimum of 1 GB RAM required (*)
- 1 GHz processor speed

(*) Although ChipInspector will run on the listed hardware, it may not complete a position-based analysis or the analysis of large chip sets. For these kinds of analysis 2 GB RAM is needed. It is essential that ChipInspector be reinstalled after upgrading RAM.

If you do not have any of these operating systems, or if you are not sure about your operating system, please contact the Genomatix customer support (support@genomatix.de).

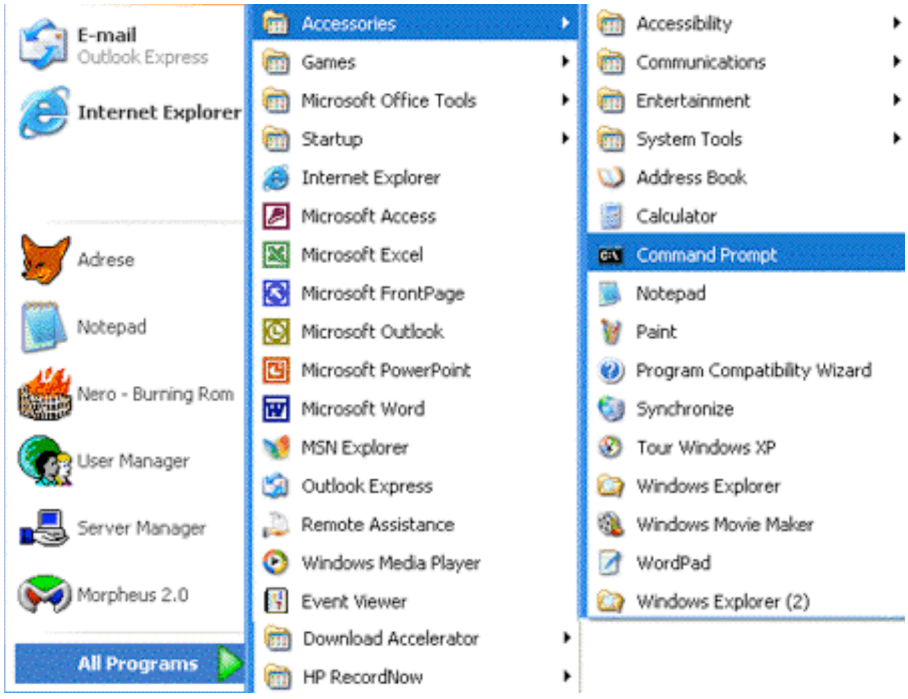
Java Runtime Environment

In order to run the ChipInspector application, you will need **Java 1.5.0** or higher.

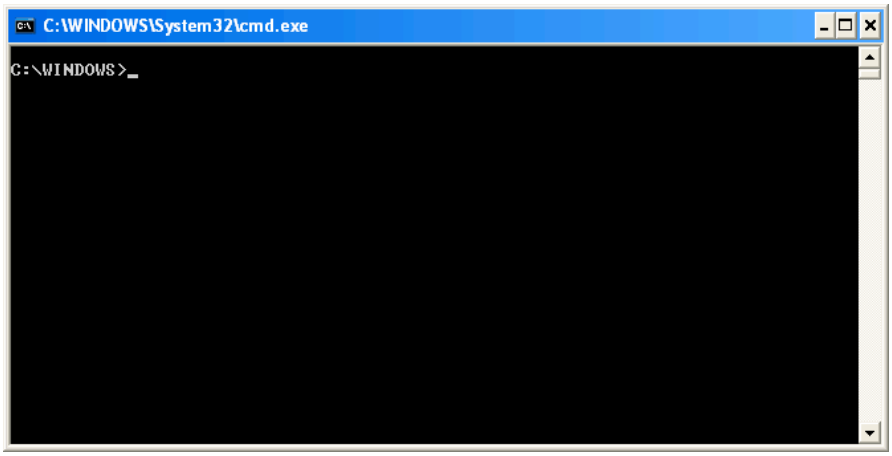
To test if you have an appropriate Java version already installed on your system, type "**java –version**" on command line.

Here is an example for windows users how to check the installed java version:

Click on Start/All Programs/Accessories/Command Prompt (see screenshot below).



A command window will pop up:



Type in ***java –version*** and press Enter.

If Java is installed, you will get an output like:

```
C:\>java -version
java version "1.5.0_06"
Java(TM) 2 Runtime Environment, Standard Edition (build 1.5.0_06-b05)
Java HotSpot(TM) Client VM (build 1.5.0_06-b05, mixed mode, sharing)
```

If Java is **not** yet installed on your computer, or if you have a Java version older than 1.5.0, please follow the link http://www.java.com/ to download and install the newest version of Java (at least version 1.5.0).

Browser

ChipInspector is a Java program which can be run without an internet browser but provides links to the Genomatix tools which use the W3C standard SVG for graphical output.

To fully explore the interactive SVG output of Genomatix tools (currently available for ElDorado, Gene2Promoter, MatInspector, and FrameWorker), the **Adobe SVG Viewer 3.0 is necessary**. Older versions will not work, as several v3.0 specific features are used for the graphics.

If the graphics are not displayed properly, please follow the links below to get more information about installation of SVG for your computer system:

- Windows: Adobe SVG Viewer 3.0 release notes for Windows (PDF)
- Macintosh Adobe SVG Viewer 3.0 release notes for the Mac (PDF)
- Linux/Unix beta versions of Adobe SVG Viewer 3.0 available for the **RedHat Linux 7.1** and **Solaris 8** systems.

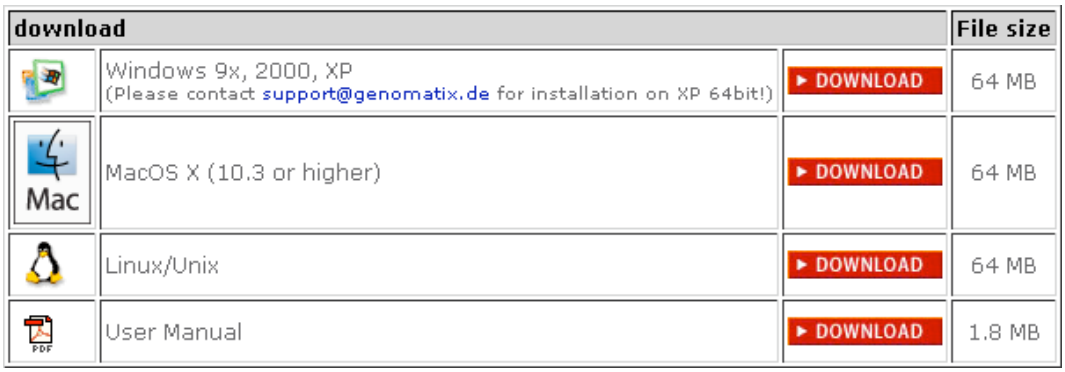
Installation and Configuration of ChipInspector

ChipInspector is a JAVA program which must be installed locally on your computer. Please proceed for download and installation as follows.

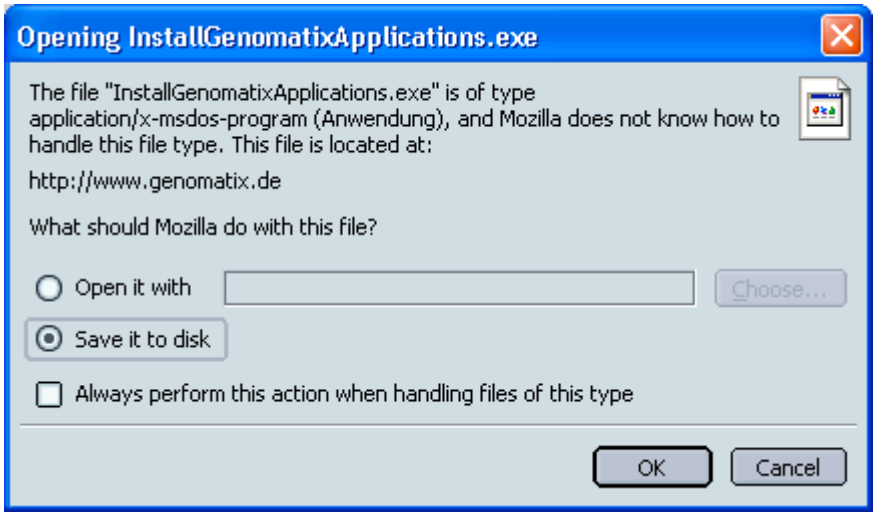
Download

To download ChipInspector, please follow the following steps:

1. Create a folder on you hard disk where you want to store the installer
2. Switch to http://www.genomatix.de/products/ChipInspector/ChipInspector6.html
3. Choose your operating system from the download
4. Click on the download button next to your operating system



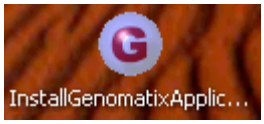
Clicking on the download-icon will result in the following screen:



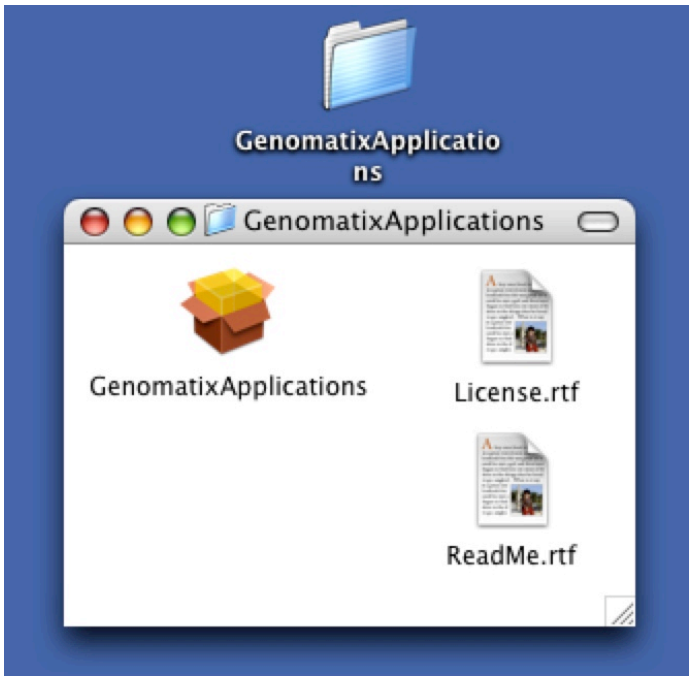
Choose the option “save to disk” and click “ok”

A window will show up, where you can choose a folder to save the file. Choose the folder where you would like to save the installer and press ok.

If the installer is successfully downloaded, windows users should see the following icon with the subtitle “InstallGenomatixApplication.exe”



Mac users will find a folder named "GenomatixApplications" on their desktop or in their designated download folder.It contains an installer package, a ReadMe and the license file. Double clicking the "GenomatixApplications" installer package will start the installation of the software.



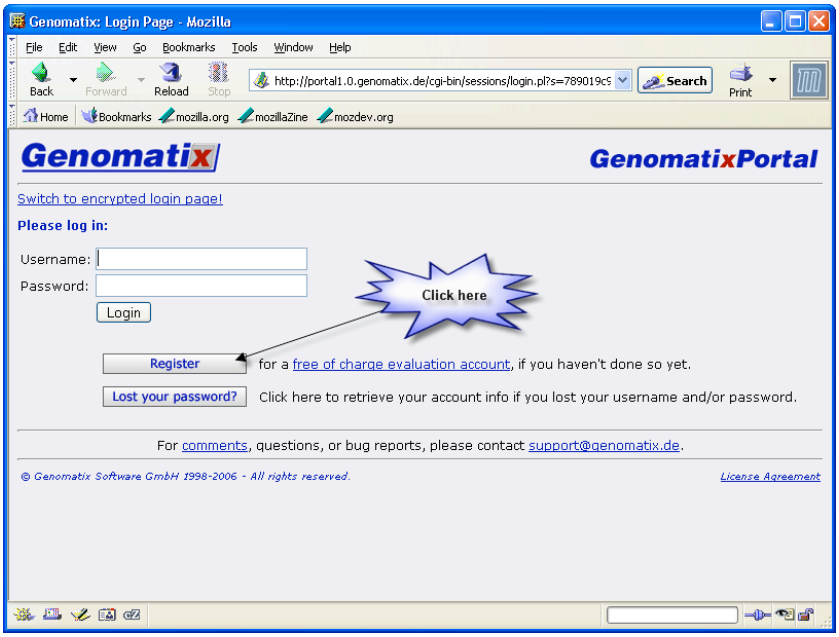
Get Login and Password

To apply the ChipInspector application you need a login and a password. Registration is **free of charge**. An e-mail with your personal username and password will be sent to you right away.

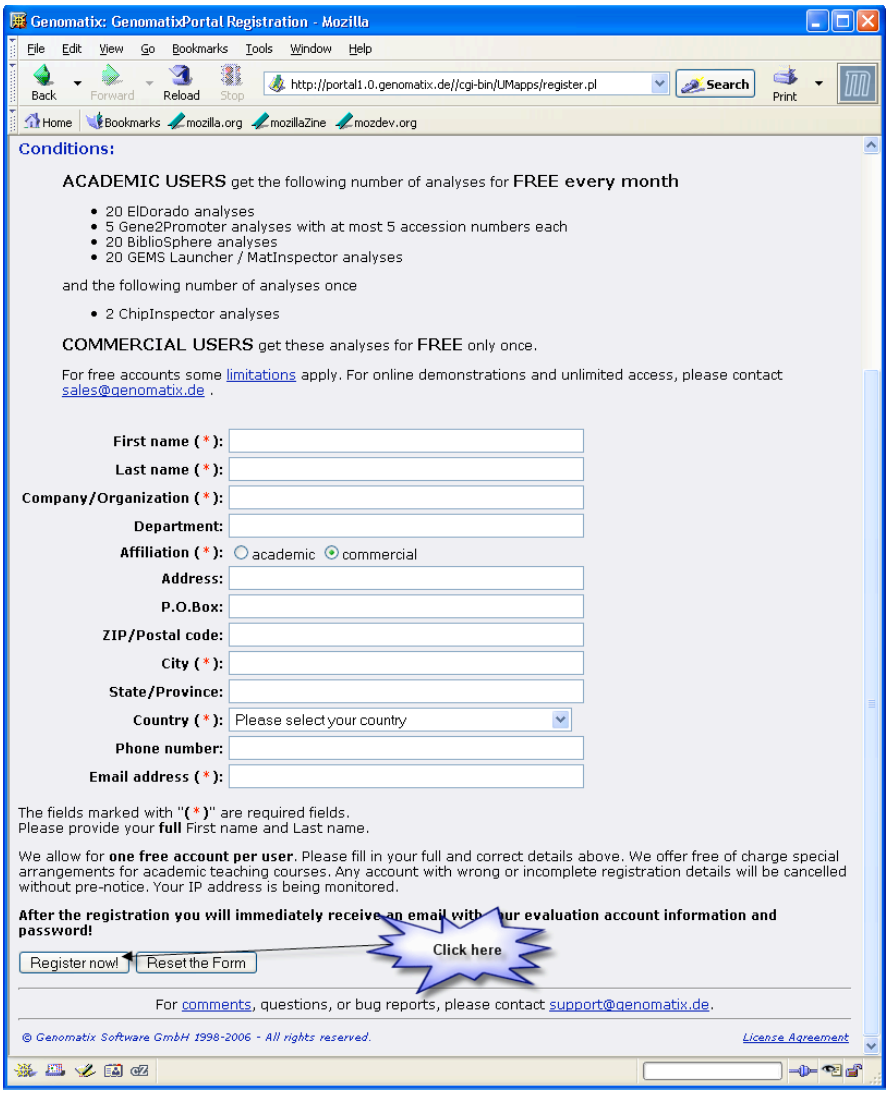
Registration

Open your internet browser and switch to www.genomatix.de.
Click on “Login” in the left frame of the webpage.

If you do not have an account yet, please click on “Register”.



Fill in the form – please enter your e-mail correctly.



