MapMan Guide Overview

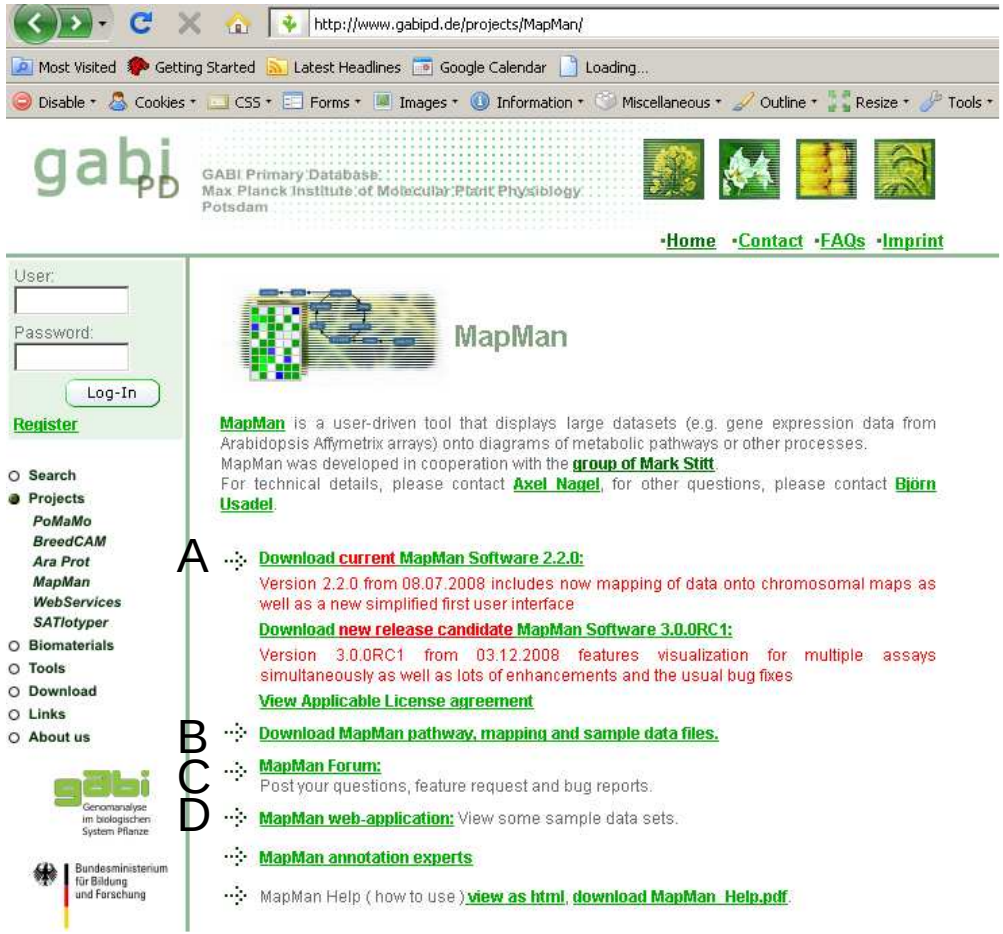
Please check
http://www.gabipd.de/projects/MapMan
For the latest version of the UserGuide

Chapter I
Obtaining MapMan and the MapMan Website

This Chapter will introduce you to the GABI MapMan Website and where to obtain MapMan

Navigate to http://www.gabipd.de/projects/MapMan/
The website offers the following:

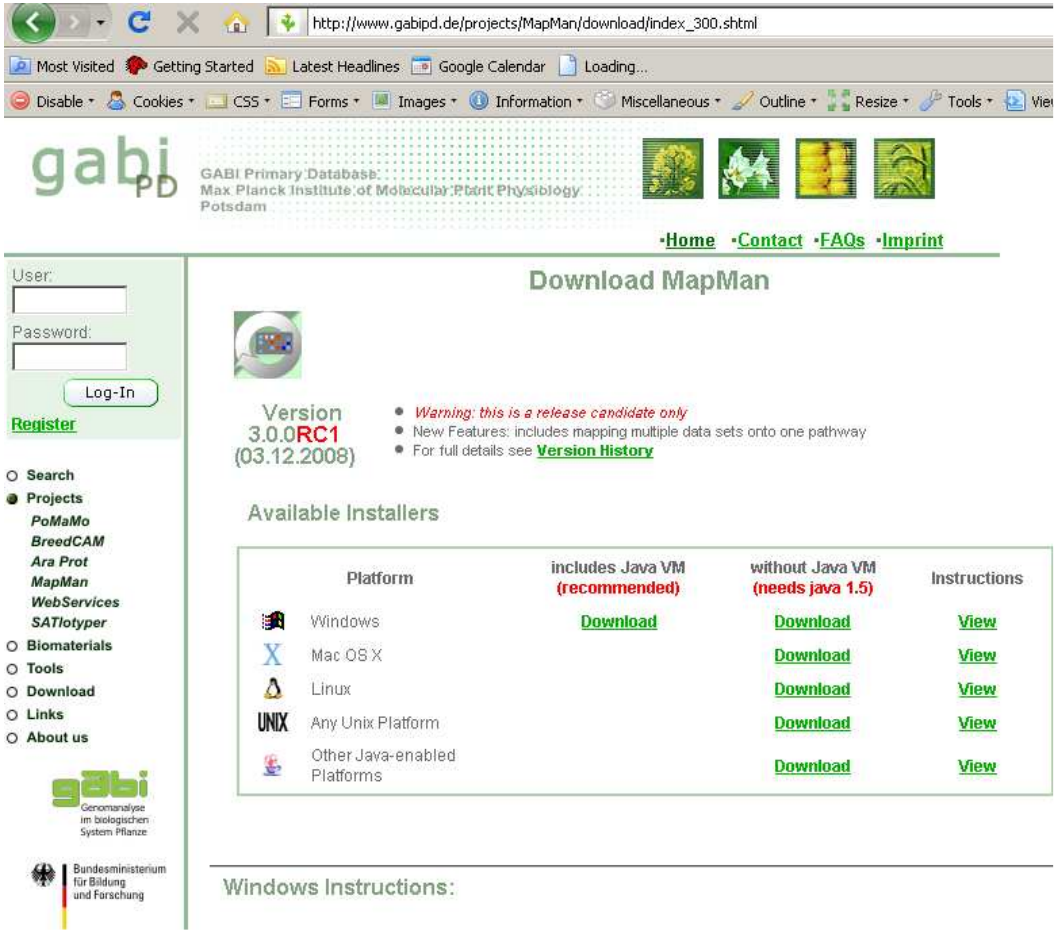
- A download of the MapMan software (A)
- A download of Mapping files and (pathway) maps (B)
- A forum where users can ask for help, share experiences and ask for new features (C)
- A web version of MapMan (D)



A Downloading MapMan

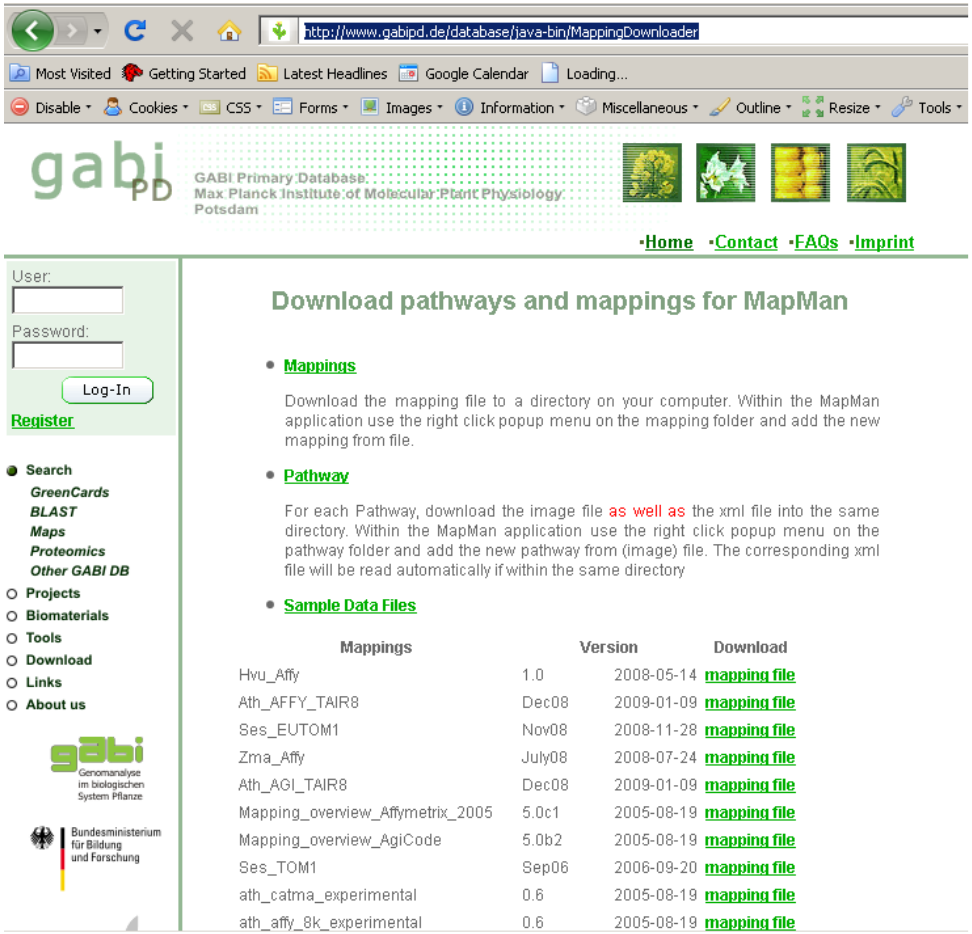
MapMan is based on Java and is thus available for all common operating system platforms. If Java is not installed, the version including Java should be downloaded.

We recommend downloading the version including the java as this combination has been tested.



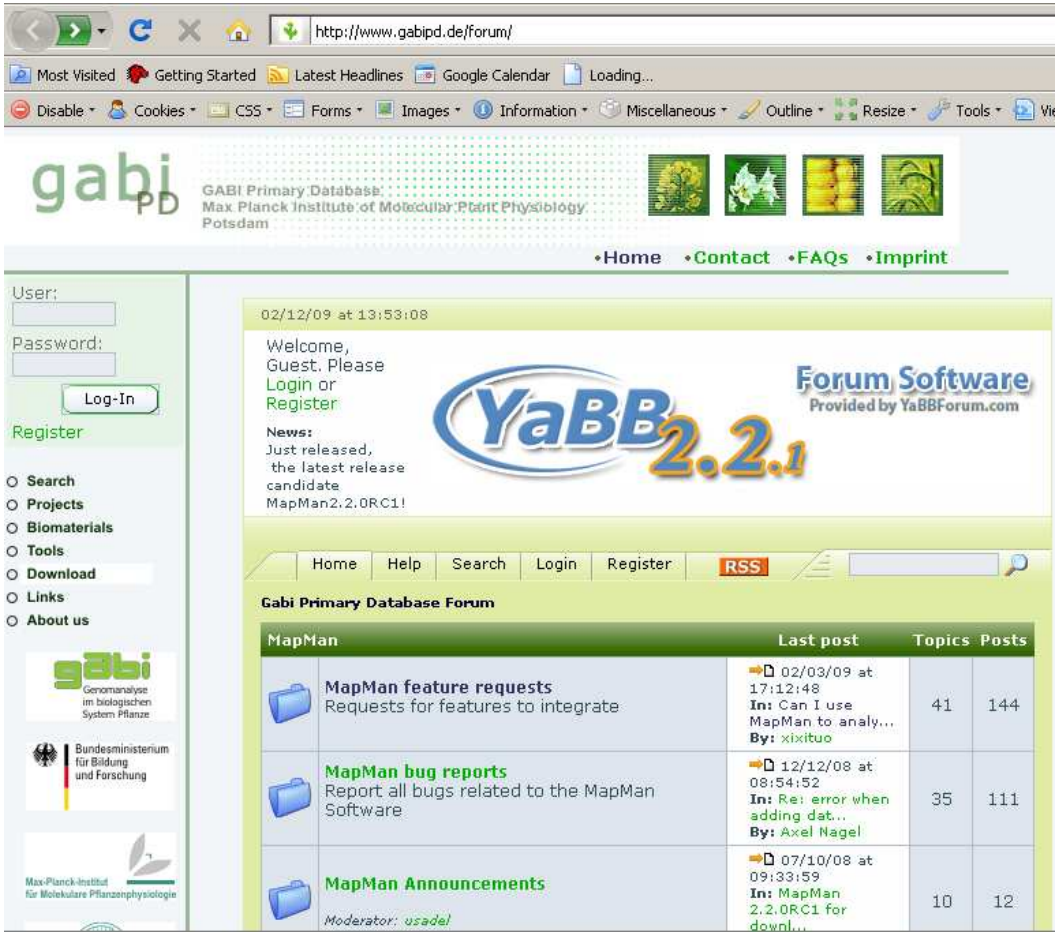
B Downloading additional data

Usually, all files (Pathway Images, Mapping Files and Sample Data files) can be downloaded and updated from within MapMan. However, for reference and demonstration purposes, individual files can be downloaded from the website directly. (This is helpful in cases where downloading is not possible from within MapMan. This could be the case if a firewall is set too restrictively.)



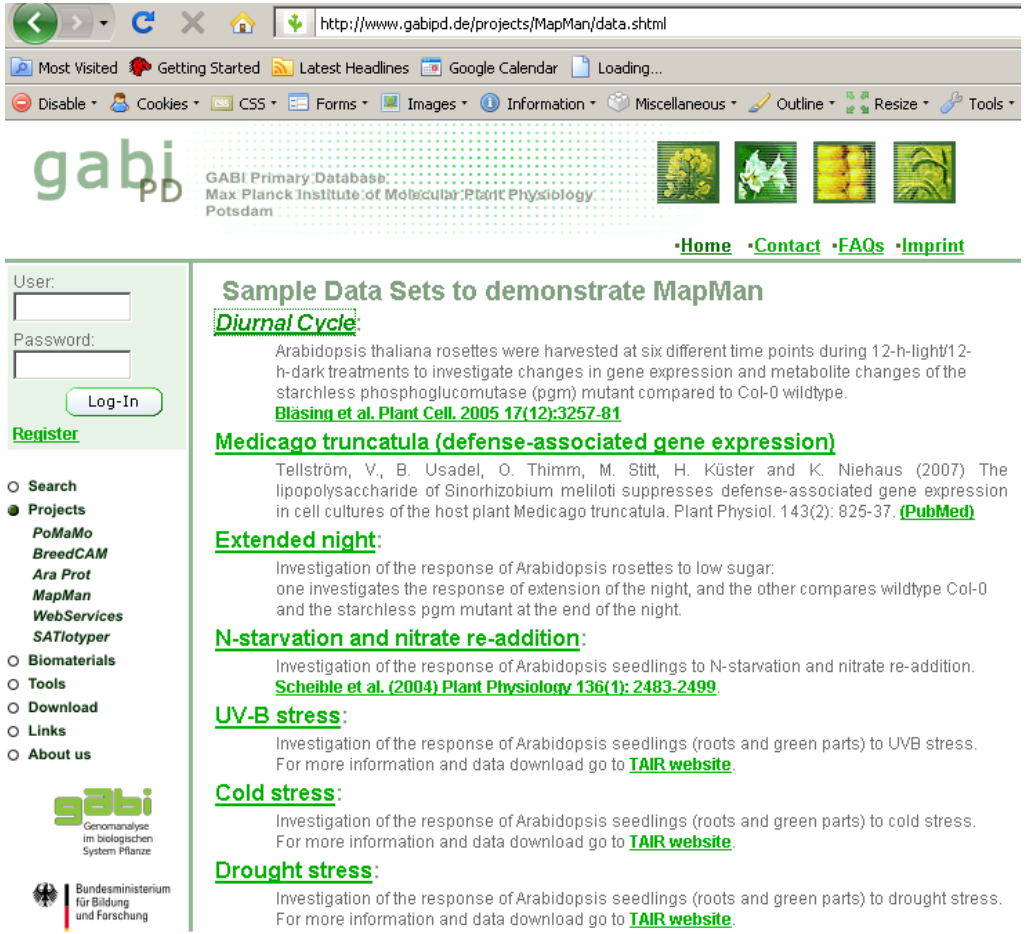
C MapMan Forum

In the forum, questions or problems can be stated. Furthermore, solutions to old questions or issues can be browsed.
However, individual help is available at any time via email.



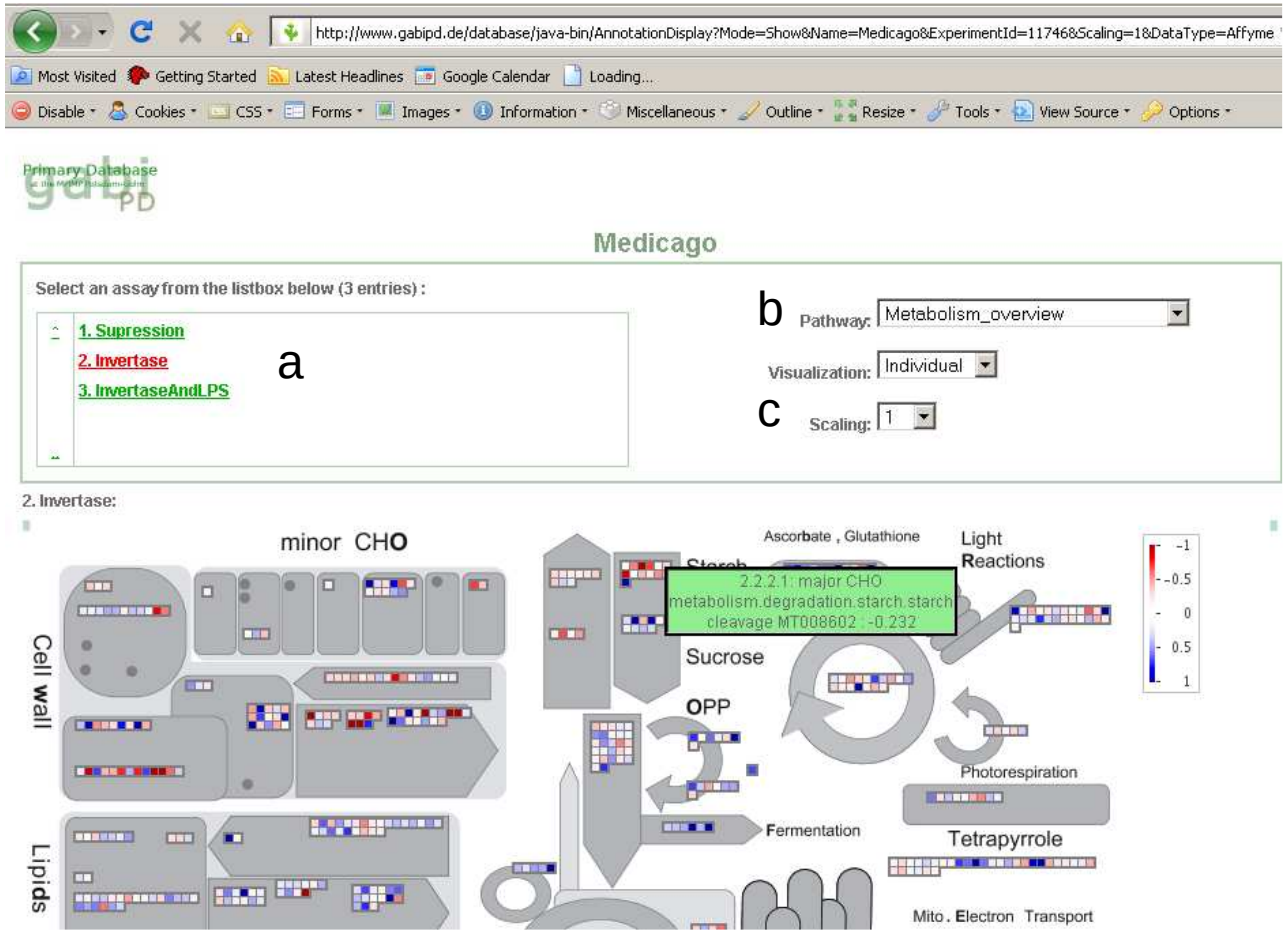
D MapMan Web-application I

The MapMan Web application illustrates some of the functionality and showcases several experiments. Clicking on a link visualizes this particular experiment.



D MapMan Web-application II

Data to be inspected (a) from an experiment comprising several arrays, the map to be displayed (b) , and the Scaling (c) can be selected interactively. Limited interactivity is provided by a mouse over function.



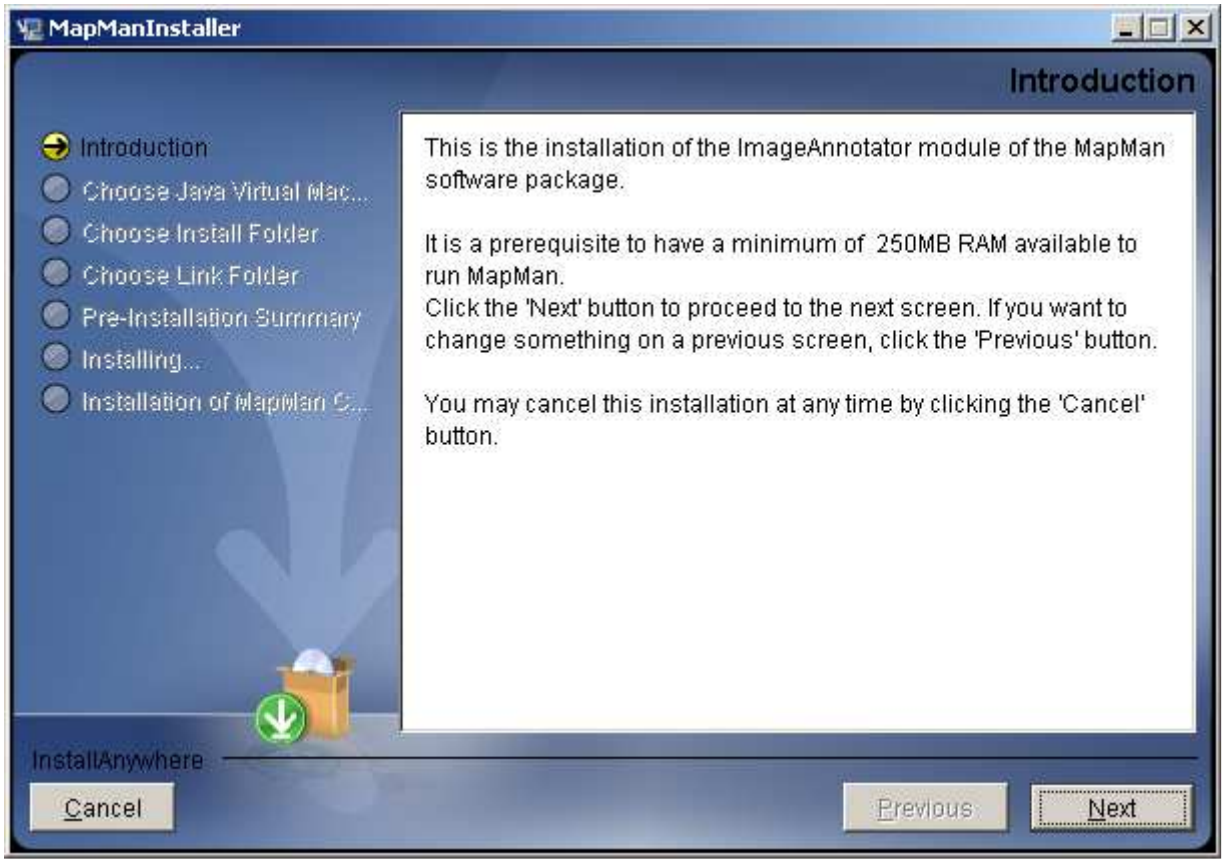
Chapter II
Installing MapMan

This Chapter guides you through the Installation of MapMan

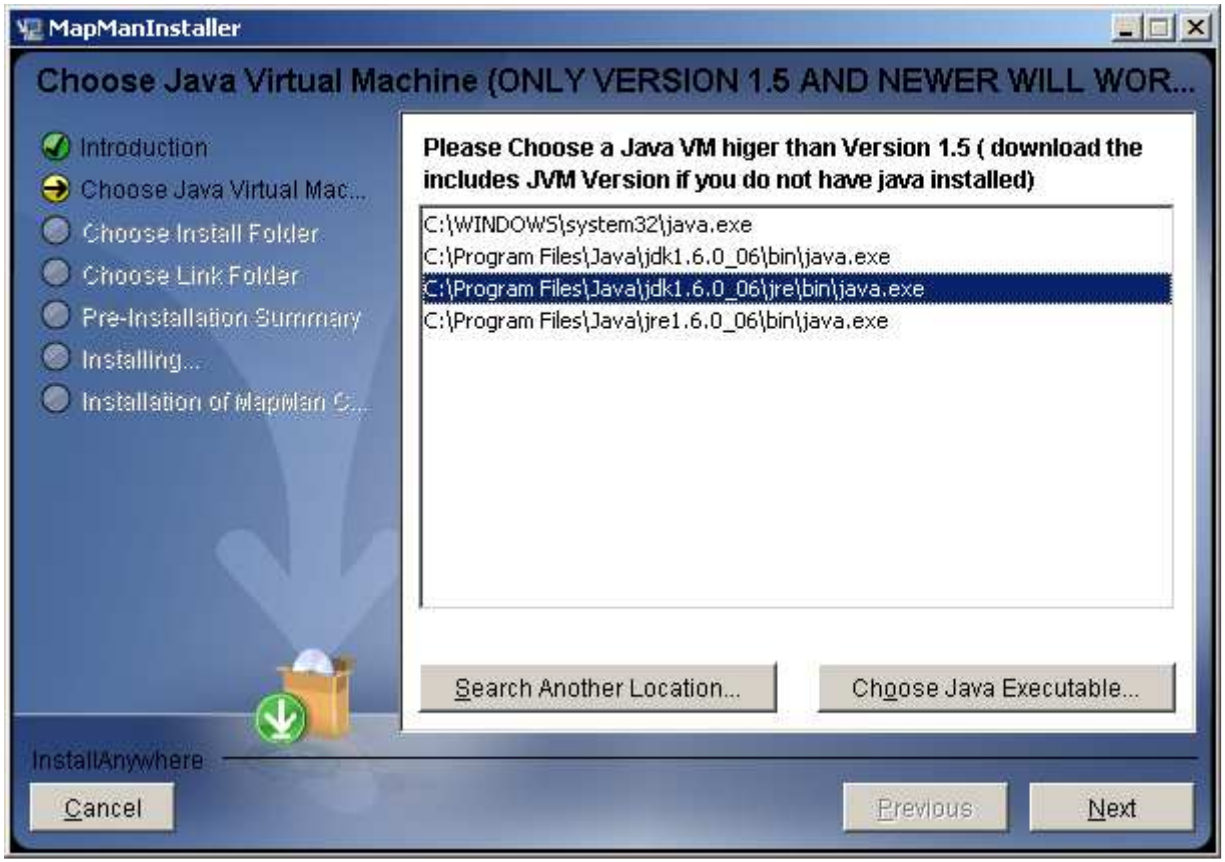
After downloading and running the MapMan installer, the user can switch the language of the installer. This doesn't affect the language of MapMan which is available in English only.

Preferences set in previous versions of MapMan will be kept in new installs, this includes files that you have linked into MapMan. If you want to start freshly, delete the file called ".ImageAnnotator.xml" from your user folder. This folder is usually called C:\Documents and Settings\yourname or it is just the "~" (/home/yourname) folder in OSX and Linux systems.

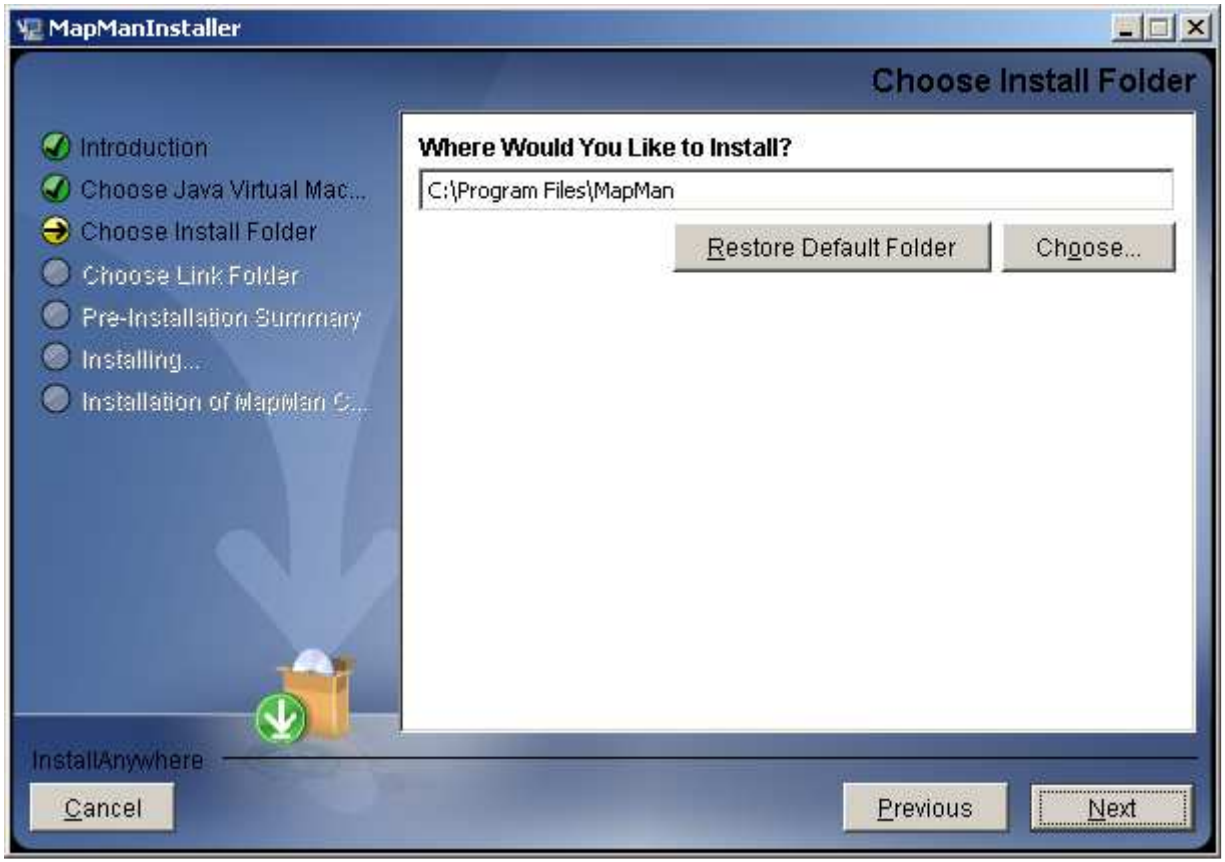
The User is greeted with a screen detailing the minimum requirements for MapMan.



