KaPPA-View 4

The Kazusa Plant Pathway Viewer, Version 4.0

Advanced Manual

ver. 1.0

Kazusa DNA Research Institute

Legal Notices

We have done our best to ensure that the material contained within this document is both useful and accurate. However, please be aware that errors may exist in this document and that the information contained in this manual is subject to change without notice. We make no warranty of any kind with regard to this material. We shall not be held liable for errors contained herein and/or for direct or indirect damage that might be incurred as a consequence of using of KaPPA-View4.

Table of Contents

1. Introduction

1-1. Background and History

The application of DNA array technology and chromatographic separation techniques coupled with mass spectrometry (MS) to transcriptomic and metabolomic analyses in plants has resulted in the generation of considerable quantitative data related to transcription and metabolism. The integration of "omic" data is one of the major concerns associated with research into identifying gene function. Thus, we developed a web-based tool, KaPPA-View, (The **Ka**zusa **P**lant Metabolic **P**athway **View**er), for representing quantitative transcript and/or metabolite data using comprehensive metabolic pathway maps for Arabidopsis thaliana (Tokimatsu et al. 2005, Tokimatsu et al. 2006, and Sakurai et al. 2006). We prepared about 130 leaves of original metabolic pathway maps in Scalable Vector Graphics (SVG) format to represent the values of gene and compound on the map symbols dynamically according to the experiment data.

In 2006, we released the second version of the tool, KaPPA-View2, which implemented a function to overlay curves on a pathway map representing gene-to-gene and metabolite-to-metabolite relationships such as co-expression and co-accumulation correlation coefficients. This novel representation of the data facilitated better understanding of metabolic regulation within the genes of a pathway map. In KaPPA-View3 which released in 2008, we expanded this function to draw the curves across maps. In addition, users able to analyze correlations between genes on the maps and non-metabolic genes such as transcription factors and genes for signal transductions by creating simple maps which are automatically generated according to the users input of arbitral gene IDs. Furthermore, KaPPA-View3 can deal with metabolic maps of any organisms, means these pathway analyses are applicable not only to Arabidopsis but also various kind of plants.

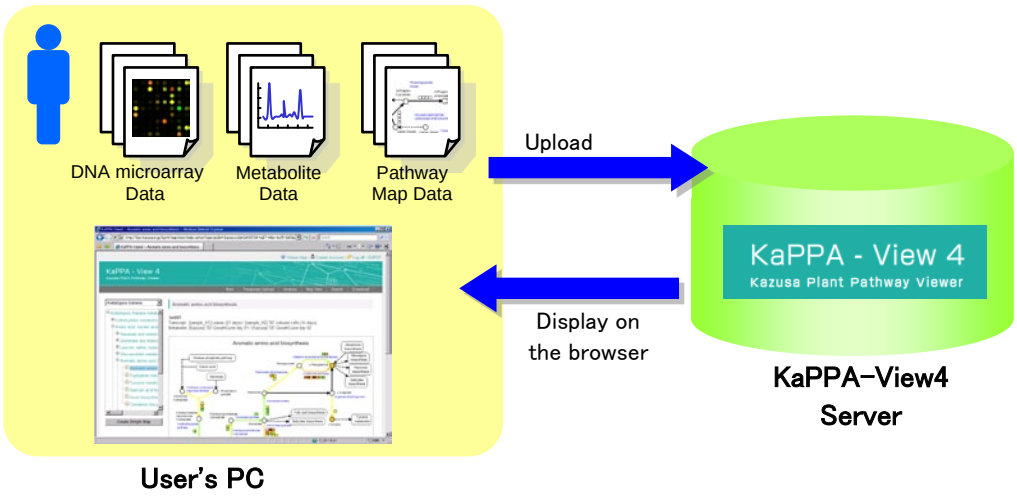
As the size of the correlation data become huge, the performance of the data analysis was not so good when correlation curves were drawn on the map(s) with a lot of gene and compound symbols. Therefore, we reviewed the database system, web application programs, and the representation procedures based on the SVG. By using Flash technology, a drastic speed up has accomplished for the representation of the maps. Thus we released the latest version of the system, KaPPA-View4 in 2010 (Sakurai et al. 2011, http://kpv.kazusa.or.jp/kpv4/). The version is also improved on the function to access from the external systems, by which developers of other databases or applications can easily incorporate a viewer function to view their omics data on KaPPA-View4. In addition to the KaPPA-View4 system with the original pathway maps (KaPPA-View4 Classic), we provide another version of KaPPA-View4 where pathway maps from KEGG PATHWAY (http://www.genome.jp/kegg/) are available (KaPPA-View4 KEGG, http://kpv.kazusa.or.jp/kpv4-kegg/). We wish the KaPPA-View4 families would be a powerful tool on your desk side to discover novel gene functions.

What is "KaPPA"?

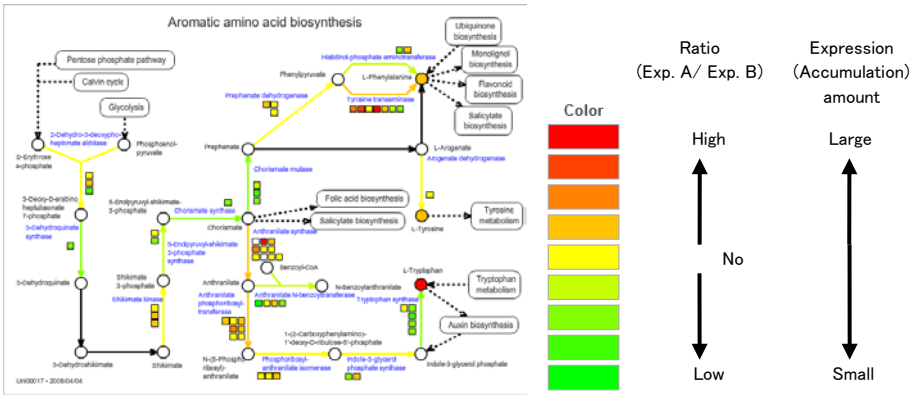
"KaPPA" stands for the real name of the tool "Kazusa Plant Metabolic Pathway Viewer". It also stands for the name of an imaginary creature in Japanese fairy story. "Kappa", which looks like a man but with a dish on his head and a large shell on his back, lives in a stream, such as a turtle does, and catches game from the surrounding world. KaPPA is always looking at transcriptome and metabolome to catch big game from the plant world.