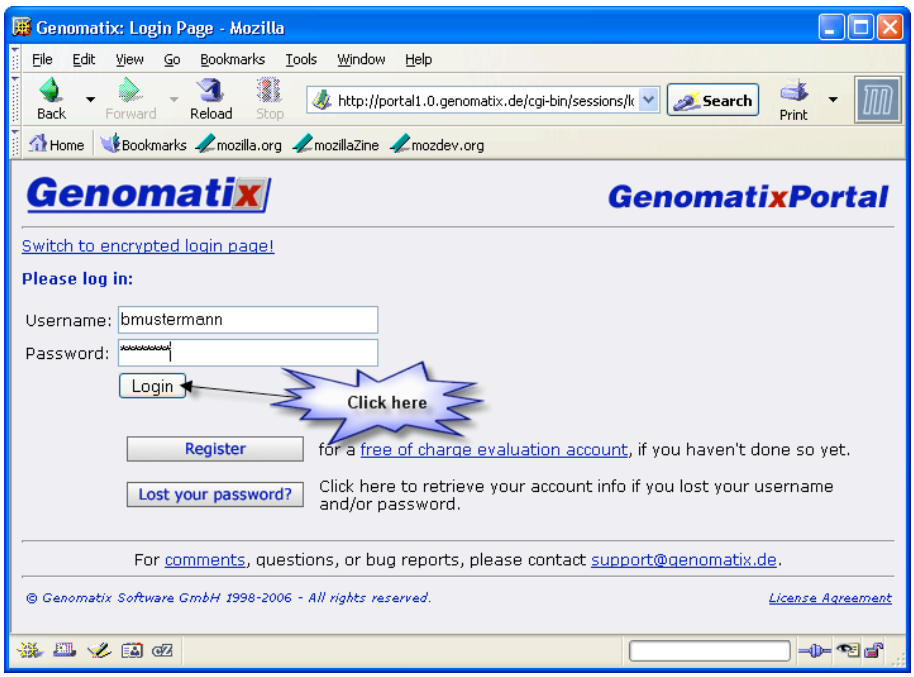
Check your e-mail. A mail with your login data should be sent to you right away.

```
Dear Bernd Mustermann,


your registration for a Genomatix evaluation account was successful!
Here are your account details:

      login name: bmustermann
        password: nUDbjPnW

Please note that username AND password are case-sensitive.

You will find the GenomatixSuite at
   http://portal1.0.genomatix.de/cgi-bin/./eldorado/main.pl
You can change your password using the "change password" link
at the top of the program page or on your "Personal profile" page.

You can also subscribe to unlimited MatInspector access on your
"Personal profile" page.

If you should run into trouble logging onto our server, please first
read
http://portal1.0.genomatix.de/online_help/help/techfaq.html

After that do not hesitate to contact us in case of any problems or
questions.


Best regards,

Your Genomatix support team


~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
For comments or suggestions please contact
          Genomatix Software
      at support@genomatix.de !
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

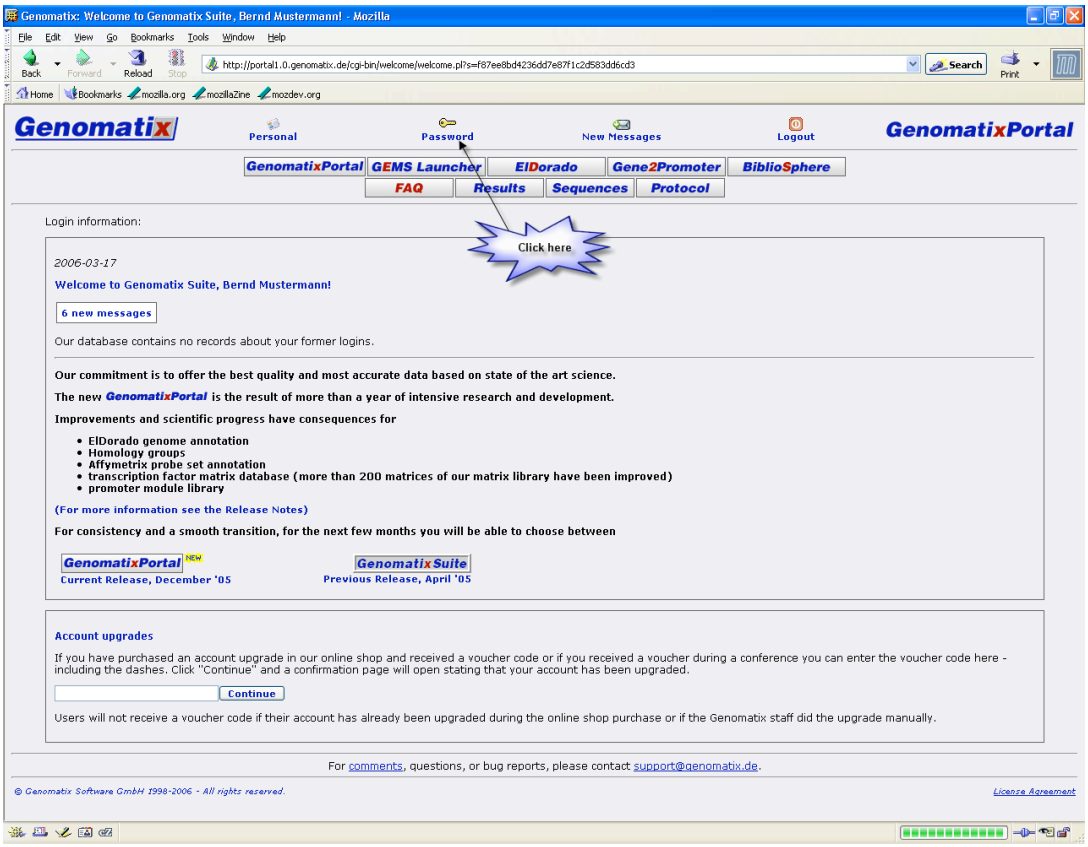
The login and password is not only valid for ChipInspector but for all Genomatix products.

Change Password

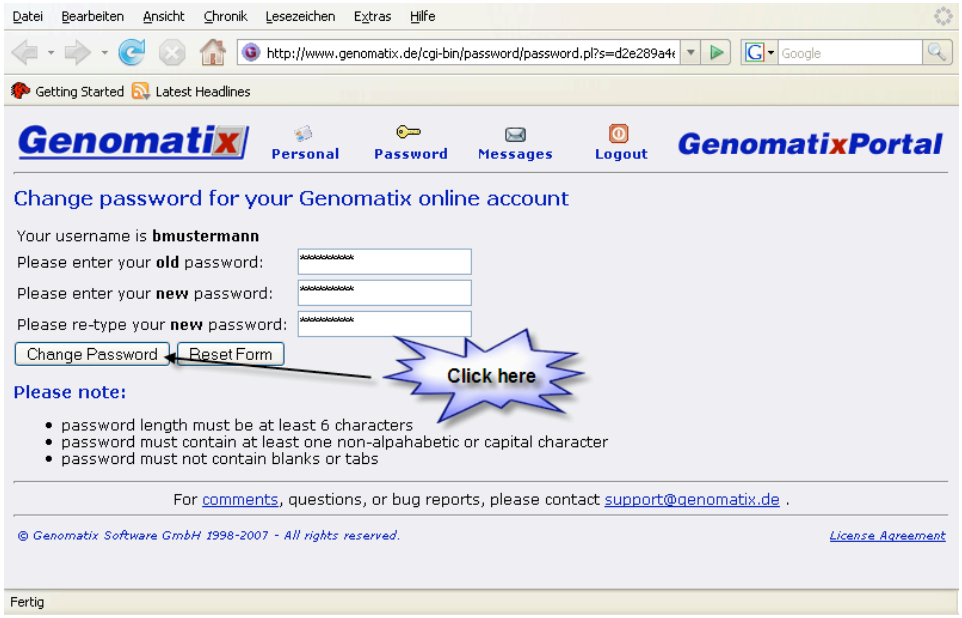
Open your internet Browser and switch to www.genomatix.de.
Click on “Login” in the upper right corner of the webpage (see above)
Enter your login and password which was sent to you via e-mail.



After login you will see the following page. Click on “Password”.



Fill in the form and click on “Change Password” to change your password.



Password Policy

Genomatix’s password policy requires all passwords to be at least 6 characters long and must contain at least one non-alphabetic or capital character. No blanks or tabs are allowed.

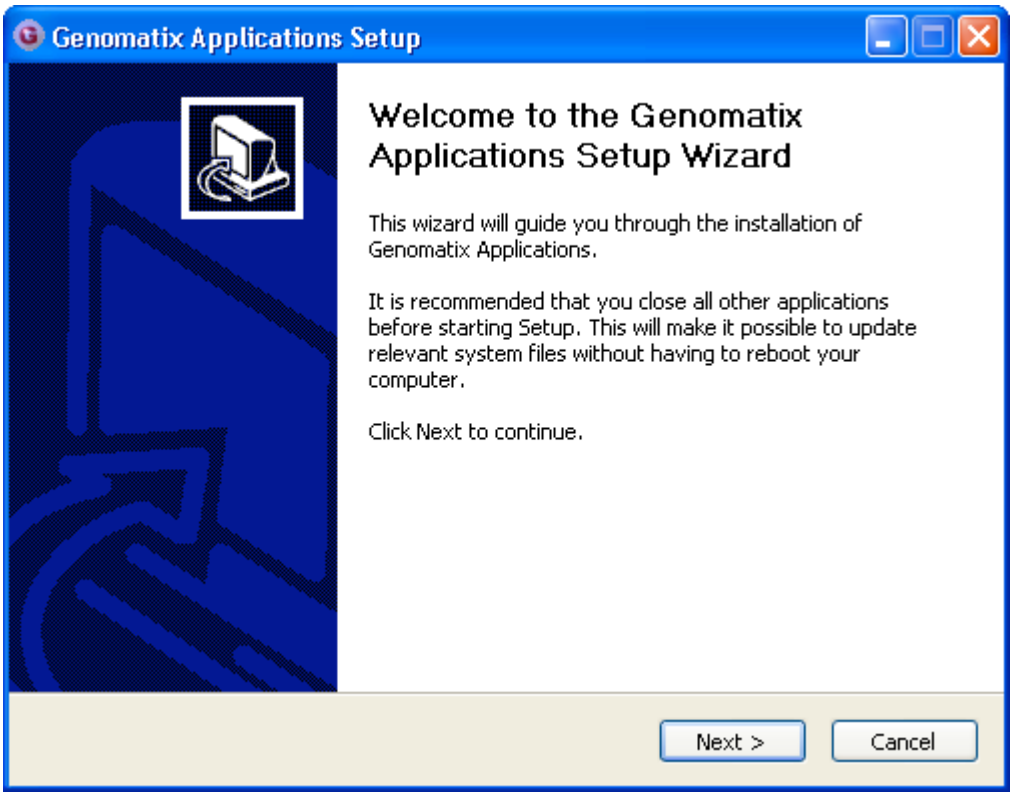
Installation

Switch to the folder on your hard disk where the installer was saved. Execute the installer (see below) and follow the instructions. By default, the installer will install both ChipInspector and BiblioSphere PathwayEdition.

install		
	Windows 9x, 2000, XP	doubleclick on the Genomatix install icon in the folder (requires administrator privileges!)
Mac	MacOS X (10.3 or higher)	run the GenomatixApplications installer from the downloaded disk image/folder (requires administrator privileges!)
	Linux/Unix	type **"java -jar InstallGenomatixApplications.jar"**

Please note that the Genomatix licensing model for ChipInspector is a single-user floating license. This means that you may install the program on any number of machines, however not run several instances of the program at the same time. If a second instance of ChipInspector is started while another instance is running, the user is given the choice of ending the concurring session. This can lead to data loss on the first instance if the analysis results have not been stored yet.

If you run a windows system, the following screen will pop up:

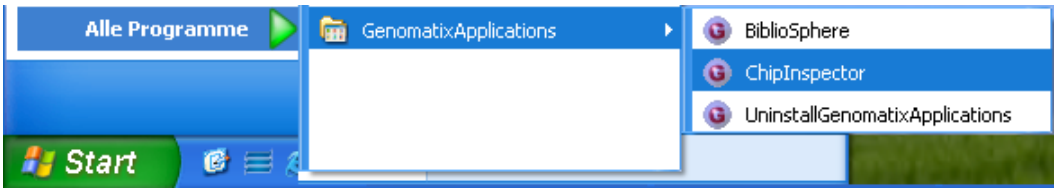


Click “Next >” and follow the instructions.

After ChipInspector is installed successfully, you can start the application in different ways:

1. Start ChipInspector from the program group

After successful installation, windows users should have a new Program Group “Genomatix Applications” with an executable “ChipInspector”. Click “Start”, ”All Programs”, ”GenomatixApplications”, ” ChipInspector”.



2. Start ChipInspector from desktop

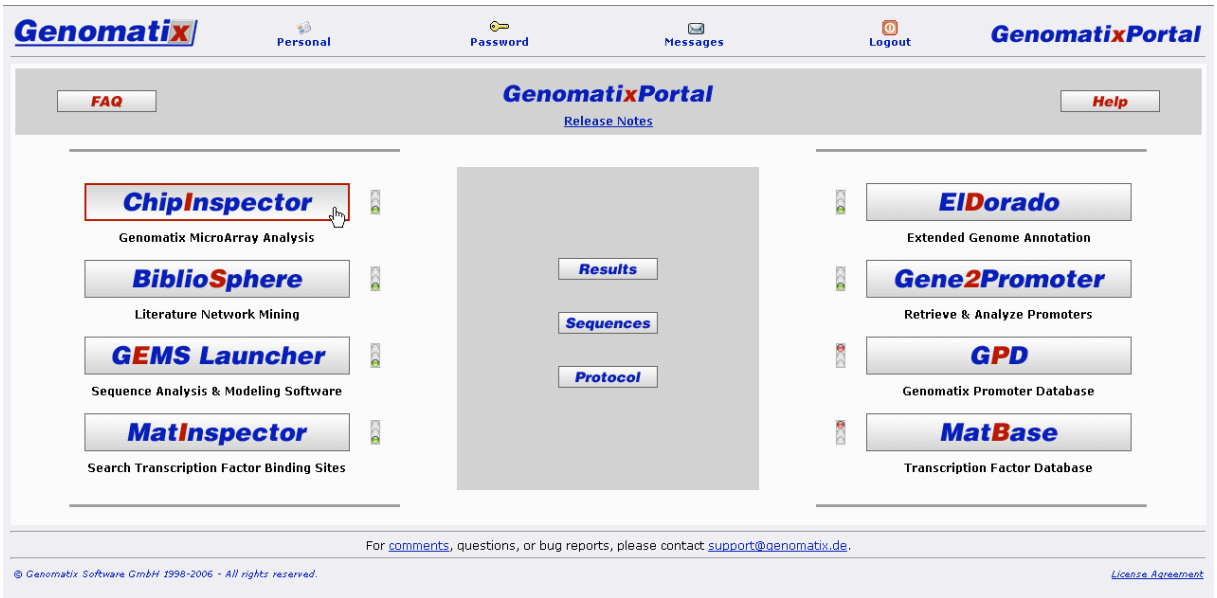
After installation you should find an Icon on your desktop:

A double click on the icon will launch the ChipInspector application

3. Start ChipInspector per batch file (MS Windows only)

On Windows systems, if ChipInspector does not start when you double click the desktop icon, you can use a batch file that you find the in a subdirectory of your Genomatix installation directory. The default location is C:\Program Files\GenomatixApplications\apps\chipinspector\conf\chipinspector.bat. Double click on the file in your windows explorer or, in the Windows start menu, choose “Execute…”, type in the complete file name including the path and click OK.

4. Start ChipInspector from the Genomatix Portal (see below)



Configuration of ChipInspector

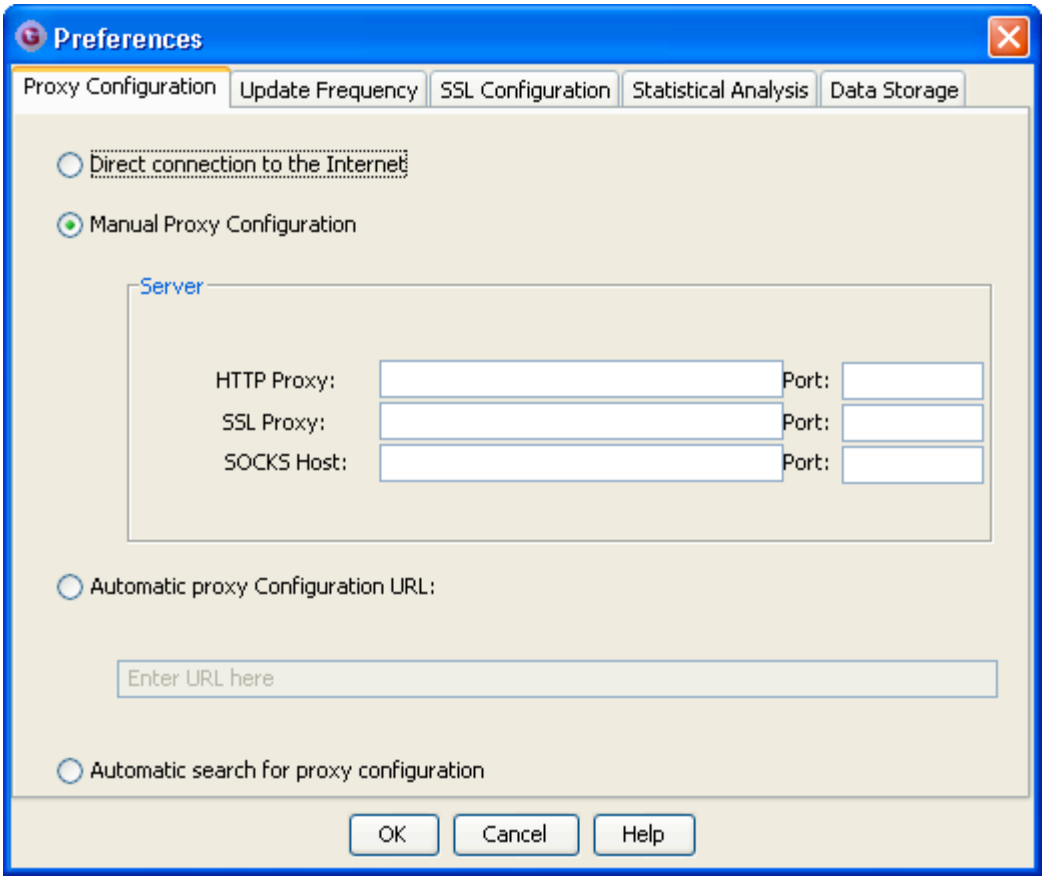
Before you start working with ChipInspector you should configure the ChipInspector concerning

- Proxy configuration (for internet access)
- Security configuration (for secure information transfer over the internet)
- Application update (to get the latest version of ChipInspector online)

ChipInspector offers a form for configuration which can be accessed as follows:

In the ChipInspector application, go to menu **"Extras"** and select "**Preferences**" to launch the preferences configuration dialog

You will get the following dialog which consists of three forms for the different configurations:



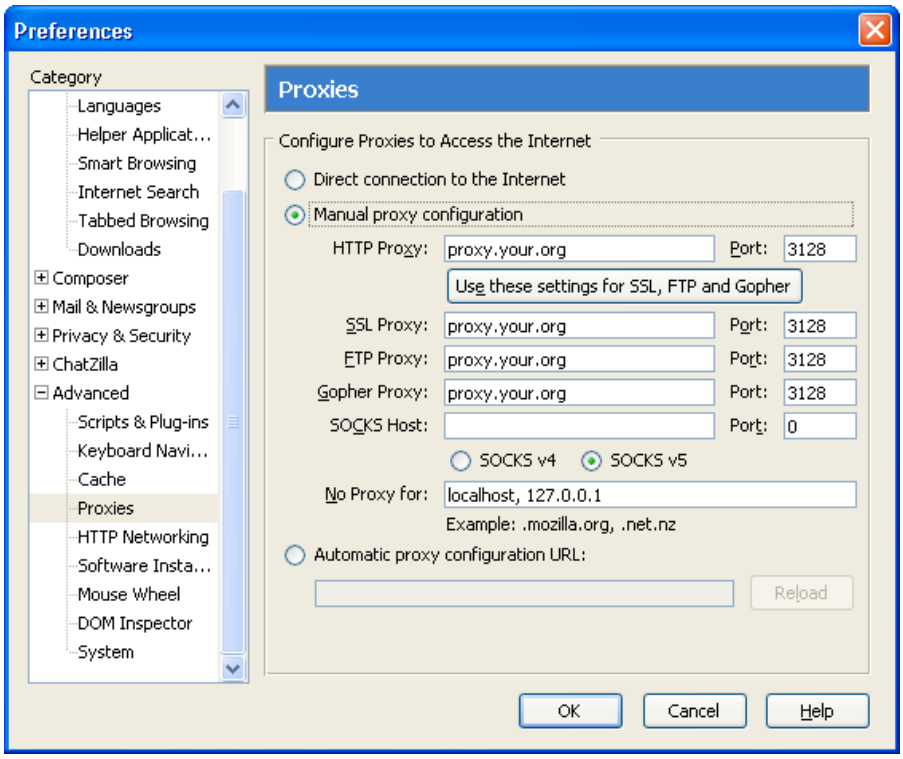
Proxy Configuration

Many companies and institutions use proxies and firewalls for secure and fast access to the Web. Thus you need to configure the ChipInspector application to get through your proxy or firewall.

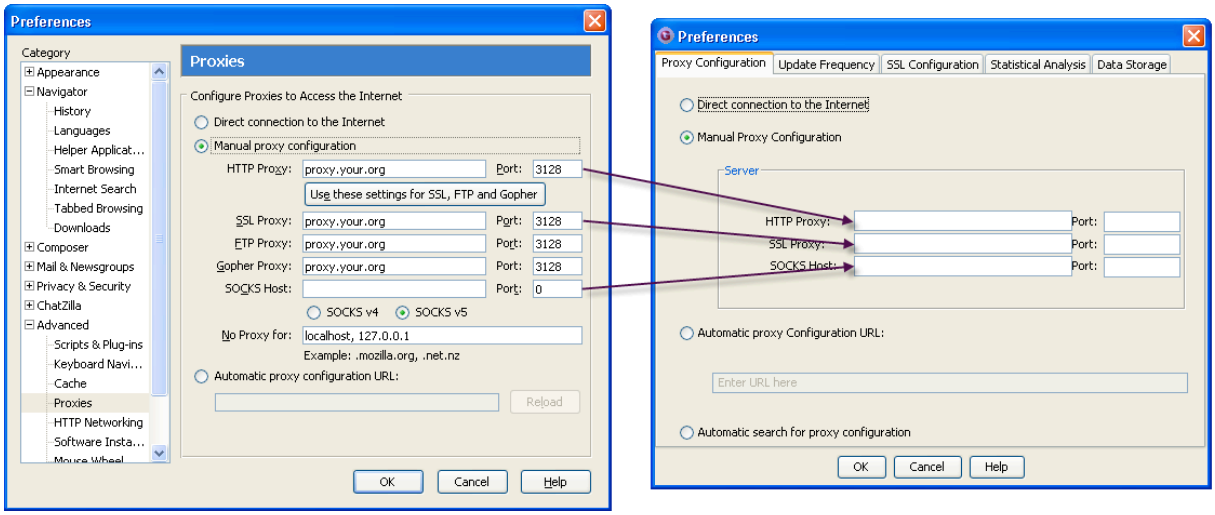
Please proceed as follows:

Get the **proxy settings** from your internet browser.
If you use internet explorer: Go to: Tools->Internet Options->Connections->**LAN settings**
If you use Netscape or Mozilla: Go to: Edit->Preferences->Advanced->**Proxies**

Below you see an example for the Mozilla browser



Configure the settings according to the configuration of your browser and press **"ok"**. Below you see an example for manual proxy configuration.



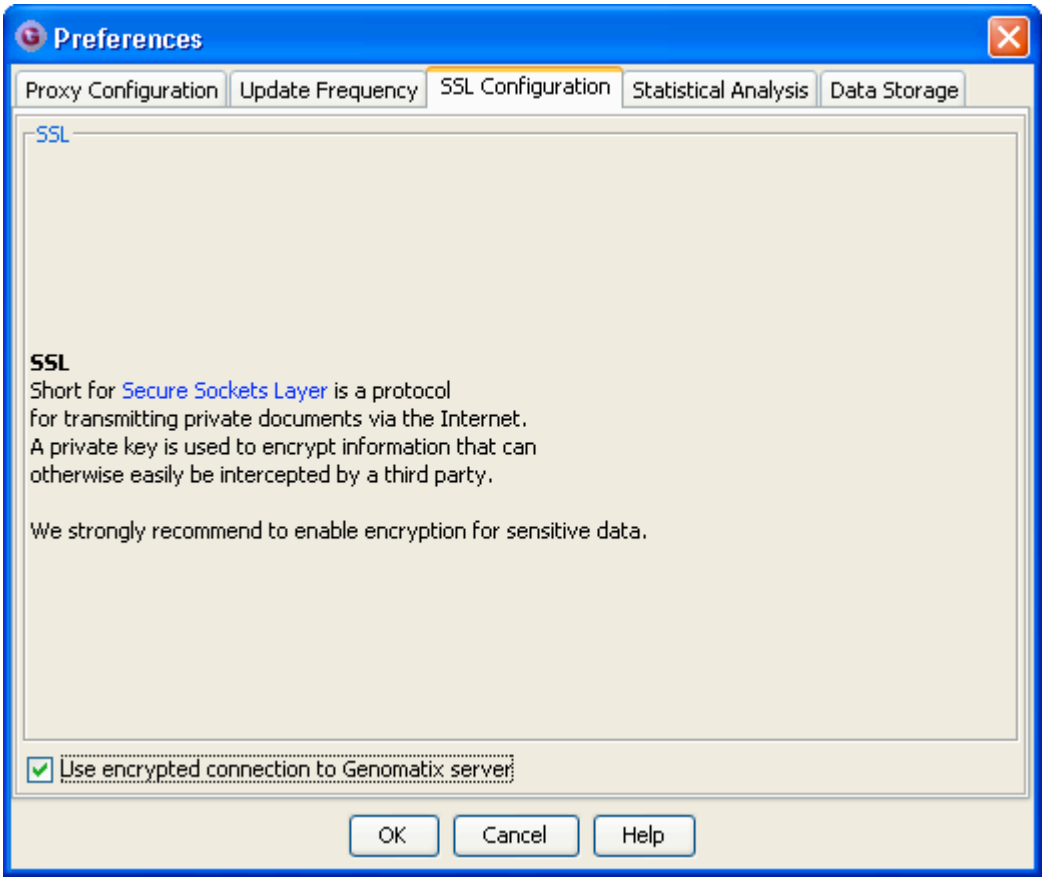
SSL Configuration

ChipInspector allows for encrypted communication with the server via internet via Secure Socket Layer (SSL). If you would like to use the encrypted protocol proceed as follows:

Start ChipInspector (see above)

Go to menu **"Extras"** and select "**Preferences**" to launch a preferences dialog for proxy configuration

Click on “SSL Configuration”:



Check the box next to “Use encrypted connection to Genomatix server” and then click “ok”.

Check for Updates

Periodically Genomatix provides important ChipInspector updates. The Genomatix Update Service helps you to keep your application current.

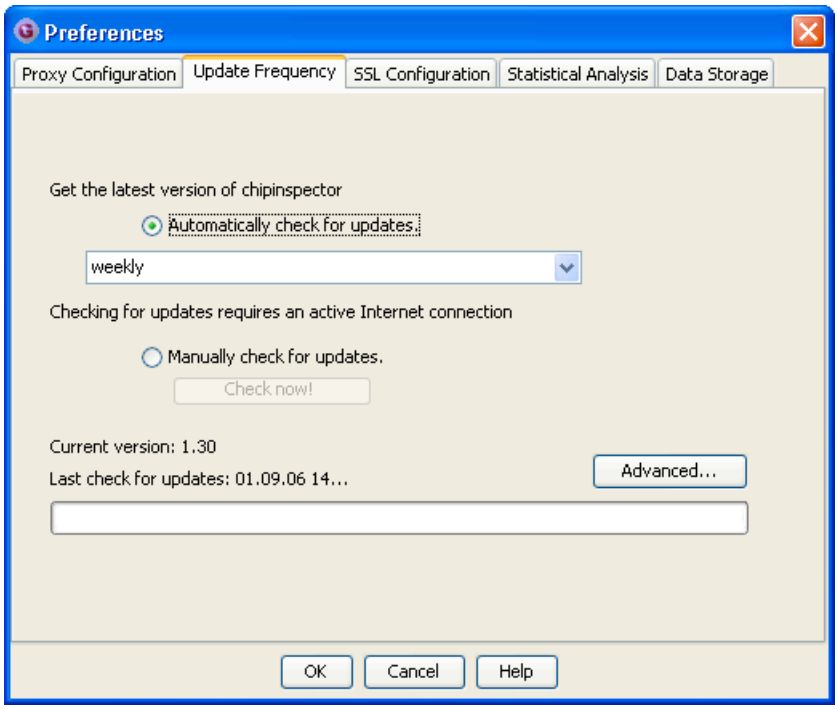
Click on “Update Frequency” in the Configuration dialog.

There are two modes for update: “Automatically check for updates” and “Manually check for updates”:

Turning on Automatic Update Notification