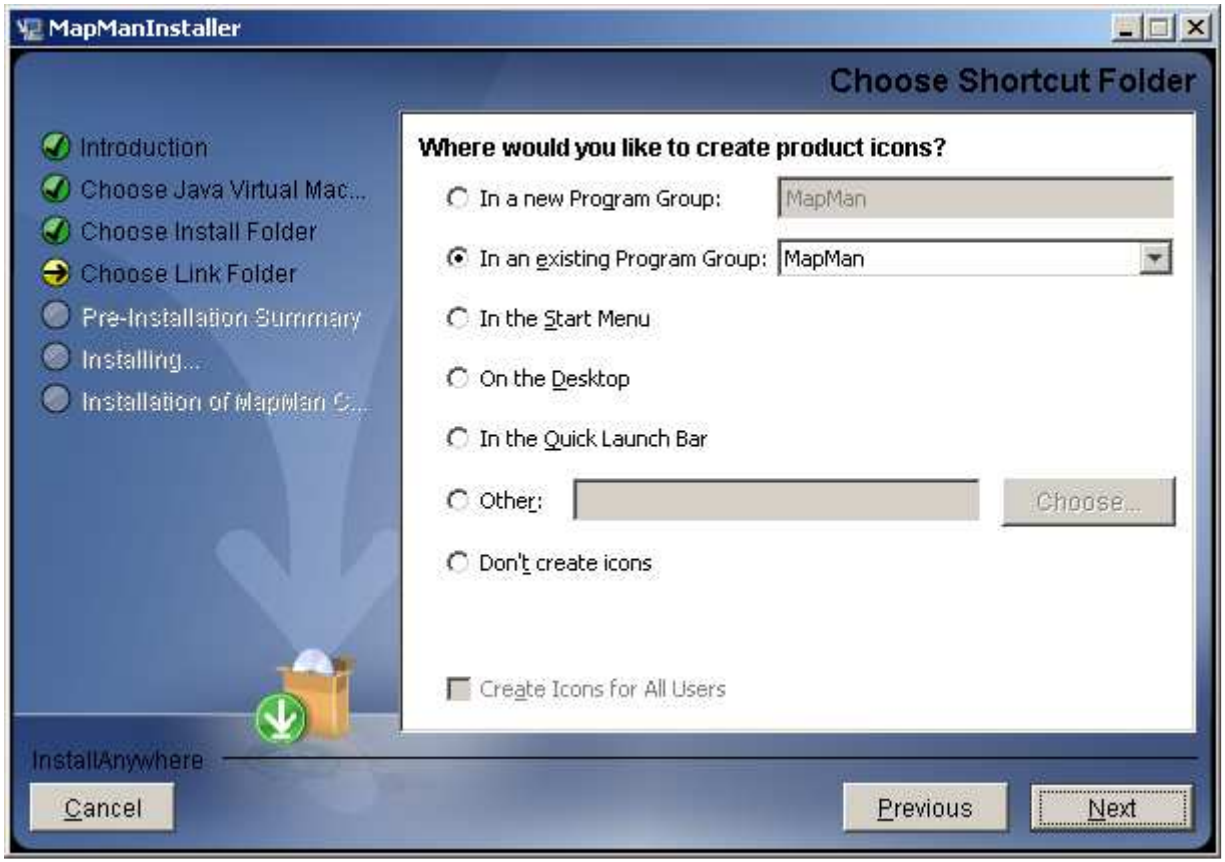
Then the user can choose an installed JAVA version. The installer locates versions of JAVA that can be used automatically. (We recommend downloading the MapMan version including JAVA and using the included JAVA. In any case JAVA version 1.5 is required as of 1/2009)



A prompt asking where to install MapMan gives the user the option to install in a non-standard or custom program folder

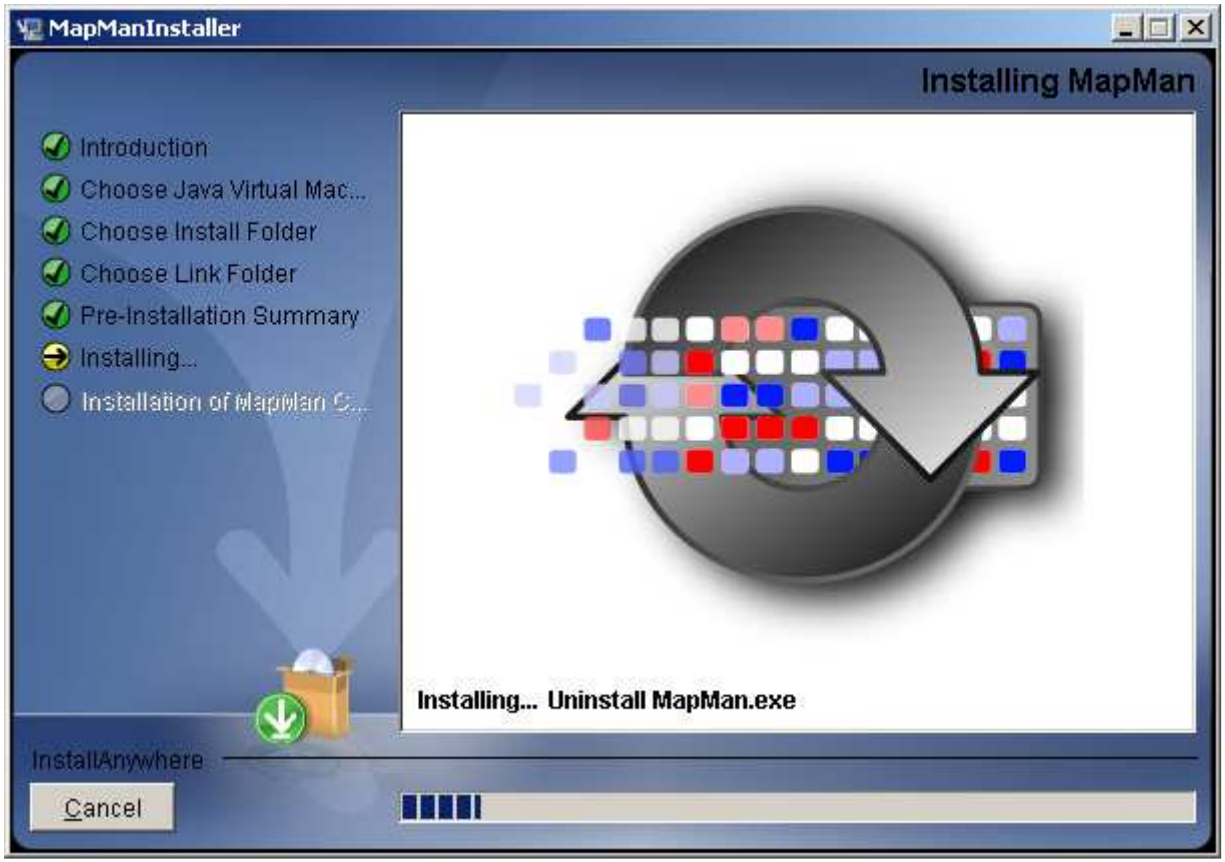


In a next step, the user can choose the program group where to install the MapMan application.

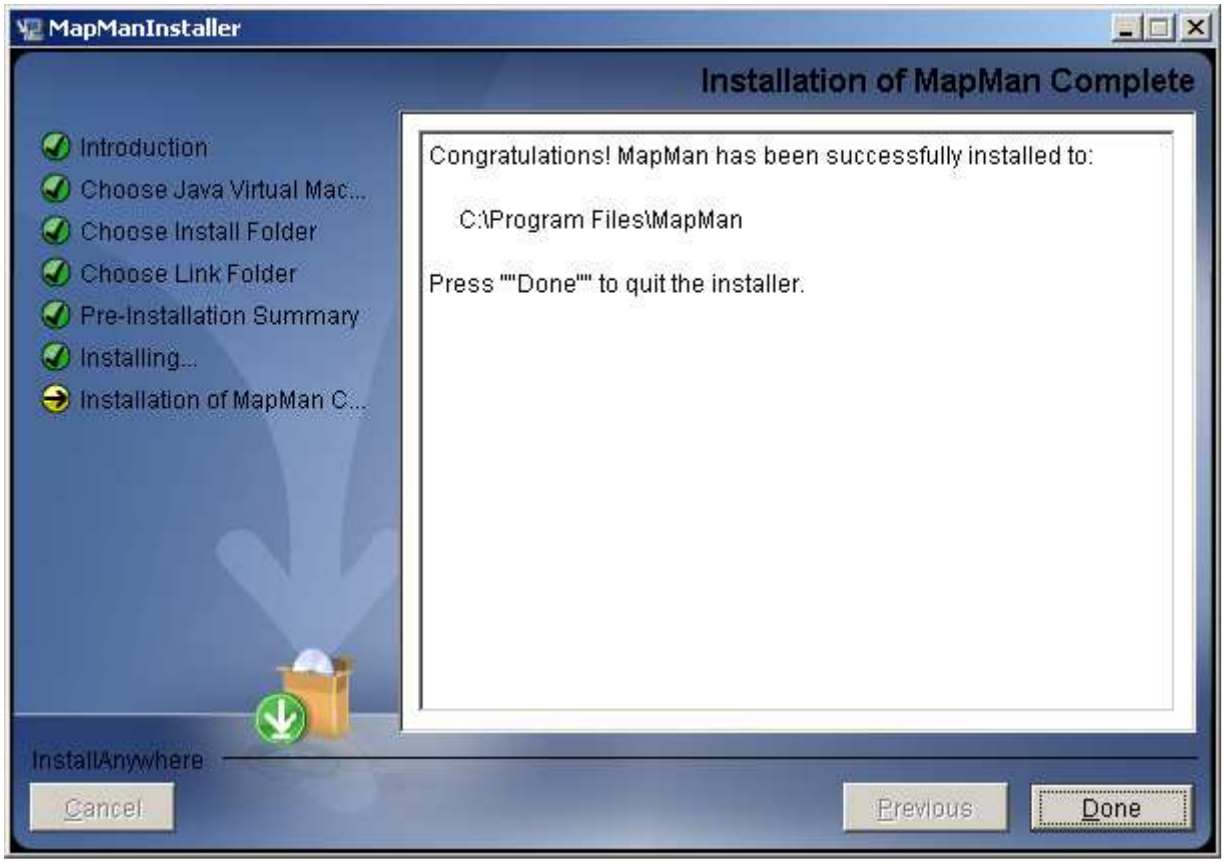


All options are then summarized.

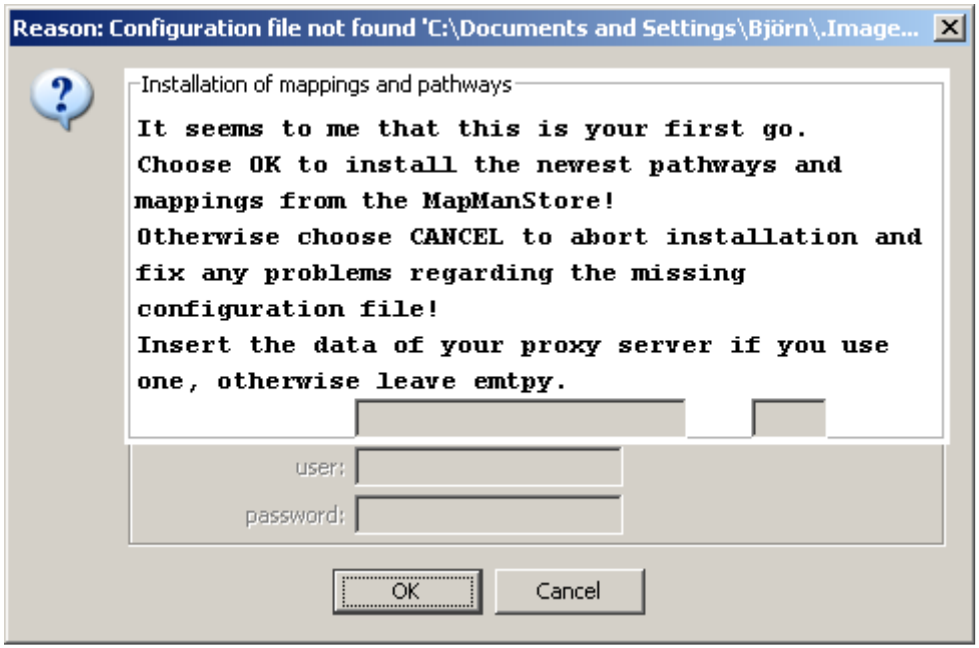
Finally, MapMan is installed on the hard disk.



Upon success, the installer displays a success message.

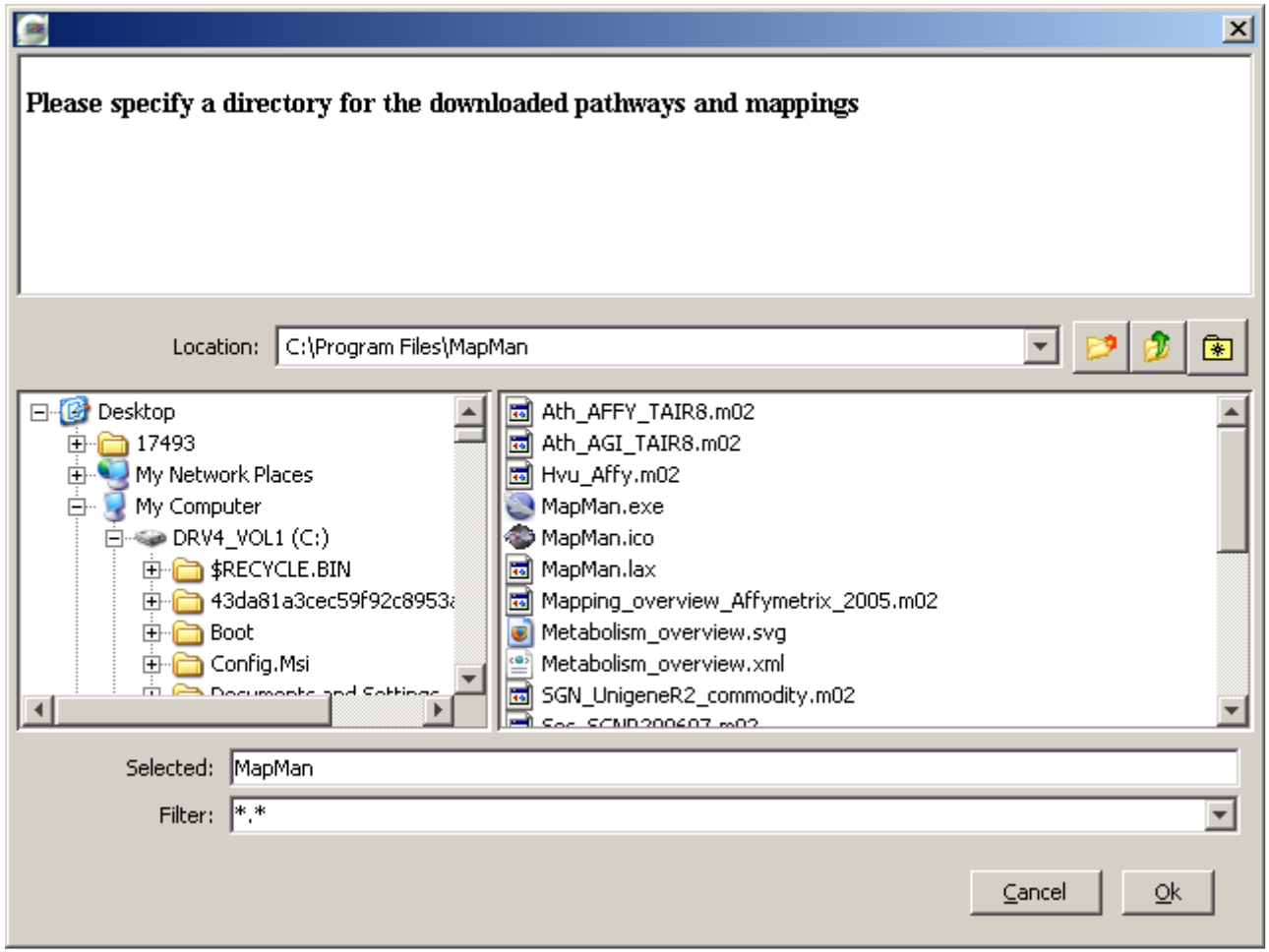


When MapMan is run the first time (and if no previous installation is found), the user is prompted for a proxy server. Most often this message can be ignored. However, some institutes use proxy-servers to channel internet traffic and to cache certain websites locally. In Internet Explorer you can go to Tools->Internet options->Connections->Lansettings to inspect your proxy settings.

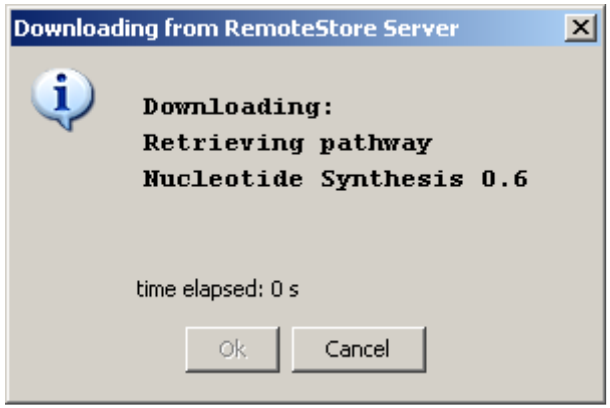


MapMan then prompts the user where to store pathway and Mapping files. As these are used for day to day usage of MapMan, MapMan won't run if these become unavailable.

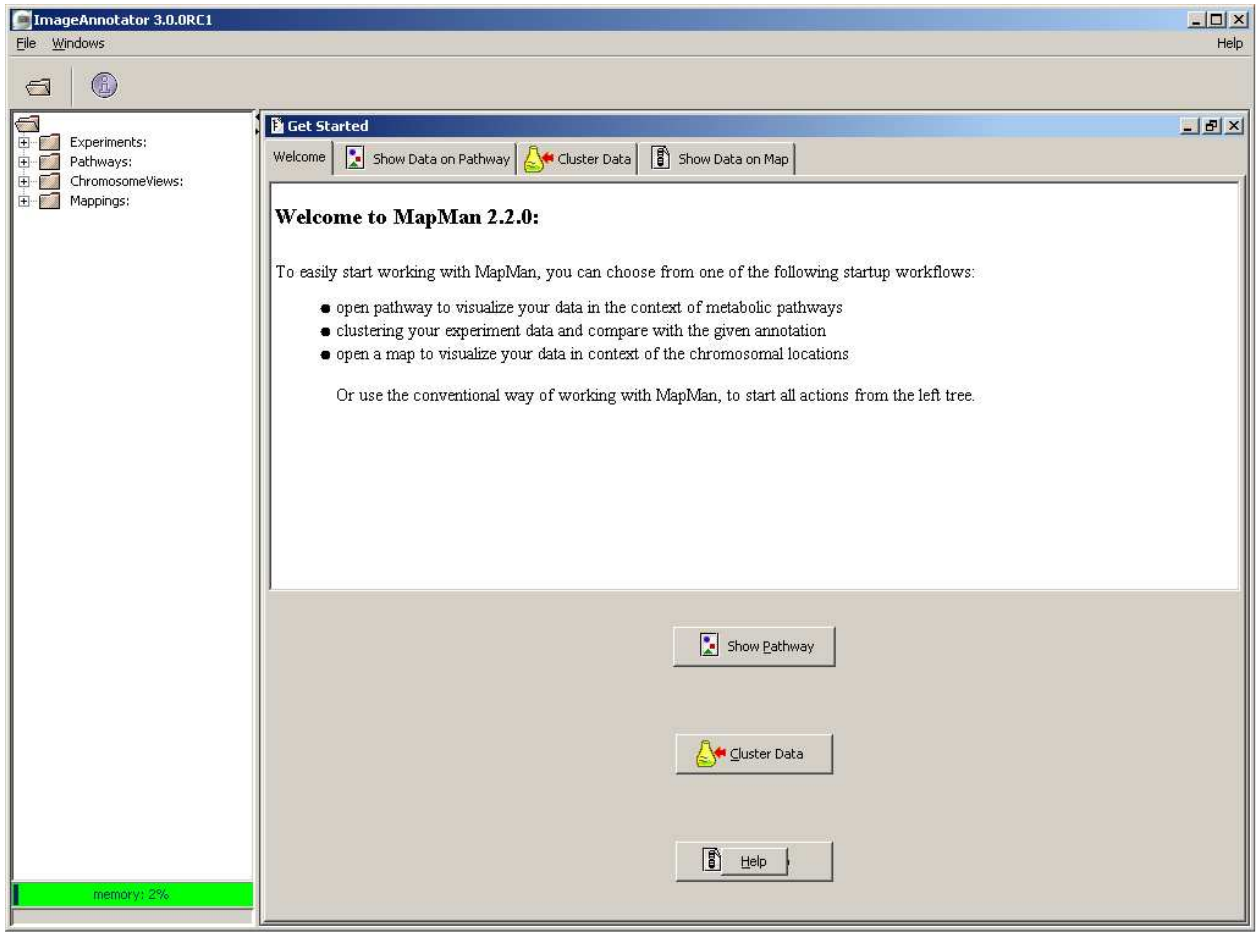
If MappingFiles are installed on a networked drive, MapMan doesn't run properly when these files are offline and cannot be accessed. Therefore, files should be put in clearly named folders and not be moved or deleted.



MapMan is then retrieving up-to date information, Mapping Files and Pathways (maps) from the MapMan website.



MapMan is ready to be used.



Chapter III MapMan Start Up

This chapter provides a basic overview over MapMan.

It demonstrates only functionalities that are immediately available after download of the ImageAnnotator software start up package, and uses only sample files contained in this package.

To inspect your own data, see Chapter V.

However, for first time user, we suggest you first train on the maps, mapping files and experimental data files provided in this start up package.

After following this chapter you will know:

- How to display built in data in ImageAnnotator
- How to adapt the display
- How to export data

Contents of ImageAnnotator after Download

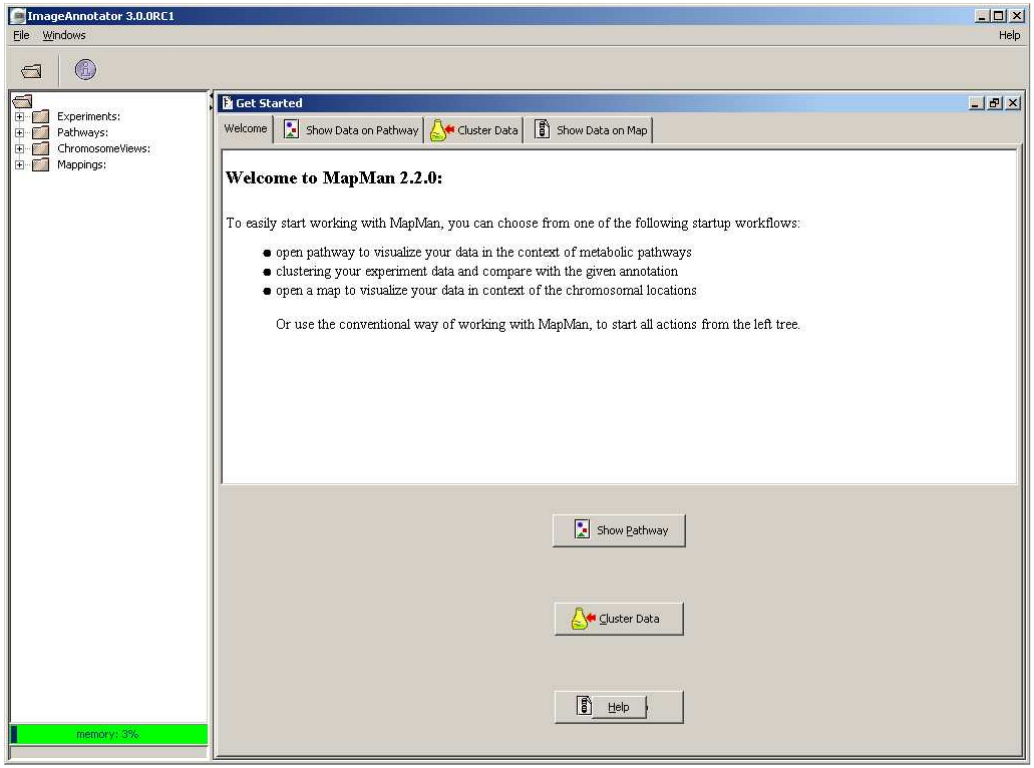
- Experimental Data Files
  - Arabidopsis files for Nitrogen, Carbon Starvation experiments
- Pathways >60
- Mapping Files
  - Arabidopis, Maize, Barley Tomato, Medicago, Potato

More content is available by right (apple) clicking on Pathways or Mapping Files

Upon start up the user is presented with the ImageAnnotator interface. On the left hand side there is a tree-like browser structure. On the right hand side data is being displayed.