1-2. Overview of KaPPA-View4

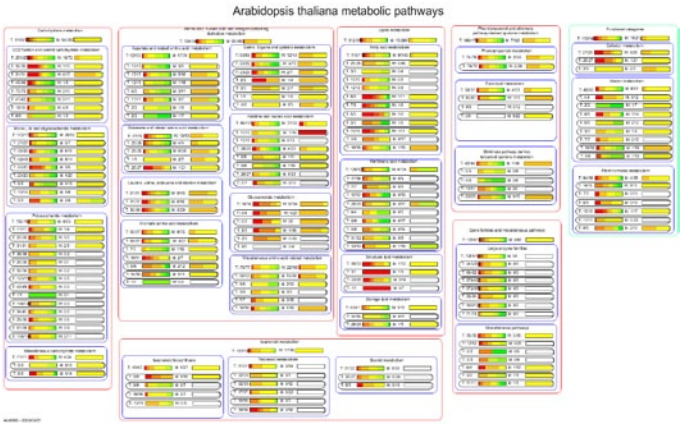
When you upload your own DNA microarray data and/or metabolite data to the system through the web browser, KaPPA-View4 displays the fitting of data for each gene or compound on the metabolic pathway maps.



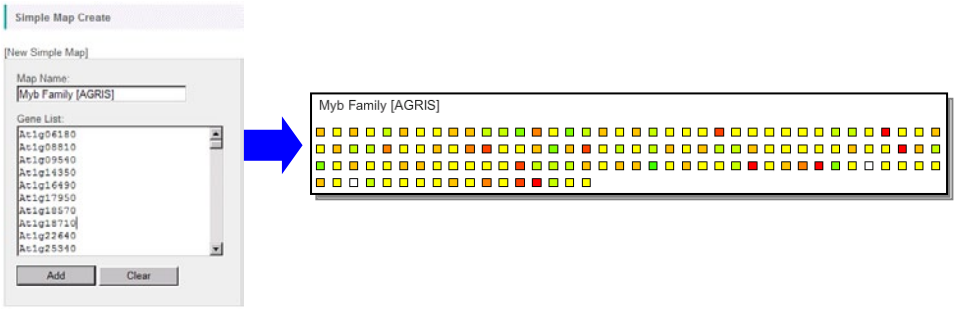
On the pathway maps, genes and compounds are represented by squares and circles respectively. The symbols are painted in different colors depending on the values such as changes in the ratio between two experiments and the amounts detected in one experiment.



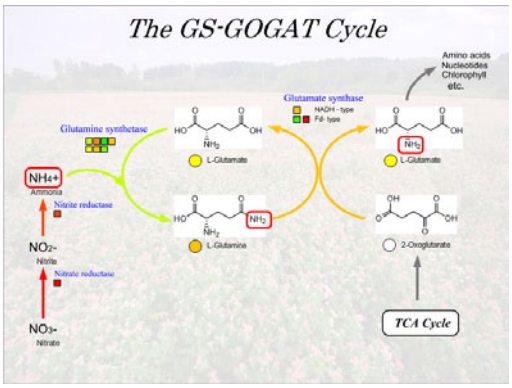
On the "Bird's-Eye Maps", you can view the summarized values for all maps and find out the pathways which have changed considerably.



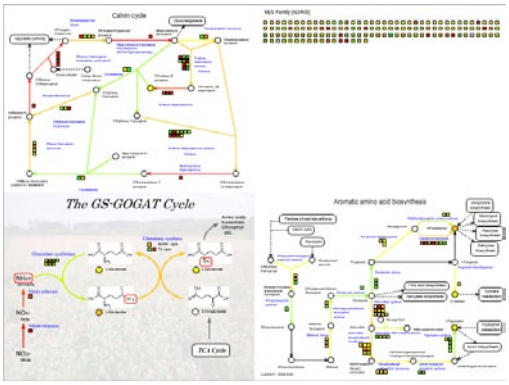
Non-metabolic genes which do not exist on the maps - such as transcription factors - can be analyzed. By entering the gene IDs, you can create simplified maps.



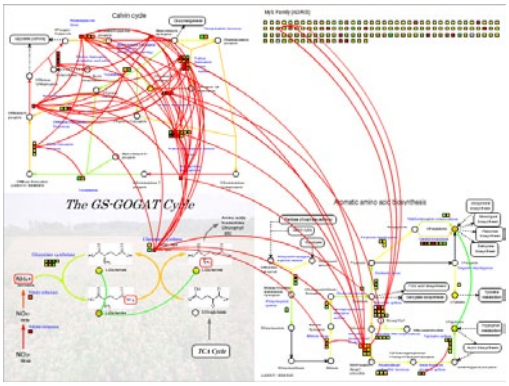
In addition, the users can also use the pathway maps prepared by themselves.



Up to four maps can be viewed at once in a single browser window.



Furthermore, gene-to-gene and/or metabolite-to-metabolite relationships such as co-expression correlations of genes can be displayed on the maps. This is the distinctive feature of KaPPA-View and will help you, for example, to analyze the relationships between metabolic genes and transcription factors that control their expressions.



KaPPA-View4 can handle multiple species, and genes of several species can be displayed side by side on the maps. The system also provides functions to upload and view the omics data from external applications.

1-3. User Setup

As KaPPA-View4 is a web-based system, it works well with major web browsers (Microsoft InternetExplorer, Firefox, Google Chrome, Safari and Opera) on any operating systems (Windows XP/Vista/7, Mac OS X and Linux).

Although the Adobe Flash Player plug-in (ver.9 or higher) is required to display the pathway maps, it is already installed in your browser in the most cases. If your browser doesn't have it, please install it according to the following site.

http://www.adobe.com/products/flashplayer/

Please be sure to install the latest version of java plug-in in your browser. The older version might be cause of unexpected trouble in page constructions.