The Automatic Update Service checks for updates at regular intervals. Any time a product update becomes available, you receive a notification. Once you receive the notification, the Update Service guides you toward the download and installation of the updates you need. The Automatic Update Service is activated as follows:

Select **"automatically check for updates"** and choose your preferred update frequency (choices are **"daily“, "weekly"** and **"monthly"**). Then press the **"ok"**-button.

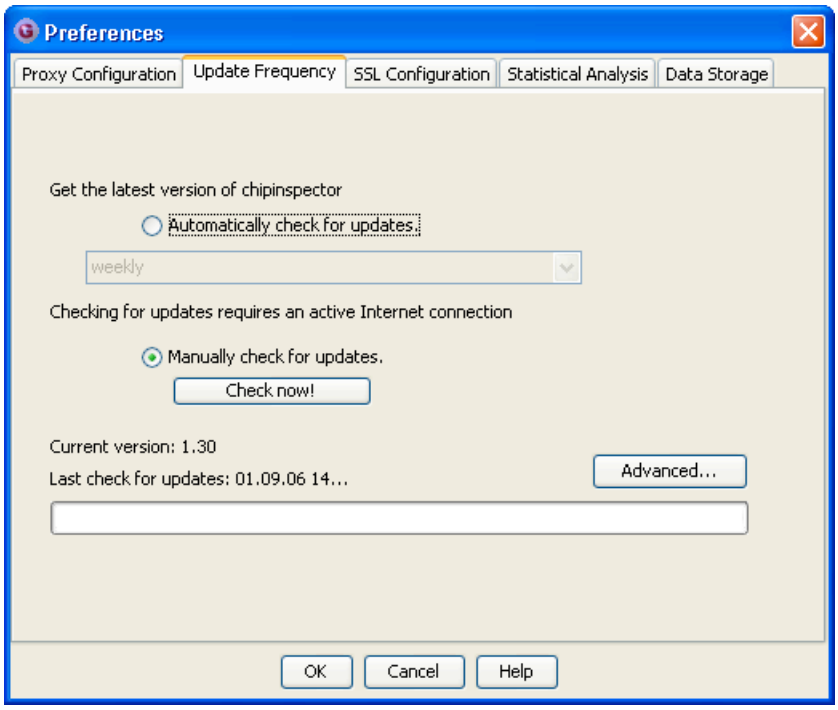


Updating your Application Manually

In some situations, you might want to update your application manually.

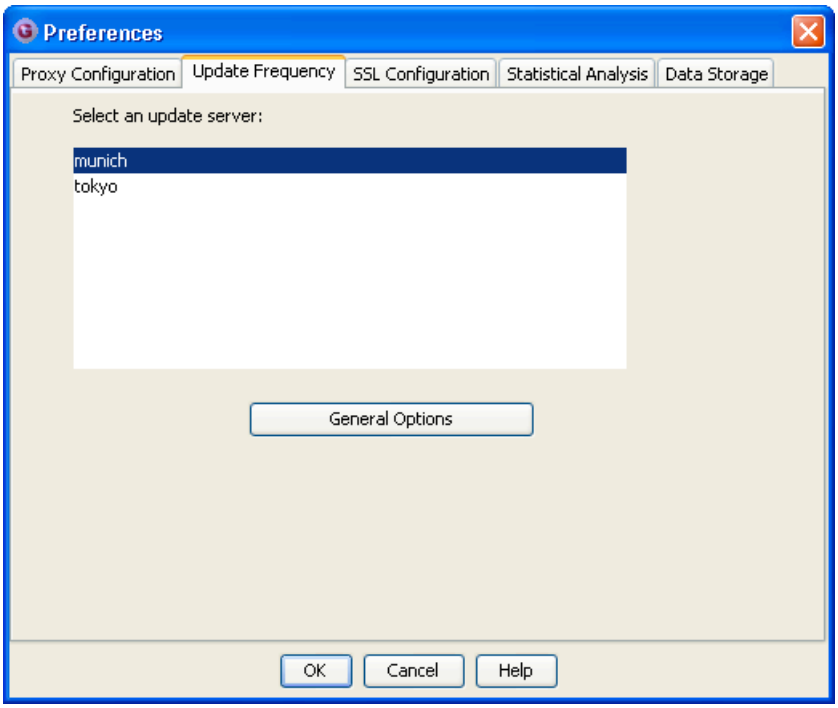
Select **"Manually check for updates".** This will activate the **"Check now"**-button.

Press the **"Check now"**-button. If an update is available the update service will guide you through the update process.



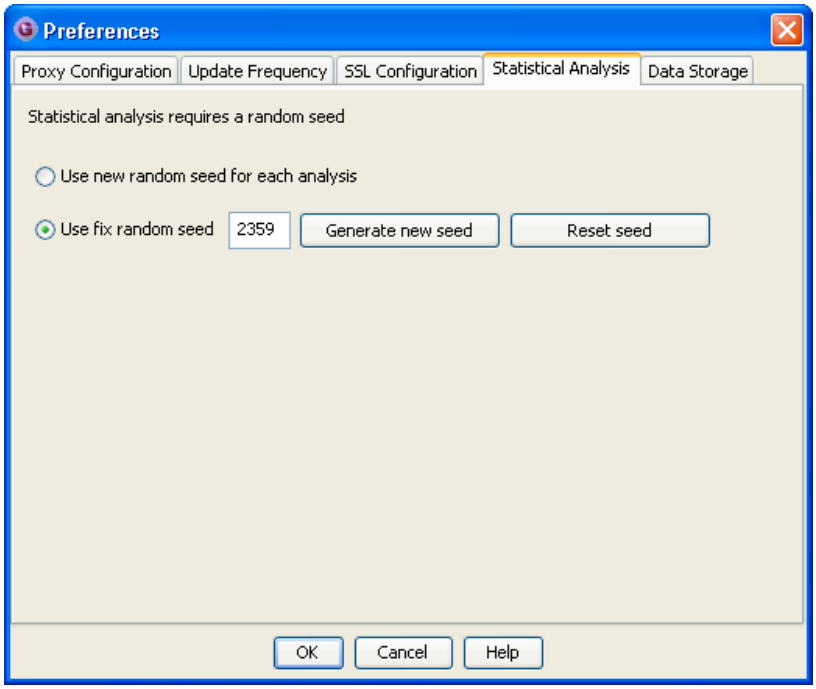
Selecting an Update Server

If update speed is slow, click the “Advanced...” button in the Update Frequency panel and select a different update server from the list. To go back to the main panel, click the “General Options” button.

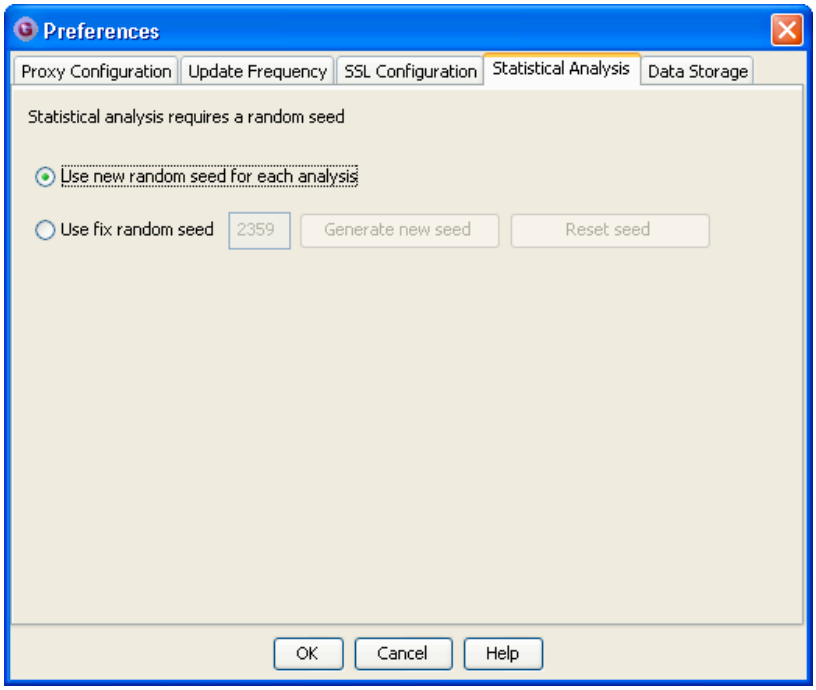


Configuration of Statistical Analysis Behavior

To ensure reproducibility of analysis results, the random seed used by the algorithm is a fixed default number used for all analyses. To use a different number, select the “Statistical Analysis” tab in the Preferences dialog and either enter a value in the number field, or generate one randomly by clicking the “Generate new seed” button. The “Reset seed” button sets the value to default. Changes to this “fixed random” number are reflected in the protocol to facilitate retrieval of any number seed for later use.

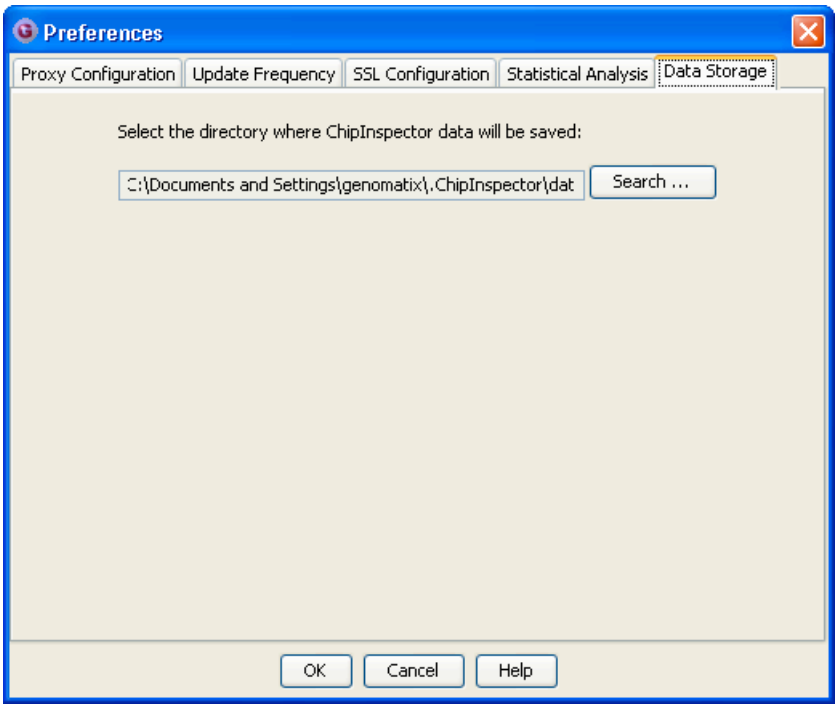


Alternatively, you can select an altogether different behavior by activating the “Use new random seed for each analysis” option. The fixed number will be ignored, and for each analysis, a newly generated random number will be used as a seed.

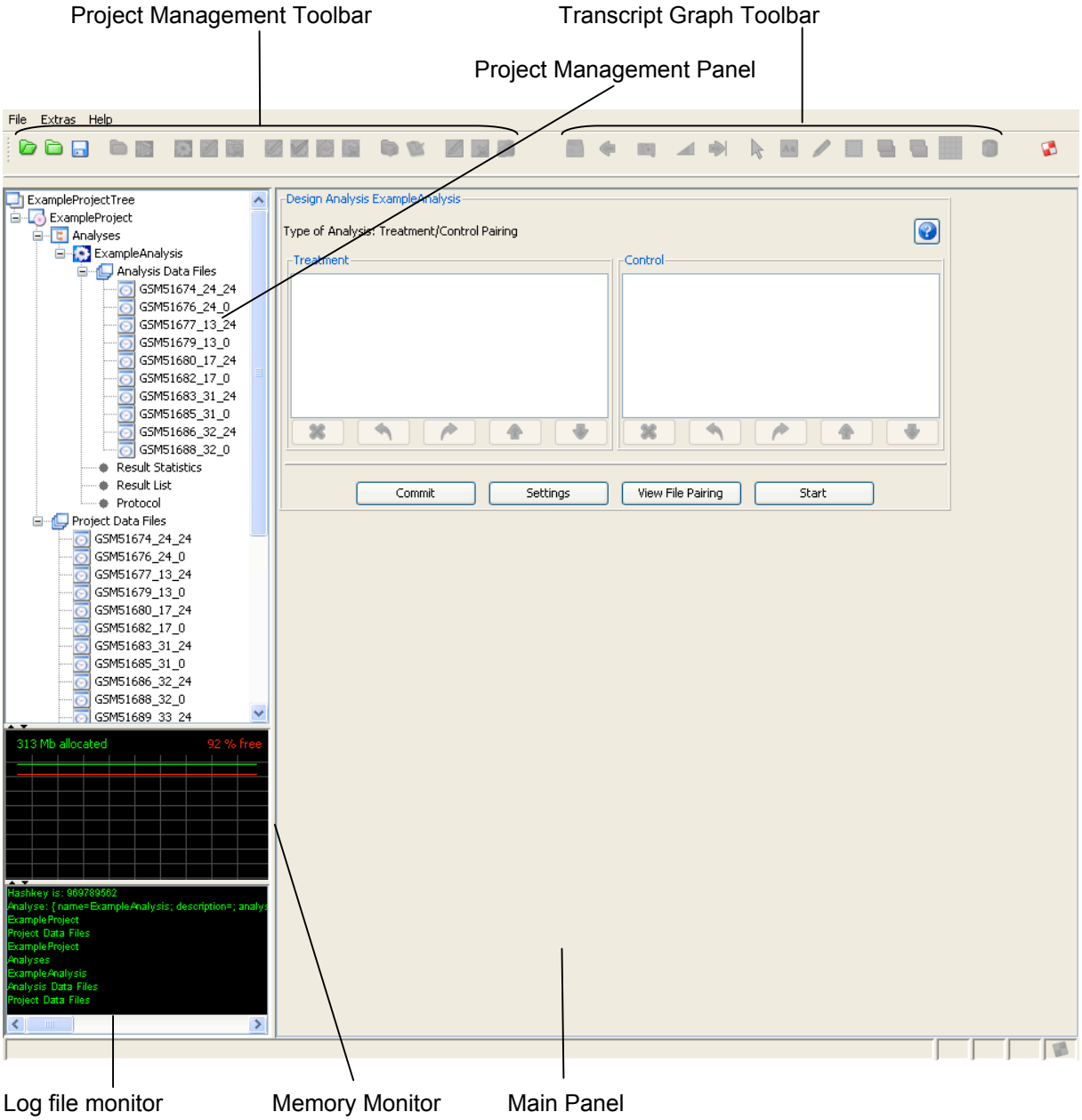


Data Storage Settings

You can set the directory where ChipInspector data will be saved in the Data Storage panel. The default is the subdirectory \.ChipInspector\dat in your home directory. To change it, click the "Search…" button to open a dialog and select a different directory.



The ChipInspector Workspace Area



Project Management Panel

The project management panel on the left side of the screen shows the projects and analyses in a tree structure. Right-clicking on an item in the tree opens a context menu for performing actions on the respective object. Only the menu items with a meaningful function for the current state of the object will be activated.

Main Panel

The main panel provides input forms for actions and displays analysis results.

Memory Monitor

The memory monitor displays information on recent memory usage by ChipInspector on your computer. The green line indicates the allocated memory, and the red line shows the percentage of memory that is free. Current values are displayed numerically.

Log File Monitor

The log file monitor displays the latest entries in the log file. To display the whole log file, double click the log file monitor

Toolbar Elements

Only the buttons that are functional in the current program state will be activated.

Project Management

	Open Project Tree	Open an existing project tree
	New Project Tree	Create a new project tree
	Save	Save the project tree
	New Project	Add a new project to the project tree
	New Batch Job	Add a new batch job to the project tree
	New Analysis	Add a new analysis to the selected project
	Edit Project	Edit the selected project
	Delete Project	Delete the selected project including all its analyses
	Edit Analysis	Edit the selected analysis
	Design Analysis	Change the design of the selected analysis
	Redo Analysis	Redo the selected analysis
	Delete Analysis	Delete the selected analysis
	Import Data Files into Project	Import CEL raw data files into the selected project
	Data Quality Overview	Show data quality overview of all project data files
	Edit Data File	Edit properties of the selected data file
	Delete Data File	Remove the selected data file from the list
	Calculate Data Quality Statistics	Calculate data quality statistics for selected data files

Transcript Graph

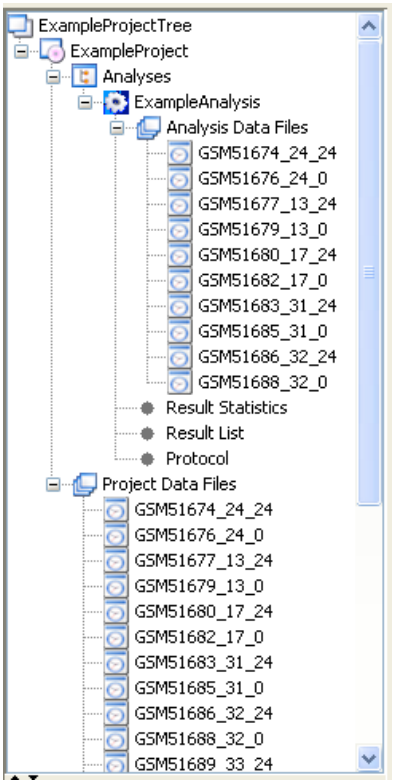
	View Locus in ElDorado	View the currently selected locus in ElDorado
	Go Back to the Result List	Display the result list
	Export Graph	Export the graph
	Change Scale	Change the length scale of the graph
	Fit Graph to Window	Set the graph's length scale to fit the window
	Select	Select an element or a region in the graph
	Add Text	Add text to the graph
	Mark a Region	Draw a labeled marker frame
	Export this Region	Export the marked region
	Bring Item to Front	Bring the selected item in the graph to the front
	Send Item to Back	Send the selected item in the graph to the back
	Toggle Grid	Toggle display of the grid on/off
	Delete Item	Delete the selected item

Other

	Free Unused Computer Memory	Free unused memory on your machine

Project Tree

Your analyses are grouped into projects; projects are grouped in a project tree. At any time, one project tree can be open in the program.



Creating a New Project Tree

To create a new project tree, select File – New Project Tree from the menu, or click on the New Project Tree ( ) toolbar button. Any other open project tree will be closed automatically.

Opening an Existing Project Tree

To open a previously saved project tree, select File - Open Project Tree from the menu, or click on the Open Project Tree ( ) toolbar button, and select a project tree file from the dialog.

Saving the Project Tree

To save the project tree, including all projects, analyses and results, select “Save Project Tree” from the File menu, or click on the Save ( ) toolbar button.

Deleting a Project Tree

To delete a project tree, including all projects, analyses and results, select “Delete Project Tree” from the File menu.

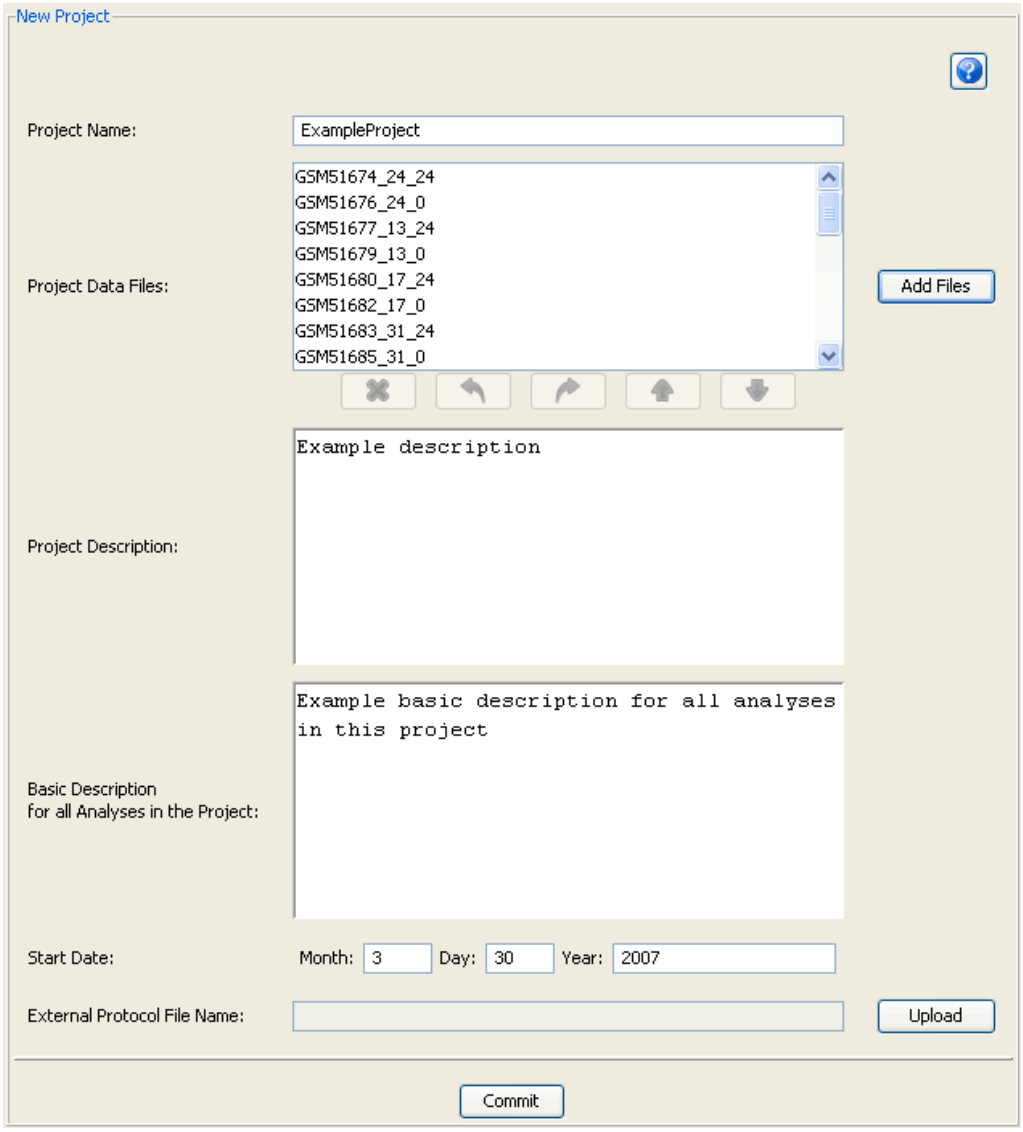
Projects

Creating a New Project

To create a new Project in the project tree, right-click the tree's root node and select "New Project" from the context menu. You can also use the New Project ( ) button in the toolbar.

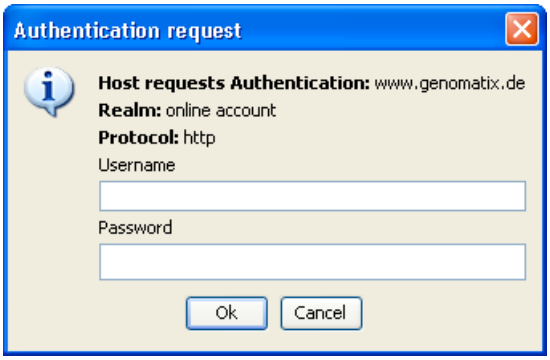
The main panel will display a form that allows you to enter a name for your project and add data files for analysis. Clicking the "Add Files" button opens a dialog for data file selection.