The browser-like structure comprises the following items:

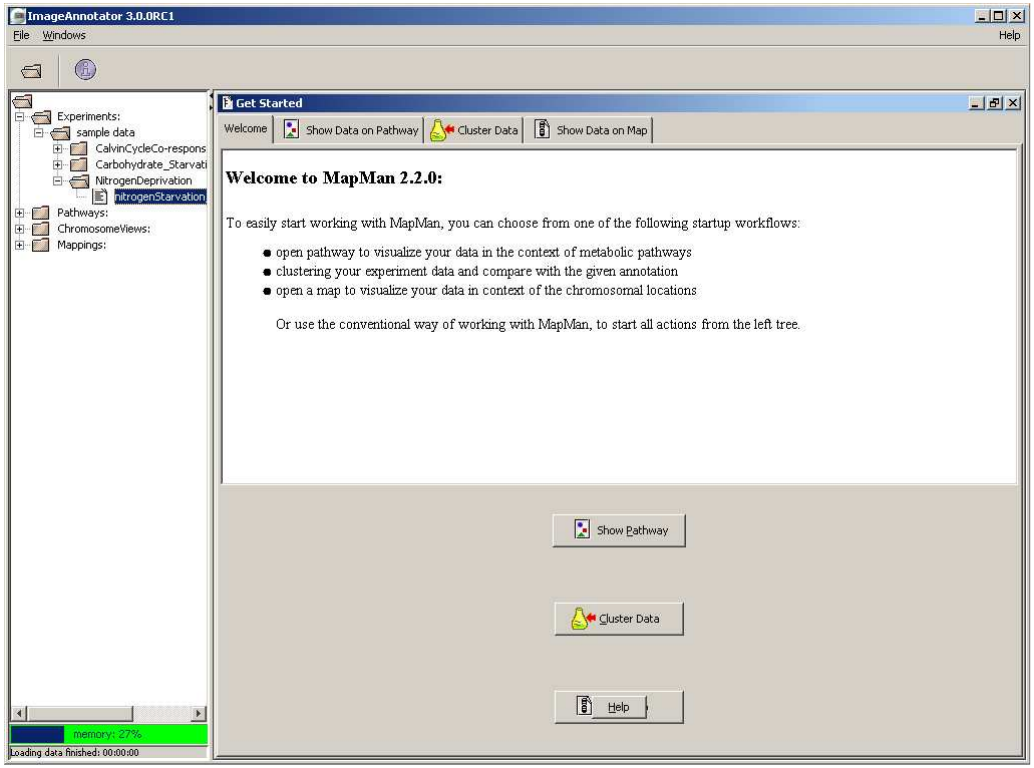
- Experiments (sample experimental data, or data imported by the user)
- Pathways (Biological Pathways or processes)
- Chromosome Views (Display of genes on Chromosomes, only for sequenced plants)
- Mappings (Files classifying transcripts, metabolites into functional classes i.e. BINS)



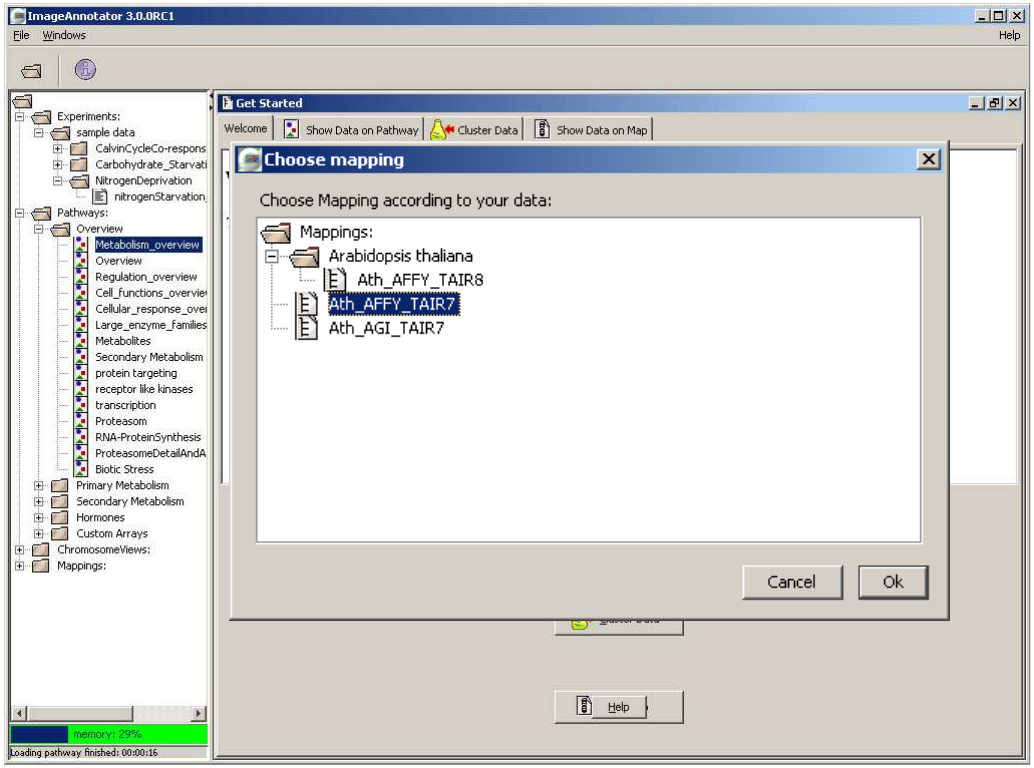
The tree structure can be browsed. As an example one might want to inspect the Experiment “nitrogen response” by expanding the Experiments folder and then the NitrogenDeprivation Folder

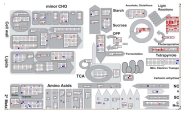
The browser like structure comprises the following items:

- Experiments (sample experimental data, or data imported by the user)
  - NitrogenDeprivation
    - Nitrogenstarvation versus full nutrition

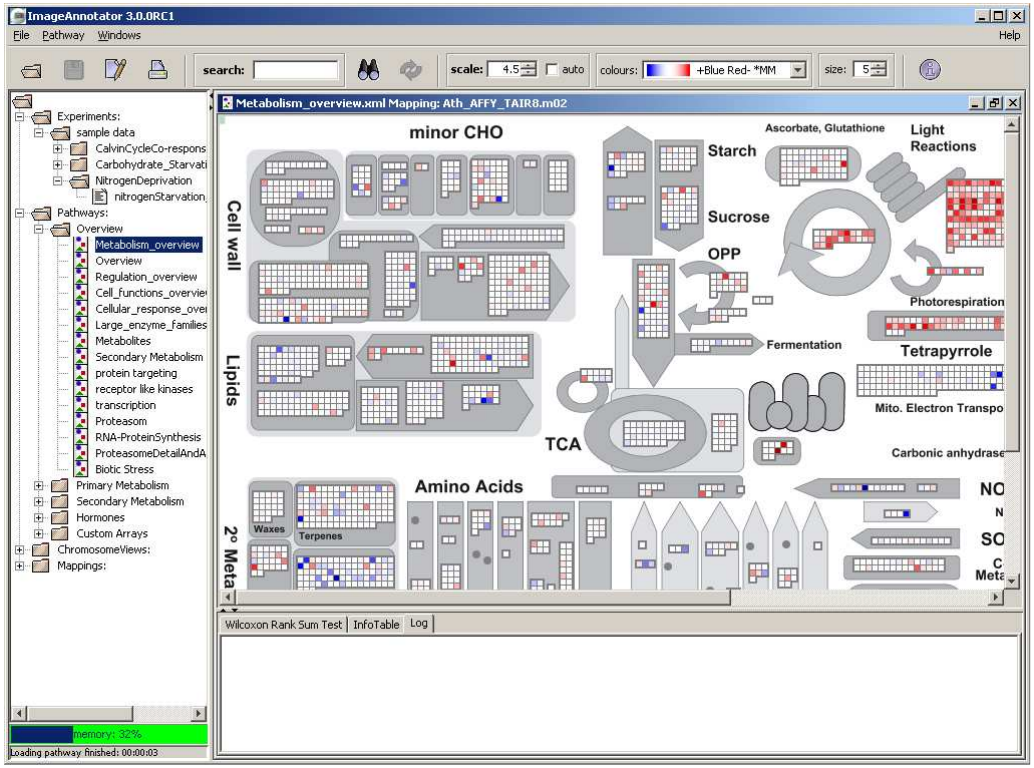


Most often, the Metabolism overview map is a good start. This map can be found in Pathways / Overview / Metabolism overview.
Double Clicking on this pathway map, brings up a dialog, asking for a mapping file. In the illustration this is an Arabidopsis thaliana (ath) Affymetrix based file, so choosing Ath_AFFY_[VERSION] is appropriate. If a particular mapping file is missing (many plants and major array platforms are supported as of 2009) one can import this into MapMan by right (apple)-clicking on Mappings and then selecting “new Mapping”. At the prompt one can choose download. If a particular platform is not available for download, either visit the forum or email usadel@mpimp-golm.mpg.de.



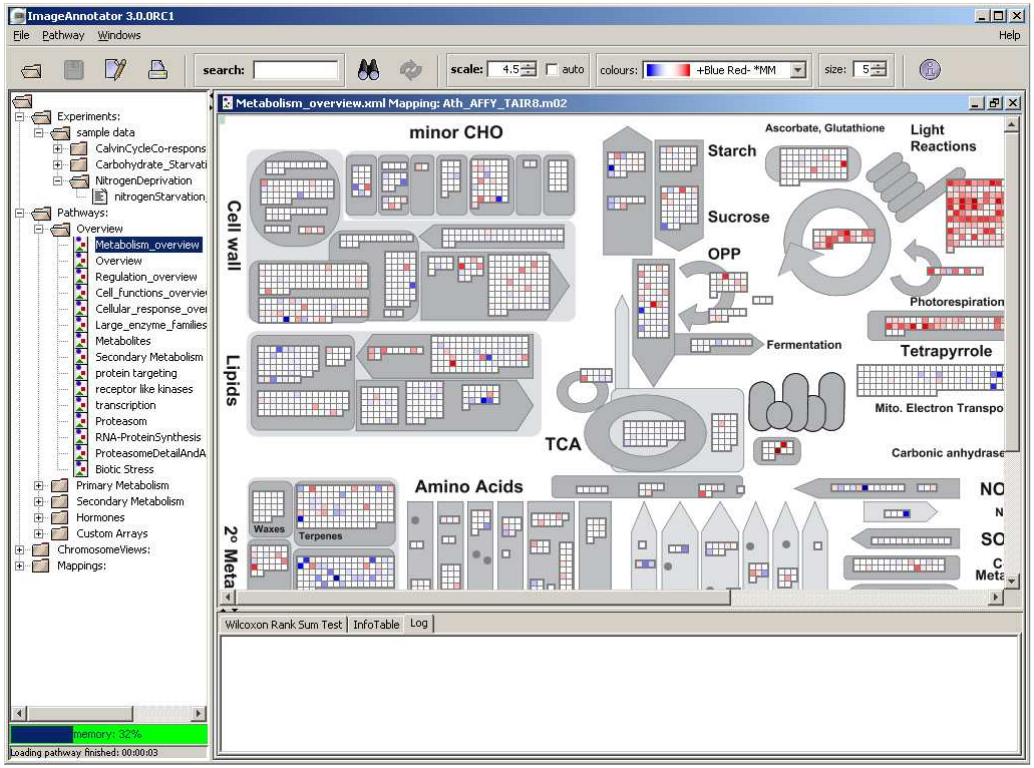
Now the experimental data is being displayed. The user is presented with the pathway display  and a log window below this display. 

In the display, each BIN or subBIN is represented as a block where each transcript is displayed as a square which is either colored blue ■ if this transcript is up- or red ■ if this transcript is down-regulated. Metabolites would be displayed as circles ● and proteins as triangles ▲.

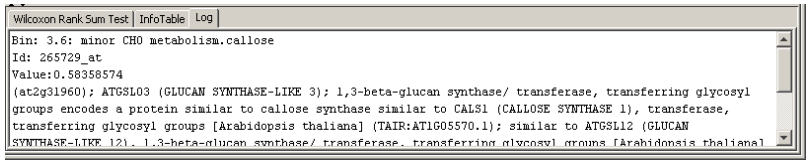


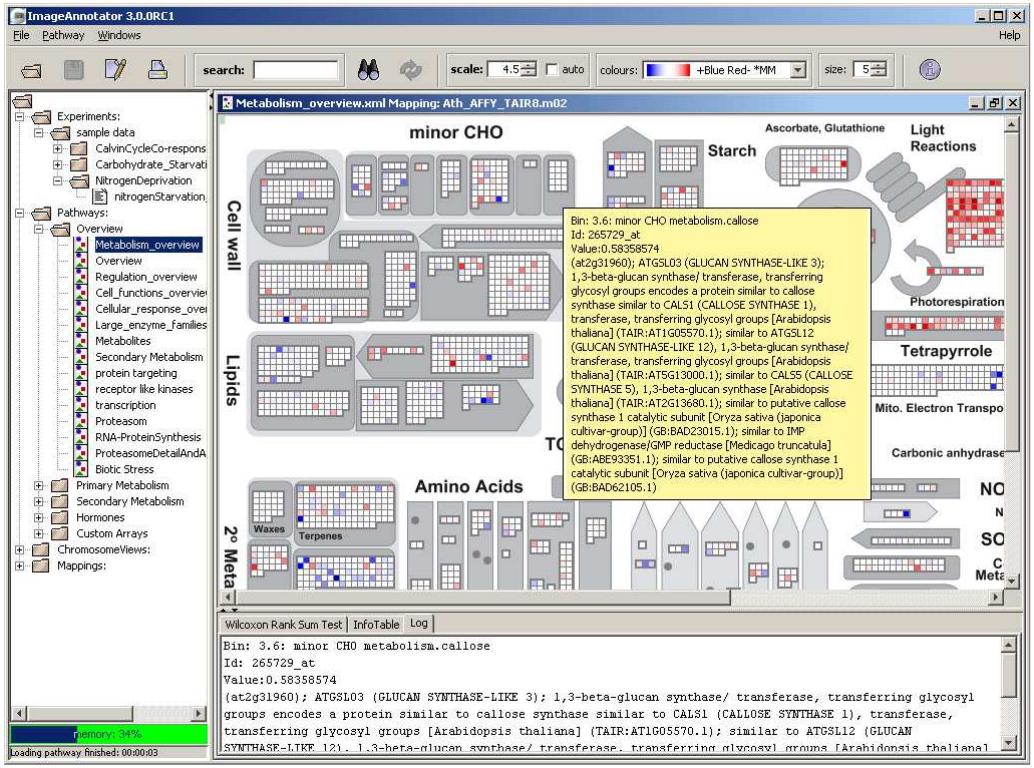
The appearance of the data can be modified by the toolbar above the pathway display 

. The user can select a different **scale** influencing colour saturation as well as the general **colour** scheme (e.g. green-red). Choosing a scale value of 4 would result in the colours getting saturated at values of 4 or -4. Finally, the **size** of each item can be decreased or increased.

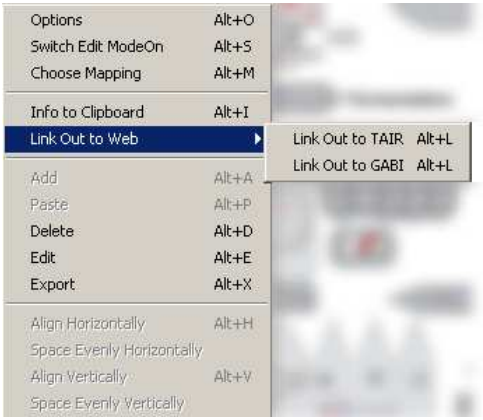


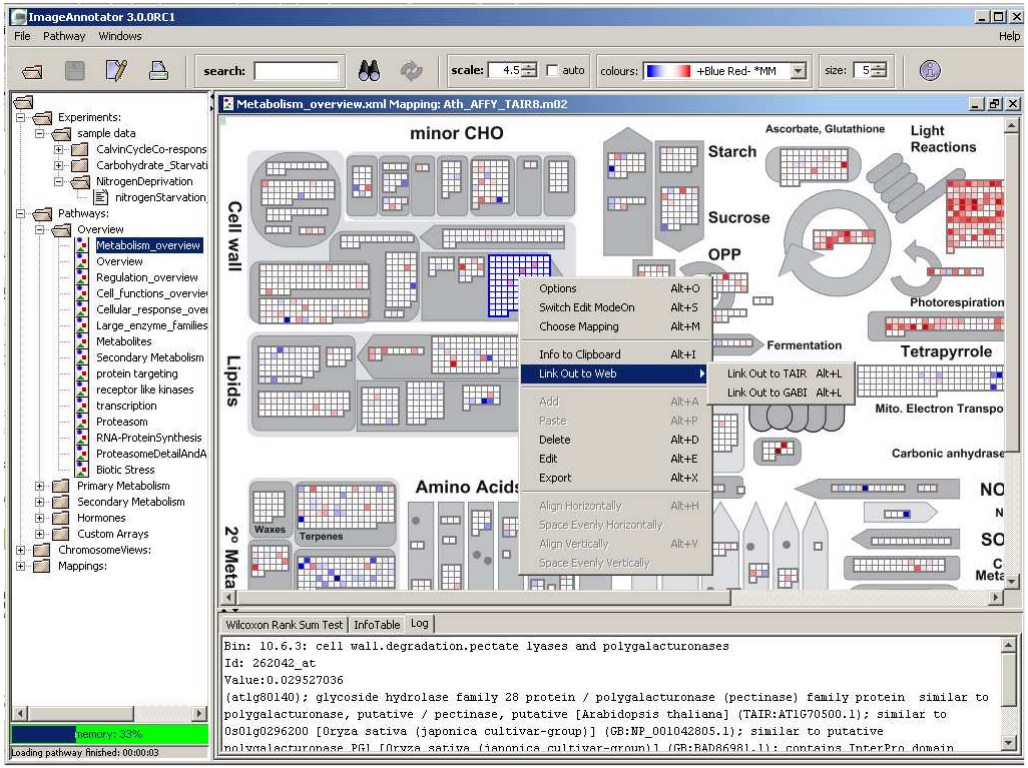
By hovering the mouse over a specific item, more information about this item is brought up in a pop-up window. Here, the specific identifier, its value as well as its annotation are displayed. This same information is available for export in the log window after clicking on a data point





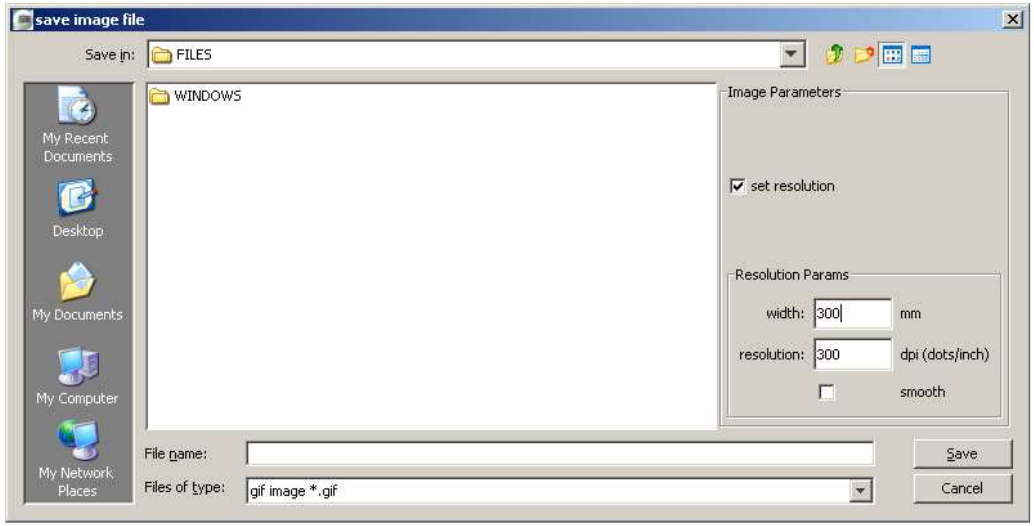
Right clicking (Apple-clicking on OSX) brings up a dialog, which enables the user to open information about this particular gene or probe (set) in a web browser. For this, relevant data sources such as TAIR (Arabidopsis) or SGN (tomato, potato) can be chosen.



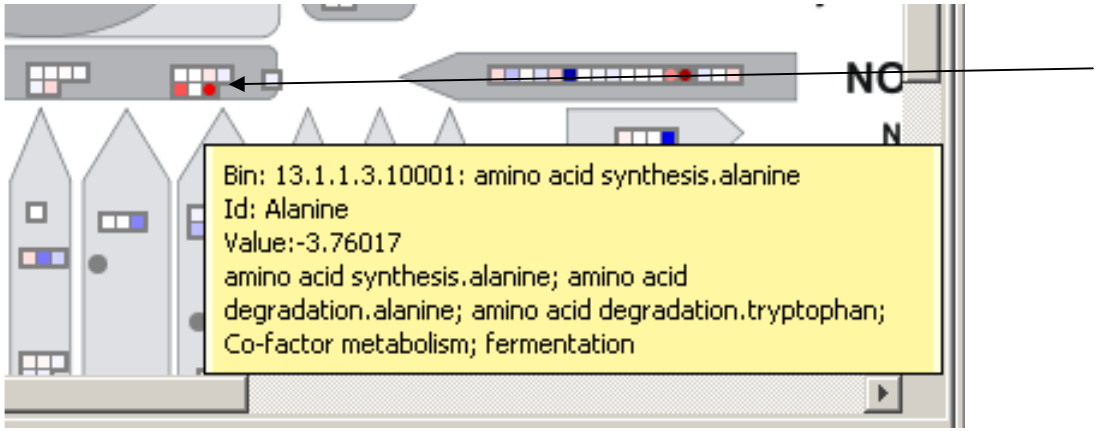


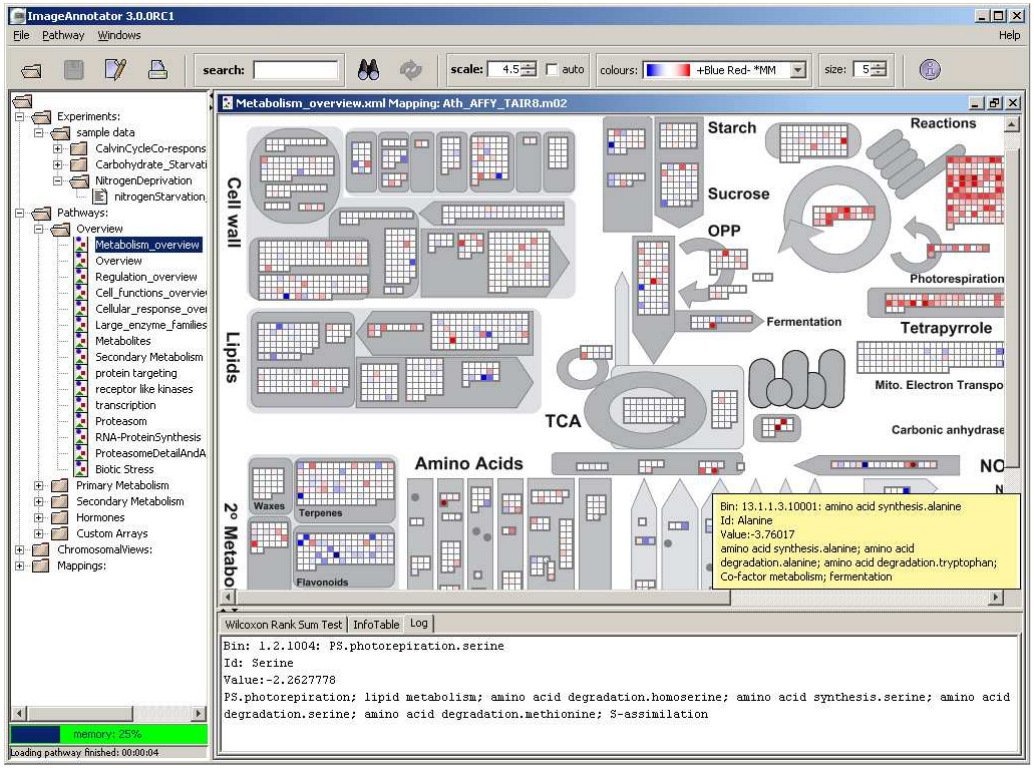
Selecting Pathway-> Save as Image brings up an export dialog, for saving the currently displayed image. Publication quality images can be generated by checking “set resolution” and giving dimension and resolution (e.g. 600 dpi).

Depending on the source of the background image (e.g. scanned textbook images or photomicrographs) the resulting figure might still look jagged which can, to some extent, be controlled by applying the “smooth” filter. Most simple Pathway images in MapMan are either available in a format that allows ultra-high resolution output or are being made available in such a format.



Combining different types of omics data. Different kinds of omics data can be mixed. In fact, the sample data already contains transcripts and a limited amount of metabolite data displayed as circles.





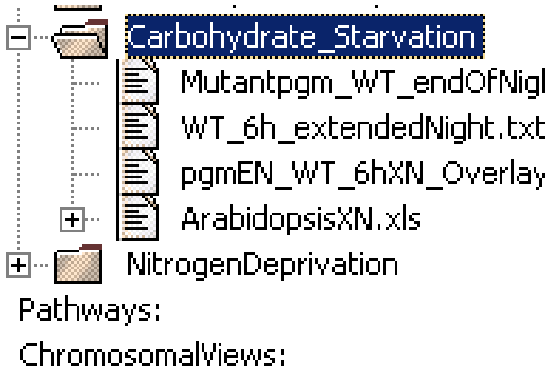
Chapter IV MapMan Statistics

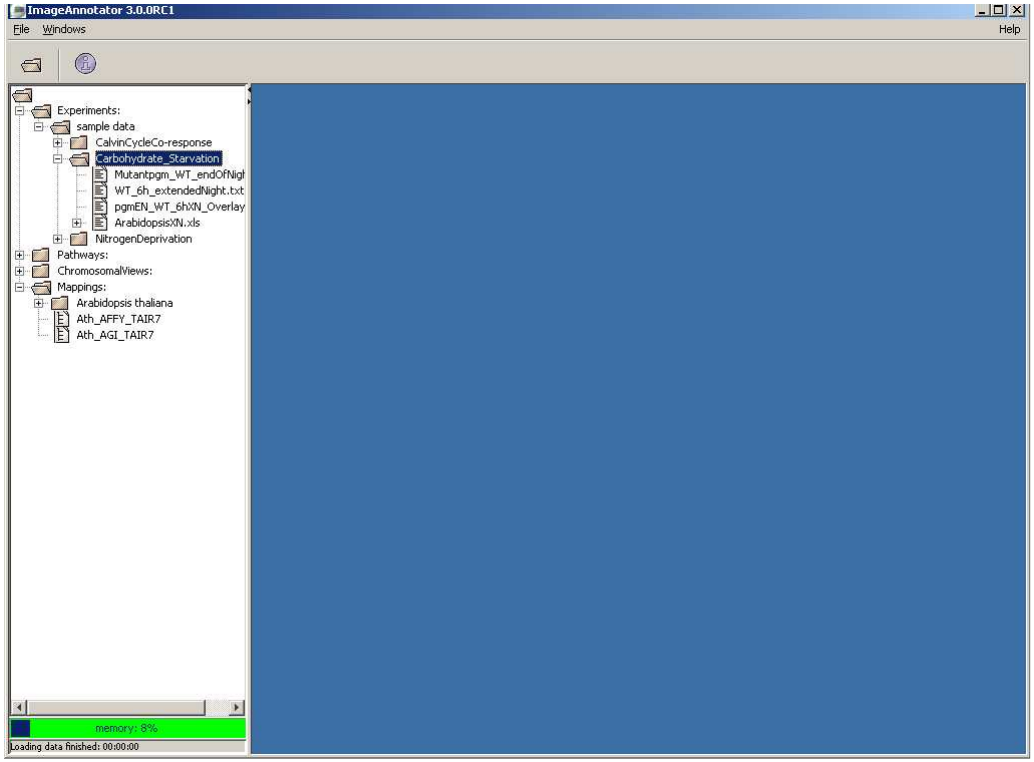
This chapter provides a basic overview of ImageAnnotator's capability, how to handle data for filtering, and how to use ImageAnnotator for calculating your own statistical tests and how to cluster data

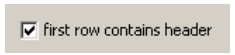
After following this tutorial you will know

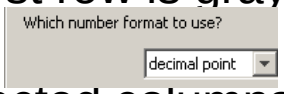
- How to ***configure filters*** based on e.g. p-values
- How to ***cluster*** data using KMeans
- How to perform a ***Wilcoxon Test on categories***

Within the Experimental data, navigate to the Experiment ArabidopsisXN in the folder Carbohydrate_Starvation. In front of this Experiment a little plus indicates, that this file contains multiple experimental conditions. Right click (Apple-click) on this file and select “Configure Dataformat”.

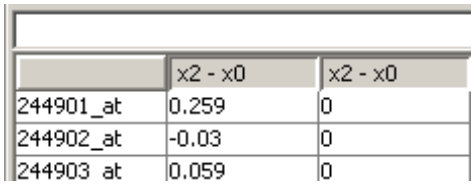




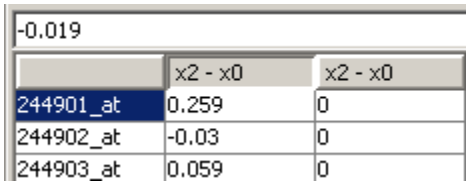
The configuration dialog shows the first few lines of the loaded file. MapMan tries to auto-configure the loaded experimental file. Here, MapMan recognized that a header is present (first row contains header 

and the first row is grayed out) and that the data format is using a "." as numerical separator 

. Furthermore, all columns are selected (the header of selected columns is pressed in). They can be de- and reselected by clicking on their header.

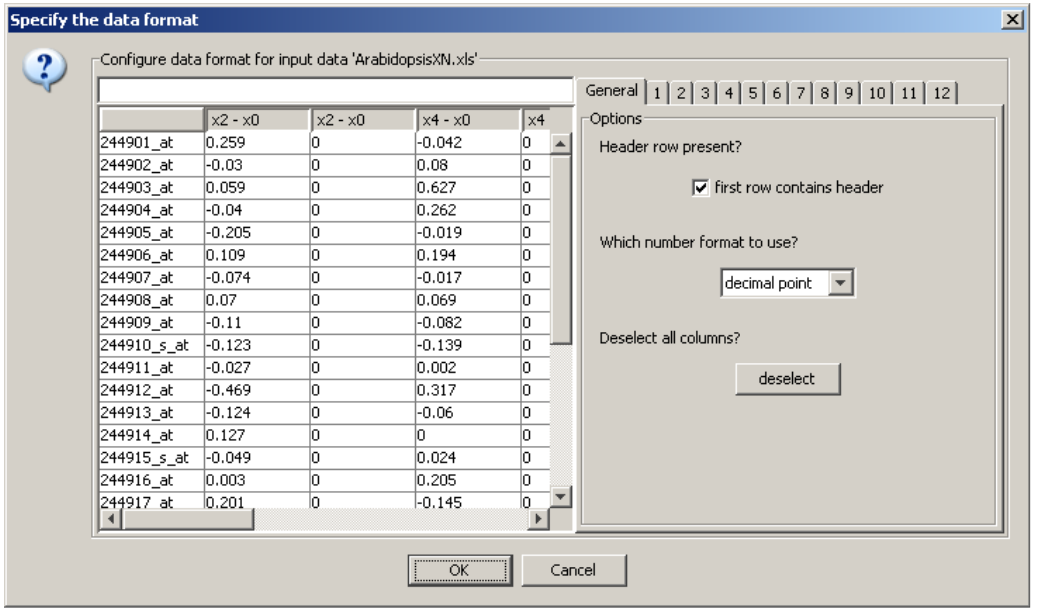


	x2 - x0	x2 - x0
244901_at	0.259	0
244902_at	-0.03	0
244903_at	0.059	0



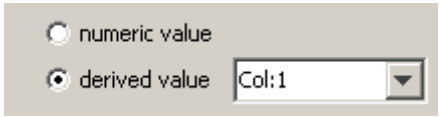
-0.019

	x2 - x0	x2 - x0
244901_at	0.259	0
244902_at	-0.03	0
244903_at	0.059	0



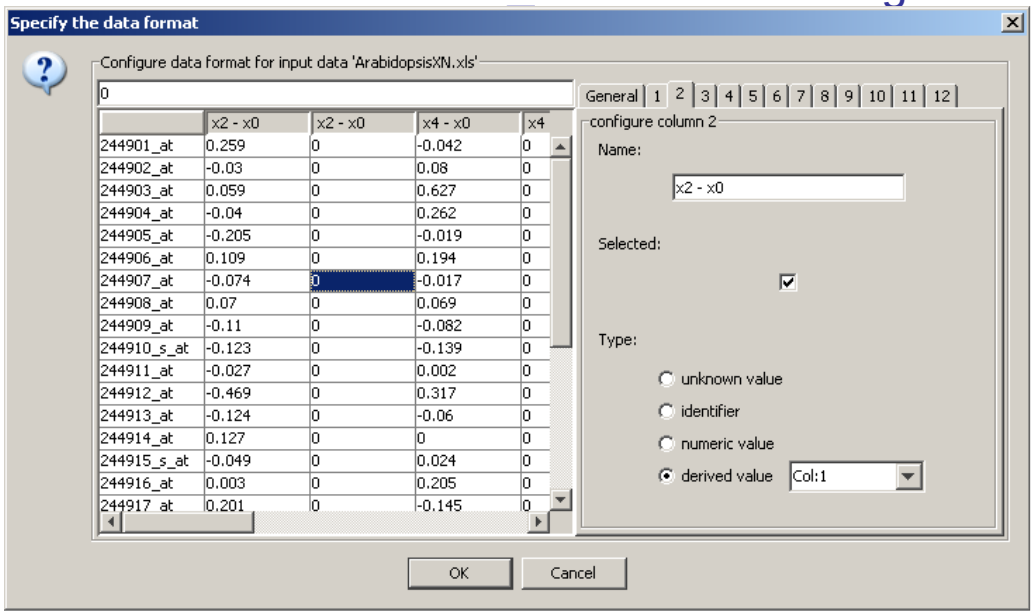
	x2 - x0	x2 - x0	x4 - x0	x4
244901_at	0.259	0	-0.042	0
244902_at	-0.03	0	0.08	0
244903_at	0.059	0	0.627	0
244904_at	-0.04	0	0.262	0
244905_at	-0.205	0	-0.019	0
244906_at	0.109	0	0.194	0
244907_at	-0.074	0	-0.017	0
244908_at	0.07	0	0.069	0
244909_at	-0.11	0	-0.082	0
244910_s_at	-0.123	0	-0.139	0
244911_at	-0.027	0	0.002	0
244912_at	-0.469	0	0.317	0
244913_at	-0.124	0	-0.06	0
244914_at	0.127	0	0	0
244915_s_at	-0.049	0	0.024	0
244916_at	0.003	0	0.205	0
244917_at	0.201	0	-0.145	0

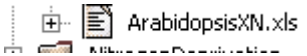
A column can be activated by clicking on any of the measured values. Typical values for columns are “numeric” which are used to display data or “derived value” which are used to filter out data in other columns. This is set-up by selecting the derived value option and then selecting which column it should be derived from. In the present case the order in the file is always: log2 Fold-change, flag for significance. E.g. the first column is the log2 fold change of an extension of the night by 2h versus the end of the night and the second column is a flag giving if this is significantly changed (1) or zero (0) otherwise. Thus, one always sets every second column as derived of the immediately preceding column.