The operation of KaPPA-View4 was tested in the following settings.

OS	Windows XP / Vista / 7 (Microsoft)	Mac OS X (Apple)
Browser	Internet Explorer 6, 7, 8 Mozilla Firefox 3.0.10, 3.5.2, 3.6.10 Google Chrome 3.0, 6.0	Safari 4.0.4 Mozilla Firefox 3.5.6, 3.6.10 Opera 9.63, 10.10

*In the case of Opera on Mac OS X, the full screen view of the maps does not work.

*Disable the pop-up blocking function of your browser. It is enabled in default for Safari, Opera, and recent Firefox and Google Chrome.

1-4. Other Manuals

This manual describes all the functions of KaPPA-View4. Refer to "Manual for Beginners" to learn the main functions and work flow of the analyses. The procedure to create User Maps using free-software "Inkscape" is described in detail in "Manual on User Map Creation". These manuals are also available from the top page of the KaPPA-View4 site.

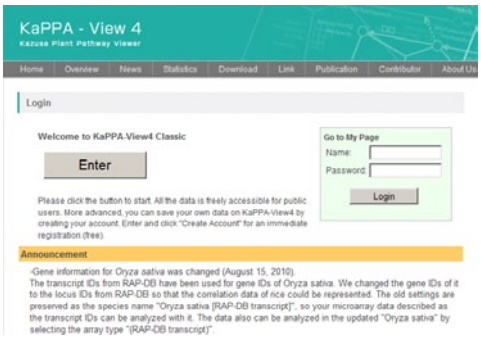
2. Login, Logoff and User Privileges

For the first step of the analyses, we introduce the way to login, main menu and how to upload the experimental data to KaPPA-View4. There are two types of users in KaPPA-View4, Guest Users and Power Users.

2-1. Access to KaPPA-View4

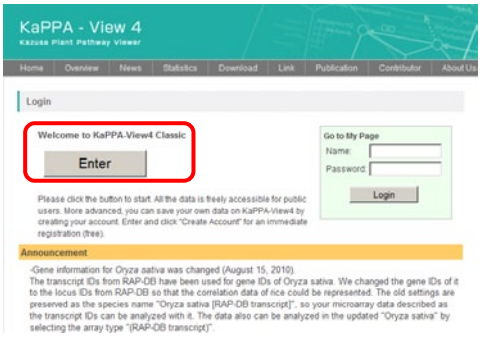
Visit to the URL below to show the top page of KaPPA-View4.

http://kpv.kazusa.or.jp/kpv4/

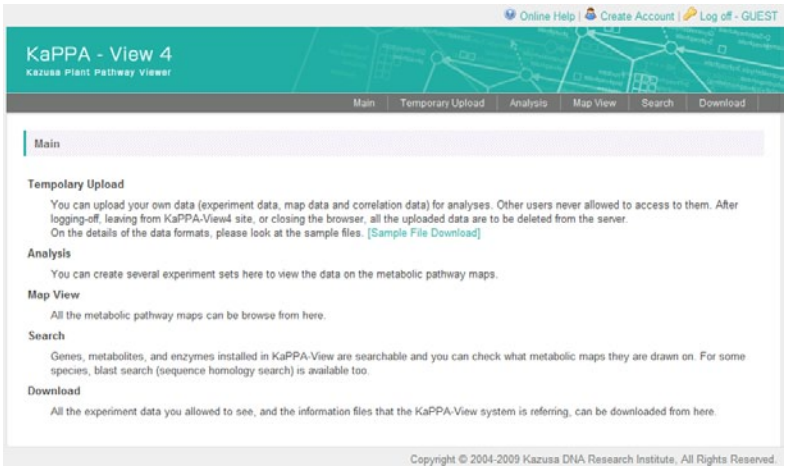


2-2. Login as a Guest User

Click "Enter" button at the top page.



After logging-in, you can see the main window.



2-3. Main Menu

The main menu is places at the top of the window.

- ***Main***

Return to the main window.

- ***Temporary Upload***

Uploading your experimental data, User Map data and correlation data for your analyses is operated through this menu. All the uploaded data is going to be deleted completely after you are logging-out.

- ***Analysis***

Uploaded data is displayed on the pathway maps through this menu. It serves the central function of KaPPA-View4.

- ***Map View***

You can browse plain pathway maps with no data from here.

- ***Search***

You can search genes, metabolites and enzyme reactions from here, and access to the pathway maps which they are on. Homology search function by blast to find genes is also provided.

- ***Download***

The default experimental data publicly available on the KaPPA-View4 and information data for genes, metabolites, reactions and maps for each species are downloadable as text files.

2-4. User privileges

2-4-1. Guest and Power Users

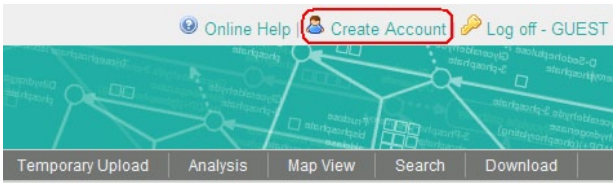
There are two types of users, Guest User and Power User. Power Users can save their experiment data on the KaPPA-View4 server, facilitating to start analyses immediately after logging-in. There are no differences on the analysis functions between the user types.

Guest User	The data uploaded to the KaPPA-View4 is removed when the user logging-off, therefore data have to be uploaded at the beginning of each analysis.
Power User	The data can be registered on the KaPPA-View4 server, therefore they are immediately utilized for the analyses.

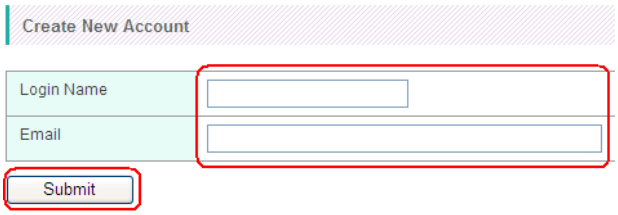
The Guest Users can create a Power User account by simple steps.

2-4-2. Create User Account

Click on "Create Account" at the top right of the window after logging-in.