You can also add description for your project, as well as a general description for the analyses in the project. The project start date can be selected; the default is the current date. ChipInspector automatically generates a protocol file for each analysis. You can upload any external file into this protocol (e.g. a MAS 5.0 .RPT file). If the file is in ASCII format, the contents will be added to the protocol of each analysis in the project; otherwise, only the path and file name will be embedded.

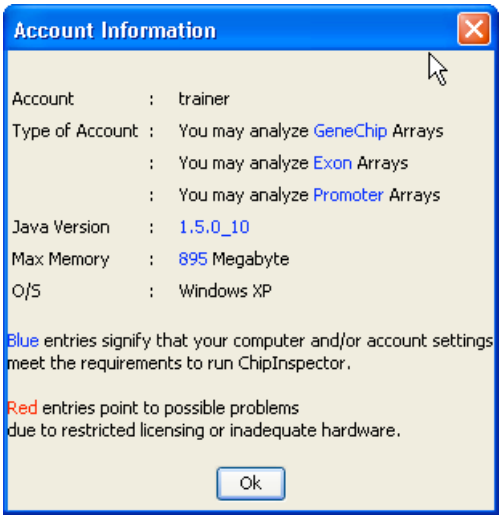


Click on "Commit" to save your input.

If you are not logged in yet, you will be asked to do so now. Please enter your username and password and click OK.



A dialog opens, which displays information on your account and system.



Importing Data Files into a Project

After creation of a new project, the specified data files are loaded. For data files based on Affymetrix chips, this is done automatically, without any more user interaction. To import chip data of other manufacturers, a data import interface is used. A progress bar informs you about the status of the process. You can add more files to a project at any time by selecting it in the project tree, right-clicking and choosing “Import Data Files into Project” from the popup menu. Alternatively, click the Import Data Files into Project ( ) toolbar button to open the import dialog.

After the file import is completed, a data quality overview is displayed. For each file, the percentage of legible data and the expression average is shown. Non-legible data are e.g. expression values that cannot be interpreted as a number. If legibility of a file is below 99% or the expression average differs markedly (> 8 standard deviations) from the mean of the loaded files, the entry will be highlighted in red. This overview can be also accessed by selecting the “Project Data Files” node of a project an clicking the Data Quality Overview ( ) button in the toolbar or, alternatively, right-clicking and selecting the according item from the context menu.

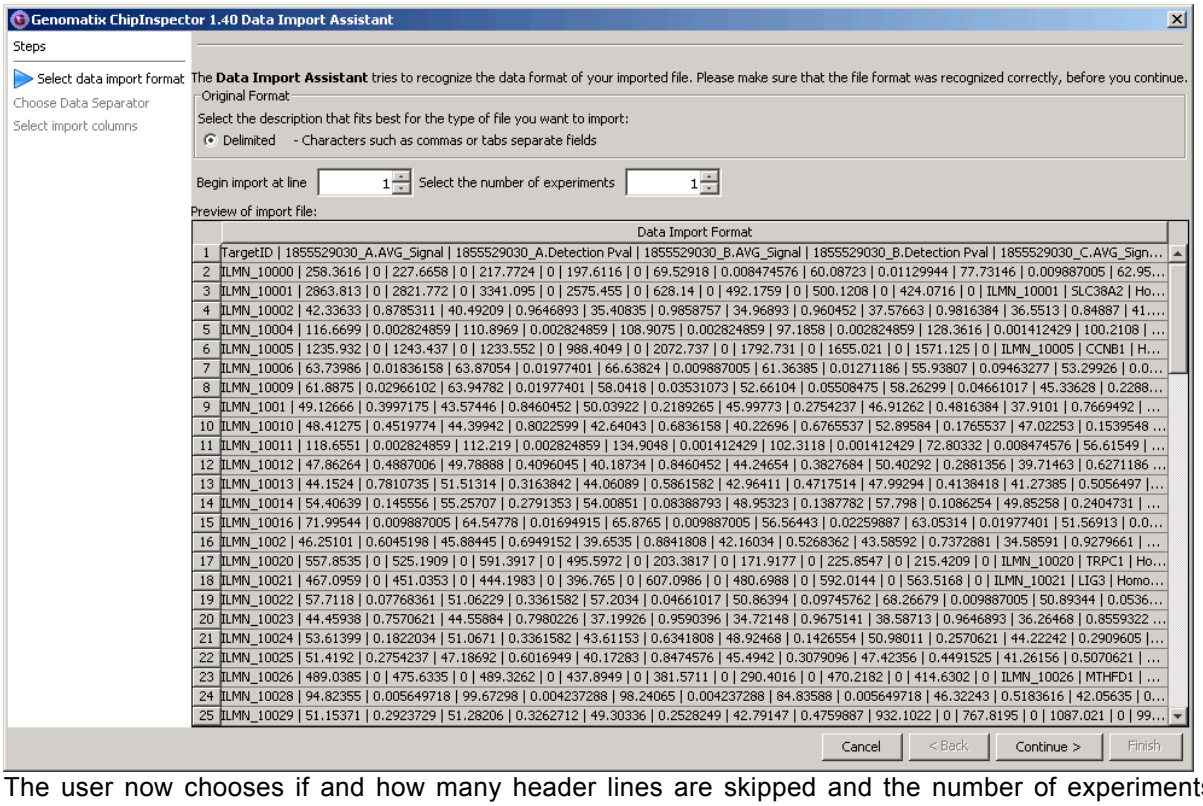
Data Quality Overview

	File	Legibility (%)	Expr. Av.
1	GSM51712_2_0	100.0	187.65
2	GSM51710_2_24	100.0	167.05
3	GSM51709_25_0	100.0	210.65
4	GSM51707_25_24	100.0	214.27
5	GSM51706_20_0	100.0	128.65
6	GSM51704_20_24	100.0	111.32
7	GSM51703_43_0	100.0	164.78
8	GSM51701_43_24	100.0	144.42
9	GSM51700_40_0	100.0	154.64
10	GSM51698_40_24	100.0	119.09
11	GSM51697_38_0	100.0	182.4
12	GSM51695_38_24	100.0	177.41
13	GSM51694_37_0	100.0	206.55
14	GSM51692_37_24	100.0	174.1
15	GSM51691_33_0	100.0	172.89
16	GSM51689_33_24	100.0	189.39
17	GSM51688_32_0	100.0	182.57
18	GSM51686_32_24	100.0	137.77
19	GSM51685_31_0	100.0	145.26
20	GSM51683_31_24	100.0	143.6
21	GSM51682_17_0	100.0	157.36
22	GSM51680_17_24	100.0	143.14
23	GSM51679_13_0	100.0	177.19
24	GSM51677_13_24	100.0	157.95
25	GSM51676_24_0	100.0	258.63
26	GSM51674_24_24	100.0	203.03

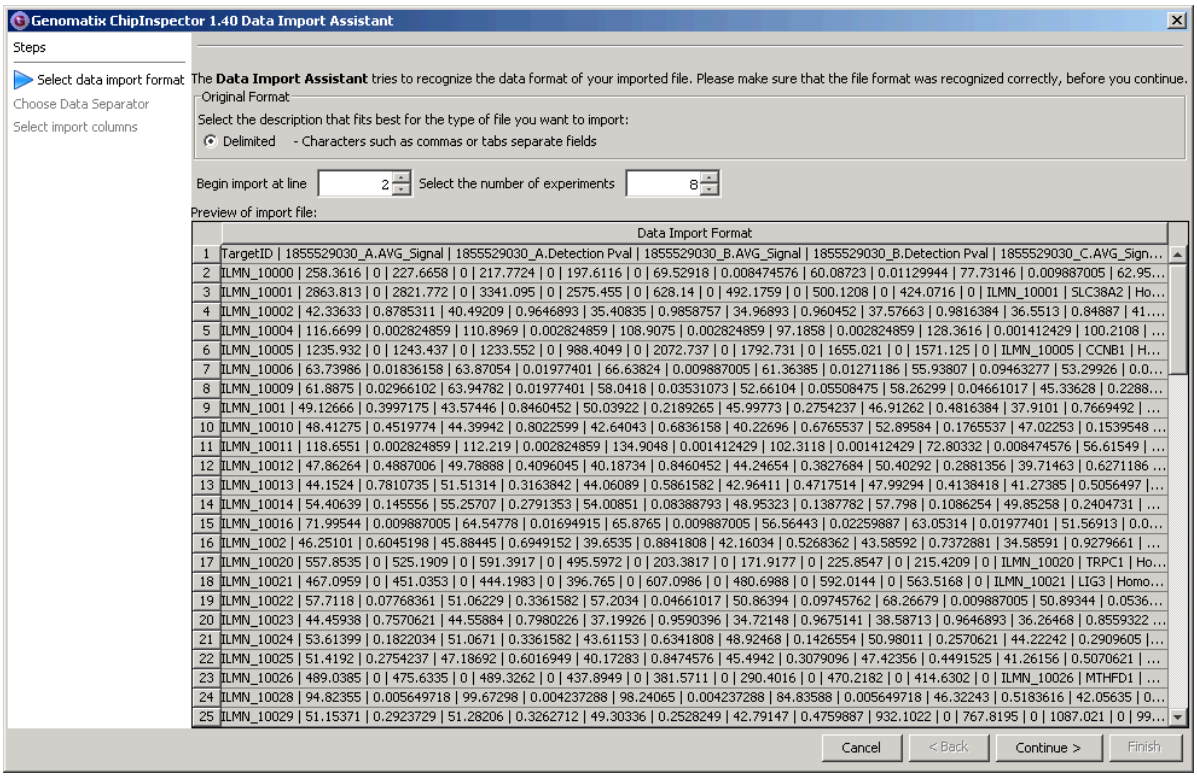
To continue with creating analyses, click OK.

Data Import Interface

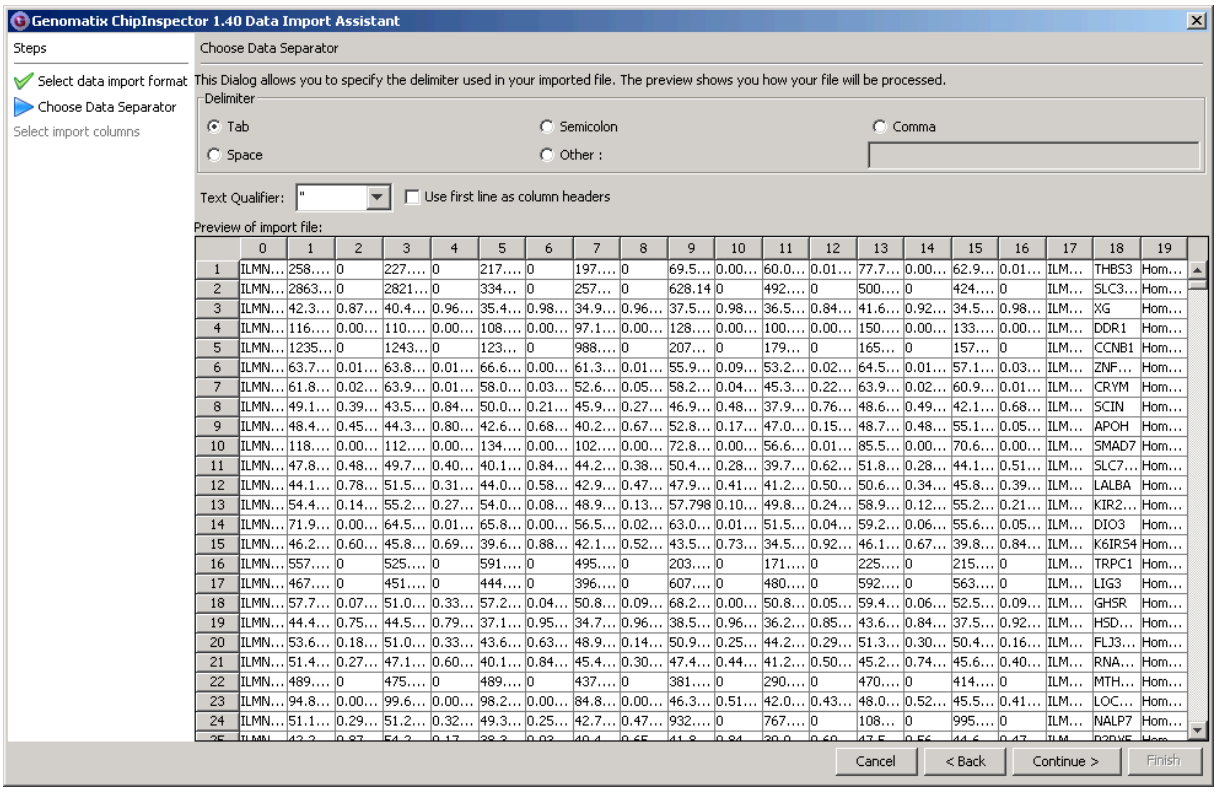
For array results from Illumina BeadStudio and Agilent, a data import interface is employed. ChipInspector will open any unrecognized file with the following view:



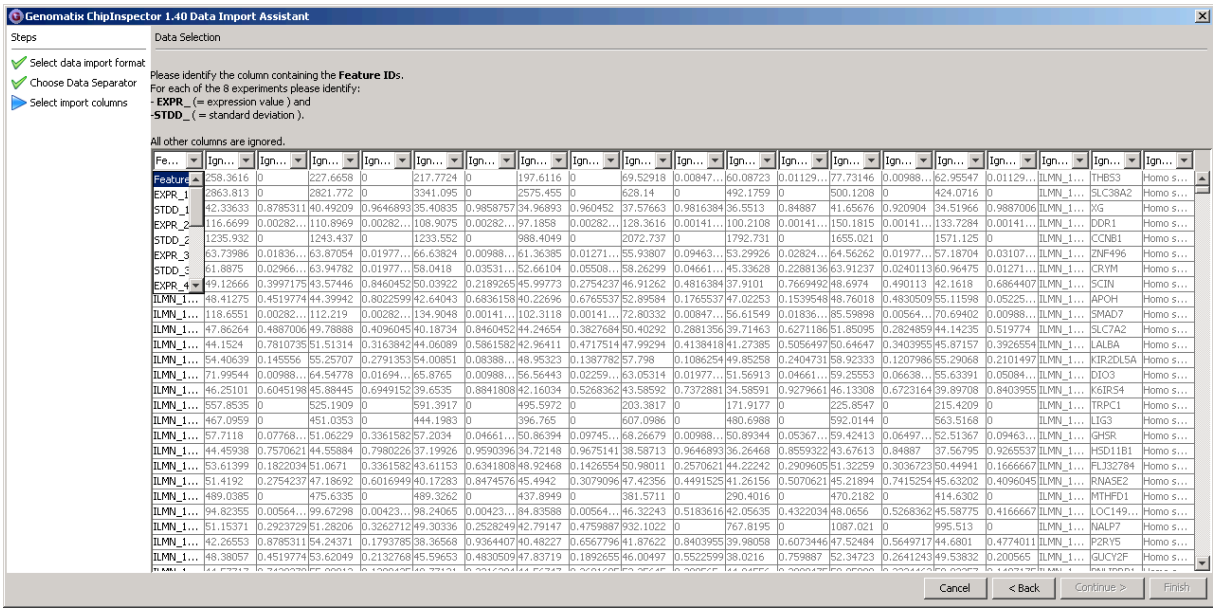
The user now chooses if and how many header lines are skipped and the number of experiments contained in this file.



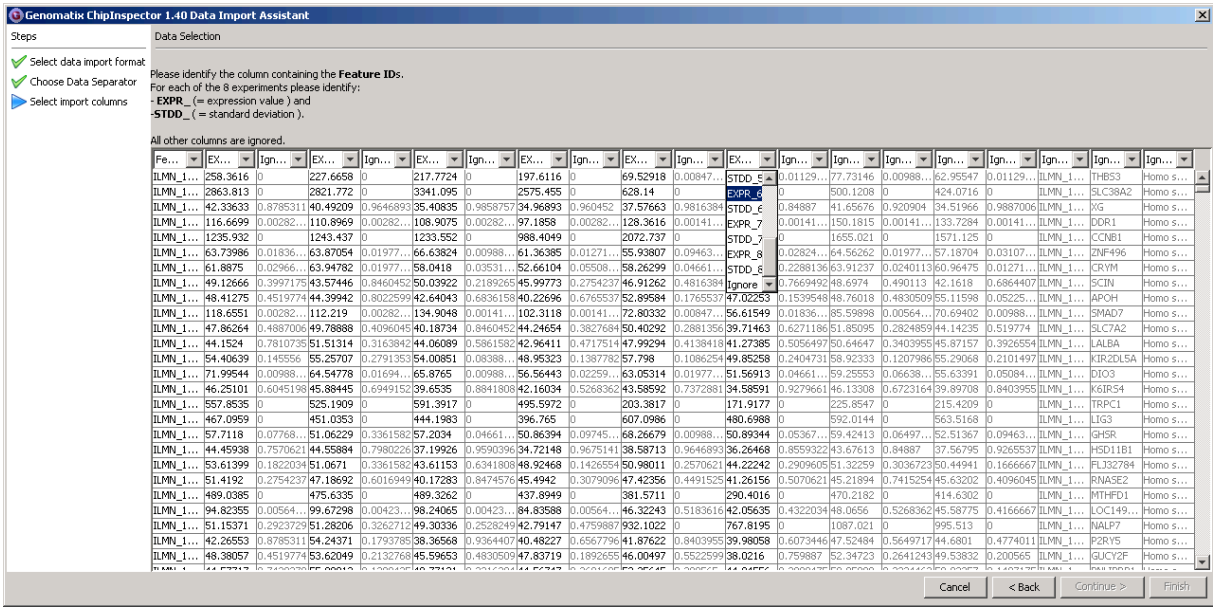
After clicking "Continue", the type of delimiter (Data Separator) needs to be determined. The resulting file setup is shown in the interface.



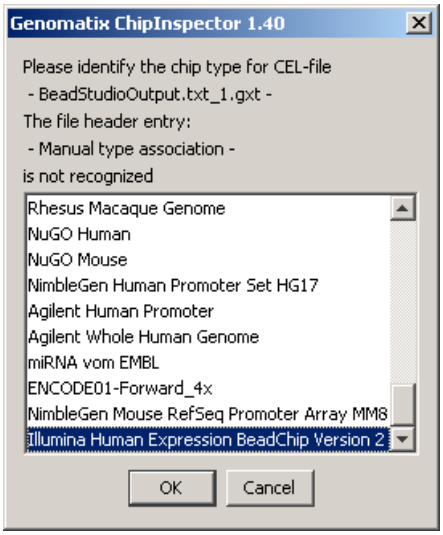
After clicking "Continue" again, the feature ID for the experiment needs to be chosen. For Illumina BeadStudio output (raw data without normalization), the column with the ILMN number is the recognized feature ID, for Agilent arrays, the corresponding identifier needs to be selected:



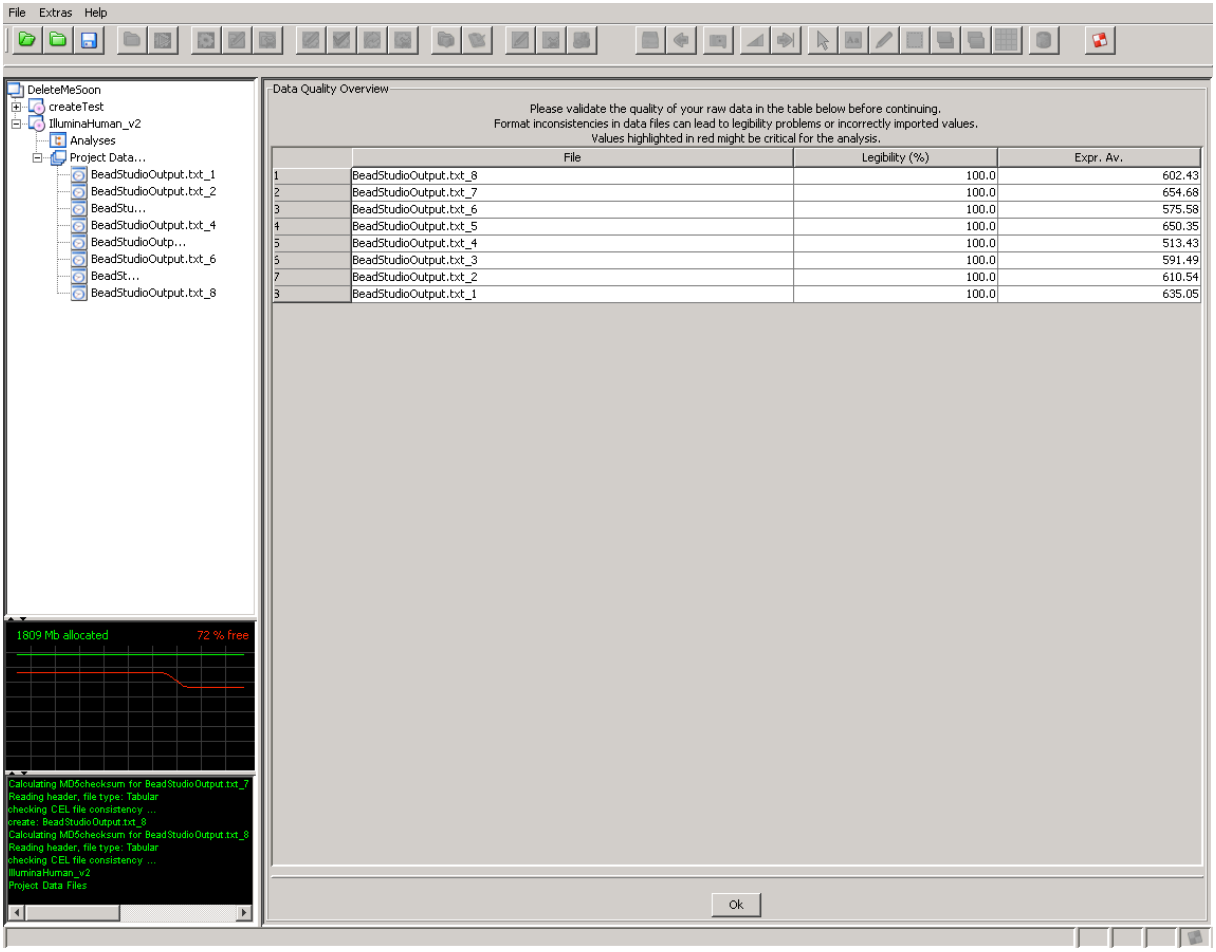
For every experiment in the file, the expression value must be identified along with an optional standard deviation value.



Finally, the user needs to identify the chip type. Please refer to the above list of accepted chip types.



The file import proceeds normally from here.



Editing a Project

Right-click on a project node in the tree structure and select “Edit Project” from the context menu, or click on the Edit Project( ) button in the toolbar to open the Edit Project panel. This looks much the same as the New Project panel and allows you to edit the project data.

Deleting a Project

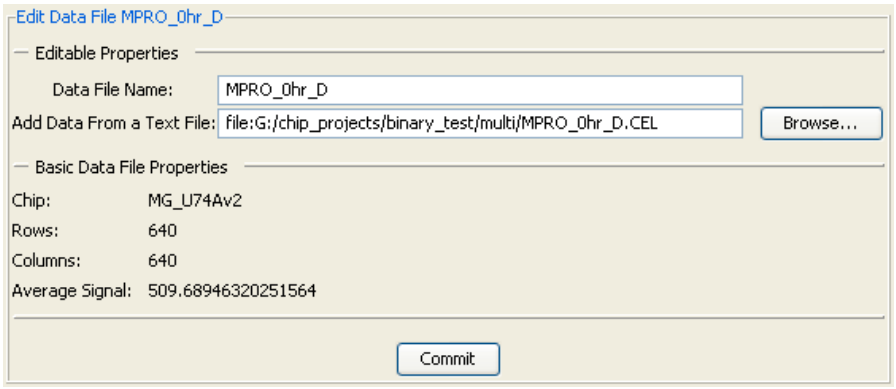
To delete a project, choose the “Delete Project” option from its context menu, or click the Delete Project ( ) button in the toolbar. The project will be deleted, including all its analyses and data file associations. The data files themselves will not be deleted from the file system.

Project Data Files

The imported data files are available in the project manager under the node of the project they have been imported to.

Editing a Data File

Choose “Edit Data File” from the context menu of a data file, or click on the Edit Data File ( ) toolbar button to open the data file editing panel, which displays the file’s properties and allows you to edit some of them.

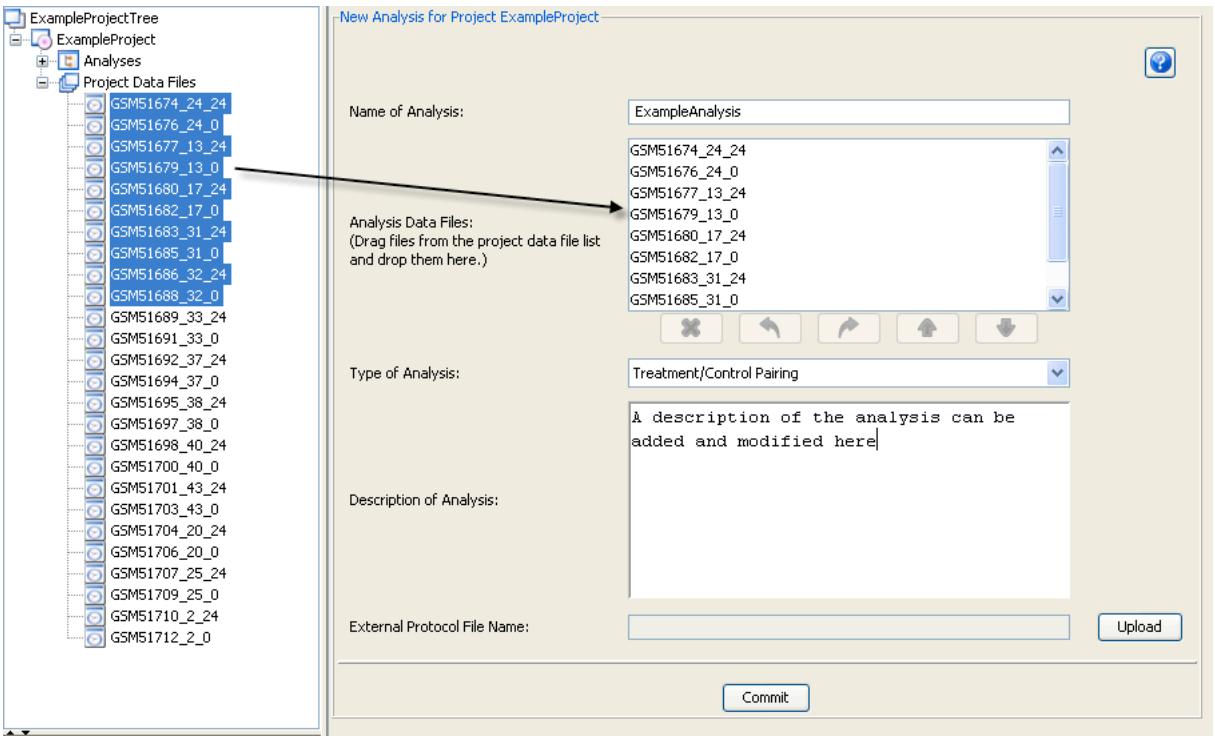


Deleting a Data File

To delete a data file from the list, right-click and choose “Delete Data File” or left click and click on the Delete Data File ( ) button in the toolbar.

Analyses

Creating a New Analysis



To create a new analysis, click on a project node and click on the New Analysis ( ) button in the toolbar, or choose the New Analysis item from the project’s context menu. This will open the New Analysis panel, where you can enter a name for your analysis, choose the data files you want to use in this analysis from the list for the project, and add them to the File Subset list by dragging and dropping. Here and in any other list in the program, you can change the order of entries by selecting one and clicking on the Up ( ) and Down ( ) buttons, remove entries with the Remove ( ) button, and undo/redo your last changes with the Undo ( ) and Redo ( ) buttons.

Choose an analysis type from the selection; available types are:

Treatment/Control Pairing

Use this option if you want to perform a one-class analysis, which compares a treated sample to a control. A single sided permutation T-test analysis is performed.

Time Course/Titration Experiment

Select this option if you want to compare a set of data points in a multi-class analysis and perform a cluster analysis on the results. In this case a multi-class permutation T-test analysis is performed.

Presence/Absence Calling

Choose this option if you want to measure expression values relative to the average expression on the chip, e.g. for gene expression values in one specific tissue. In this case a permutation T-test analysis detecting probes which are significantly above the experiment average is performed. Biological replicates with n>=2 are still required.