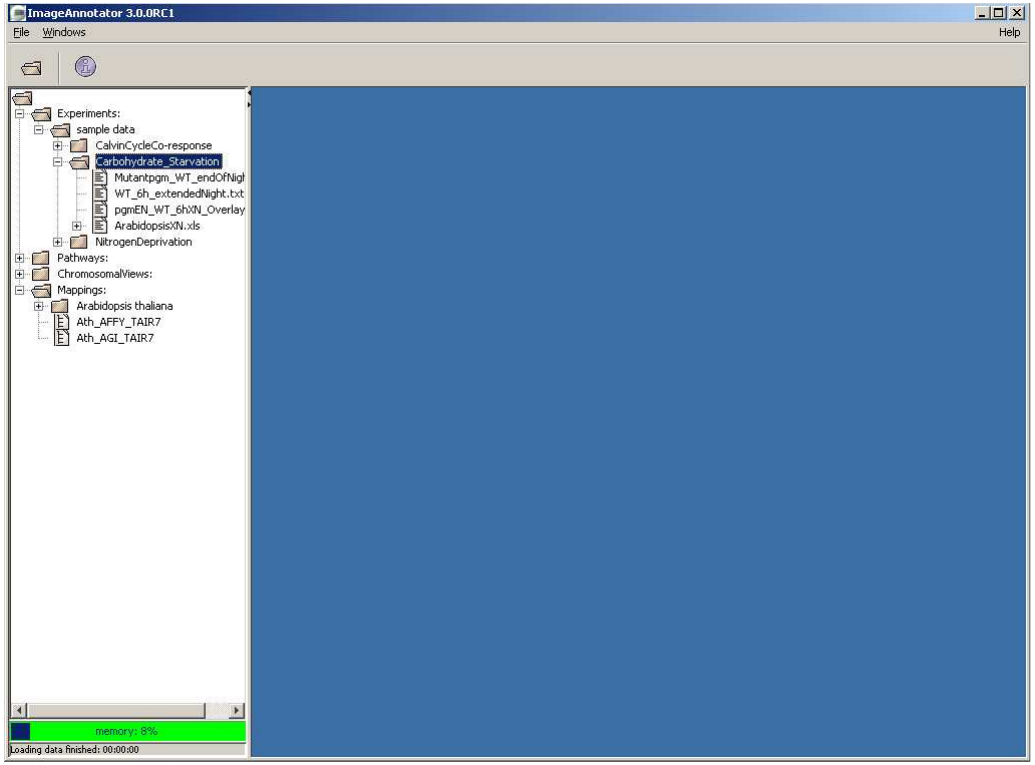


To avoid setting up the file: MapMan, recognizes columns as derived if they follow a numerical value and if they have the same name as the one before preceded by “d_”. Here, one could have used “x2-x0” for the values and “d_x2-x0” for the flag.

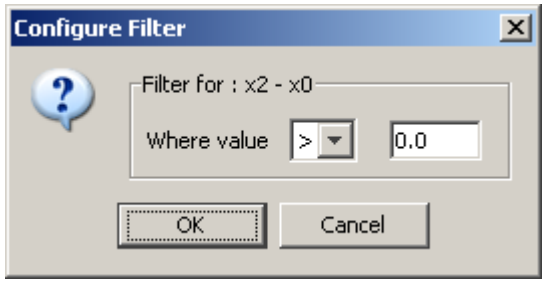
x2 - x0	x2 - x0
0.259	0
-0.03	0
0.059	0
-0.04	0
-0.205	0
0.109	0
-0.074	0

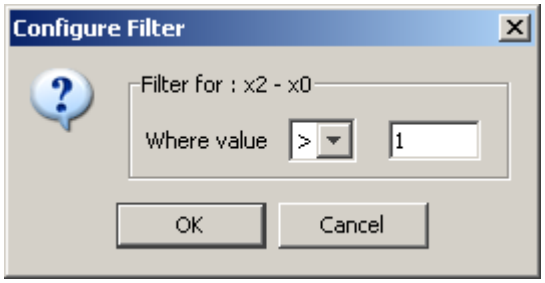


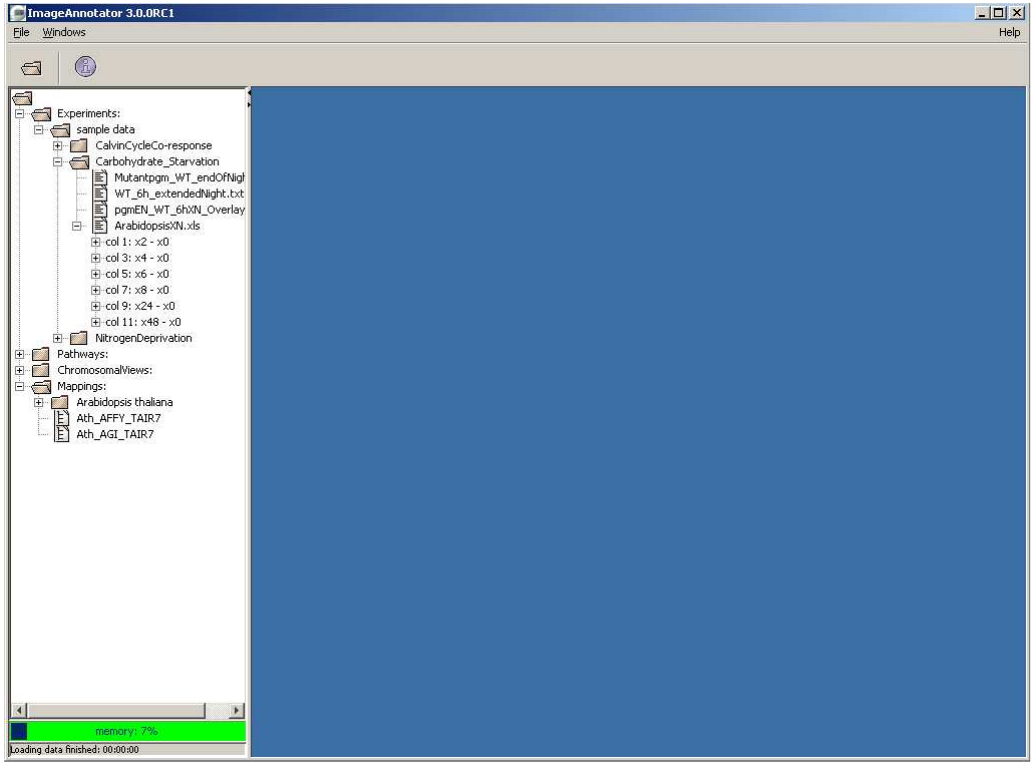
After defining the columns, a click on the (+)  left of the experiment ArabidopsisXN reveals the underlying log2 FC columns. Every one of them is marked with a plus (+), which reveals a filter icon upon clicking.



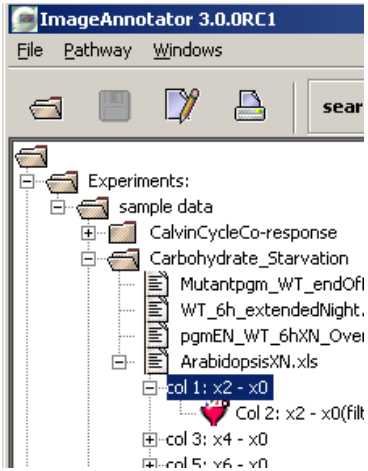
Right Clicking (Apple-clicking) on the filter columns brings up a dialog, specifying a filter that can be activated. Pressing Cancel here will delete activated filters. Let's set the filter to 1.

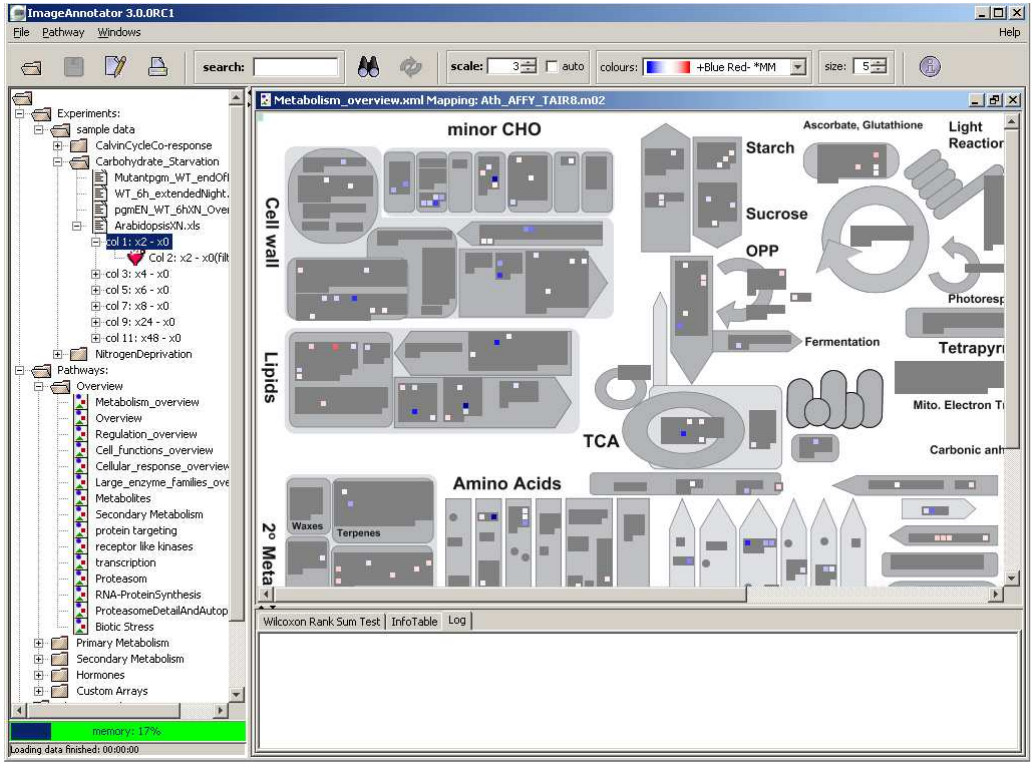




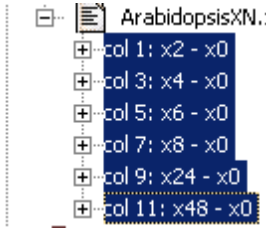


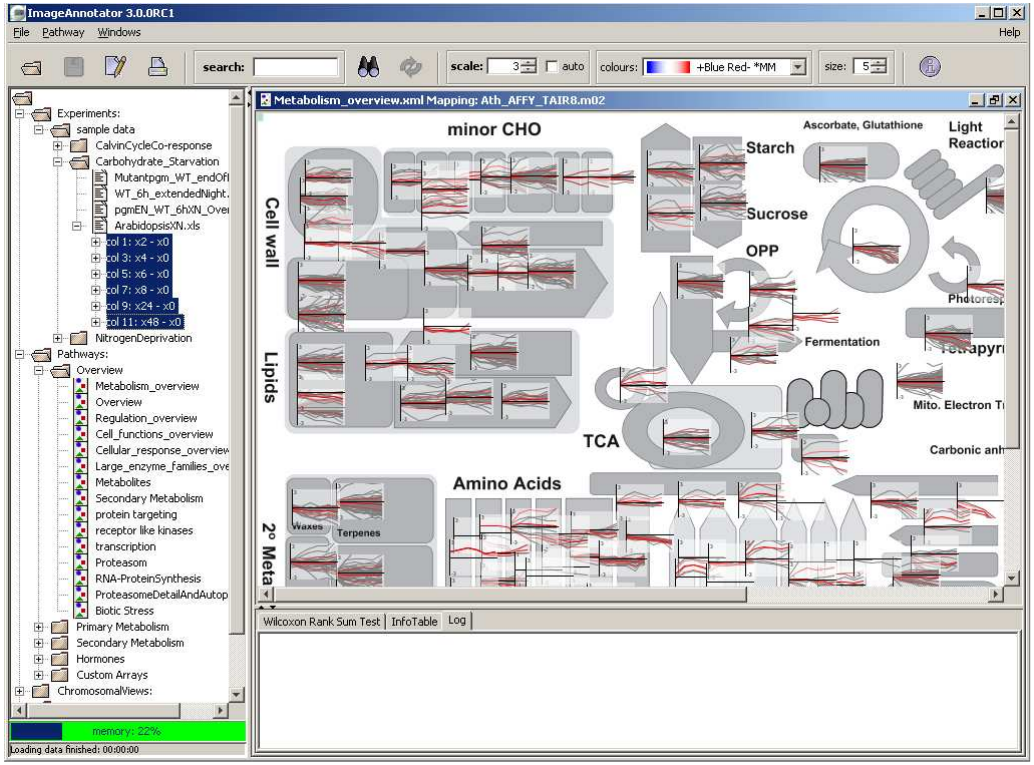
The activated filter is marked in red to indicate its activity. Now, on the Metabolism overview “map” genes which had not been flagged as significant are greyed out.



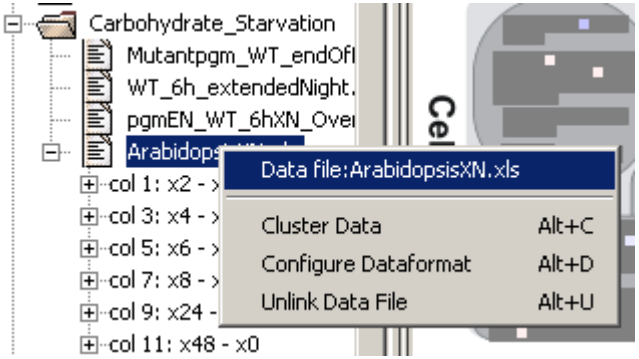


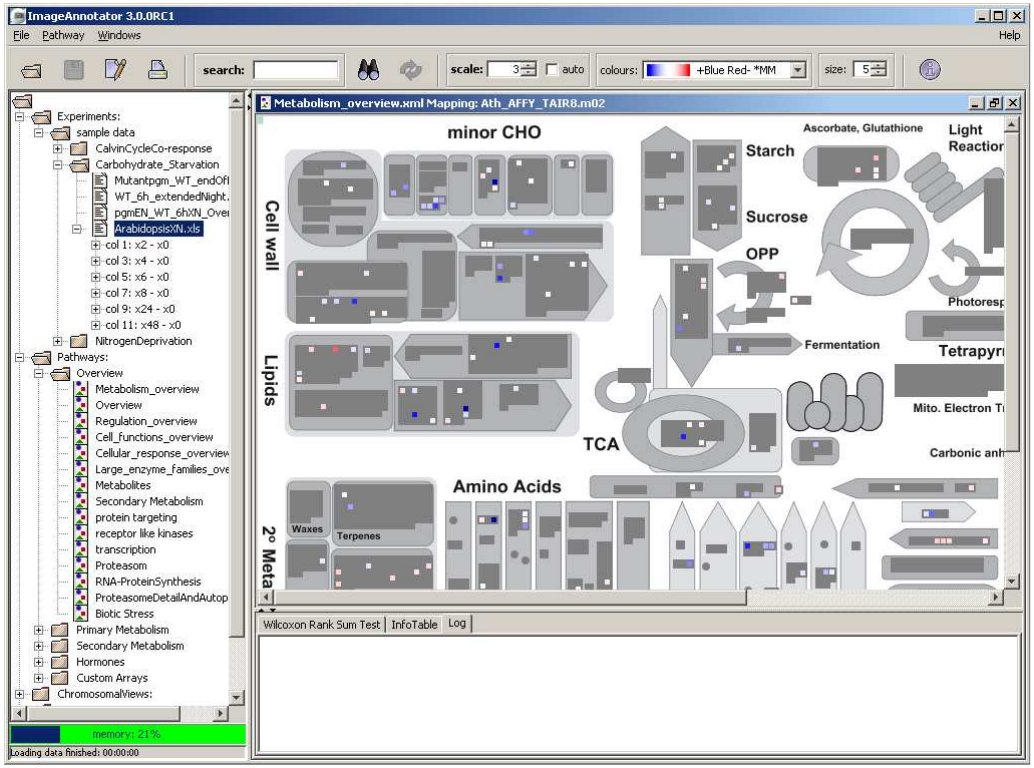
If the experiment is set up properly, multiple data points can be displayed at the same time as well. This is done by simply selecting multiple columns by shift-clicking / Ctrl-clicking on them.



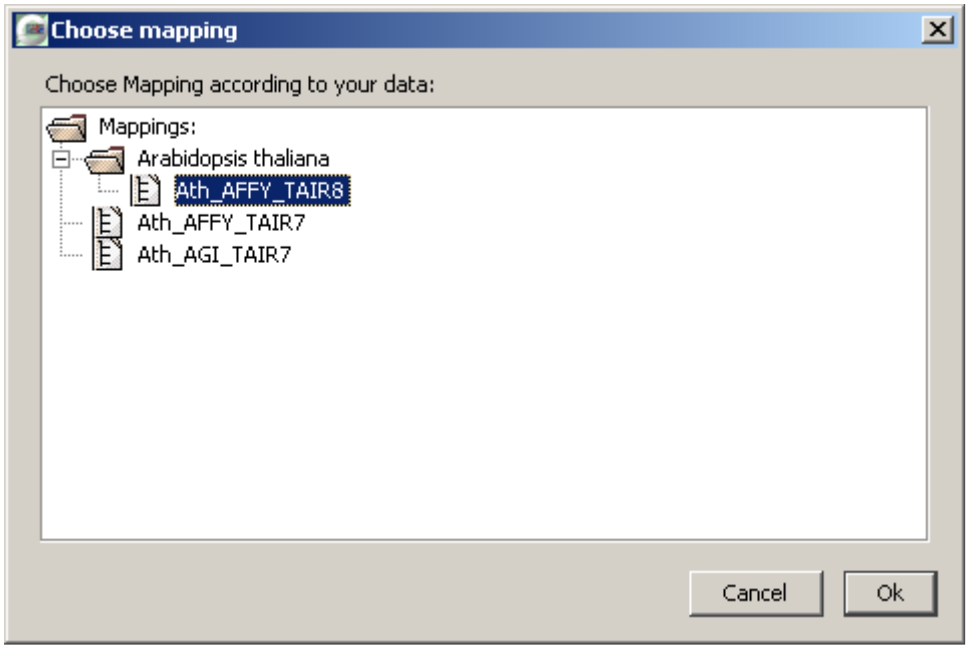


If an experiment is selected which contains multiple contrasts (log2 Fold changes), these can be used to cluster the data. Simply right-clicking (apple-clicking) on the data file e.g. ArabidopisXN brings up a pop-up menu, where one can select “Cluster Data”

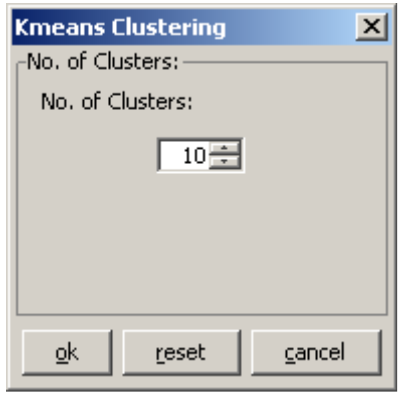




Even for clustering, a mapping file is necessary, as MapMan will display BIN assignments for each gene.

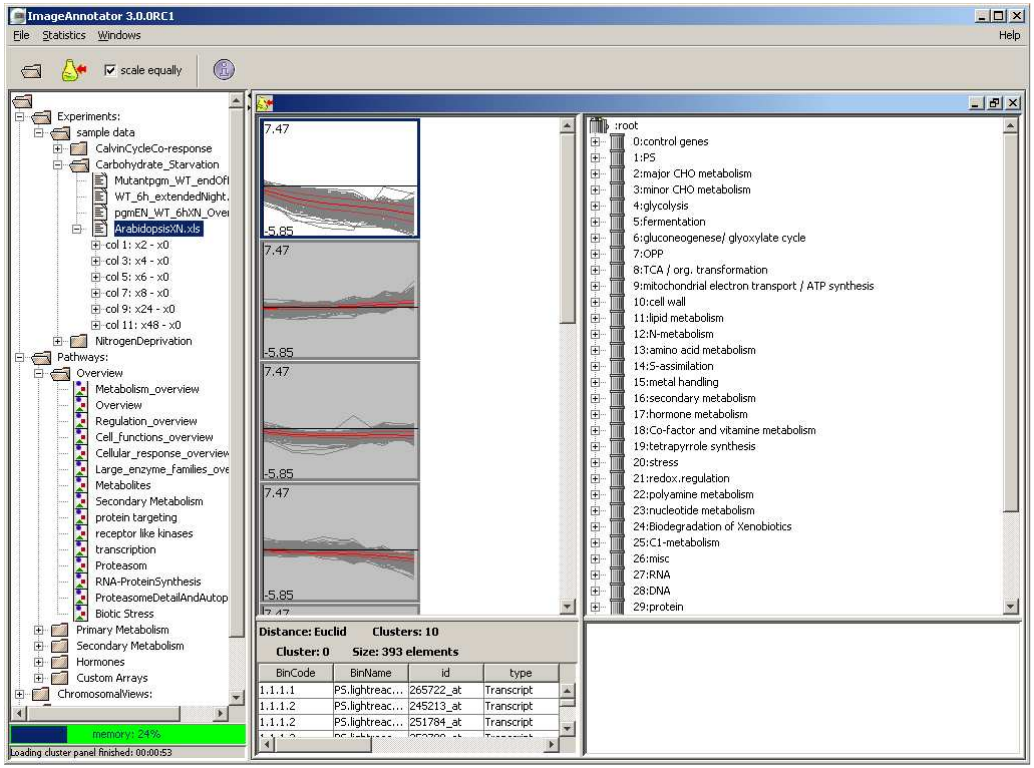


After having selected a mapping file, the user can select the desired number of clusters.

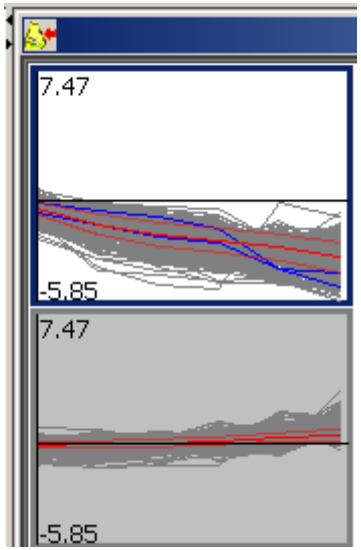


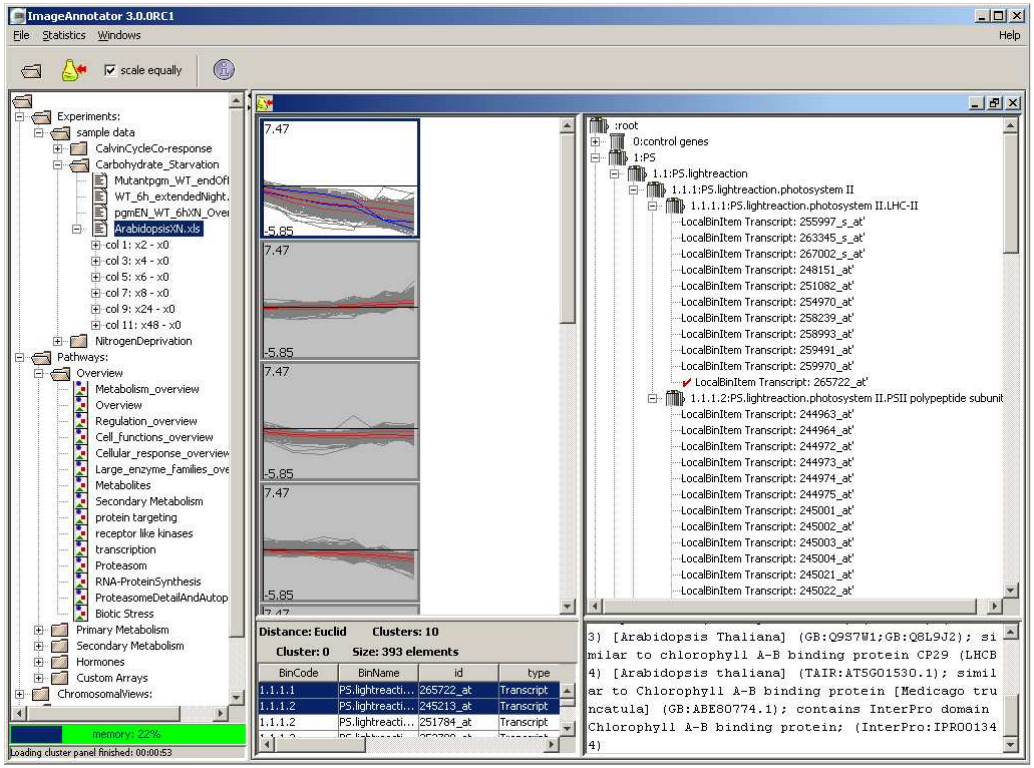
After a few seconds, a cluster view is shown in ImageAnnotator. This view is split into four parts.

- The upper left part displays the clusters and the genes in this cluster.
- The lower left part shows a tabular view of the genes in a selected cluster.
- The upper right part shows the BIN ontology
- The lower right part shows information for selected genes

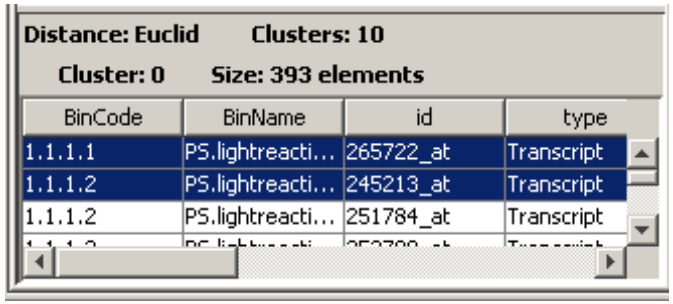


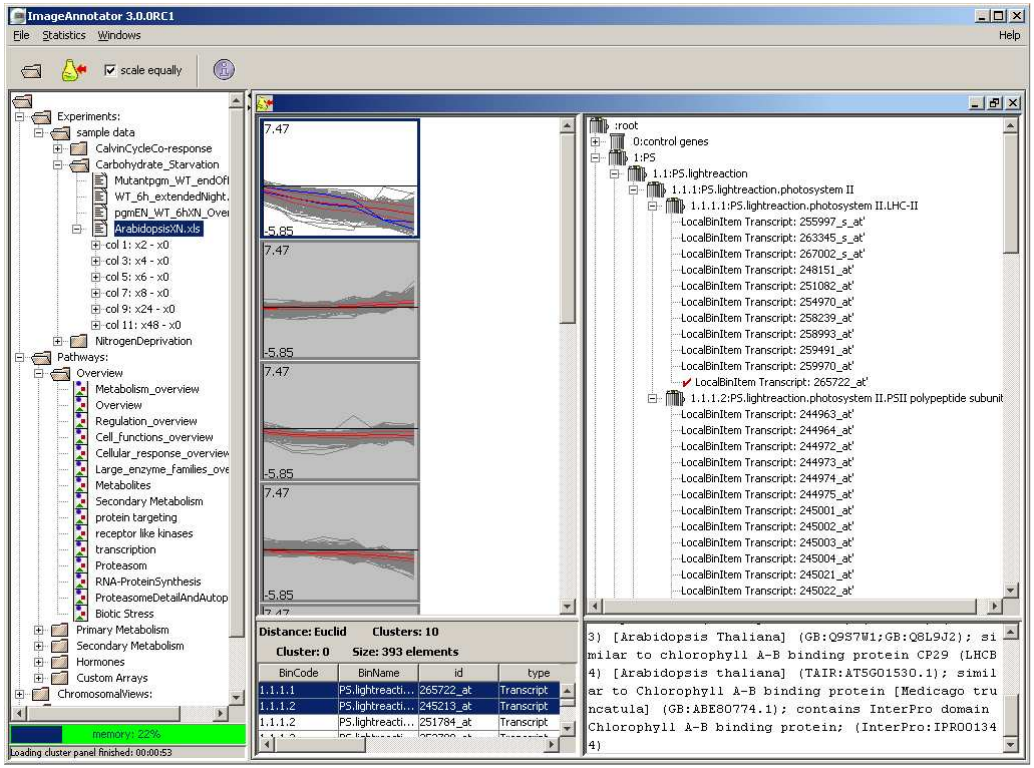
The Clusterview shows each individual gene in grey and the mean behaviour of all genes within this cluster as a red line. The Mean behaviour ± one standard deviation is depicted by orange lines. The currently selected cluster is marked by a blue frame. Selected genes in a cluster are marked by blue lines.