A pop up window appears. Enter your favorite name and e-mail address, then press "Submit".



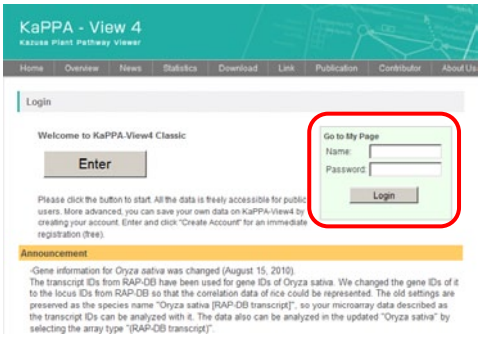
An e-mail informing the login password will be sent immediately to the address.

2-4-3. Expiration of Power User

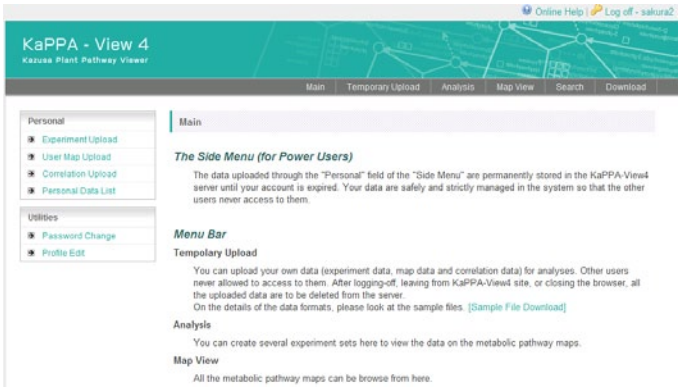
The Power User account will be removed automatically if the Power User have not logged-in to the system for 30 days. The data uploaded by the Power User will be deleted too. An e-mail alerting this will be sent to the Power User a week before the expiration date.

2-5. Login as a Power User

Enter user name and password at the top page of KaPPA-View4 and press "Login" to enter as a Power User.

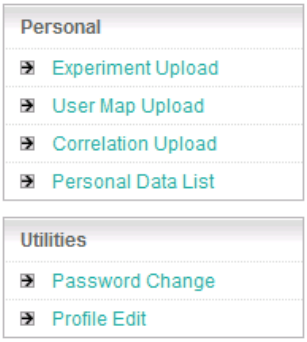


The main window for Power Users appears.



2-5-1. Power User Menu (Side Menu)

An additional menu for Power Users is visible on the left hand side of the main window.



- ***"Personal" Column***

Uploading of experiment data, User Maps, and correlation data can be operated. The data uploaded through the side menu are stored in the KaPPA-View4 server. Editing and removing of the data can be done on the "Personal Data List". The data are securely managed in the system and never accessed by the other users.

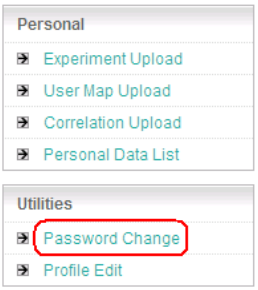
- ***"Utilities" Column***

The password and user information such as e-mail address can be edited.

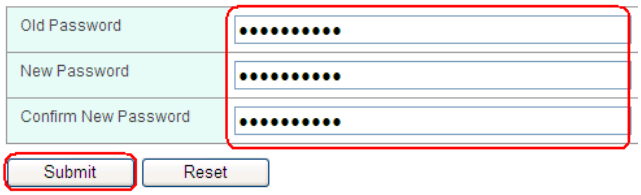
2-5-2. Changing the Password

We have done our best to manage securely the Power Users' data, but we recommend changing the password periodically to ensure it more secure.

Click "Password Change" on the Side Menu.



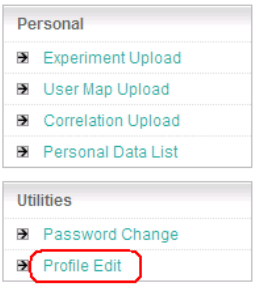
Enter old and new passwords and click "Submit".



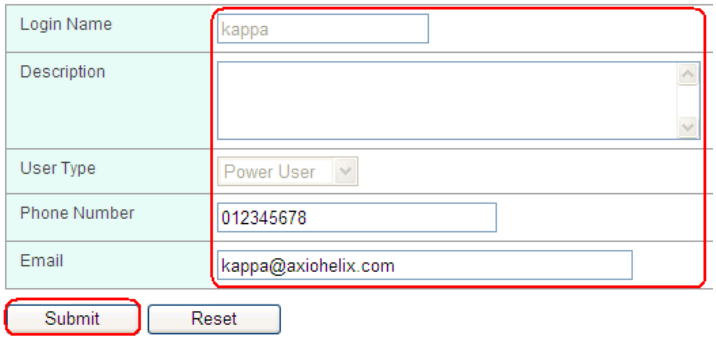
2-5-3. Editing Power User Information

The Power User can edit the user information. The information is never used for any propose without a permission of the Power User other than recognizing the user and sending announcement e-mails by the KaPPA-View4 administrator.

Click "Profile Edit" on the Side Menu.

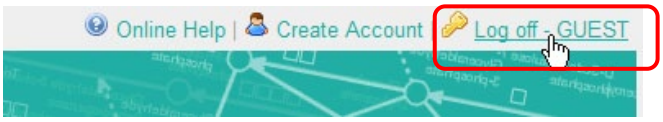


Edit the information and press "Submit".



2-6. Logoff

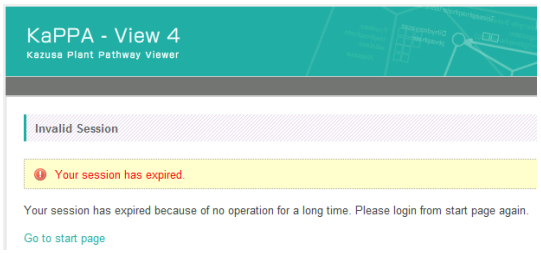
You can log-off from the system, by clicking "Log off" on top-right of the main window.



All your data uploaded according to the next chapter will be deleted from the system after logging-off.