You can also enter a description of your analysis, and upload an external protocol file, whose content will be added to the analysis protocol generated by ChipInspector if it is ASCII readable (otherwise only the file name will be embedded).

To get to the next step, analysis design, click the Commit button.

Editing an Analysis

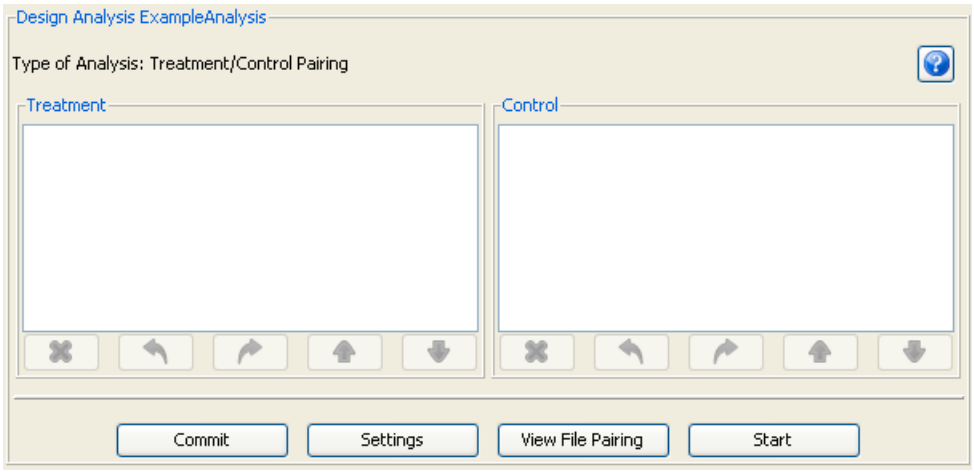
If you want to edit an existing analysis, right-click its symbol in the project tree and choose “Edit Analysis” from context menu, or click the Edit Analysis () toolbar button to open the Edit Analysis panel, which offers the same editing options as the New Analysis panel.

Designing an Analysis

After the editing of analysis is committed, or if you choose “Design Analysis” from the analysis context menu or click on the Design Analysis () toolbar button, the Analysis Design panel opens. The design options depend on the type of the analysis:

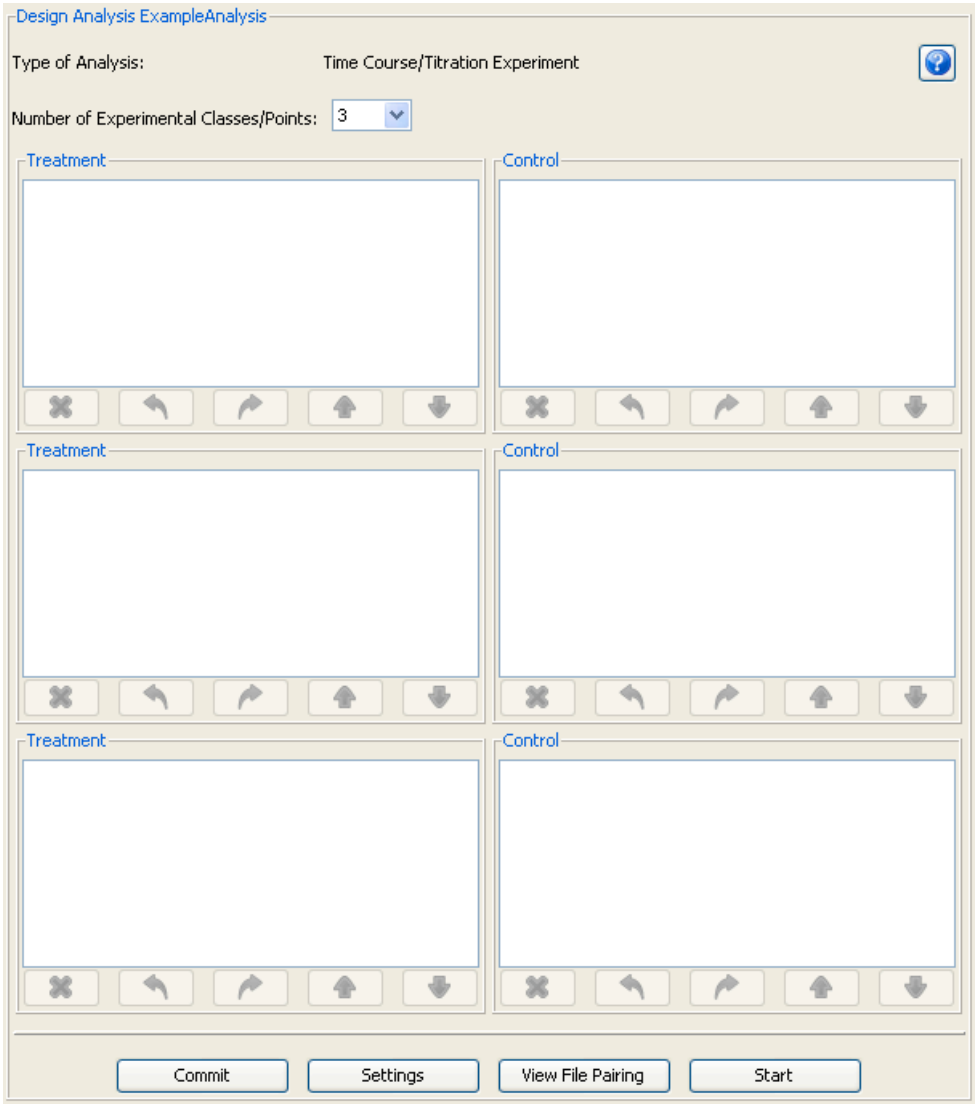
Treatment/Control Experiment

From the Analysis Data Files list in the project manager, drag the files you want to use as treatment and control onto the respective list fields. You can choose the file combinations and view the pairings resulting from your choice; see “File Combinations” for details.


Design Analysis ExampleAnalysis
Type of Analysis: Treatment/Control Pairing
Treatment
Control
Commit
Settings
View File Pairing
Start

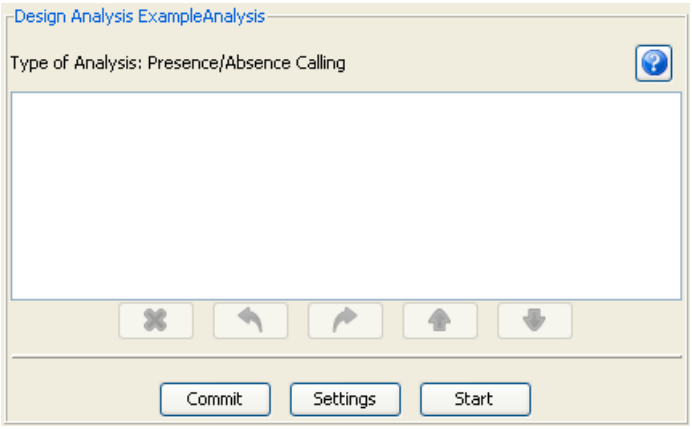
Time Course Experiment

Choose the number of experimental classes from the selection; the matching number of list fields will be displayed. From the Analysis Data Files list in the project manager, drag the files you want to use as treatment and control for each class onto the respective list fields. You can choose the file combinations and view the pairings resulting from your choice; see chapter "File Combinations" for details.



Presence/Absence Calling

From the Analysis Data Files list in the project manager, drag the files you want to analyze into the file list.

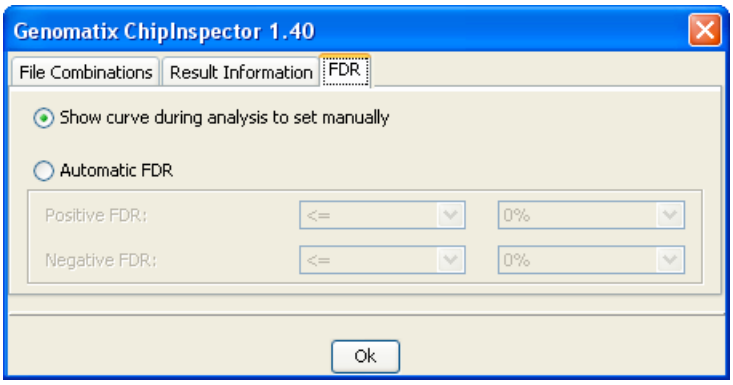


Analysis Settings

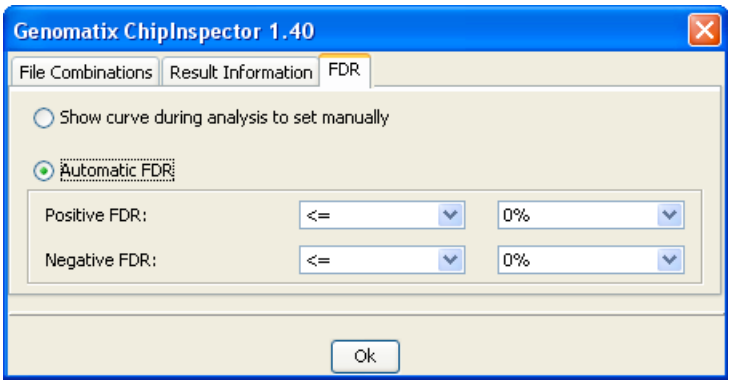
Clicking the Settings button in the Analysis Design panel opens a dialog that allows you to set further analysis parameters:

FDR

The False Discovery Rate (FDR) can either be set manually during the analysis, or alternatively, pre-selected and then automatically set in the analysis.



If automatic FDR is selected, the FDR for up-regulated features (Positive FDR) and down-regulated features (Negative FDR) can be set separately. Negative FDR setting is only available for treatment-control experiments. The FDR percentage can be set to values between 0 and 99 in increments of 1. The exact selected value might not be attainable in a specific analysis; you can set the behavior of the program for this case: with (<=), the largest value that is smaller than or equal to the selected value will be used, whereas (~) uses the value with the smallest absolute difference to the selected value.

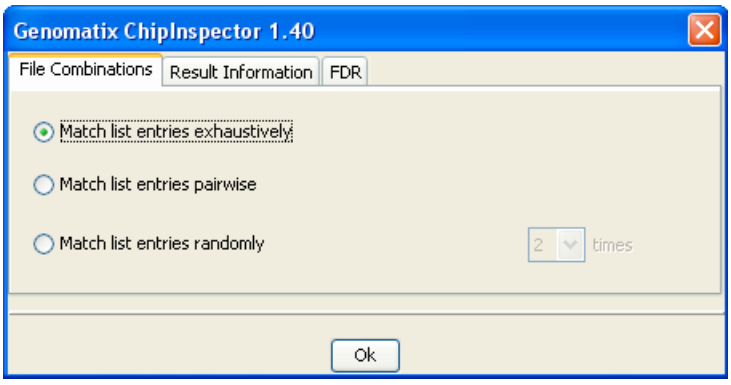


File Combinations

Click on the File Combinations button to open a dialog where you can select the desired manner of file combinations. The View File Pairing button displays a list of the resulting pairs.

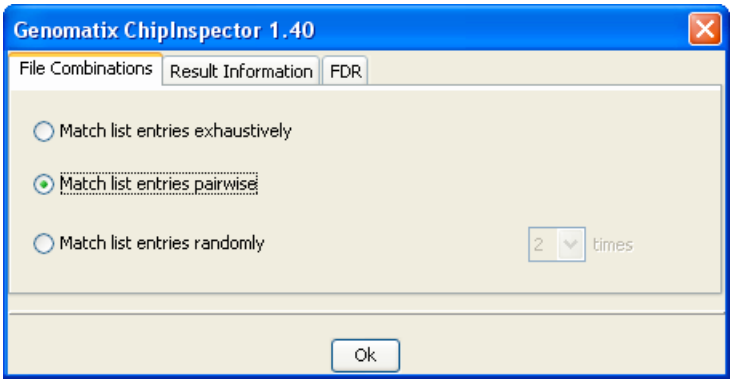
Exhaustive Matching

Exhaustive matching combines every one of the files from one experimental class with every file in the respective control group. This is the default setting.



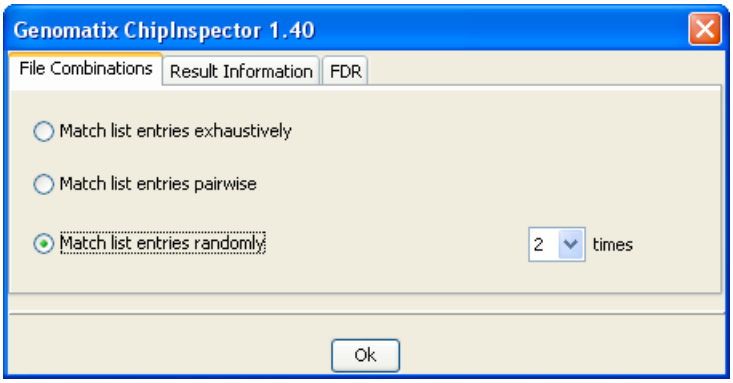
Pairwise Matching

Pairwise matching combines the files one by one in the order they appear in the lists.



Random Matching

Random matching compares a randomly selected set of control files to each of the files in the experimental classes. The number of control files in a set can be selected (between 2 and 10).



File combination examples:

Exhaustive matching

Matching Table

Analysis Option Type: EXHAUSTIVE	
Treatment	Control
MPRO_8hr_A	MPRO_0hr_A
MPRO_8hr_A	MPRO_0hr_B
MPRO_8hr_A	MPRO_0hr_C
MPRO_8hr_A	MPRO_0hr_D
MPRO_8hr_B	MPRO_0hr_A
MPRO_8hr_B	MPRO_0hr_B
MPRO_8hr_B	MPRO_0hr_C
MPRO_8hr_B	MPRO_0hr_D
MPRO_8hr_C	MPRO_0hr_A
MPRO_8hr_C	MPRO_0hr_B
MPRO_8hr_C	MPRO_0hr_C
MPRO_8hr_C	MPRO_0hr_D
MPRO_8hr_D	MPRO_0hr_A
MPRO_8hr_D	MPRO_0hr_B
MPRO_8hr_D	MPRO_0hr_C
MPRO_8hr_D	MPRO_0hr_D

Pairwise matching

Matching Table

Analysis Option Type: EACH_LIST_ENTRY_ONCE	
Treatment	Control
MPRO_8hr_A	MPRO_0hr_A
MPRO_8hr_B	MPRO_0hr_B
MPRO_8hr_C	MPRO_0hr_C
MPRO_8hr_D	MPRO_0hr_D

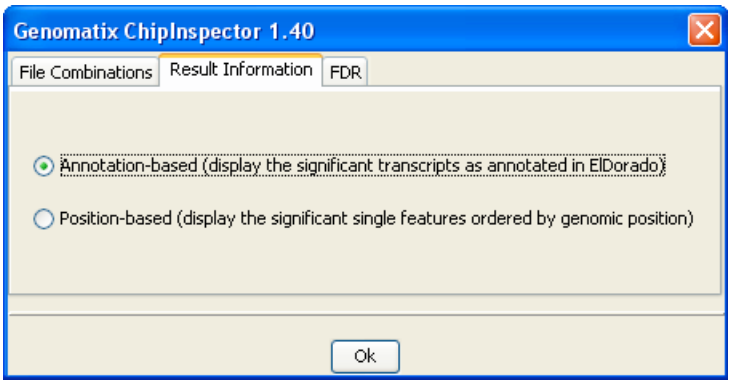
Random matching (2 controls each)

Matching Table

Analysis Option Type: RANDOM Random Number: 2	
Treatment	Control
MPRO_8hr_A	MPRO_0hr_D
MPRO_8hr_A	MPRO_0hr_B
MPRO_8hr_B	MPRO_0hr_A
MPRO_8hr_B	MPRO_0hr_D
MPRO_8hr_C	MPRO_0hr_B
MPRO_8hr_C	MPRO_0hr_B
MPRO_8hr_D	MPRO_0hr_A
MPRO_8hr_D	MPRO_0hr_C

Result Information

Depending on the chip type, the analysis results can be displayed in different ways.



Annotation based

This view is available for all chip types. It displays the significantly regulated transcripts as annotated in ElDorado. The statistical analysis is based on the single probes that map to the exons (for expression arrays and exon arrays) or the promoter region (for promoter arrays) of the annotated transcripts.

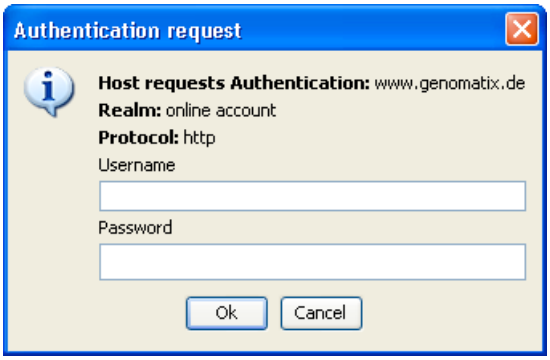
Position based

This option is available for promoter and exon arrays. The significant probes are displayed in the order in which they appear on the genome.

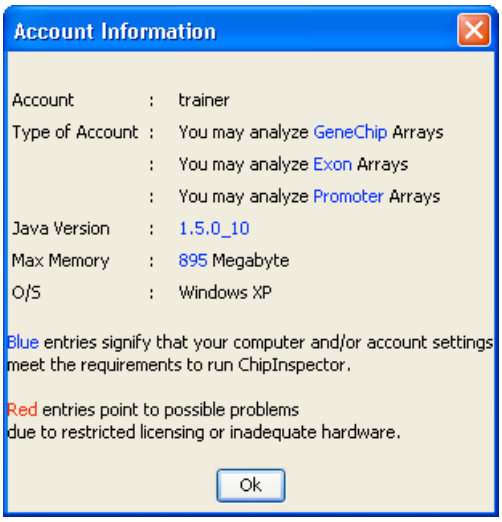
Starting an Analysis

To start the analysis on the fly, click the Start button in the analysis panel.

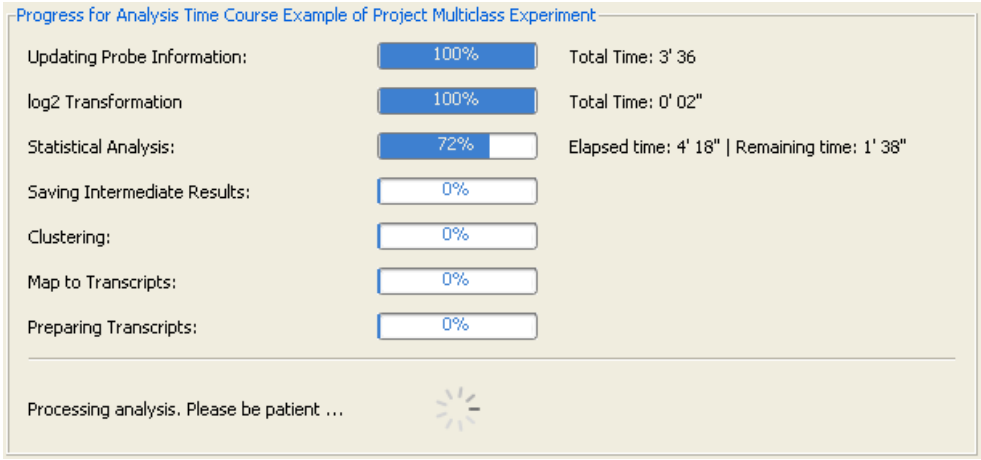
If haven't logged in yet, you will be asked to do so now. Please enter your user name and password in the login dialog, and click OK.



A dialog opens, which displays information on your account and system.

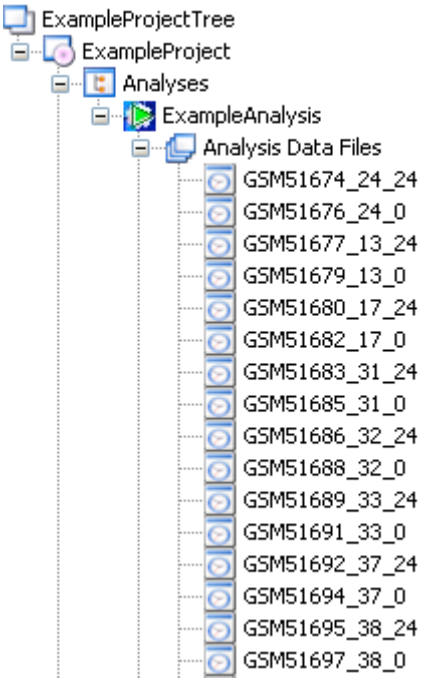


While the analysis is running, its progress is displayed (the exact layout may vary with the analysis type):



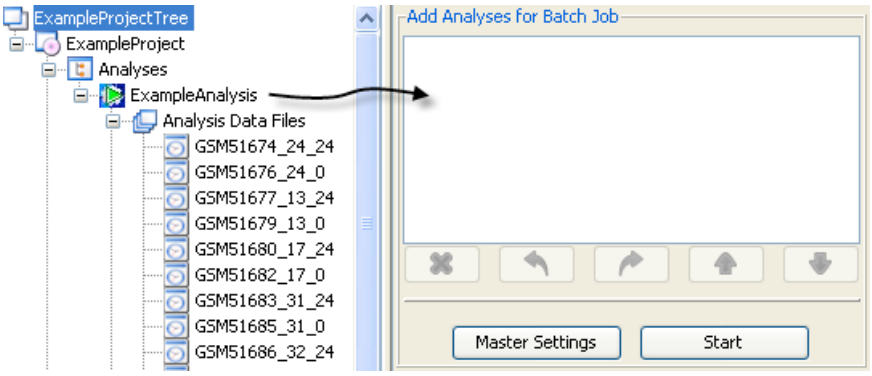
Committing an Analysis for a Batch Job

Clicking the Commit button in the Analysis Design panel does not start the analysis immediately, but saves the design and flags the analysis in the project manager panel with a green arrowhead symbol for later addition to a batch job (see below).



Creating and Starting a Batch Job

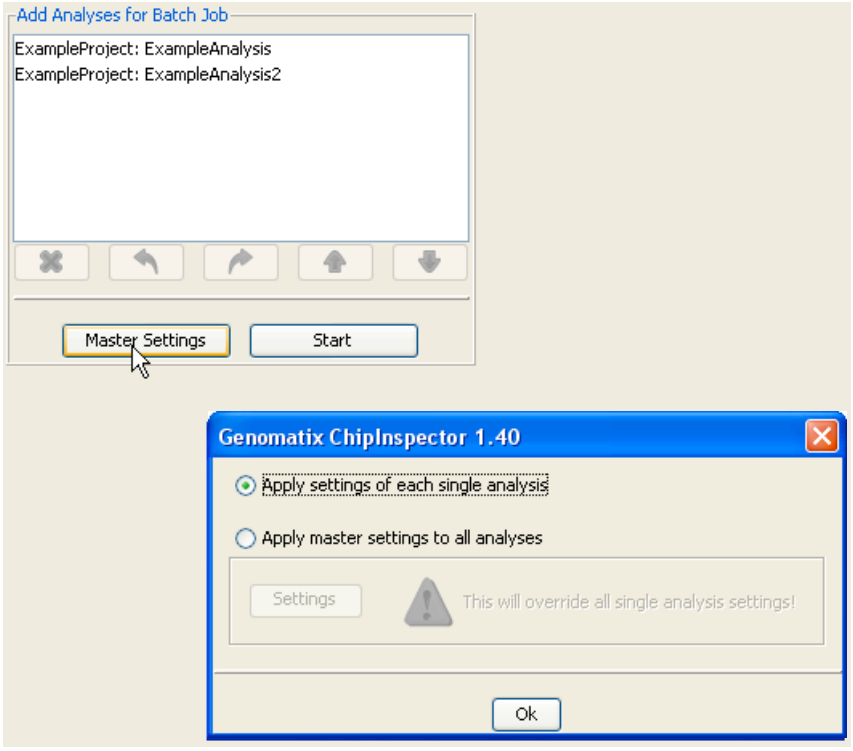