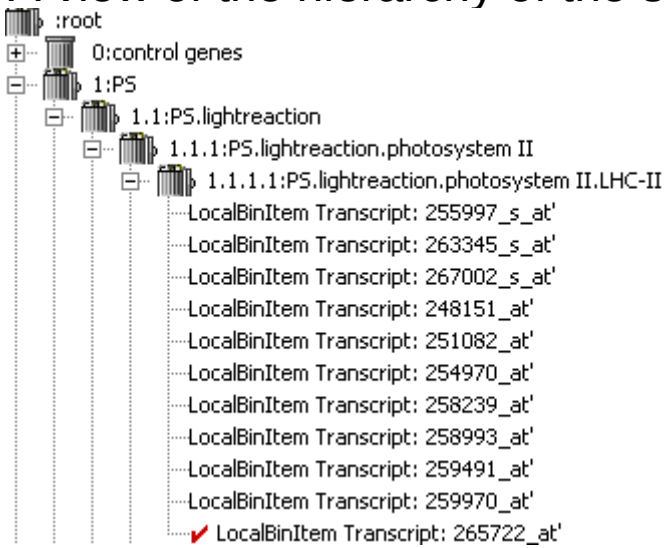


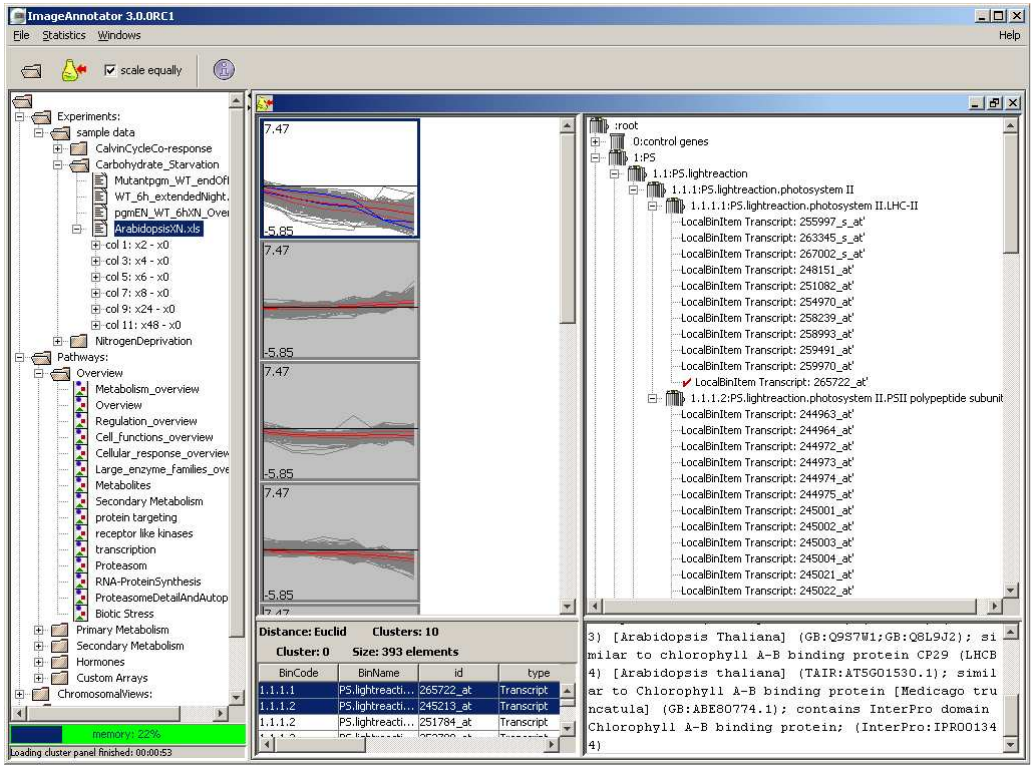
The tabular view of the cluster shows the BINs represented in each cluster, the corresponding items (transcripts, metabolites etc.), their description and their individual values. Items can be selected by clicking on them. Multiple items can be selected by shift or ctrl-clicking. The selected items are then highlighted in blue and their behaviour is displayed in blue in the clusterview.



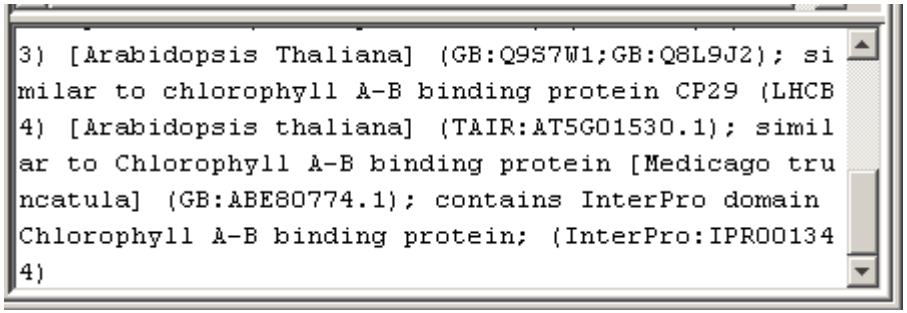


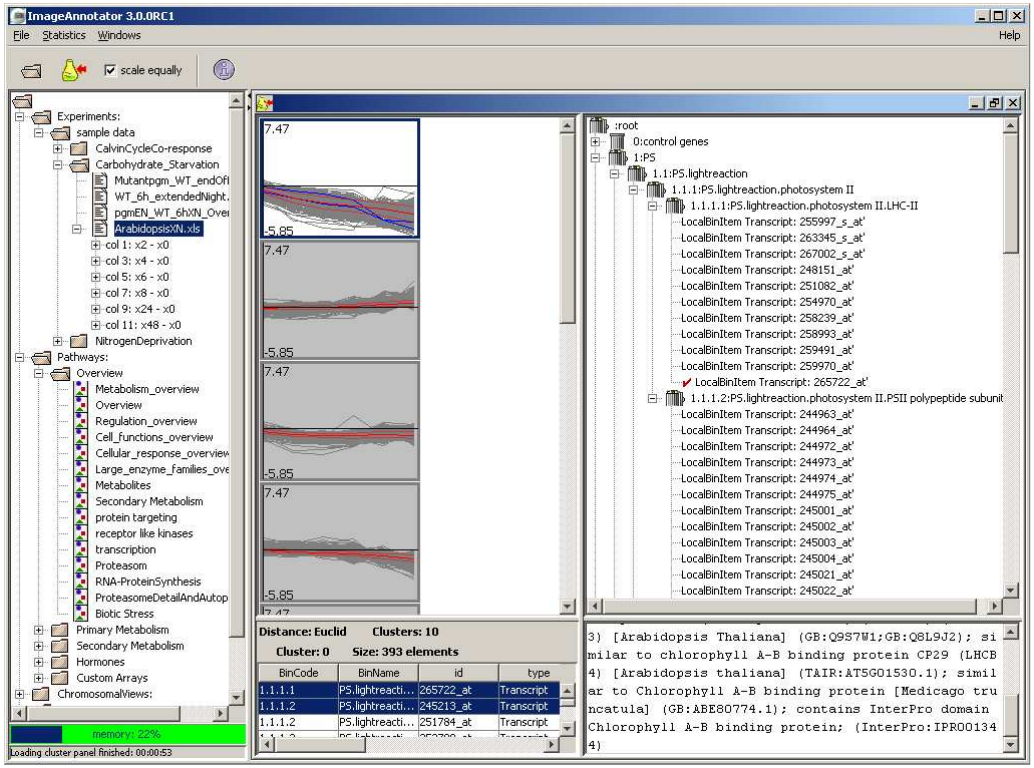
A view of the hierarchy of the selected items is display in the upper right corner



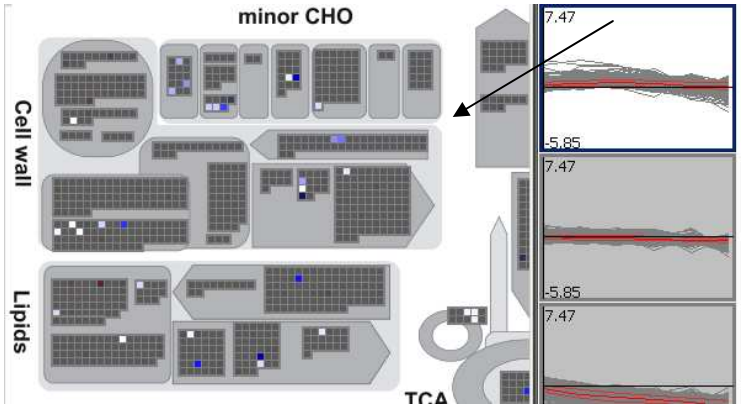


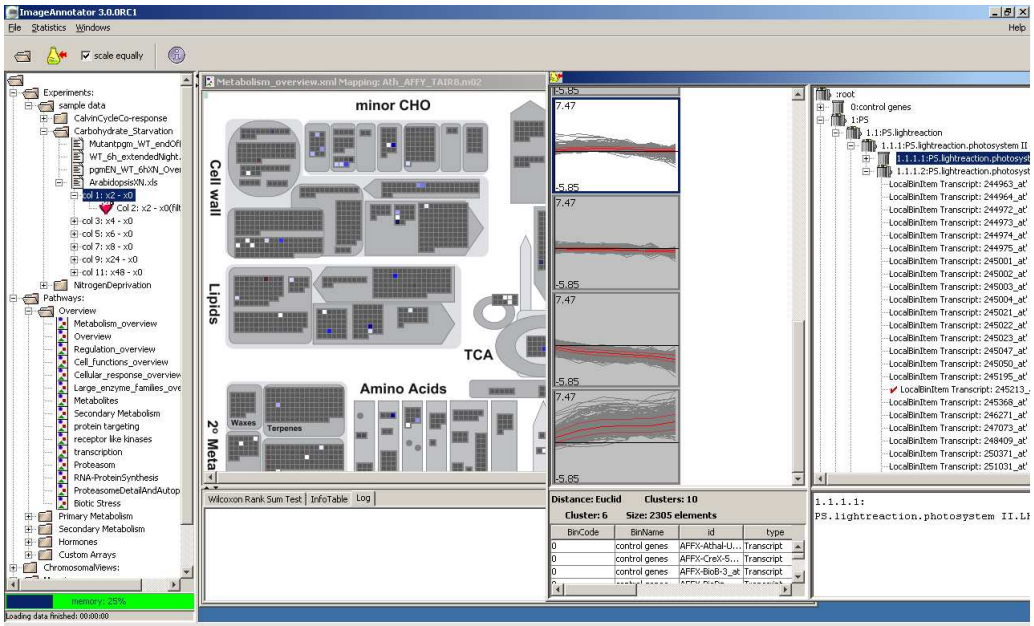
Currently (1/2009), the annotation of the item that was last selected is displayed in the lower right corner.



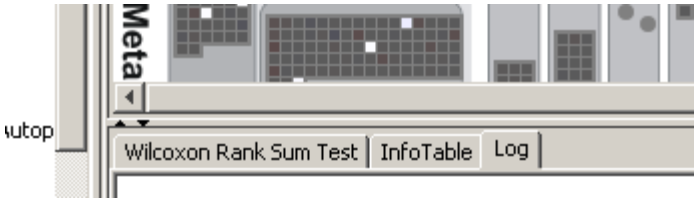


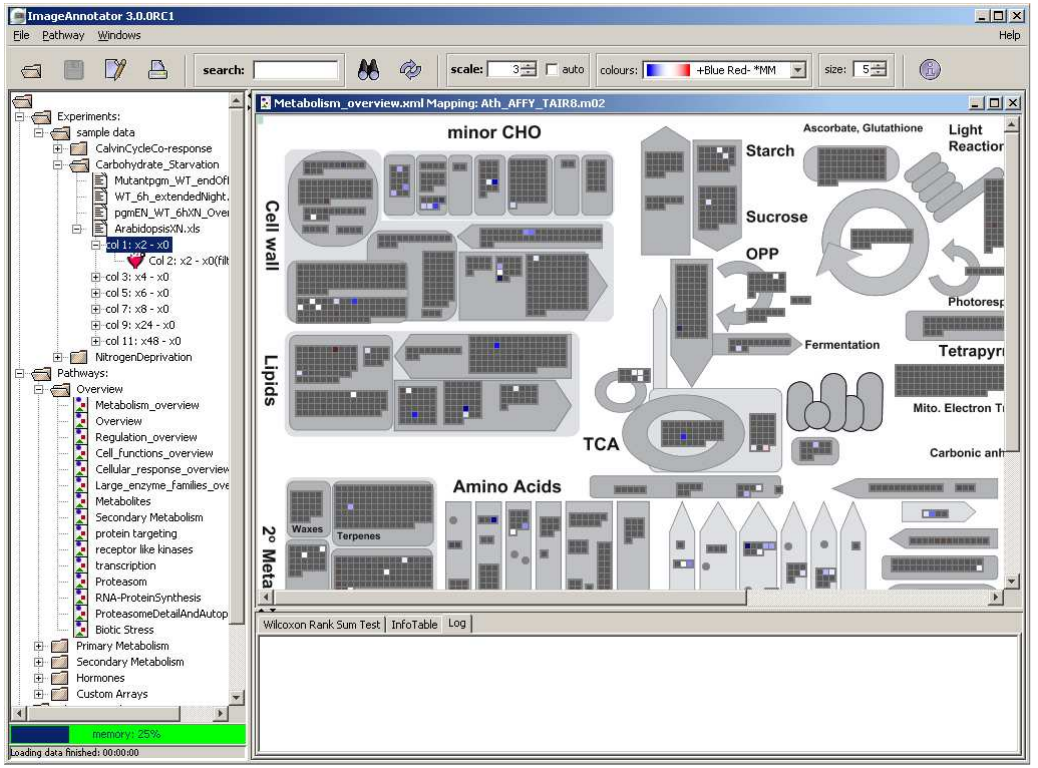
One can also “drag” clusters (by clicking on them and holding down the mouse button whilst moving the mouse) into Pathway pictures. In this case, all genes not in this cluster will be darkened. Clicking on the two-arrow symbol  displays all genes again.





Back in the Pathway View, clicking on the Wilcoxon Rank Sum Test tab, performs a Wilcoxon Rank Sum test. This is done for each BIN and subBIN separately.



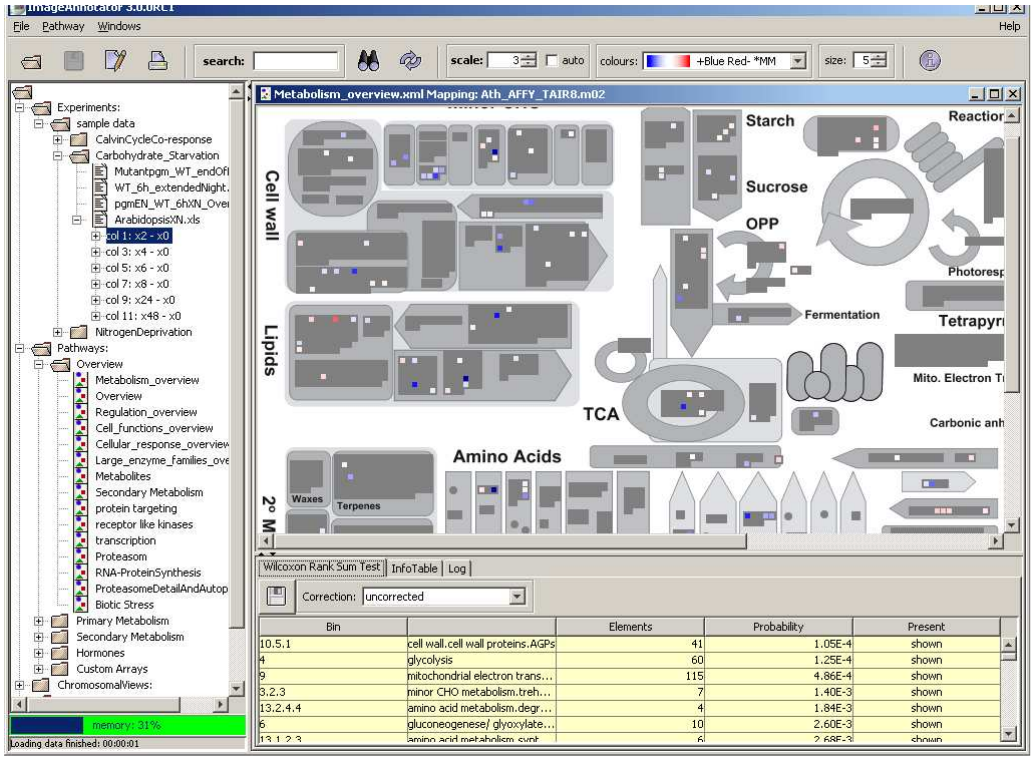


Within the table, for each BIN and sub BIN the p-value (the significance that it behaves differently) is given. The test can be corrected for the multitude of tests performed at the same time. (Using a p-value of 0.05 is not appropriate when one performs more than one test at a time. ImageAnnotator performs more than 1000 tests at once, as there are more than 1000 BINs)

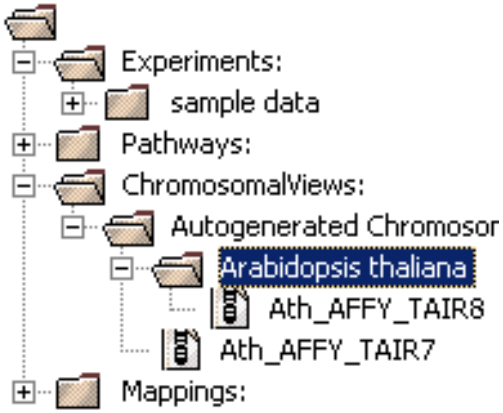
Wilcoxon Rank Sum Test | InfoTable | Log

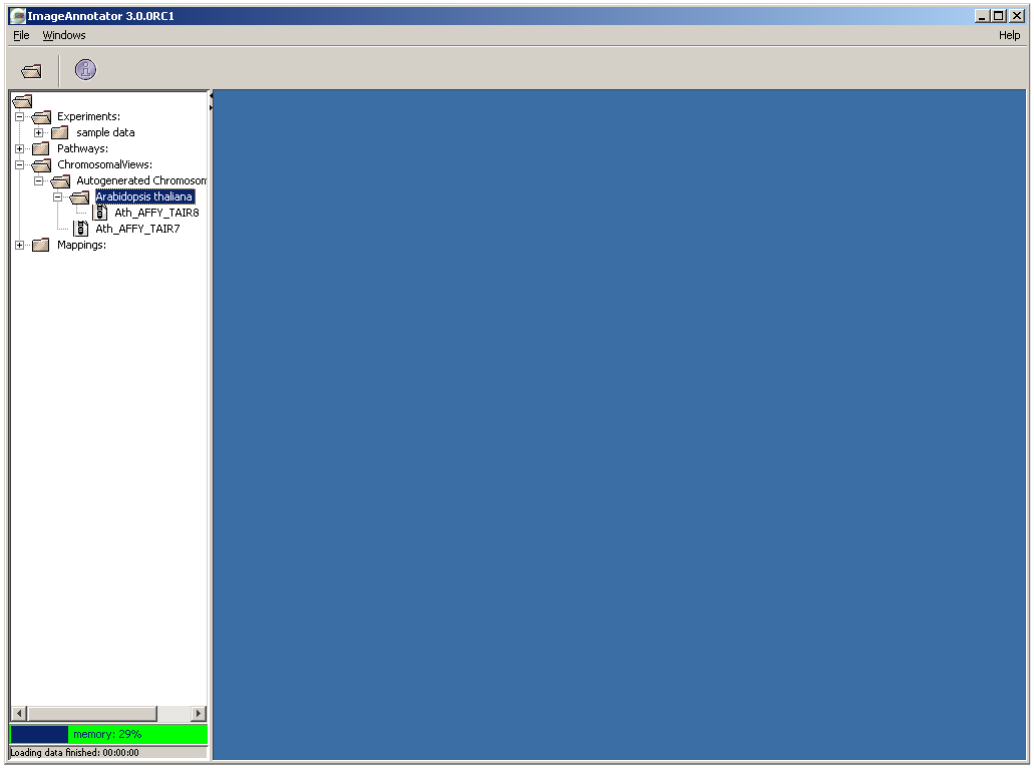
Correction: uncorrected

Bin		Elements	Probability	Present
10.5.1	cell wall.cell wall proteins.AGPs	41	1.05E-4	shown
4	glycolysis	60	1.25E-4	shown
9	mitochondrial electron trans...	115	4.86E-4	shown

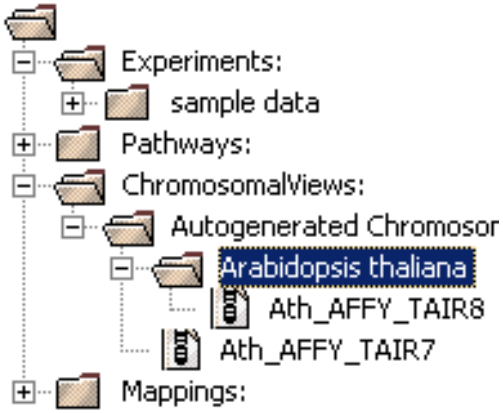


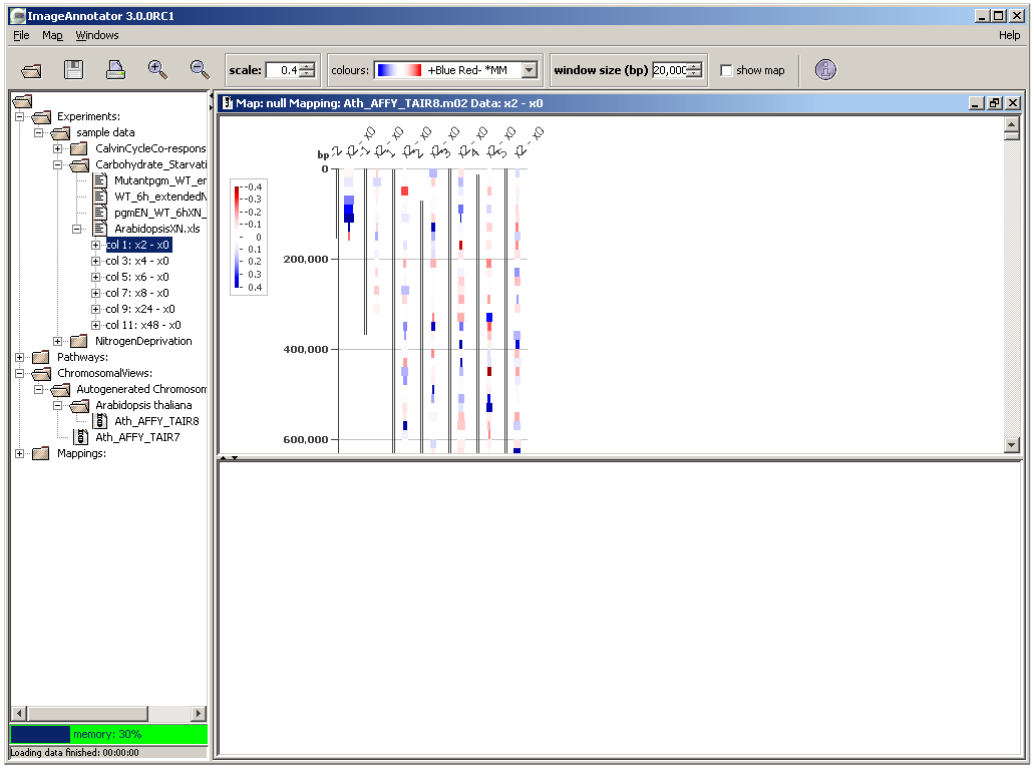
If you are dealing with a sequenced species, mapping genes (or better probesets) to their physical location in the genome is possible. Therefore, instead of using a pathway based display, one could also display probes by their physical location. In order to do so MapMan offers the folder ChromosomalView, where such displays are automatically generated based on the physical location extracted from the MappingFile.



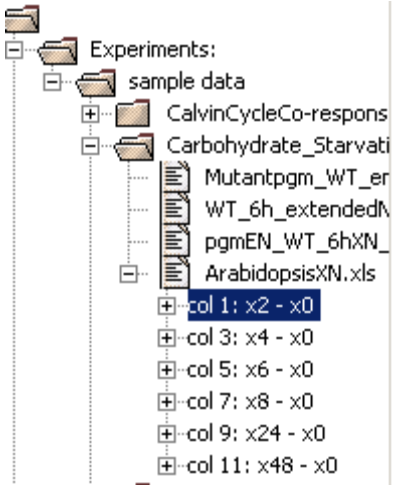


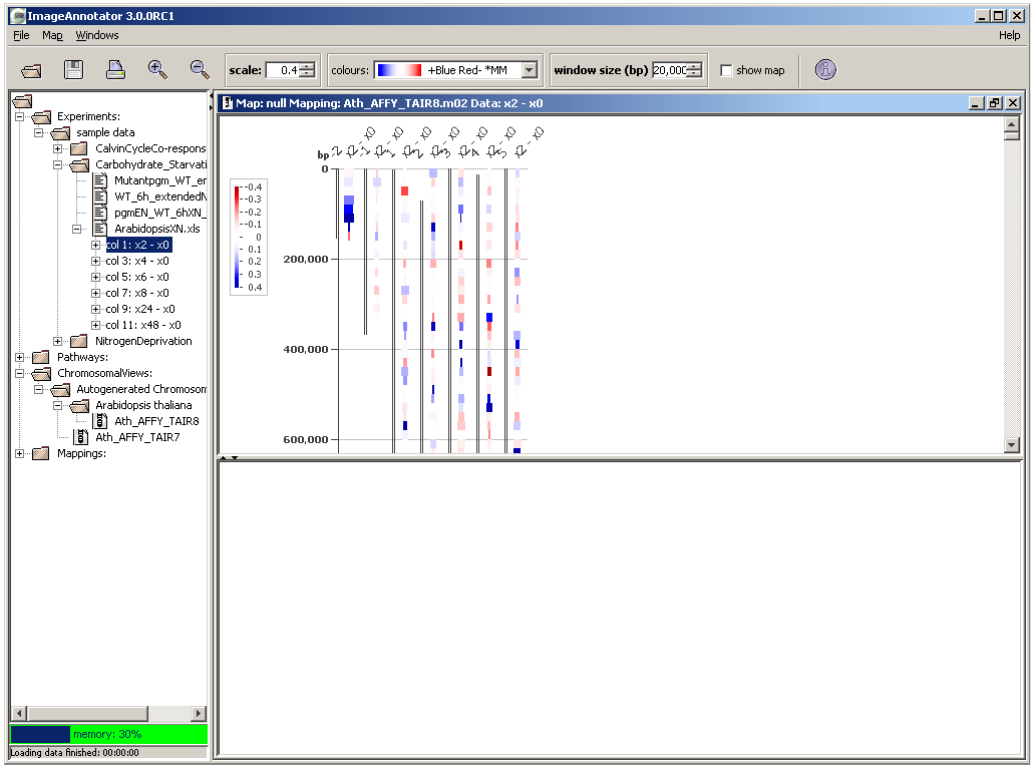
Opening such a view by double clicking brings up a display where the chromosome scaffold is shown. Please note that also mitochondrially and plastid encoded probe sets are shown, which when poly-A primed (common for many arrays), DO NOT FAITHFULLY REPRESENT THE TRUE EXPRESSION STATE

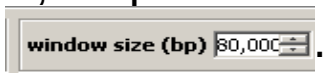




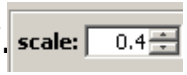
After the chromosomal view is activated, one can display single or multiple experiments by selecting them from the experiment tree by clicking on them (or by shift/control clicking to select multiple experiments).

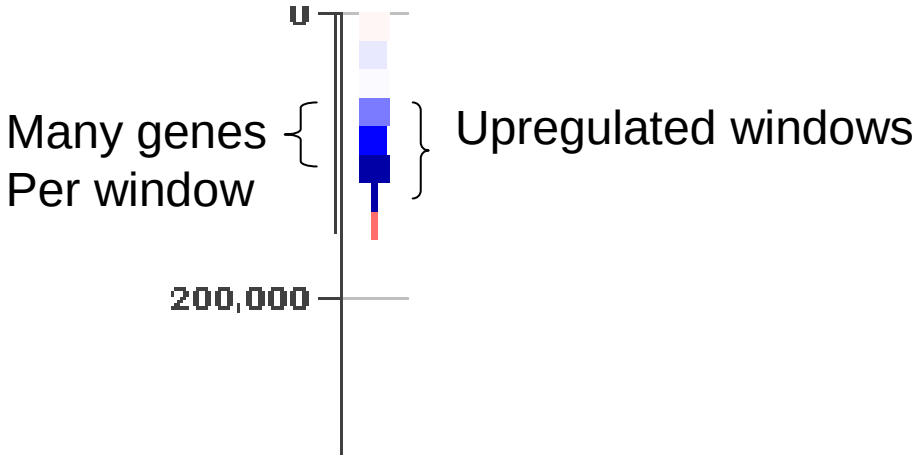


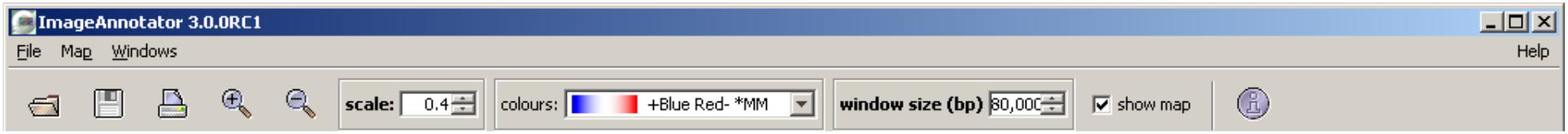


As it is not possible to show all genes at once, MapMan uses a window in which the signals (or log2 fold changes) are averaged 

. The size of this window can be adjusted. Again values are visualized using a false color scale which can be modified by the color selection bar 

. The number of genes within each windows is shown by the width of the bar representing the expression. Toggling the “show map” switches display of gene ids on and off. Scale controls the color intensity. 