2-6-1. Automatic Logoff

If you don't do any operations for 60 min after you log-in, the system regards as you are log off. The message below appears when you do an operation again.



You are automatically log off when you close all the browser's windows too.

3. Uploading and Management of Data

The users can upload their own experiment data, map data, and correlation data to KaPPA-View4 and can analyze them on it. This chapter covers the procedures for uploading and management of these user's data. Refer to the **9. File Format** for the details of the uploading files.

3-1. Temporary and Permanent Uploading

There are two types of data uploading procedures as follows.

Temporary Uploading	The uploading is performed through the "Temporary Upload" function on the Main Menu. The uploaded data are treated as temporary data and deleted after logging-off.
Permanent Uploading	The uploading is performed through the Side Menu of Power Users' main window. The uploaded data are stored in the KaPPA-View4 server and can be utilized by the user immediately after logging-in. The data are removed from the server when the Power User's account is deleted.

The uploading procedures are essentially the same, except the menu to access them differs.

3-2. Uploading Experiment Data

Experiment data are those such as gene expression data obtained by DNA microarrays and metabolite data detected with metabolomics technologies. The experiment data uploaded to the KaPPA-View4 can be accessed through the "Analysis" function and utilized for representation on the pathway maps.

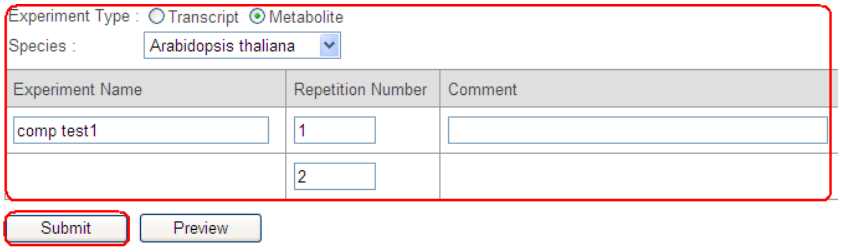
Follow the step below to show the Experiment Data uploading interface.

Temporary Uploading	Click "Temporary Upload" on the Main Menu and then select "Experiment" tab.
Permanent Uploading	Click "Experiment Upload" on the Side Menu.

Click "Browse" to select the uploading file on your PC, and press "Upload".

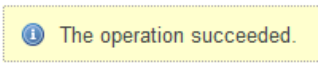
The selected file is scanned and the the names and repetition numbers of the experiment are listed. Select a type of data (Transcript or Metabolite) and species name, and change experiment names and repetition

numbers if it needed, and then click "Submit" to start uploading and registration the data to the KaPPA-View4 server.

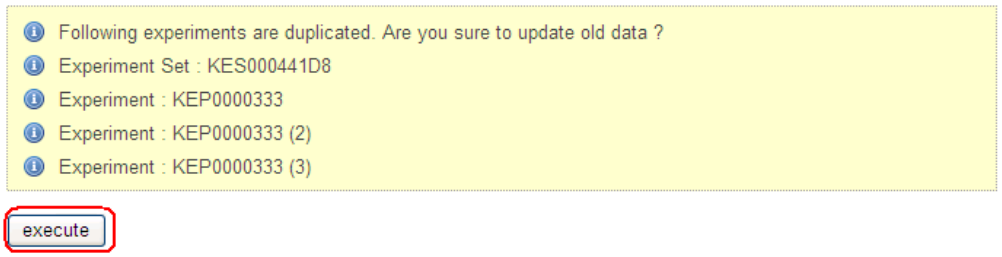


*Click "Preview" to confirm the data by listing first 100 lines of the file before submitting.

After completing the process, you can see the following message. It takes a few tens of seconds.



If the file contained experiment IDs that are already used in the KaPPA-View4, the following message appears. By clicking "execute" the old data will be overwritten by new one.



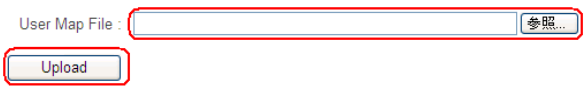
3-3. Uploading User Maps

In addition to the metabolic pathway maps that are provided by KaPPA-View4 as defaults, users can upload their own pathway maps (User Maps) and utilize them for their analyses. The free software "Inkscape" is used to create the User Maps in scalable vector graphics (SVG) format. Refer to "Manual on User Map Creation" for the details of User Map preparations.

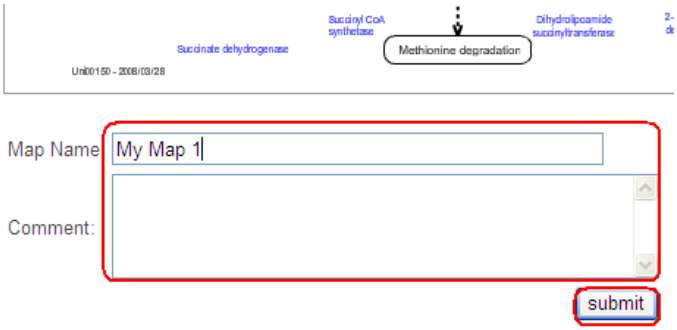
Follow the step below to show the User Map uploading interface.

Temporary Uploading	Click "Temporary Upload" on the Main Menu and then select "Map" tab.
Permanent Uploading	Click "User Map Upload" on the Side Menu.

Click "Browse" to select the User Map file (.svg) on your PC, and then press "Upload".

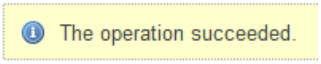


A preview appears under the interface. Enter the map name and press "Submit" button. The "Comment" can be left blank.



The map name will be displayed on the Metabolic Pathway Tree on the "Analysis" and the "Map View" functions.

After completing the process, you can see the following message.