To run one or more analyses in a batch job, commit each analysis after design (see above); then, create a batch job by right-clicking on the root element of the project tree, and selecting “New Batch Job” from the context menu, or by clicking the button in the toolbar. The Batch Job panel will be displayed.



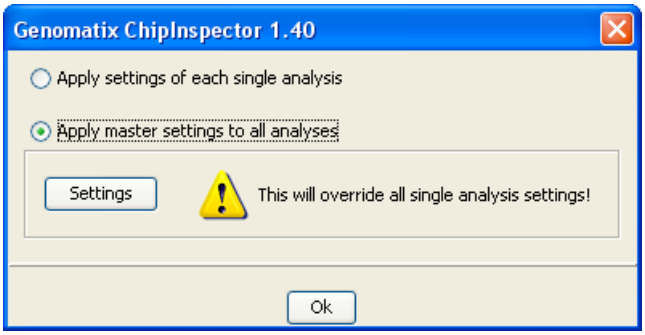
Add your committed analyses to the batch by dragging them from the project manager into the analysis list.

Master Settings

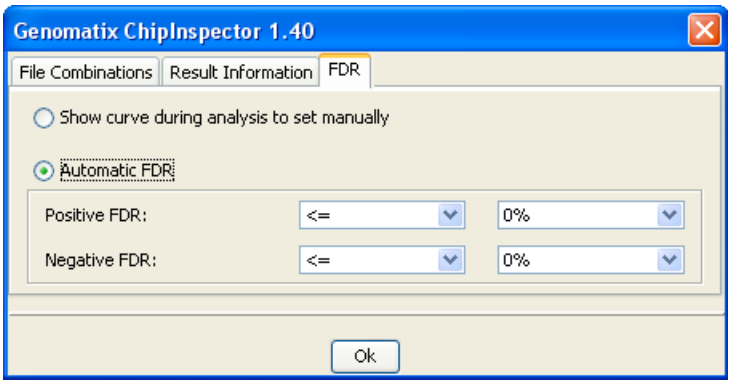
Some analysis parameters can be set uniformly for all analyses in a batch, overriding the settings in the individual analyses. By default, the individual settings apply. To define and activate master settings, click the “Master Settings” button and select “Apply master settings to all analyses” in the dialog.



Click the “Settings” button to open the settings dialog.



In the settings dialog, select the tab for each of the settings you want to define and click OK. You can change the settings for FDR and result information. Please refer to File Combinations for a detailed description of the possible parameter values.



Redoing an Analysis

Redoing an analysis means taking the same set of data files and re-analyzing them, e.g. in different combinations. In order to redo an analysis, choose “Redo Analysis” from its context menu, or select the analysis in the project tree and click the Redo ( ) button in the toolbar. You are asked to provide an extension of the analysis name; the default is the current timestamp. A copy of the analysis with the original design, but without any results, is created. Choose “Design Analysis” from context menu of the copy (or use the Design Analysis ( ) toolbar button) to open the Analysis Design panel. Edit the design and start the analysis as described in “Designing an Analysis”.

Statistics Curve

The statistics curve (blue) displays the result of the statistical analysis as a plot of the observed expression ratio over an artificial background based on randomized expression ratios (expected ratio) for each perfect match probe. As to details concerning statistics please refer to chapter “Data Analysis”.

The diagonal line passing through the origin represents observed ratio = expected ratio; two more lines represent observed ratio = expected ratio + $Delta_{(+)}$ and observed ratio = expected ratio + $Delta_{(-)}$. $Delta_{(+/-)}$ are threshold values; the change in the expression of a single probe (feature) is considered significant if observed ratio > expected ratio + $Delta_{(+)}$ (up-regulated features), or if observed ratio < expected ratio + $Delta_{(-)}$ (down-regulated features).

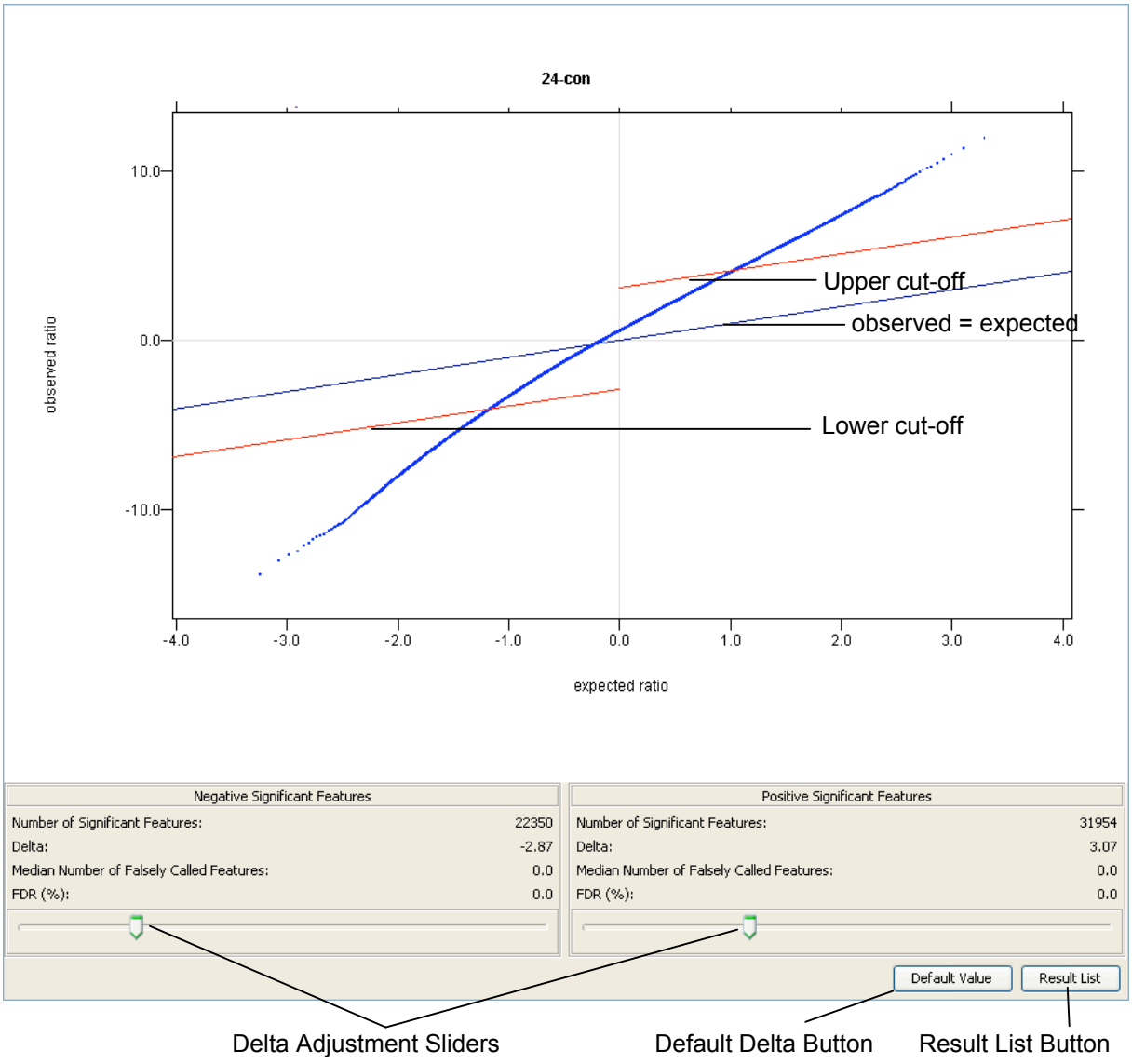
The False Discovery Rate (FDR) is estimated for a given Delta by dividing the average number of features that are called significant in the background data (the falsely called features) by the number of significant features resulting from the experimental assignment. The default Delta values for the thresholds calculated by ChipInspector maximize the number of significant features while maintaining low FDRs.

Sliders can be used to adjust the Delta values in order to change the numbers of significant and of falsely called features, and thus the FDR. You can always reset to the default values by clicking the Default Value button.

The display and the options available for further data processing vary depending on the different types of analyses:

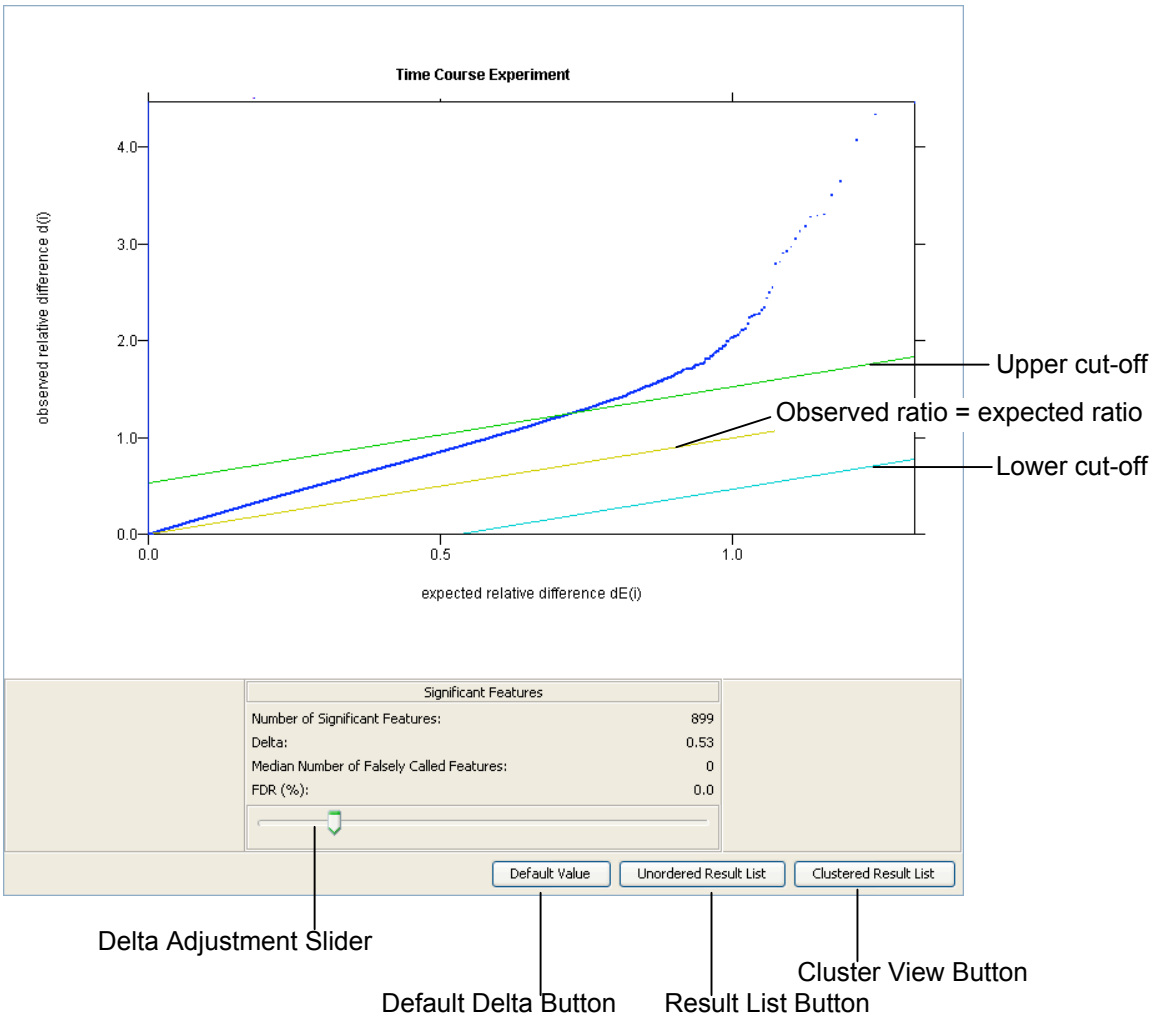
Treatment/Control Experiment or Presence/Absence Calling

The diagonal line is plotted in black, the cut-off thresholds are displayed in red. Two sliders, one each for negative and positive significant features, allow for adjustment of the cut-off values. Clicking the Continue button maps the significant features to the transcripts and displays the resulting transcript list.



Time Course Experiment

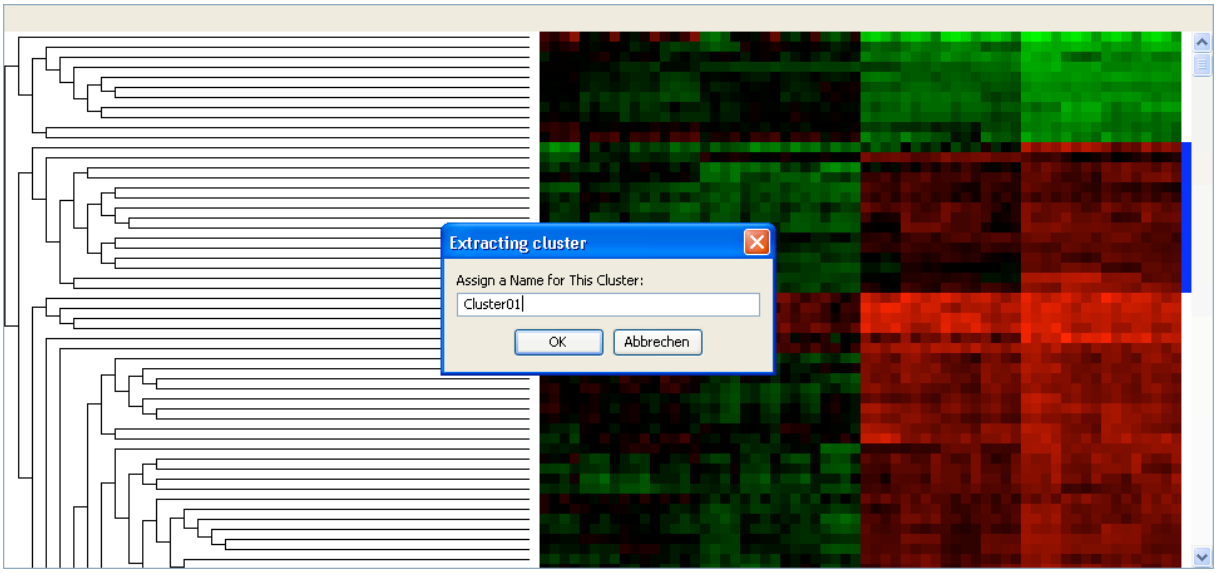
The diagonal line is displayed in yellow, the positive and negative cut-off lines in green and cyan, respectively. One slider is used to adjust the Delta value.
Clicking the Clustered Result List button maps the significant features to the transcripts and displays a cluster tree view of the transcripts and their relative expression values. Clicking the Unordered Result List button maps the significant features to the transcripts and displays a non-clustered transcript list.



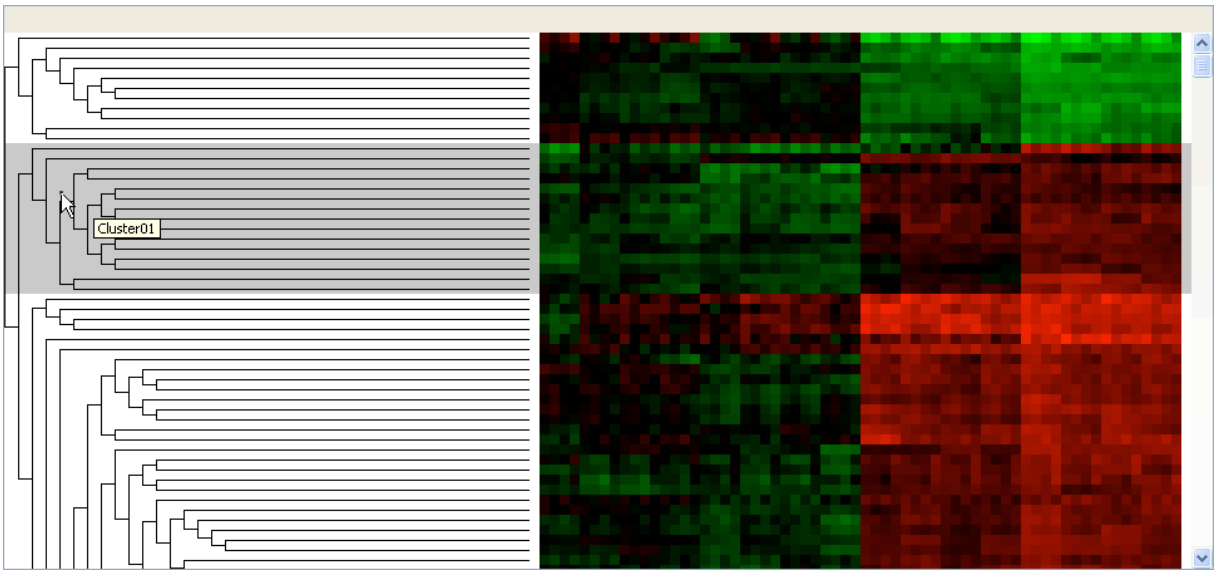
Cluster View

The results of the cluster analysis of a time course experiment are displayed as a hierarchical cluster tree. The logarithmic expression ratios for each file pairing are displayed as green (down-regulated) and red (up-regulated) squares. The degree of differential expression is indicated by the color saturation. Positioning the mouse pointer over any square displays the experimental class it belongs to. Clicking on a tree node displays a blue bar to the right that shows the range of the cluster.

You can extract a result list for a cluster of interest by right-clicking a tree node, optionally assigning a name to the cluster in the dialog box that will be displayed, and clicking OK.



The cluster will appear shaded grey in the cluster view, and a mouse-over will display its name.



Result List

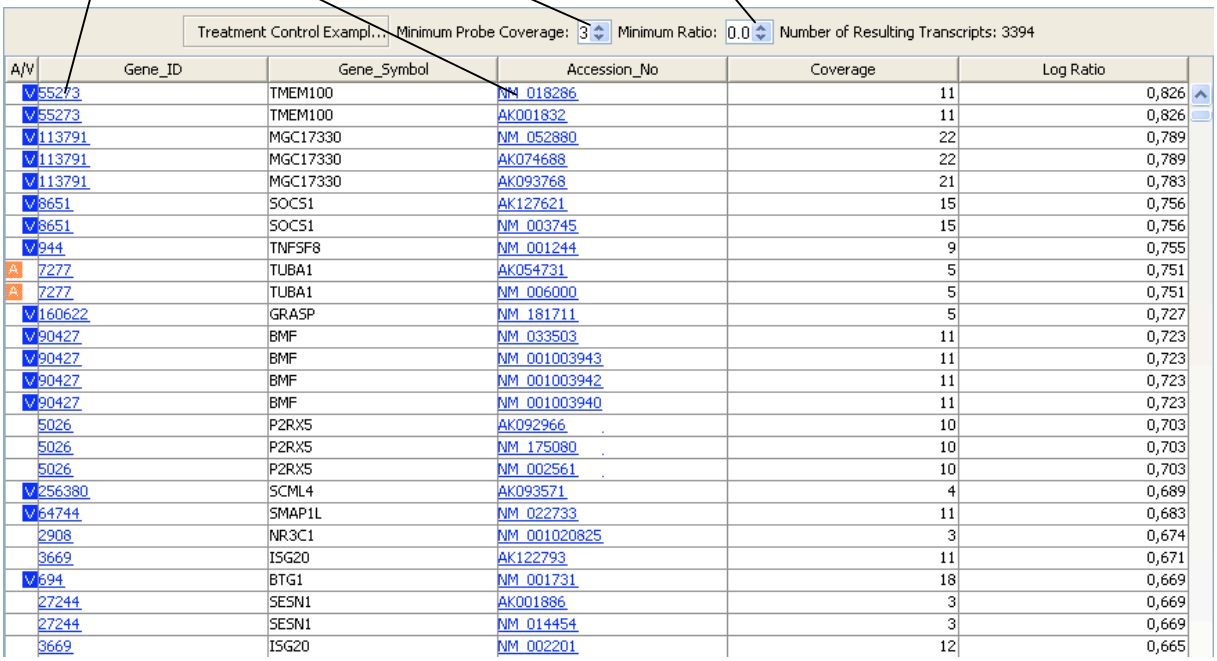
If the transcript based view has been selected in the analysis design, the result list is displayed after the mapping of the significant features. It lists the significantly regulated transcripts and their probe coverage (i.e. the number of significant probes that map to the transcript exons). You can filter the list by minimum probe coverage (default is 3).
Clicking the Gene ID displays functional information on the gene, retrieved from ElDorado. Clicking the accession number will open a detailed graph of the transcript and all other transcripts of its locus. If there is significant evidence for both up and down regulation of a transcript, two entries are listed for it, displaying the according positive and negative expression ratios.
The list can be sorted by any column by clicking on its header. Depending on the analysis type, further information is available:

Treatment/Control Analysis or Presence/Absence Calling

For each significant transcript, the base 2 logarithm of the expression ratio is displayed. You can filter by its minimum absolute value. If there is evidence for both up- and down-regulation of a transcript, the Ambiguous (A) symbol appears in the first column. The Varying (V) icon indicates that there is high variability in the expression ratios of the significant probes that map to the transcript.

ElDorado More Gene Info Link

Graph Link Probe Coverage Filter Expression Ratio Filter



A/V	Gene_ID	Gene_Symbol	Accession_No	Coverage	Log Ratio
V	55273	TMEM100	NM_018286	11	0,826