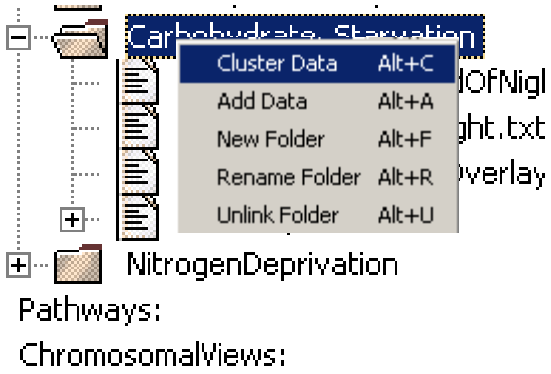
Chapter V Loading own data and Customizations

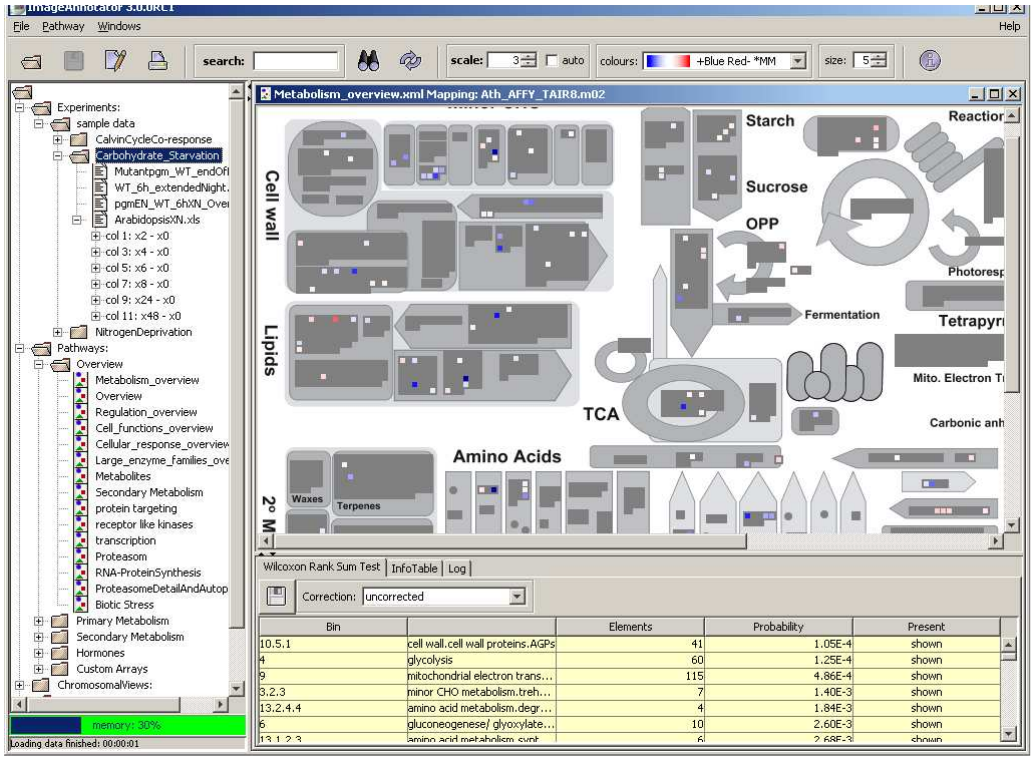
This chapter provides a basic overview of how to use the ImageAnnotator tool with your own data and how to customize MapMan.

This includes

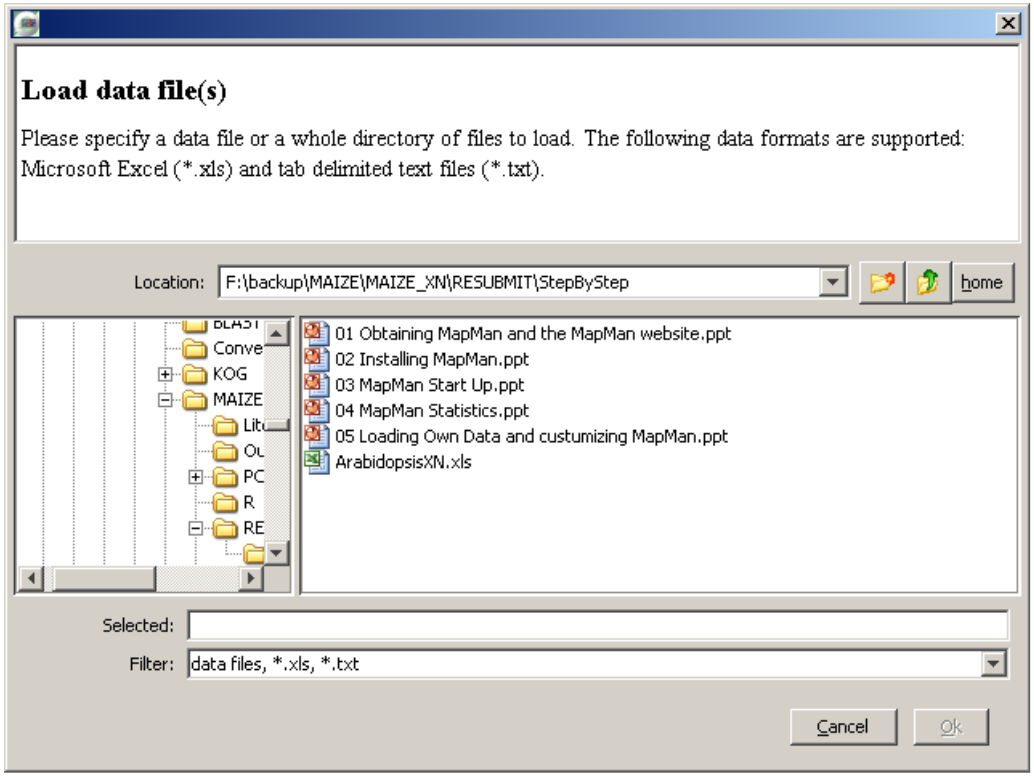
- Loading your own data
- Drawing own maps (pathways)
- Making your own Mapping files

In order to link your own data into ImageAnnotator, right click (apple click) on a folder where you want to link your data and select “Add Data”. Alternatively select “Add Data” from the File menu. Please remember that your files are only linked, i.e. it is a good idea to store them in clearly named folders on your hard-disk and not to later move or delete these files.





A Browser Window will open, allowing you to find your files. After you have selected your files, a configuration window will appear. (To configure your data see the Tutorial Statistics p. 34).

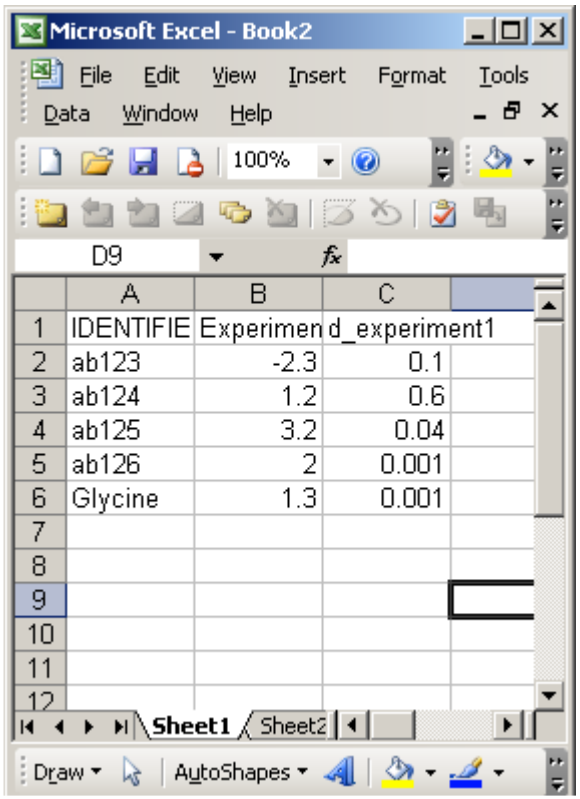


A note on the File Format.

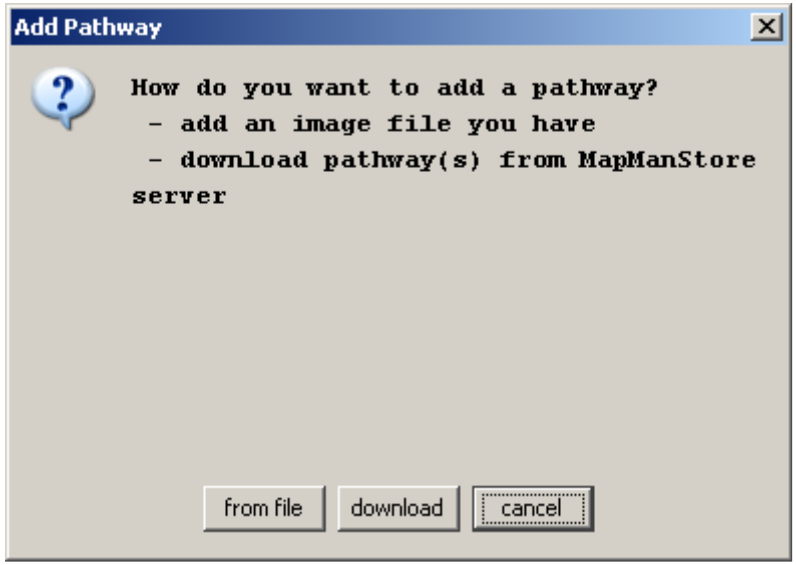
ImageAnnotator reads Excel and tab separated text files. However, Excel files are read in slower than text files.

Generally, the files should contain a header in the first row, and the first column should contain rownames. All further columns can contain log (fold change) values or derived values such as p-values. Transcript data can be freely mixed with Metabolite data.

MapMan does not support multiple values per Gene or metabolite, so these should be averaged before reading them into MapMan.

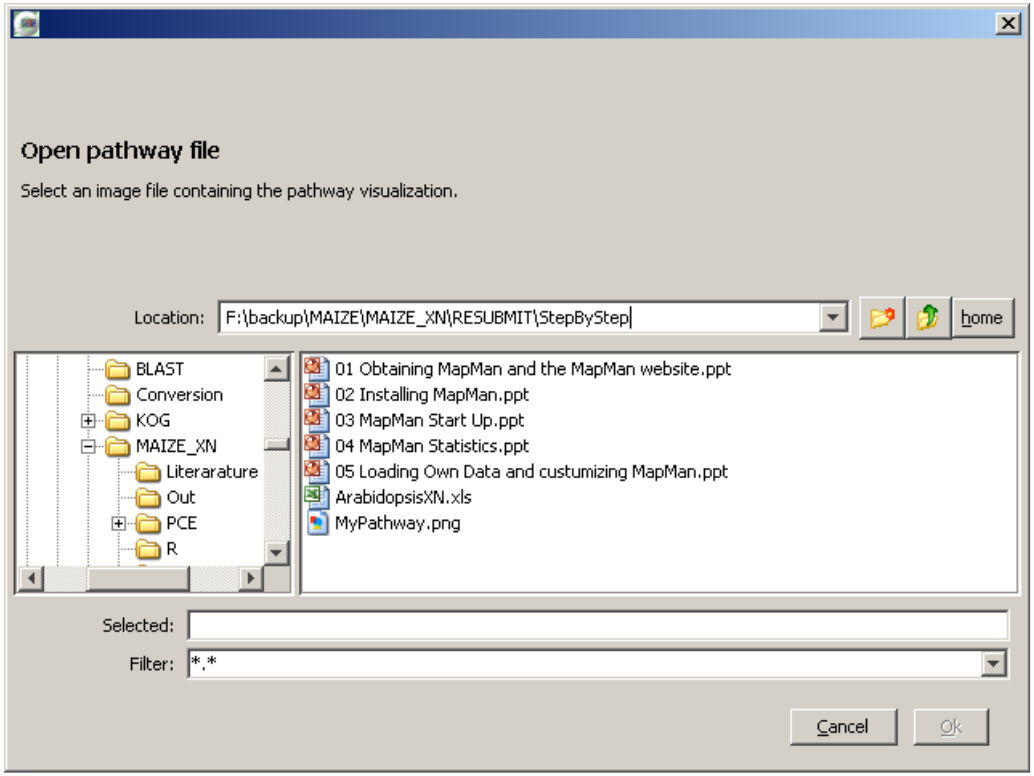


The same procedure can be used to link new pathways (maps). Again one right clicks (apple clicks) on a pathway folder and chooses “Add Pathway”. In contrast to experimental data, ImageAnnotator also allows you to download pathways from the MapMan server. (This is useful as there will be updates of pathways or new pathways available on the MapMan server and these can be downloaded in a pre-customized form making them immediately ready for use). If you want to load your own pathway, select “from file.”

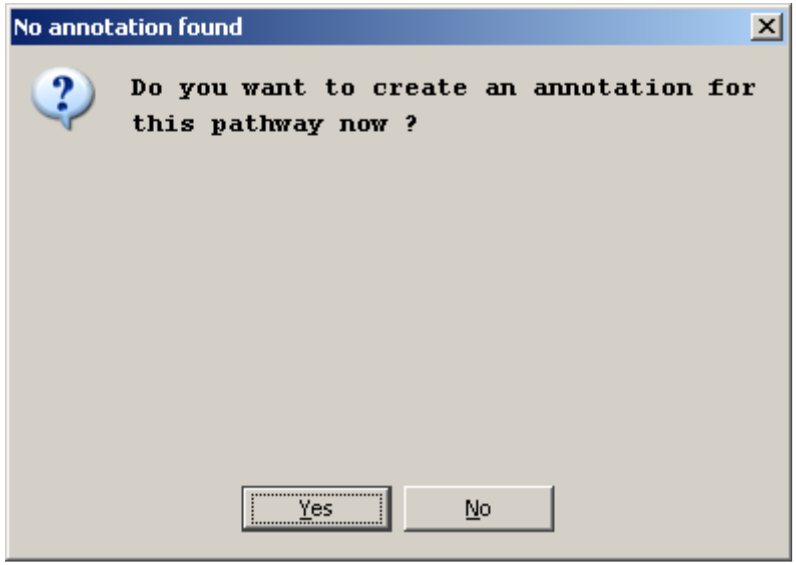


A file selection dialog will open, where many picture formats such as PNG, JPEG, TIFF GIFF and SVG can be selected.

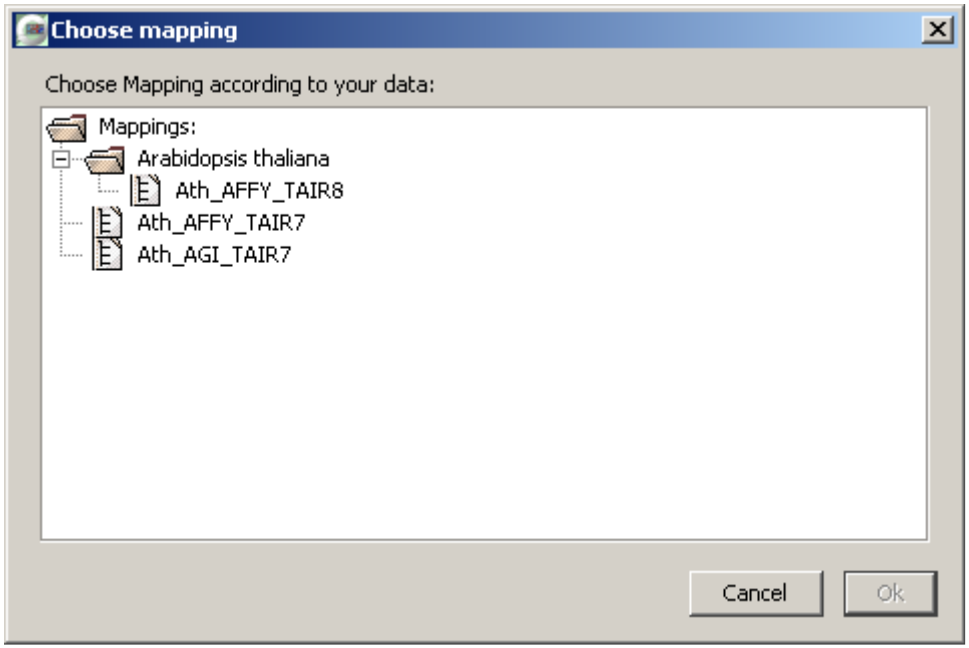
Please note: PNG, JPEG, TIFF and GIFF are all so called bitmapped formats, these come with a certain resolution and size e.g. 100 dpi (dots per inch) and 8x6 inches. (Or 800x600 pixels). Usually for publications, 300dpi or higher are required. If you save your file as SVG e.g. using Corel Draw® you can generate any resolution. E.g. 600dpi 50x40 inches without getting jagged edges.



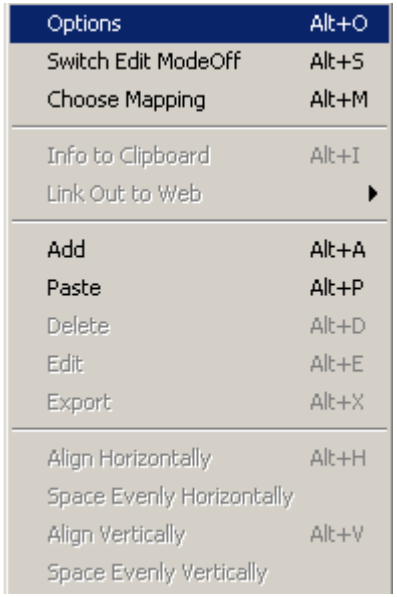
After having selected a pathway template, a dialog box pops up, asking if the user wants to customise the pathway now. When clicking yes ImageAnnotator switches into customisation mode.

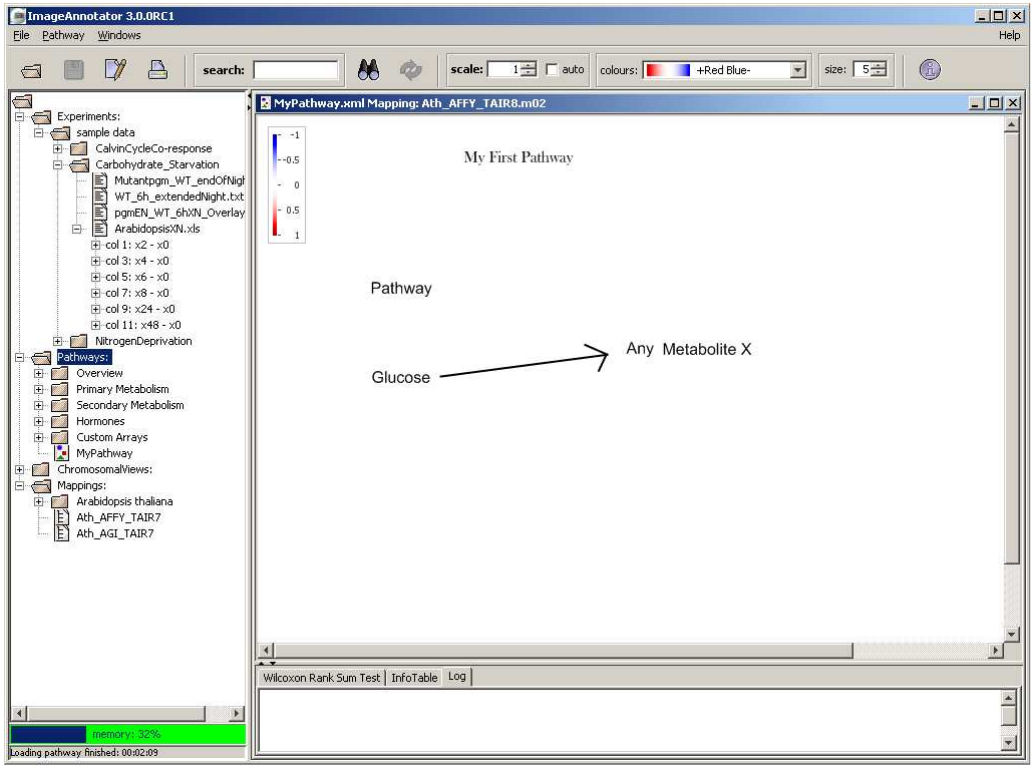


One has to choose a Mapping file to be used in customization mode.



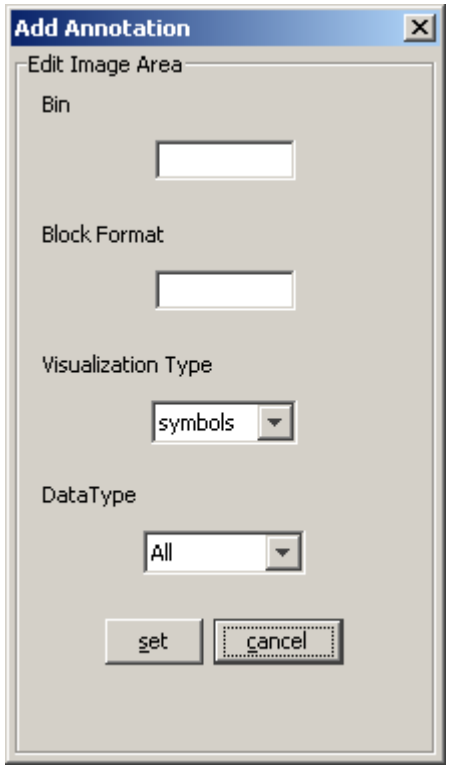
In this mode right-clicking (apple-clicking) anywhere on the pathway (map) will bring up a pop-up, where selecting “Add” will select the current position of the mouse arrow as the position where a new BIN or subBIN will be displayed.

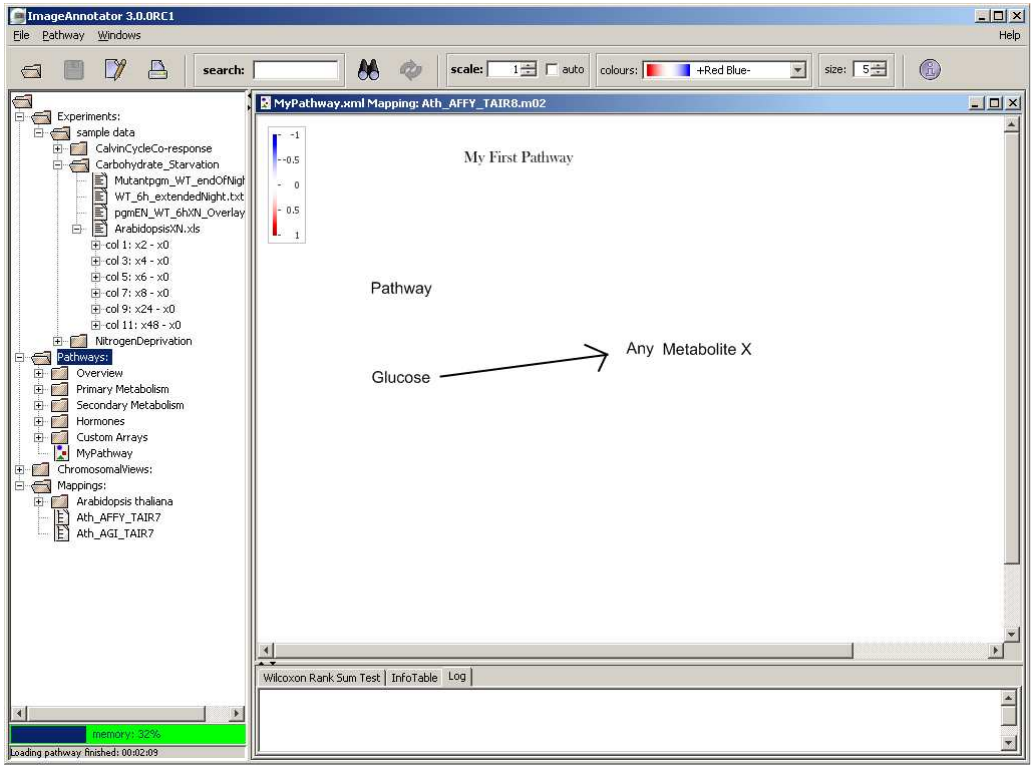




Now the user has to set certain parameters for this BIN. These are

- the BIN in question (e.g. 1.1.1.1)
- the “block format” which can be x or y and a number giving the number of columns or rows (e.g. x4 or y3). If choosing “x”, items will be added from left to right and then from top to bottom, if choosing “y” items are added from top to button and then from left to right.
- DataType will select if only metabolites, transcripts, proteins, enzymes, or all data types should be displayed.

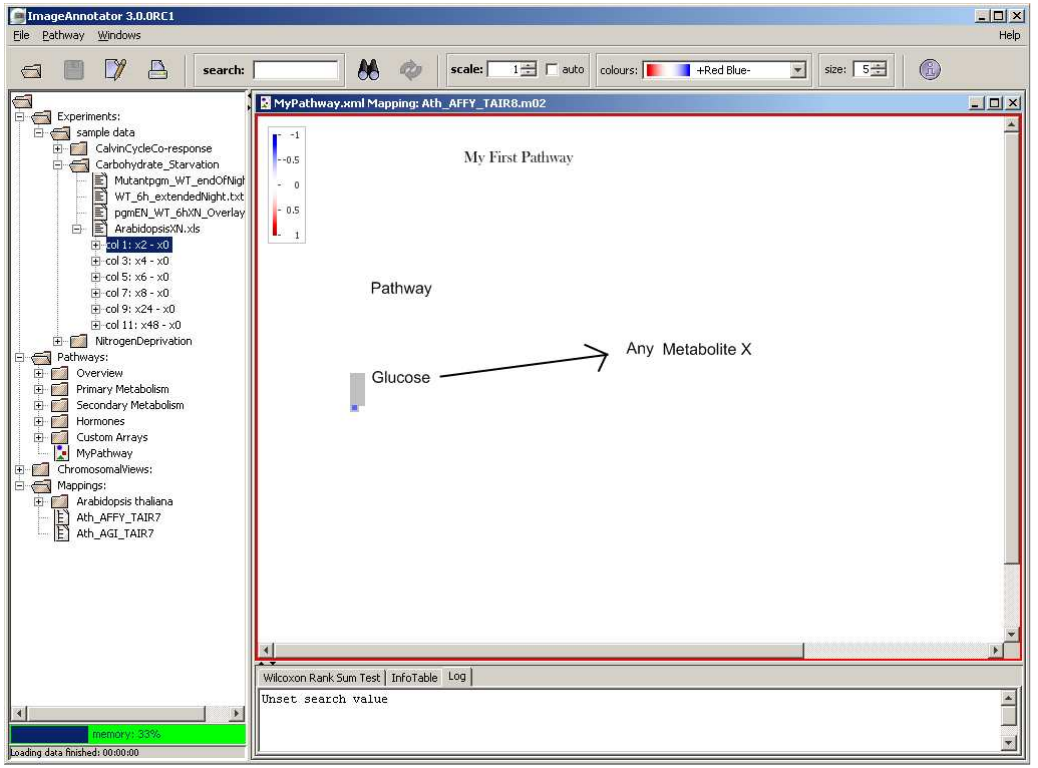




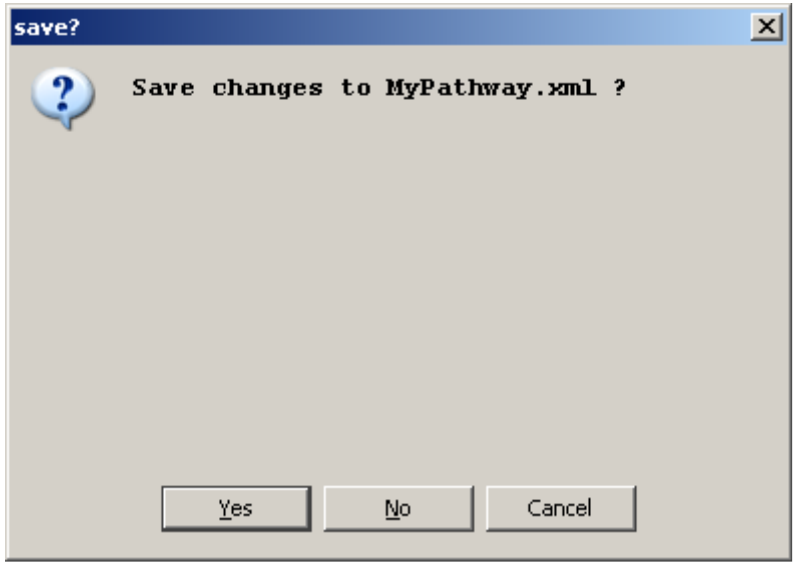
In the current example the data column x2 versus x0 from ArabidopsisXN was still active. So the data is immediately displayed.

The box displaying the data can be freely moved by dragging.

If this doesn't work click the pen and paper icon (to show that this mode is active the picture is framed in red). One can also reconfigure the box by right (apple)-clicking on it to bring up the configuration dialog)



Upon closing the pathway map, ImageAnnotator asks if the annotations of the picture should be saved. (Click Yes)



Constructing your own Mapping File

If your own species is not supported by MapMan or if you want to customise the assignment of genes to BINs you might want to develop a self made Mapping File.
A Typical Mapping File has 5 columns.

1.The BINcode (the numerical code)
2.Name (The name for the BINcode)
3.The Identifier (e.g. an Affymetrix code, or metabolite name)
4.The description for this item
5.The type of the item (T=Transcript, M=Metabolite, P=Protein, E= Enzyme)

Make sure that the Names for the BINs are consistent. E.g. 1.1 PS on one row and 1.1 SP in the next would be tagged by MapMan as inconsistent BINnames and the file would not load.
The hierarchy is inferred from these names. Level 1 is no dot (.) Level two in the hierarchy contains one dot (.) etc.