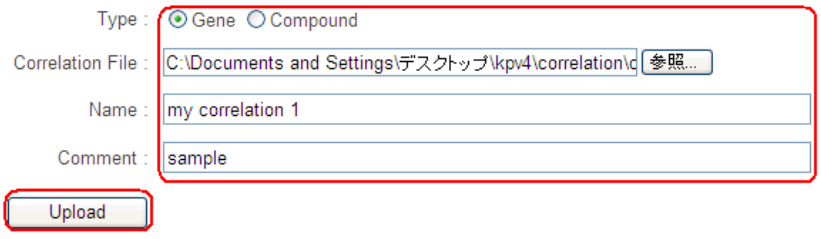
3-4. Uploading Correlation Data

Gene-to-gene and metabolite-to-metabolite relationships such as correlation coefficients of gene co-expressions and metabolite co-accumulations can be overlaid on the pathway maps. This is the unique function of KaPPA-View4, facilitates researchers to find new functions of the genes and their regulatory mechanisms.

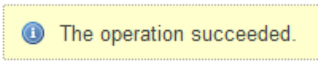
Follow the step below to show the Correlation Data uploading interface.

Temporary Uploading	Click "Temporary Upload" on the Main Menu and then select "Correlation" tab.
Permanent Uploading	Click "Correlation Upload" on the Side Menu.

Click "Browse" button to select the correlation data file on your PC. Select the data type (Gene or Compound), enter the data name and comment, and press "Upload".



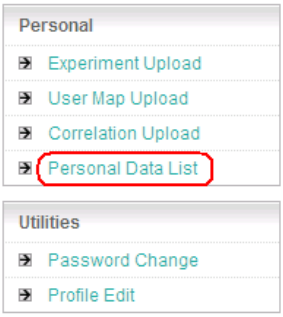
After completing the process, you can see the following message. It might take a few dozen of seconds to several minutes according to the file size.



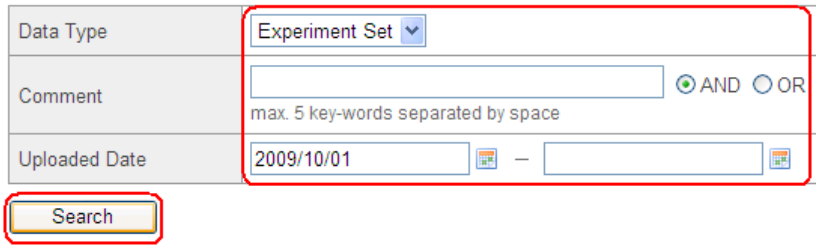
3-5. Management of Uploaded Data (for Power Users)

The data uploaded by a Power User according to the procedure for permanent uploading can be edited and deleted by the Power User.

Click "Personal Data List" on the Side Menu.

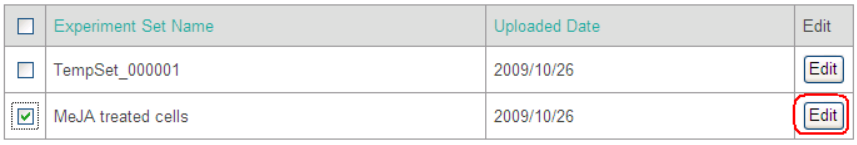


Select "Data Type" and click "Search" button to list up the uploaded files. The resulted items can be filtered by keywords and date of uploading.

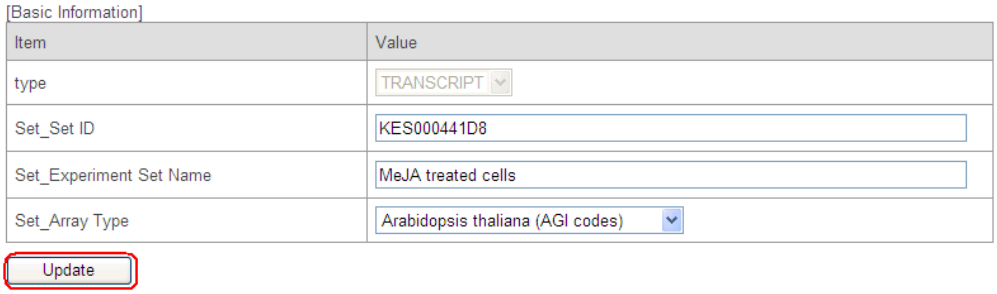


3-5-1. Editing of Data Information

"Edit" buttons appear at the right of each row of the list when "Experiment Set" was selected. Press the "Edit" button to pop up detailed information (meta data) for the experiment data.



Press "Update" after editing the meta data to reflect the alteration.

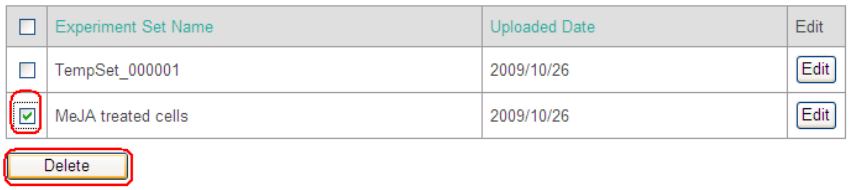

[Basic Information]

Item	Value
type	TRANSCRIPT
Set_Set ID	KES000441D8
Set_Experiment Set Name	MeJA treated cells
Set_Array Type	Arabidopsis thaliana (AGI codes)

Update

3-5-2. Removing Data

Check the items to delete and Click "Delete" button.



	Experiment Set Name	Uploaded Date	Edit
☐	TempSet_000001	2009/10/26	Edit
☑	MeJA treated cells	2009/10/26	Edit

Delete

3-6. Sending User Maps to the KaPPA-VIew4 Administrator (for Power Users)

Despite our best effort, there may be information error or deficiency in the default map. Once user maps with corrected errors or species specific pathways that default map didn't cover are completed, consideration them to the public is greatly appreciated.

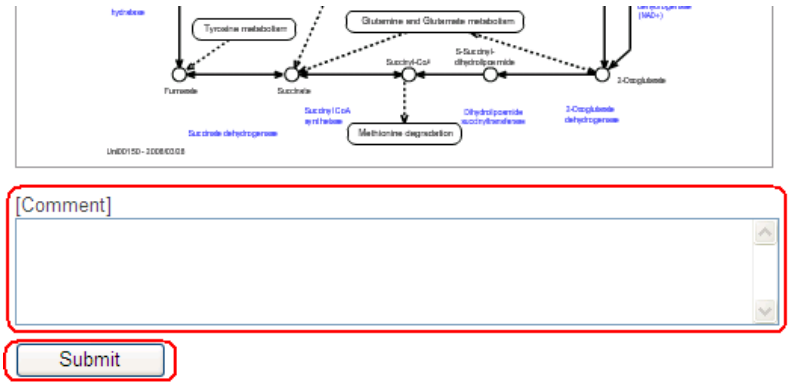
Power Users can send their user map to the KaPPA-View4 administrator with simple steps. The administrators will discuss on using the user map for improving the default map. We acknowledge the map donor with a great appreciation.