V	55273	TMEM100	AK001832	11	0,826
V	113791	MGC17330	NM_052880	22	0,789
V	113791	MGC17330	AK074688	22	0,789
V	113791	MGC17330	AK093768	21	0,783
V	8651	SOCS1	AK127621	15	0,756
V	8651	SOCS1	NM_003745	15	0,756
V	944	TNFSF8	NM_001244	9	0,755
A	7277	TUBA1	AK054731	5	0,751
A	7277	TUBA1	NM_006000	5	0,751
V	160622	GRASP	NM_181711	5	0,727
V	90427	BMF	NM_033503	11	0,723
V	90427	BMF	NM_001003943	11	0,723
V	90427	BMF	NM_001003942	11	0,723
V	90427	BMF	NM_001003940	11	0,723
	5026	P2RX5	AK092966	10	0,703
	5026	P2RX5	NM_175080	10	0,703
	5026	P2RX5	NM_002561	10	0,703
V	256380	SCML4	AK093571	4	0,689
V	64744	SMAP1L	NM_022733	11	0,683
	2908	NR3C1	NM_001020825	3	0,674
	3669	ISG20	AK122793	11	0,671
V	694	BTG1	NM_001731	18	0,669
	27244	SESN1	AK001886	3	0,669
	27244	SESN1	NM_014454	3	0,669
	3669	ISG20	NM_002201	12	0,665

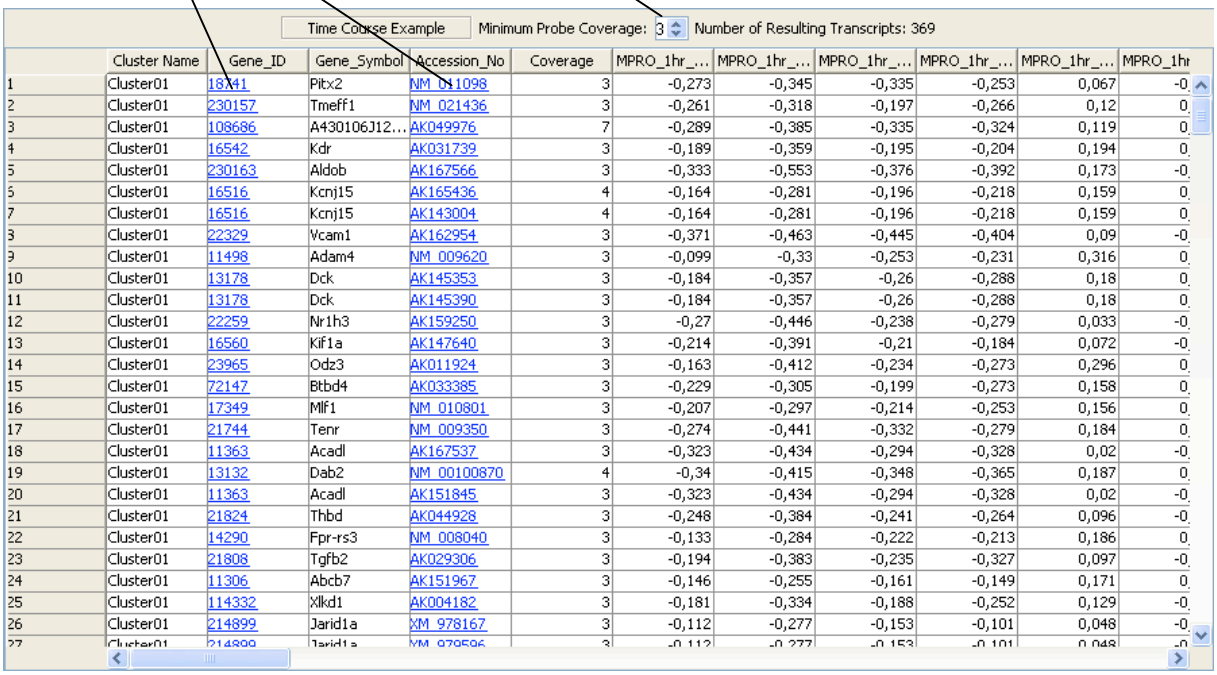
Time Course Experiment

The mean log2 ratios of every pairing in the analysis for each transcript are displayed. If the list represents a cluster in a time course experiment, the cluster's name, if set, is shown as well.

ElDorado More Gene Info Link

Graph Link

Probe Coverage Filter



Time Course Example | Minimum Probe Coverage: 3 | Number of Resulting Transcripts: 369

	Cluster Name	Gene_ID	Gene_Symbol	Accession_No	Coverage	MPRO_1hr_...	MPRO_1hr_...	MPRO_1hr_...	MPRO_1hr_...	MPRO_1hr_...	MPRO_1hr
1	Cluster01	18741	Pitx2	NM_011098	3	-0,273	-0,345	-0,335	-0,253	0,067	-0
2	Cluster01	230157	Tmeff1	NM_021436	3	-0,261	-0,318	-0,197	-0,266	0,12	0
3	Cluster01	108686	A430106J12...	AK049976	7	-0,289	-0,385	-0,335	-0,324	0,119	0
4	Cluster01	16542	Kdr	AK031739	3	-0,189	-0,359	-0,195	-0,204	0,194	0
5	Cluster01	230163	Aldob	AK167566	3	-0,333	-0,553	-0,376	-0,392	0,173	-0
6	Cluster01	16516	Kcnj15	AK165436	4	-0,164	-0,281	-0,196	-0,218	0,159	0
7	Cluster01	16516	Kcnj15	AK143004	4	-0,164	-0,281	-0,196	-0,218	0,159	0
8	Cluster01	22329	Vcam1	AK162954	3	-0,371	-0,463	-0,445	-0,404	0,09	-0
9	Cluster01	11498	Adam4	NM_009620	3	-0,099	-0,33	-0,253	-0,231	0,316	0
10	Cluster01	13178	Dck	AK145353	3	-0,184	-0,357	-0,26	-0,288	0,18	0
11	Cluster01	13178	Dck	AK145390	3	-0,184	-0,357	-0,26	-0,288	0,18	0
12	Cluster01	22259	Nr1h3	AK159250	3	-0,27	-0,446	-0,238	-0,279	0,033	-0
13	Cluster01	16560	Kif1a	AK147640	3	-0,214	-0,391	-0,21	-0,184	0,072	-0
14	Cluster01	23965	Odz3	AK011924	3	-0,163	-0,412	-0,234	-0,273	0,296	0
15	Cluster01	72147	Btbd4	AK033385	3	-0,229	-0,305	-0,199	-0,273	0,158	0
16	Cluster01	17349	Mlf1	NM_010801	3	-0,207	-0,297	-0,214	-0,253	0,156	0
17	Cluster01	21744	Tenr	NM_009350	3	-0,274	-0,441	-0,332	-0,279	0,184	0
18	Cluster01	11363	Acadl	AK167537	3	-0,323	-0,434	-0,294	-0,328	0,02	-0
19	Cluster01	13132	Dab2	NM_00100870	4	-0,34	-0,415	-0,348	-0,365	0,187	0
20	Cluster01	11363	Acadl	AK151845	3	-0,323	-0,434	-0,294	-0,328	0,02	-0
21	Cluster01	21824	Thbd	AK044928	3	-0,248	-0,384	-0,241	-0,264	0,096	-0
22	Cluster01	14290	Fpr-rs3	NM_008040	3	-0,133	-0,284	-0,222	-0,213	0,186	0
23	Cluster01	21808	Tgfb2	AK029306	3	-0,194	-0,383	-0,235	-0,327	0,097	-0
24	Cluster01	11306	Abcb7	AK151967	3	-0,146	-0,255	-0,161	-0,149	0,171	0
25	Cluster01	114332	Xlkd1	AK004182	3	-0,181	-0,334	-0,188	-0,252	0,129	-0
26	Cluster01	214899	Jarid1a	XM_978167	3	-0,112	-0,277	-0,153	-0,101	0,048	-0

Exporting Results

You can export the results in different formats for further analysis. To do so, click on "File" in the menu and choose from the available export options:

Export Values per Transcripts

This will create a list with the mean expression ratio logs for every significantly regulated transcript in every comparison done in the analysis. It can be either a tab delimited or an Excel file.

Export Values per Feature

This will create a list with the expression ratio logs for every significant single probe in every comparison done in the analysis. It can be either a tab delimited or an Excel file.

Export for BiblioSphere Analysis

This will create an MS Excel file that is compliant with the Genomatix BiblioSphere PathwayEdition format requirements for input files.

Position Based View

If the position based result view was selected in the settings dialog of the analysis design, the probes are displayed from top to bottom in their order on the genome. Significantly regulated (exon tiling arrays) or enriched (ChIP experiments with promoter tiling arrays) probes are marked with double asterisks. Move the mouse pointer over a probe to display its log2 expression ratio in a tool tip.

For each chromosomal strand, ElDorado based genome annotation in the regions covered by probes appears alongside the probes on the selected chromosome. The graph shows the chromosomal position of each probe, as well as locus annotation including promoter, exon and intron regions. Promoter regions are depicted in yellow, exon and intron regions in green and grey, respectively. Transcription start regions based on CAGE tag evidence are shown as red arrowheads.

The distances between the tiled probes covering a region are roughly similar. However, gaps in the tiling pattern can occur. In the graph, gaps of at least two probe lengths are represented by an interruption of the transcript graph. A grey separator indicates a gap of at least 1000 nucleotides.

Chromosome selection

Set the chromosome by stepping through the selector list or entering the requisite denominator. Click the arrow button to display the probes mapping to the selected chromosome.

Changing significance thresholds

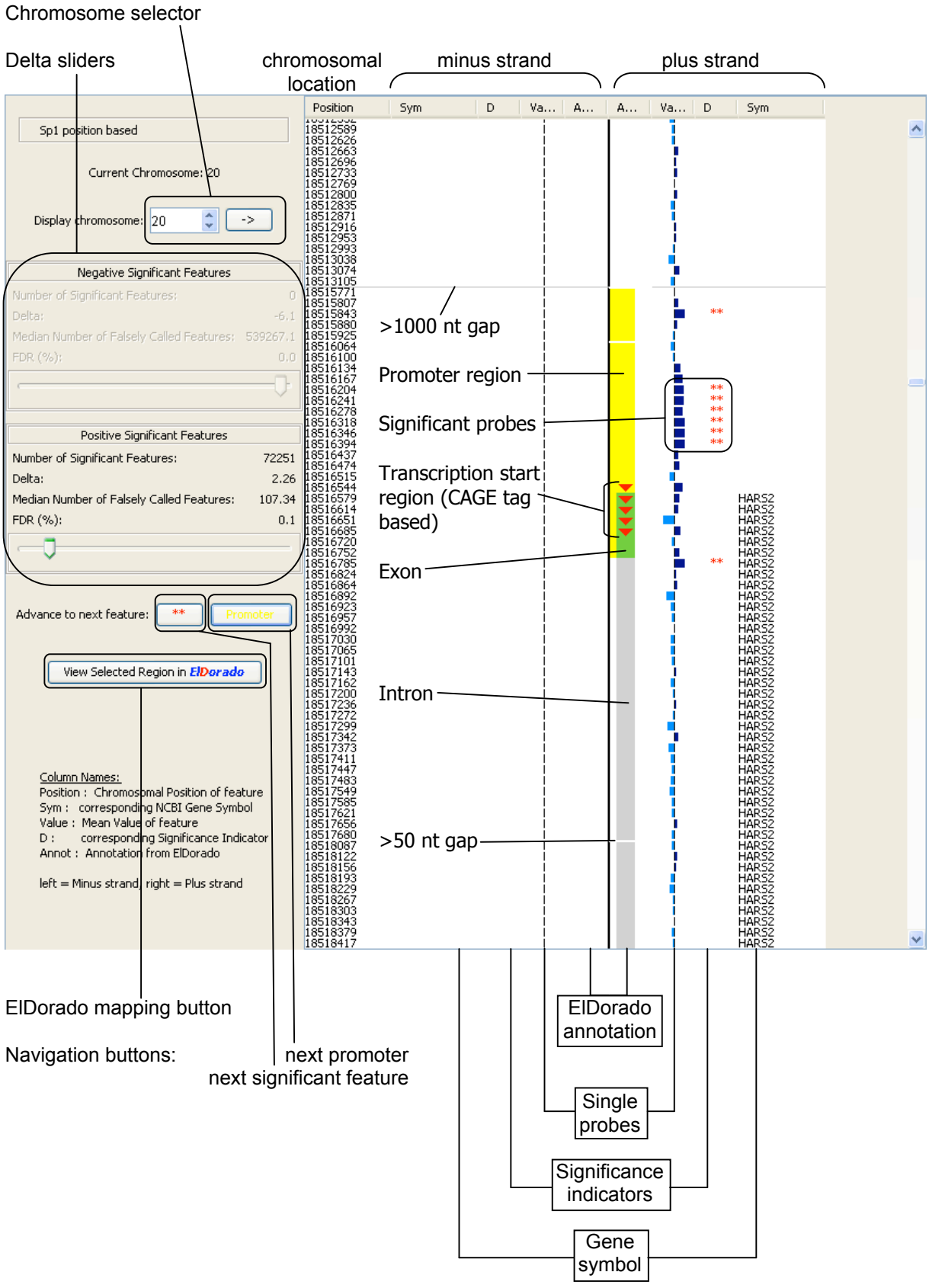
Move the sliders for the Delta values to change the stringency settings for significant features. For promoter tiling arrays, only positive features are available.

Navigation

The graph can be navigated by clicking the respective buttons for jumping to the next significant feature or the next promoter region.

Genomic mapping in ElDorado

Regions that are covered by contiguous probes can be selected by clicking and drawing the mouse. A selected region is mapped onto the genome by clicking the ElDorado mapping button and may then be viewed in ElDorado.



Transcript Graph

The Transcript Graph is only available if the annotation based result view was selected in the analysis design. It is displayed when an Accession No link of a transcript in the result list is followed. It shows a graphical representation of all transcripts that map to the locus of the selected transcript. You can modify and export the diagram.

Graph Overview

Probes

Expression values and positions of significant probes are displayed at the top of the graph as vertical red (significant probes) or blue (non-significant probes) bars. The height of a bar reflects the probe's relative expression value. Click on a probe to display its Delta value.