BINCODE	NAME	IDENTIFIER	DESCRIPTION	TYPE
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.1076.1.S1_at	chlorophyll A-B binding protein CP2//	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.1.A1_at	chlorophyll A-B binding protein /	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.2.A1_at	chlorophyll A-B binding protein / L//	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.3.S1_at	chlorophyll A-B binding protein / L//	T

Constructing your own Mapping File

When making your own mapping file, please bear in mind that the ImageAnnotator software can only display items by BIN. Thus, if you want to display individual isoforms of an enzyme you might want to give them separate numbers.
E.g.
BINcode 1.2.10001 for Isoform 1
And 1.2.10002 fro Isoform 2

Using the BINcode 1.2 would refer to both isoforms

BINCODE	NAME	IDENTIFIER	DESCRIPTION	TYPE
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.1076.1.S1_at	chlorophyll A-B binding protein CP2//	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.1.A1_at	chlorophyll A-B binding protein /	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.2.A1_at	chlorophyll A-B binding protein / L//	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.3.S1_at	chlorophyll A-B binding protein / L//	T

Constructing your own Mapping File

The description section is free text; you can use it for annotations or even for comments.

BINCODE	NAME	IDENTIFIER	DESCRIPTION	TYPE
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.1076.1.S1_at	chlorophyll A-B binding protein CP2//	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.1.A1_at	chlorophyll A-B binding protein /	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.2.A1_at	chlorophyll A-B binding protein / L//	T
1.1. 1.01	PS.lightreaction.photosystem II.LHC-II	Zm.13239.3.S1_at	chlorophyll A-B binding protein / L//	T

Making more far reaching customizations

If one had quantitative data from metabolomics experiments including their subcellular concentration, one could make separate identifiers for the metabolite in different localizations. E.g. Glucose could become Glucose_plastid, Glucose_cytosol etc.

As ImageAnnotator displays BINs one might then introduce one BIN per subcellular location, to profit from these measurements.

This could be done similarly to the following example:

BINCODE	NAME	IDENTIFIER	DESCRIPTION	TYPE
1234.1	Plastid metabolites.glucose	Glucose_plastid	Plastidial glucose concentration by non aqeuos fractionation	M
1235.1	cytosolic metabolites.glucose	Glucose_cytosol	Cytosolic glucose concentration by non aqeuos fractionation	M
1234.2	Plastid metabolites.fructose	Fructose_plastid	plastidic fructose concentration by non aqeuos fractionation	M
1235.2	cytosolic metabolites.fructose	Fructose_cytosol	Cytosolic fructose concentration by non aqeuos fractionation	M

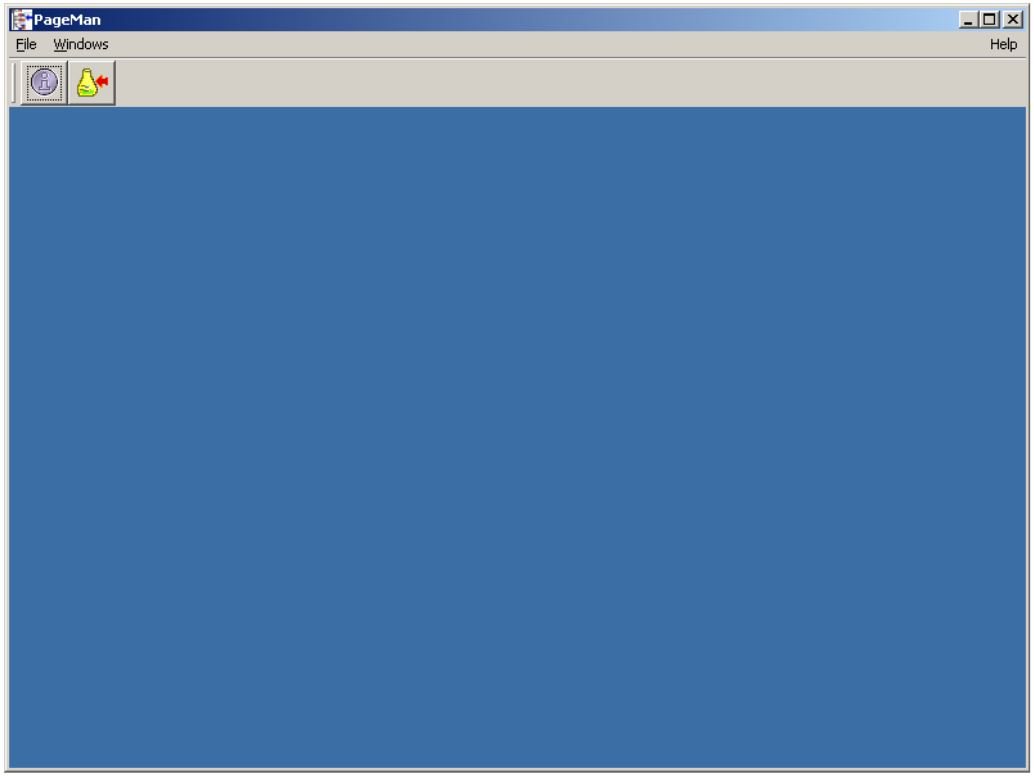
Chapter VI Compression and Visualization of Omics data using PageMan

These slides provide a basic overview of how to load data into PageMan and to use PageMan to visualize and compress data.

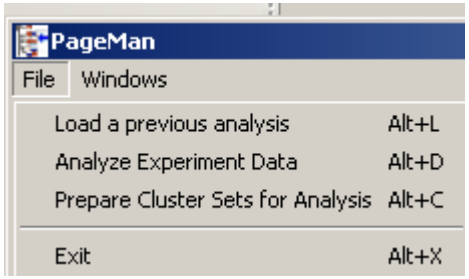
It also explains the different modes that can be used from within PageMan

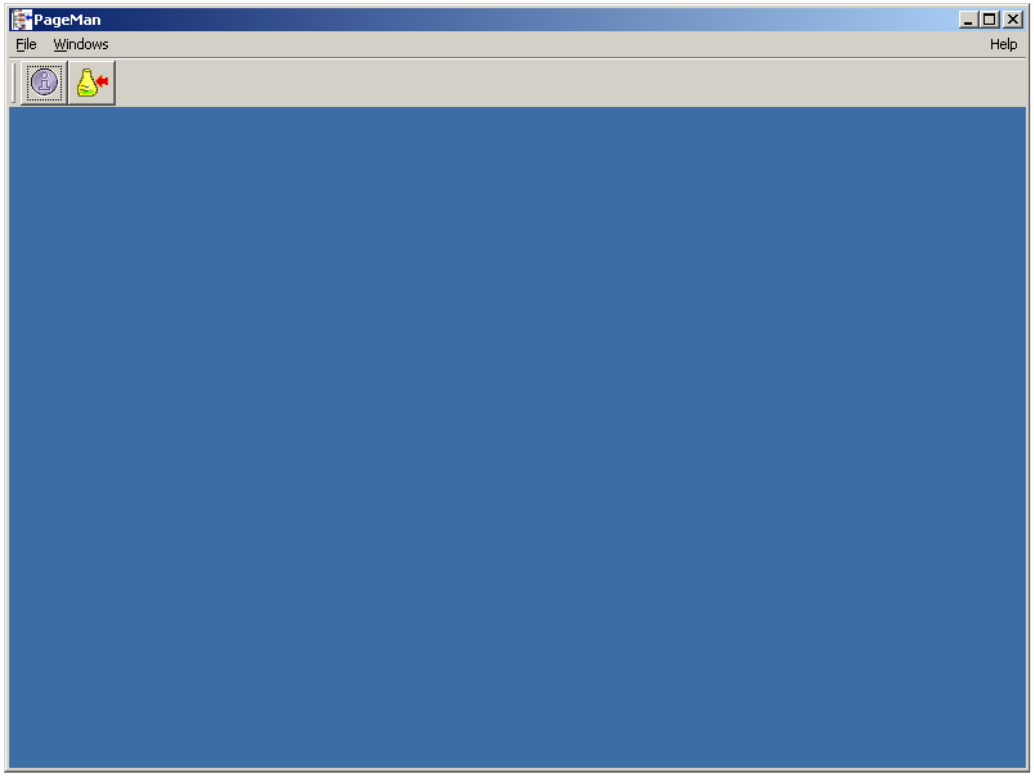
Finally it shows how to convert other ontologies into a mapping file

When you start up PageMan you are presented with a very simple interface.

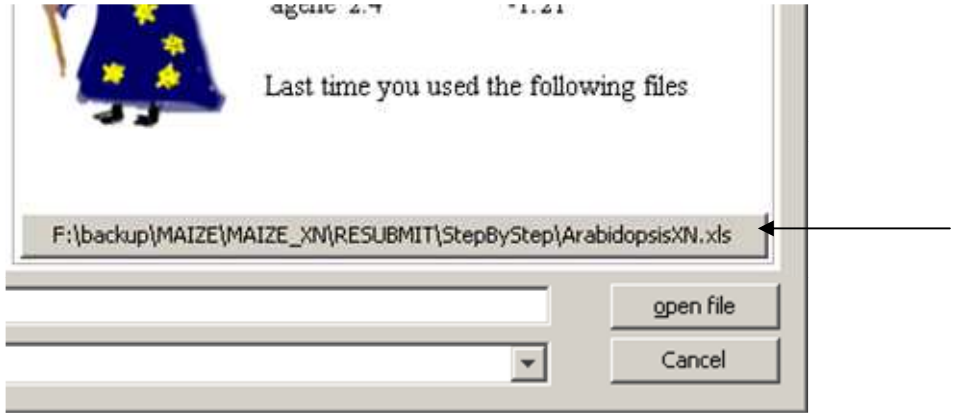


Data can be analyzed by either pressing the yellow bottle button  or by selecting “File->Analyze Experiment Data”.

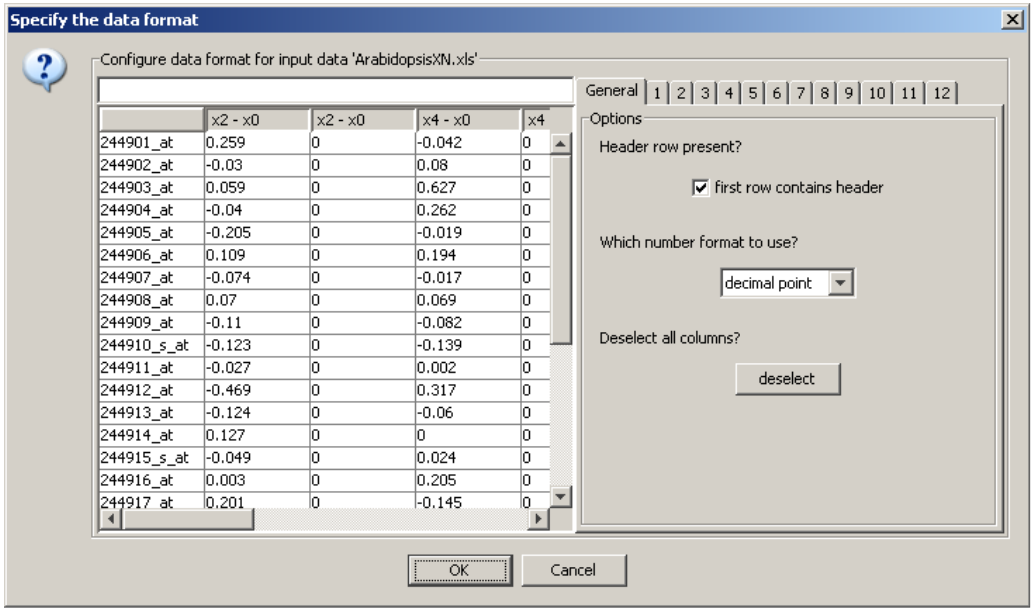




Firstly, the file carrying the experimental data has to be opened. PageMan displays the three last choices which can be reused by clicking on them.



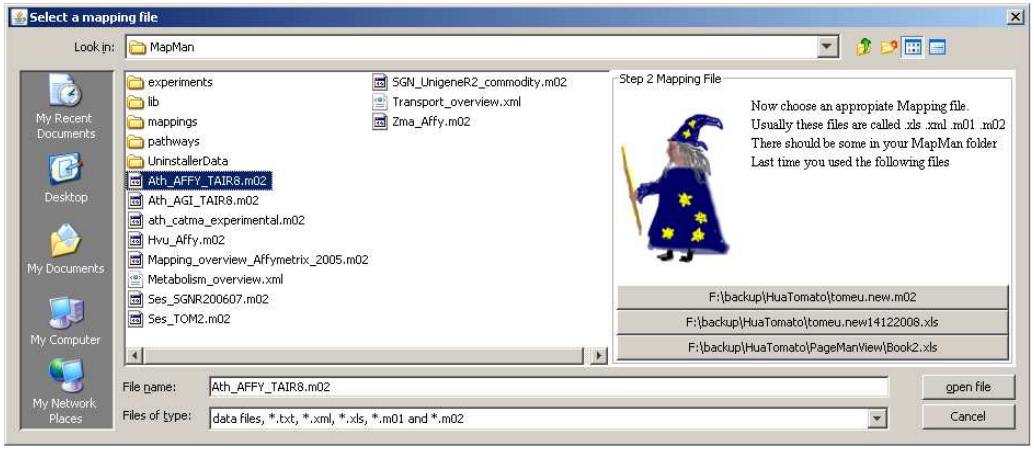
The same configuration dialog as in MapMan is being brought up. Please note, that PageMan will only consider numerical values.



	x2 - x0	x2 - x0	x4 - x0	x4
244901_at	0.259	0	-0.042	0
244902_at	-0.03	0	0.08	0
244903_at	0.059	0	0.627	0
244904_at	-0.04	0	0.262	0
244905_at	-0.205	0	-0.019	0
244906_at	0.109	0	0.194	0
244907_at	-0.074	0	-0.017	0
244908_at	0.07	0	0.069	0
244909_at	-0.11	0	-0.082	0
244910_s_at	-0.123	0	-0.139	0
244911_at	-0.027	0	0.002	0
244912_at	-0.469	0	0.317	0
244913_at	-0.124	0	-0.06	0
244914_at	0.127	0	0	0
244915_s_at	-0.049	0	0.024	0
244916_at	0.003	0	0.205	0
244917_at	0.201	0	-0.145	0

In a next step, an appropriate mapping file has to be chosen. As long as the mapping file is in MapMan format and representable as a tree, all ontologies can be used. Indeed, a converter for KEGG, GO, and MIPS Funcat is available as a separate program. (The GO DAG will be converted into a tree in the conversion process)

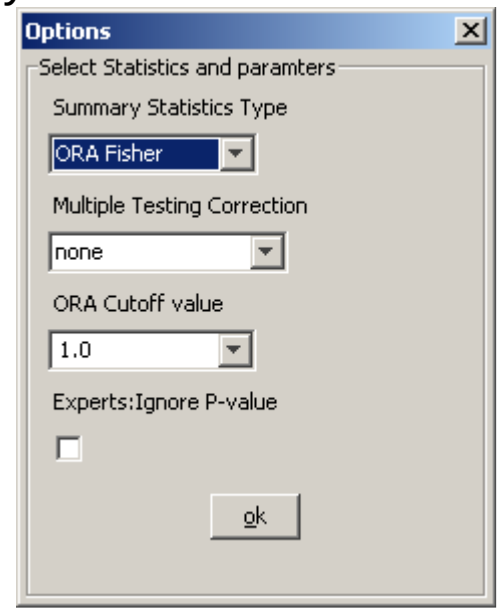
Flat relations can directly be imported into MapMan. As examples pfam families, COG, or KOG are possible examples that can be imported.



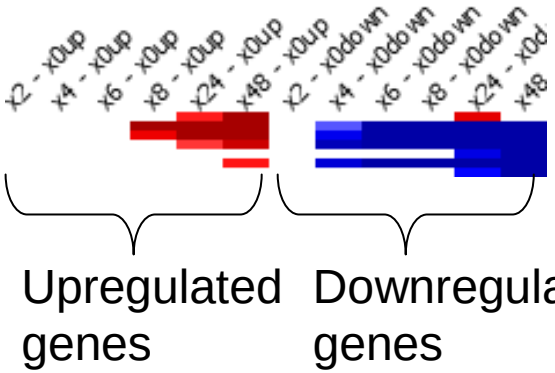
Finally, PageMan allows the user to choose which statistic to use. Possible Statistics include

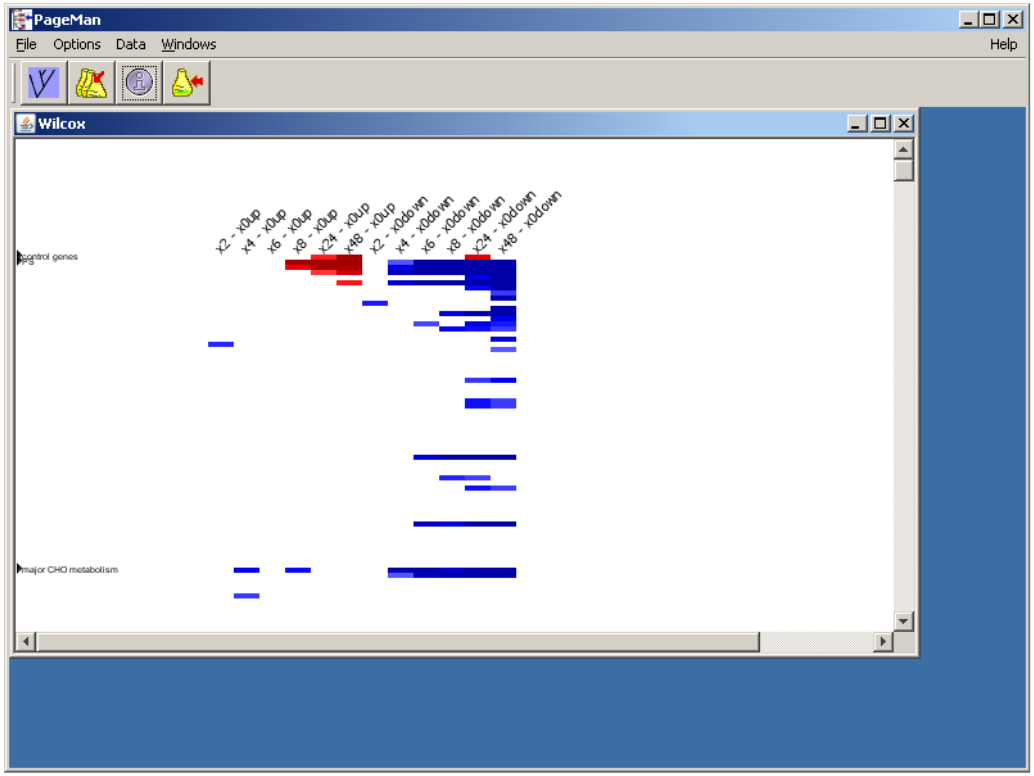
- Overrepresentation Analysis using
  - χ^2
  - Fisher's exact test (Recommended!)
  - Hypergeometric distribution
- Wilcoxon test
- Average
- Sum (soon)

PageMan will then calculate the BIN wise average or sum.
Alternatively, for the overrepresentation analyses it will calculate if certain items surpass a threshold ("ORA cutoff") more often than expected by chance. The Wilcoxon statistics just compares the values in the BIN versus all other values.



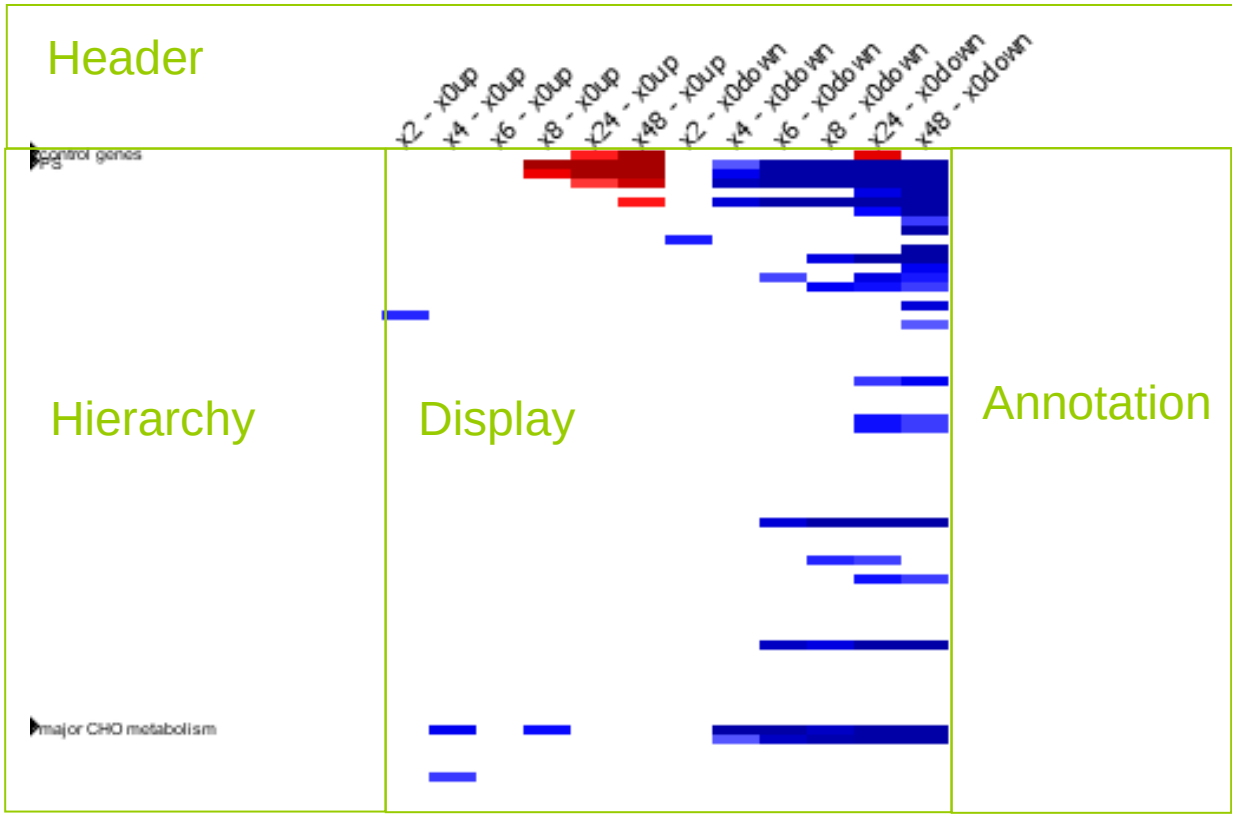
The data is then statistically analysed and visualised. In the case of a Wilcoxon test individual BINS and subBINs are either coloured blue (up-regulated) or red (down-regulated). Non significant BINS are left white by default.
In the case of overrepresentation analyses, the BIN is either coloured blue for more items than expected by chance or red if it is less items than expected by chance. In this case each experiment is evaluated twice, once for genes that are up-regulated and once for genes that are down-regulated. In an ideal case this should result in either down-regulated genes being encountered lower than expected by chance and up-regulated more than expected by chance or vice versa. However, finding genes less than expected by chance often only works with BINS containing many genes.



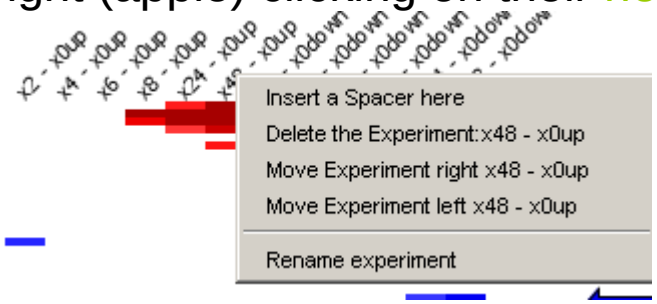


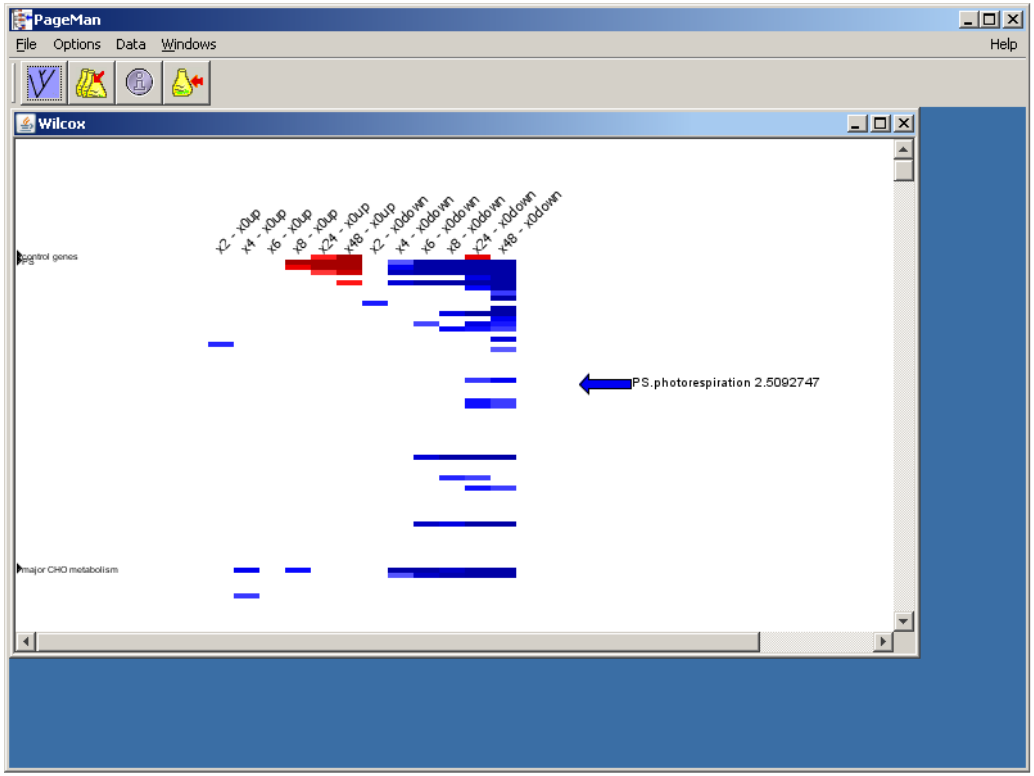
The display is partitioned in 4 areas:

- Header
- Hierarchy
- Display
- Annotation

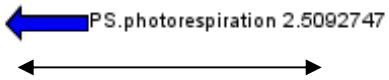


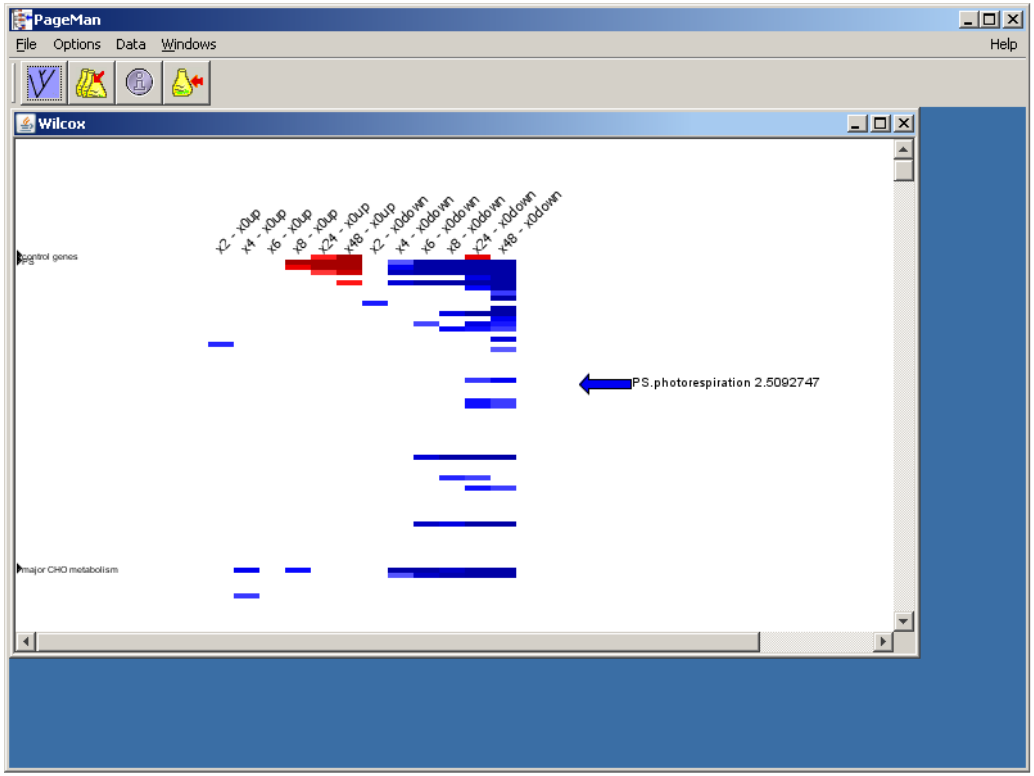
PageMan offers the user to adjust the display of the data. Clicking on a coloured item in the display area brings up an annotation next to it. A hierarchy tree can be displayed by clicking on the tree icon  . Data columns can be moved or deleted by right (apple) clicking on their headers.