The "Send to Admin" button appears when "Map" is selected as a data type. Check an item and click "Send to Admin" button.

	Map Name	Map Comment	Map Date
☑	My Map 1		2009/10/26

Delete Send to Admin

A window pops up to show the preview of the map. Enter comments to the administrators and press "Submit" to send the map data.



The message can be either in English or Japanese. The administrator will e-mail the user to discuss the treatment of the user map.

Do not forget to send the following information.

- Name

Used for specifying the contribution.

- Affiliation

Used for specifying the contribution.

- Contact e-mail address

The administrator may contact regarding the user map information.

- Use of the user map

1) Information addition to the default map

2) Curation of the default map

3) Adding a new map

Please let us know the suitable species and position on the Pathway Tree.

Note:

Convert all characters to outlines when the user map to be sent uses any fonts other than Arial.

4. Data Analysis

This chapter covers the functions of data analyses on the metabolic pathway maps which are available through the "Analysis" on the Main Menu. KaPPA-View4 provides various analysis functions as listed below.

- The genes, compounds and enzymatic reactions represented as squares, circles and arrows, respectively, are painted with proper colors according to the experiment data.

- Up to 4 maps can be simultaneously displayed and used for the analysis on a single browser window.

- Genes not drawn on the default maps can be analyzed by creating "Simple Maps" from the user-input gene IDs.

- Gene-to-gene and/or metabolite-to-metabolite relationships such as correlation coefficients of gene co-expressions and metabolite co-accumulations can be overlaid on the pathway maps.

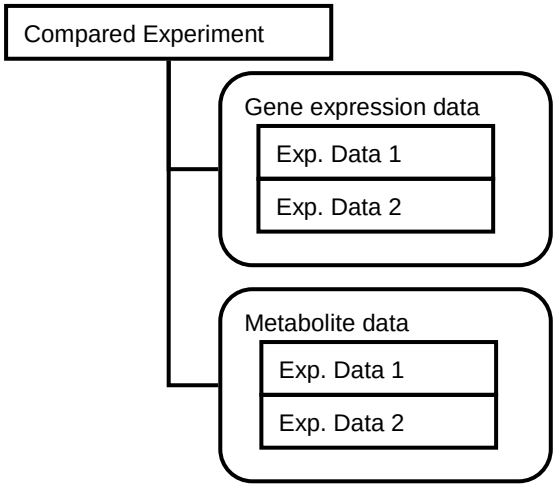
- Overview of experiment data and correlation data are available through Bird's Eye Maps.

- Experiment data can be compared between the species and the results are represented on the maps.

- Two experiment data from a species can be compared and represented on the maps.

4-1. Data Selection for Browsing

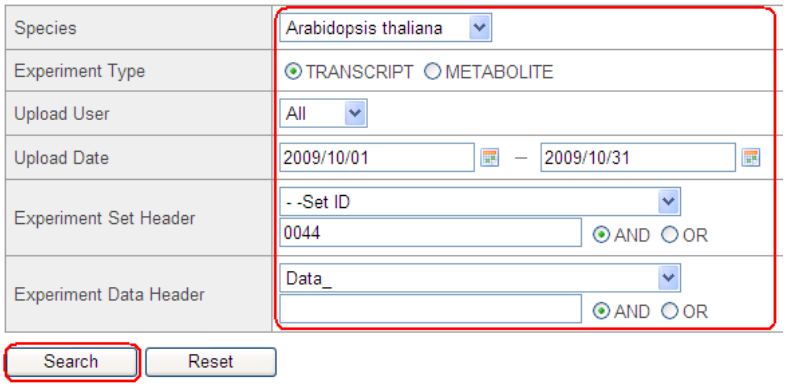
In KaPPA-View4, a unit of analyzing data is defined as a compared data between two experiments. We refer this unit as a "Compared Experiment". One Compared Experiment is comprised of a pair of gene expression data, a pair of metabolite data, or both of them.



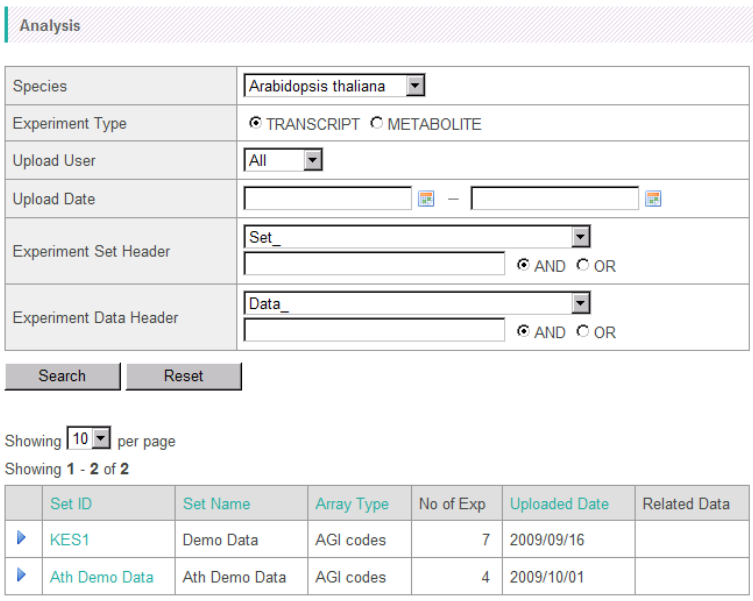
4-1-1. Creation of a "Compared Experiment Pair"

Click "Analysis" on the Main Menu.