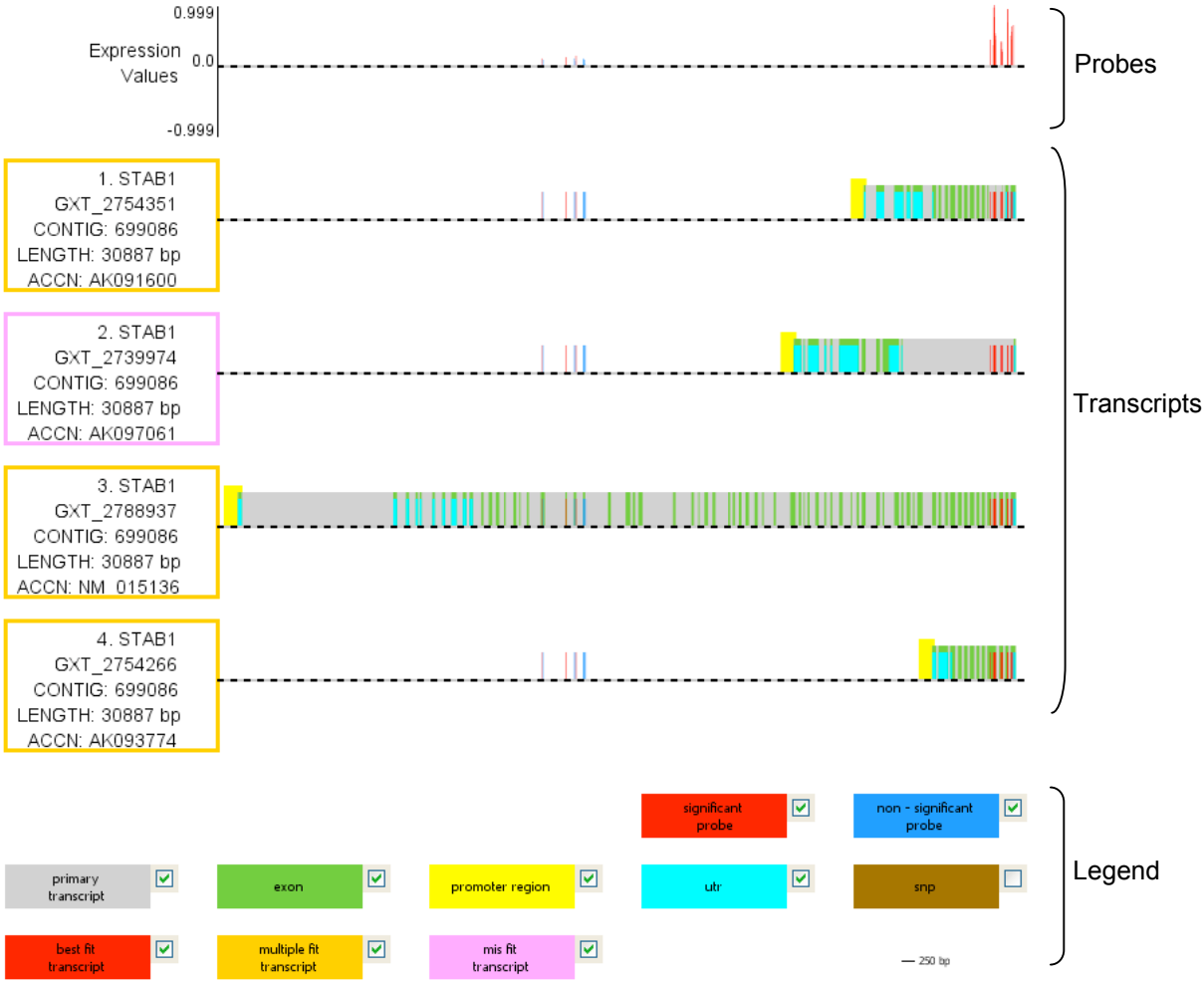
Possible probe positions depend on the chip type: for expression and exon arrays, only the probes that map to an exon of at least one transcript in the displayed locus appear, while for promoter tiling arrays, the probes mapping to an annotated promoter region or up to 1000 bp upstream of the transcription start site of a transcript are shown.

Transcripts

The transcripts of the selected locus are displayed aligned by their genomic location. The primary transcripts are grey, exons green, promoter regions yellow, and UTRs cyan. The display of SNPs (brown) is switched off by default. The gene name, Genomatix transcript ID, contig number, length, and accession number of each transcript are displayed in a frame, which is colored red if a single transcript fits the probes, orange for multiple fitting transcripts, and pink for transcripts that don't fit. The probes are displayed in-line with the transcript graph, red for significant probes, blue for non-significant probes. Clicking an element displays its start and end position, as well as its strand orientation.

Legend

The legend explains the meanings of the graphical elements. A checkbox lets you toggle the display of each element type. A scale bar is provided to facilitate estimation of the lengths of transcript elements.



Modifying and Exporting the Graph

Saving a JPEG Image

The graph view lets you export the diagram, or parts of it, in JPEG format. To export the whole diagram, click on the Export Graph ( ) button; this will open a save dialog. You can adjust the image quality in this dialog with a slider. If you want to export only a certain region of the graph, first click the Export this Region ( ) button and draw a marker frame around the region of interest. Then, use the Export Graph button as before.

Changing the Scale of the Graph

Zoom in and out of the graph incrementally by clicking the Zoom buttons. Clicking on the Change Scale ( ) button opens a dialog with a slider that lets you change the transcripts' length scale. You can use the mouse or the left and right arrow keys to use it. The Fit Graph to Window ( ) button adjusts the scale to the current window size. Scaling and zooming affects the size of the exported file.

Adding Your Own Elements

To add text to the diagram, click the Add Text () button, and then click at the position in the graph where you want to place the text. You can edit and format it in the edit panel that opens at the bottom of the window.
Clicking the Mark a Region () button and drawing a frame with the mouse adds a labeled frame to the graph. A panel at the bottom lets you format the frame and edit the label. To adjust the frame size, click and drag one of the resizing squares at its corners.
Added text or frames can be moved around in the diagram by clicking and dragging. If the grid is toggled on with the Toggle Grid () button, elements are snapped to it when you draw or drag them.

Other Formatting Options

Clicking on the Bring Item to Front () and Send Item to Back () buttons places the currently selected item in front of or behind all other elements.
If you click on a legend item, a dialog will open that allows you to change the color of all elements of the according type.

Deleting Elements

The Delete Item () button deletes the currently selected element.

Protocol

The protocol shows the parameters for the analysis, data quality assessment and any errors which may have occurred during the analysis. External protocol files can be embedded here (see “Creating a new Project” for details).

ChipInspector Online Help

Online Resources

To access the online help, click on “Help” in the ChipInspector main menu and select “Help”, or click on the Help () button in the ChipInspector main panel.

Contacting Genomatix

If you encounter any problems, please contact support@genomatix.de.

Literature

Tusher VG, Tibshirani R, Chu G (2001)
Significance analysis of microarrays applied to the ionizing radiation response.
PNAS 98, 5116-5121

Appendix: Description of the Algorithm

Probe to Transcript Assignment

The basis for ChipInspector is the Genomatix proprietary probe to transcript assignment based on mapping of all probes of a microarray against the most current version of the genome of interest. Mapping is performed on with a proprietary high-performance algorithm which is able to find exact and similar matching positions in the genome. Probes are evaluated according to the mapping result and according to the correlation analysis based on the up-to-date annotation of ElDorado. Only probes which fulfill the high-quality criteria are used for the analysis. Annotation and quality information is saved in mapping files which are an integral part of the program. The files are updated regularly with every new version of the annotation and provided for download as soon as an updated version of a genome becomes available.

Normalization

Ratios of the single probe signals are calculated and a logarithmic transformation (log2) is performed. Normalized of the ratio values are done via total intensity normalization.

Statistical Analysis

Statistical analysis is an integral part of ChipInspector. The aim of the analysis is to identify probes which show significant changes according to the experiment. Therefore at least three replicates per experiment are needed. Significant probes are discovered by a permutated T-test with false discovery rate (FDR) calculation. This approach is derived from the SAM algorithm by Tusher et al. (2001).

One Class Analysis (Experiment versus Control)

A single sided permutation T-test analysis is performed.

Multiclass Analysis

For a timeline analysis or an analysis including multiple stages a multiclass permutation T-test analysis is performed.

Presence/Absence calling

It is also possible to measure expression values relative to the average expression on the chip, e.g. for gene expression values in one specific tissue. In this case a permutation T-test analysis detecting probes which are significant above the experiment average is performed.

Projection

Significant probes are projected to transcripts using the mapping files described under 1. As default value, three significant probes are needed to detect a transcript as significant. This figure of three probes was determined empirically via spike-in experiments and proved to produce a low false positive rate while maintaining high sensitivity. However, the number of probes to define a transcript can be adapted by the user.

Cluster Analysis

For multiclass analyses ChipInspector provides an option for hierarchical clustering by calculating the Pearson distance between two data points.