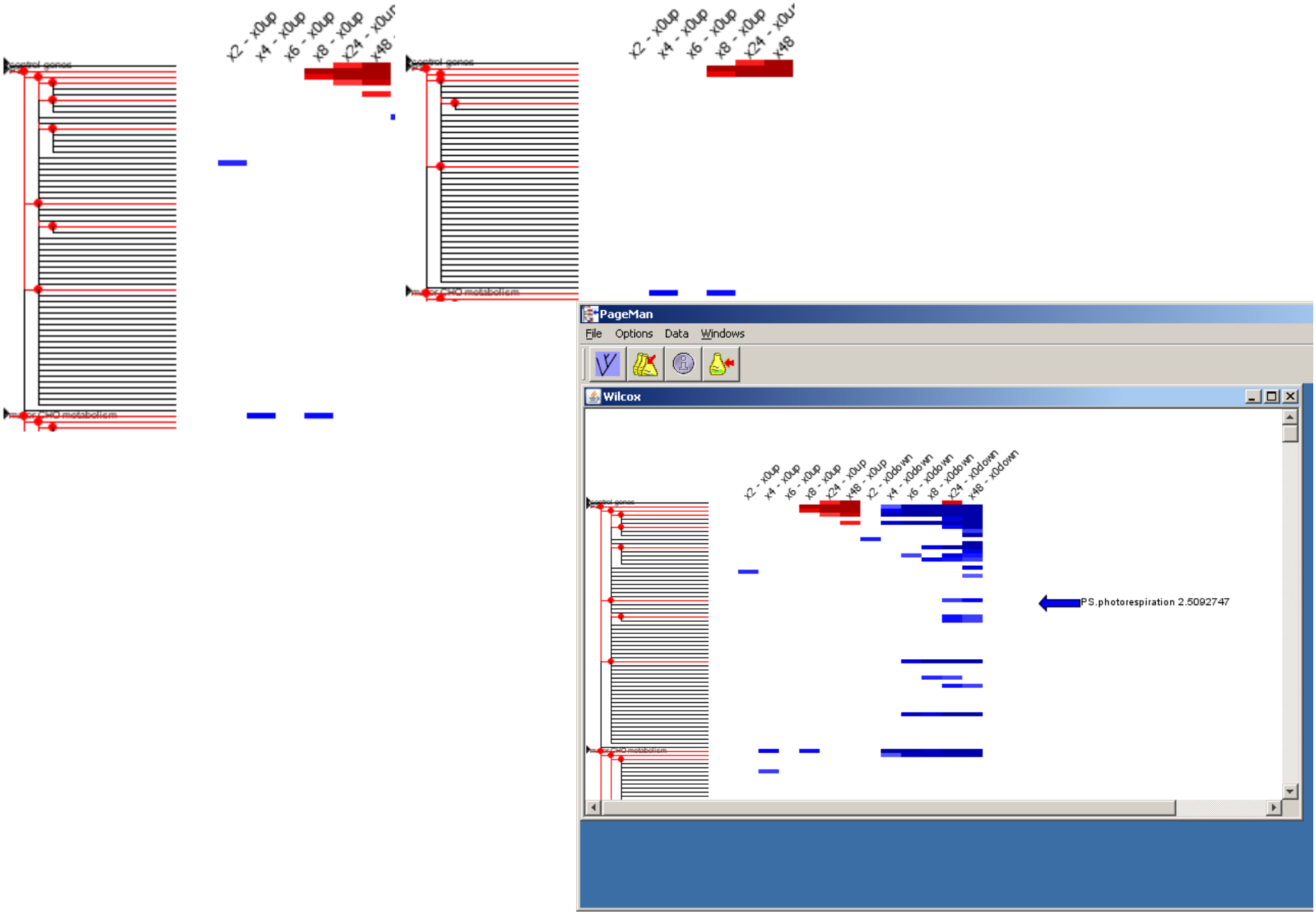


Annotations can be moved around horizontally (or vertically after deselecting “Options->Lock annotation”). Right (Apple) clicking allows to rename or delete annotations.

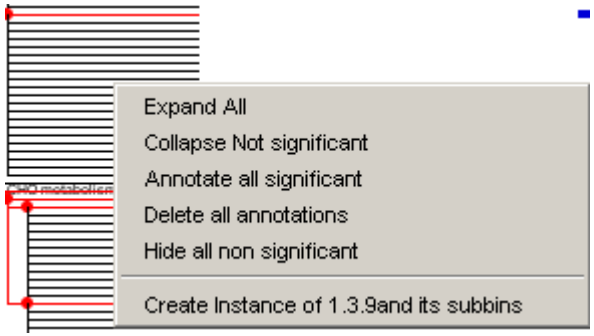


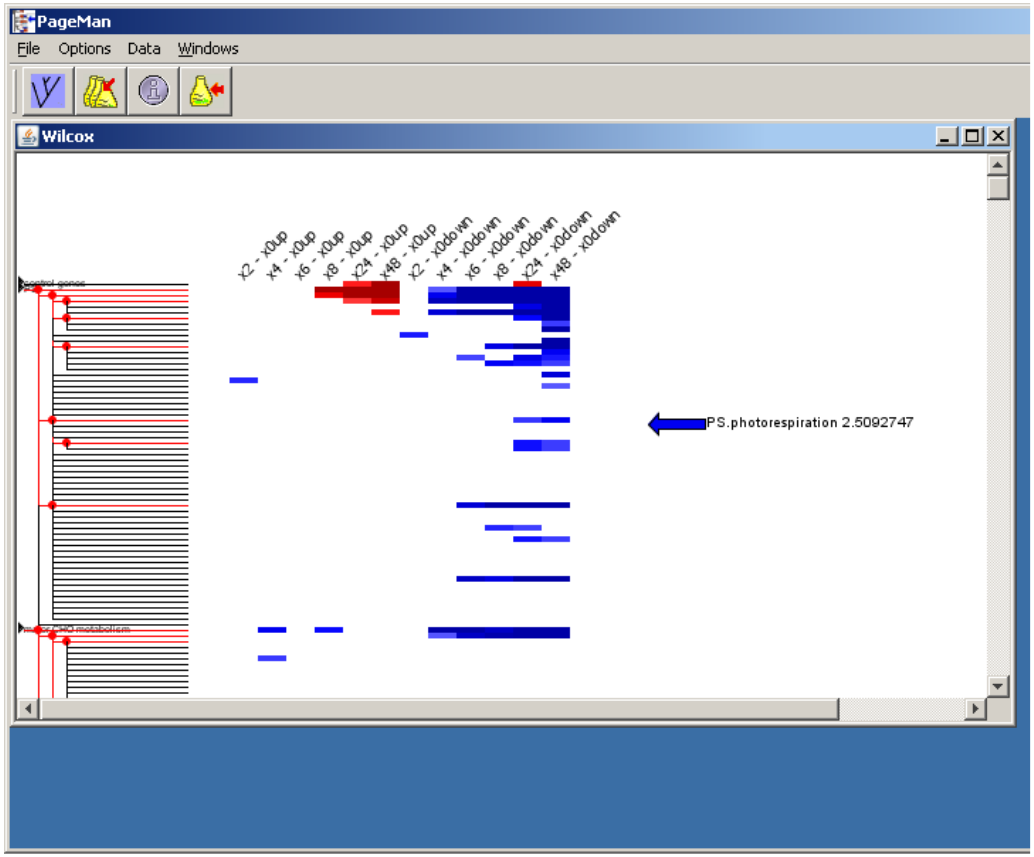


If the hierarchy is displayed clicking on nodes collapses the subBINS.

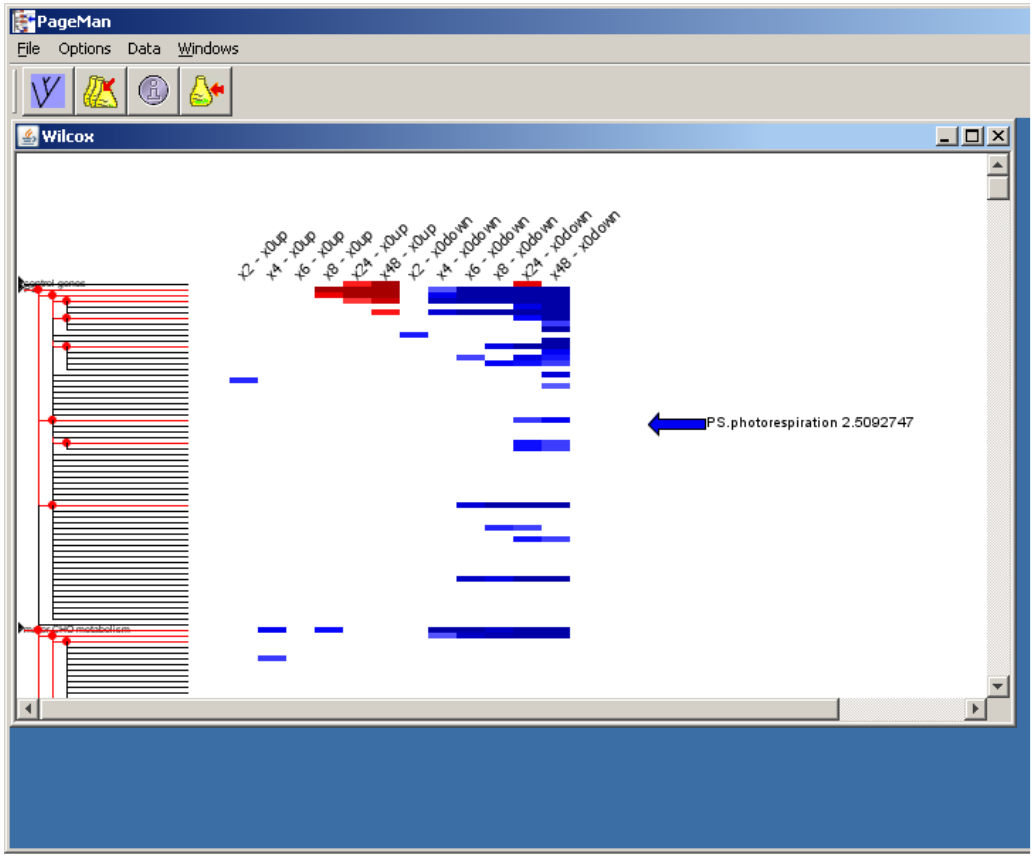


Right (apple) Clicking on the hierarchy allows to expand all nodes, annotate all significant BINs, hide all non significant BINs, collapse all nodes where no subnode is significant or to zoom in on processes.

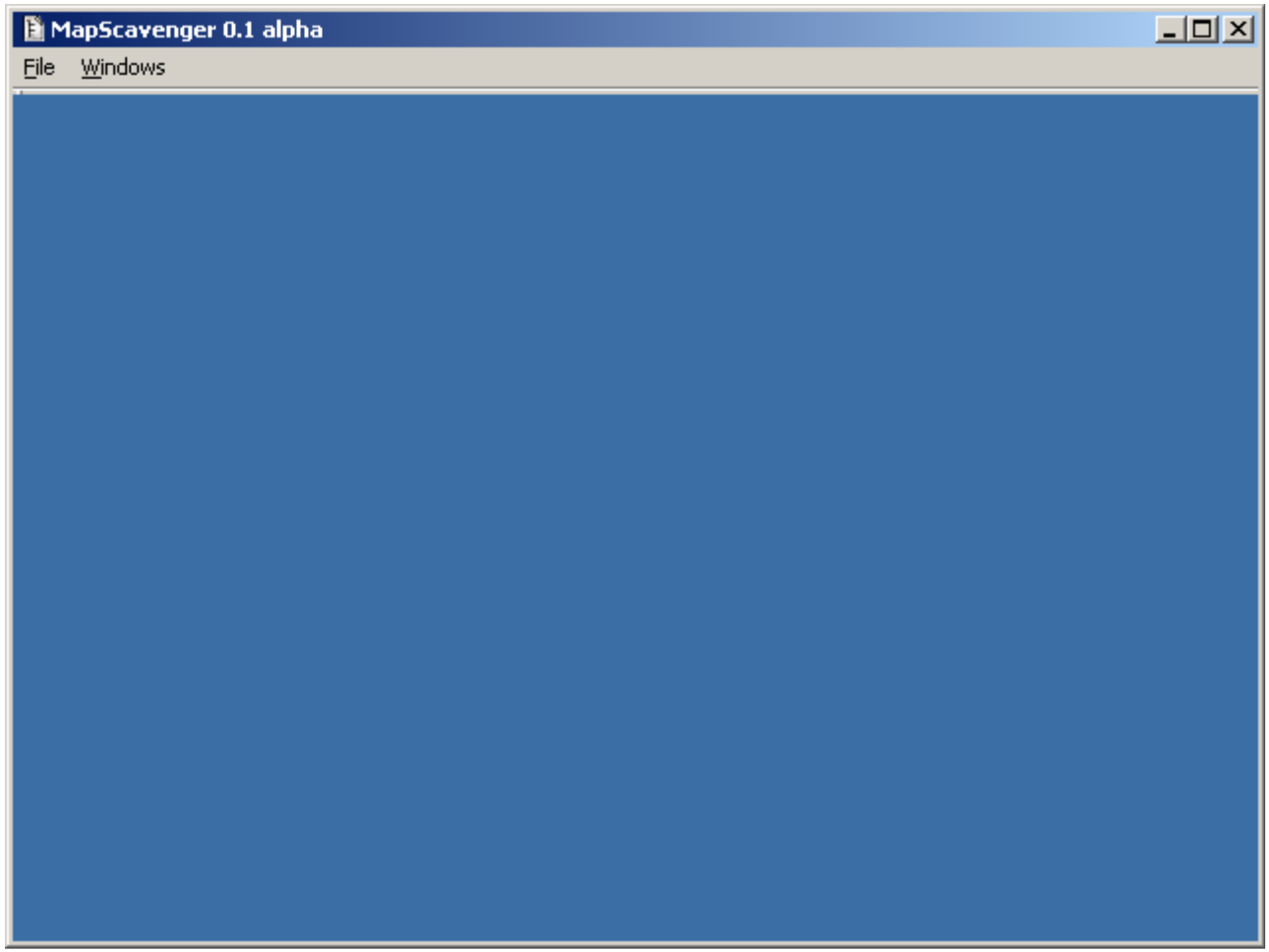




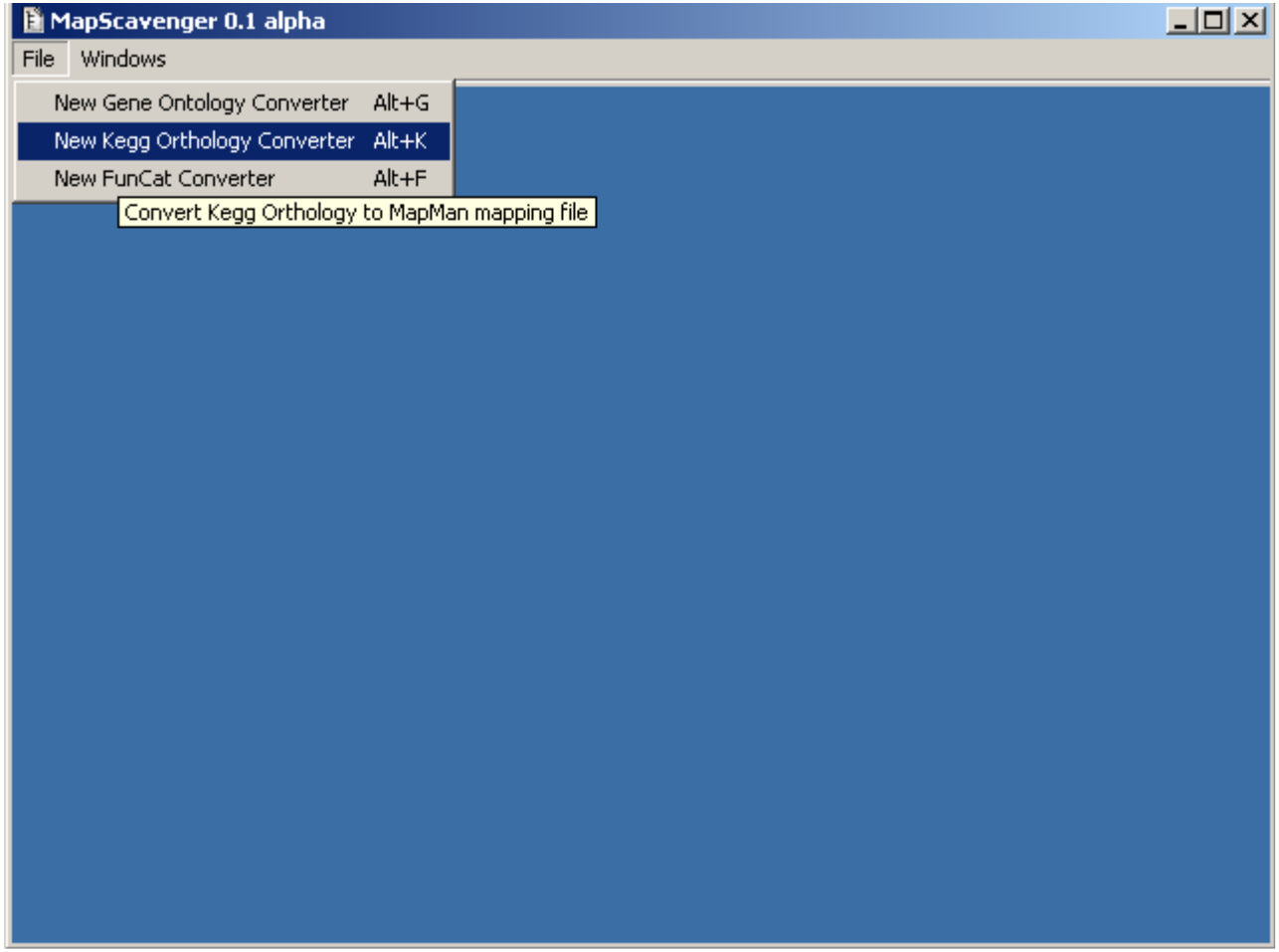
Further data can be added by clicking the “Add experiment” button.  As long as the same ontology is used, data from different species can be used as well,.



Using other biological ontologies is possible as well. To this aim, these need to be converted into MapMan mapping files. A tool called MapScavenger is provided to do so. It allows the conversion of GO, MIPS Funcat and KEGG ontology files.

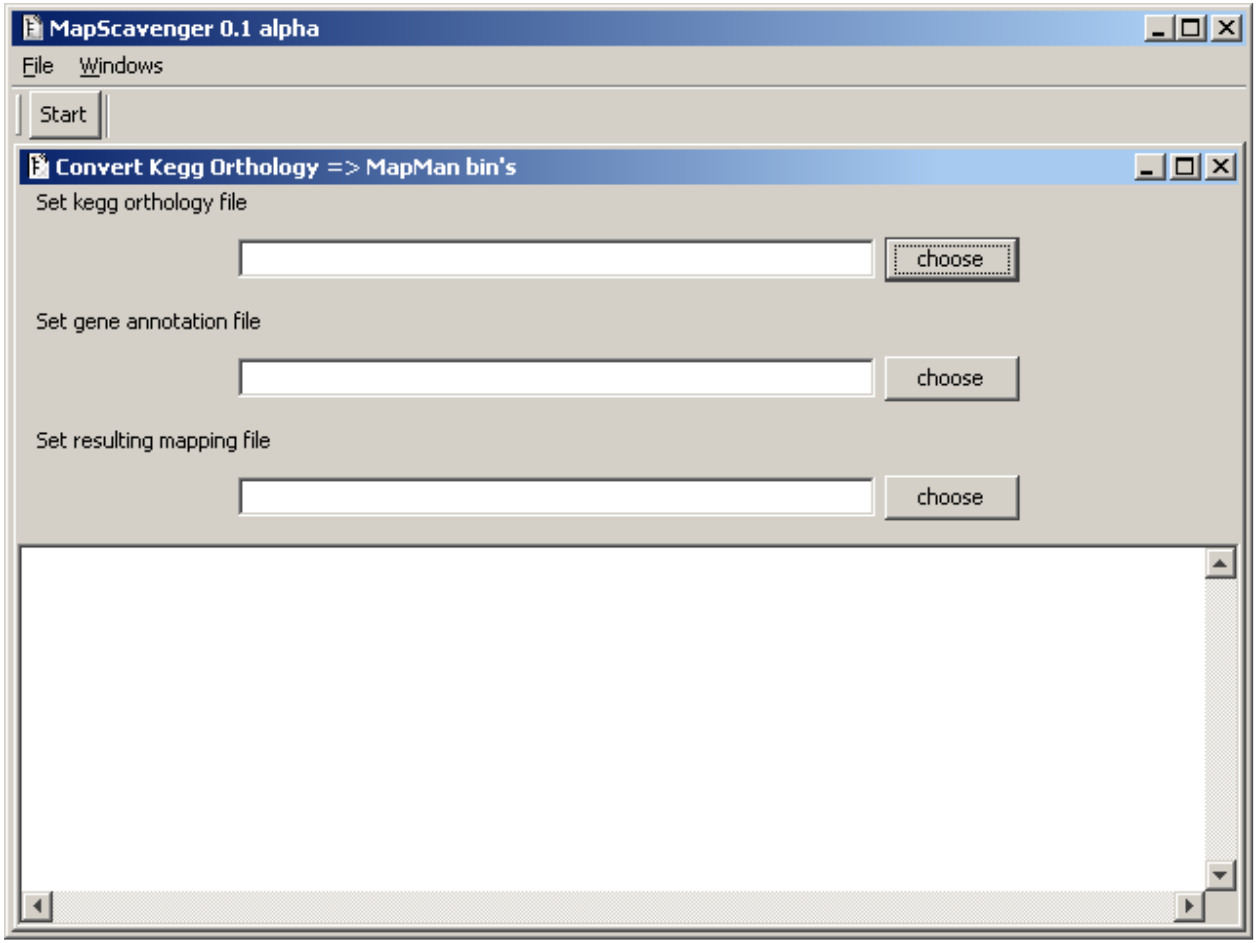


By Selecting “File-> New Kegg Orthology Converter” a KEGG conversion process is started. Please bear in mind that you might have to acquire a license in order to use KEGG.

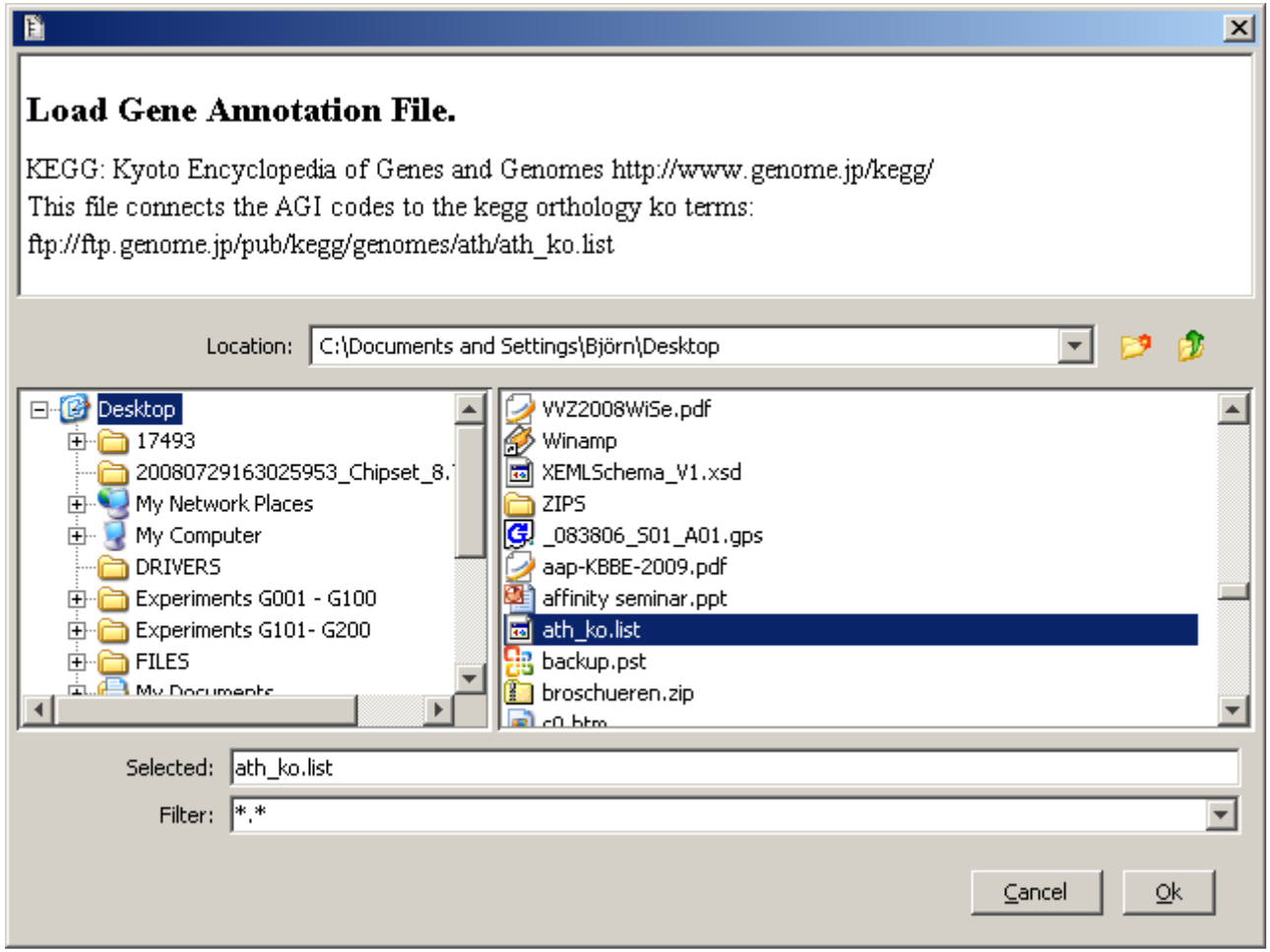


What you then need is an overview of the KEGG ontology, this is available from http://www.genome.jp/dbget-bin/get_htext?ko00001.keg+-f+F+D.
Browse to this location and save the file as html from your browser.

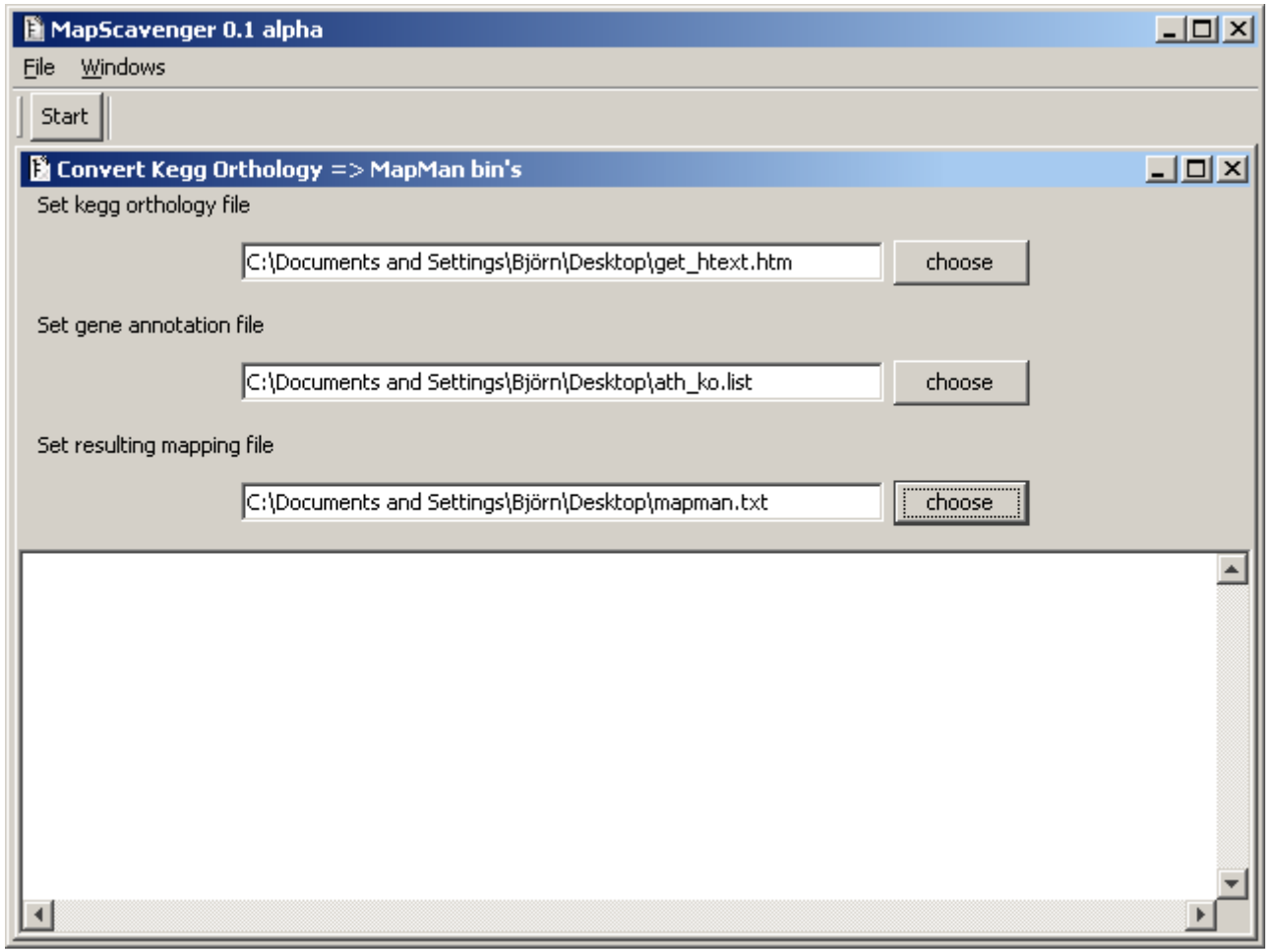
Also save the file ftp://ftp.genome.jp/pub/kegg/genes/organisms/ath/ath_ko.list for the Arabidopsis thaliana classification. Please note that the location of these files changes frequently in KEGG.



Localise the files that you downloaded and determine where to save your new MapMan file.



After all files have been loaded and set up, press the Start icon.



If everything goes well you should have a mapping file of your own choice to be used with PageMan.

As we have to interpret the data from the website that you downloaded, whenever this website changes we have to adapt the Scavenger. If you have problems running the Scavenger drop us a line and we try to adapt it to the latest KEGG websites.