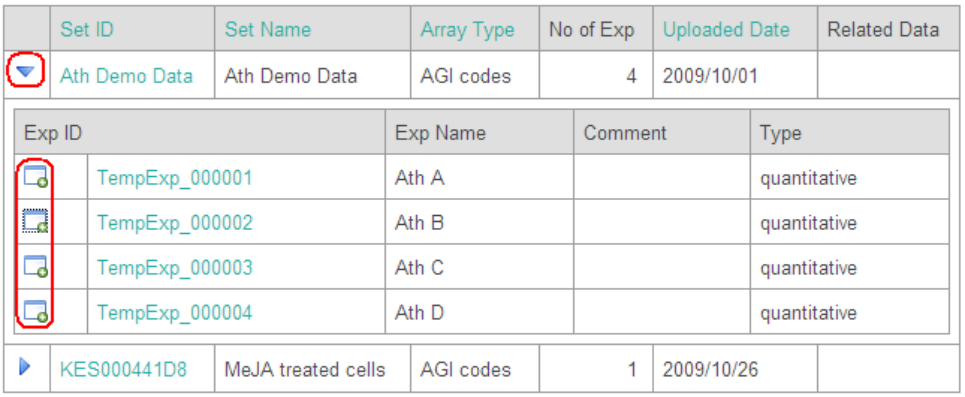
A data search window appears. Select a species, check Experiment Type (TRANSCRIPT or METABOLITE) and press "Search" button to list up the registered data. The results can be filtered by other conditions such as uploaded date if it needed.



The data uploaded by the users according to **3. Uploading and Management of Data** and the default data provided by KaPPA-View4 are available from the list.



In KaPPA-View4, several experiment data can be grouped and managed as "Experiment Set". For example, DNA microarray data obtained at several time points after a drug treatment can be registered as an experiment set. The items appeared in the list show the names of the Experiment Sets. By clicking an arrow head (▶) placed on the left of each row, each experiment experiment data belonged to the Experiment Set appears.

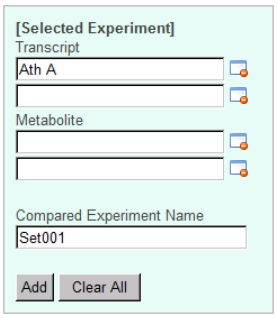


	Set ID	Set Name	Array Type	No of Exp	Uploaded Date	Related Data
▼	Ath Demo Data	Ath Demo Data	AGI codes	4	2009/10/01	

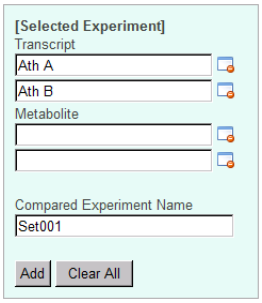
Exp ID	Exp Name	Comment	Type
TempExp_000001	Ath A		quantitative
TempExp_000002	Ath B		quantitative
TempExp_000003	Ath C		quantitative
TempExp_000004	Ath D		quantitative

▶	KES000441D8	MeJA treated cells	AGI codes	1	2009/10/26

Click a data icon ( ) to add the data on the "Selected Experiment" panel at the top-right.

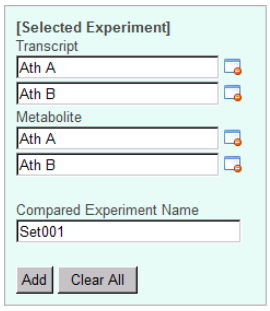


By clicking the icon of another data, the data is registered as the second data.

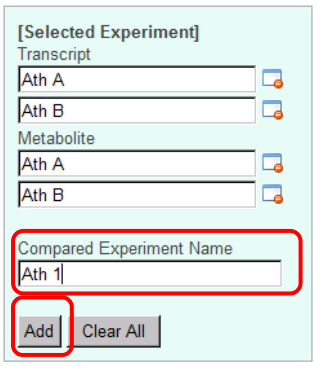


The added data can be removed by clicking the remove data icon ( ) on the "Selected Experiment" panel.

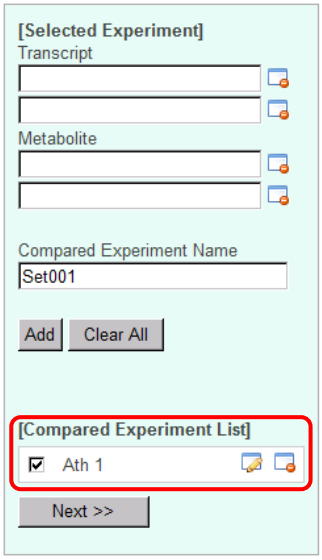
After selecting a pair of gene expression data, users can further select a pair of metabolite data as following the similar procedure described above. Select "METABOLITE" for Experiment Type for adding metabolite data.



After adding a pair of transcript data, a pair of metabolite data, or both of them, enter a name and click "Add" to register a Compared Experiment Pair.

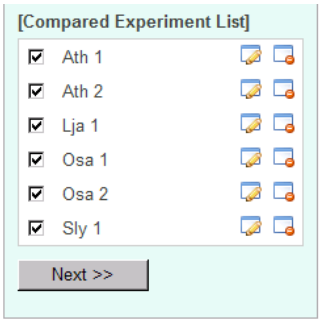


The registered Compared Experiment Pairs will be listed at "Compared Experiment List" as shown below.



Click remove icon ( ) or edit icon ( ) placed beside the Compared Experiment Pairs to delete or edit them. When the edit icon ( ) clicked, the registered information appears in the "Selected Experiment" panel. Edit the selection of the data and/or the name, and click "Add" button again to fix the alteration.

The users can register several Compared Experiment Pairs by repeating these steps.



The registered Compared Experiment Pairs are effective until logoff.

- ***Type of Experiment Data***

There are two types of experiment data, "quantitative" data and "ratio" data, which are recognized at the right most column of the experiment list.

A "quantitative" data holds quantitative values from a single experiment data. The users can register a pair of two quantitative data as a Compared Experiment Pair.

A "ratio" data holds ratio data such as gene expression ratio values detected by a 2-color DNA microarray experiment. The ratio data can be registered by itself alone without pairing to another data to create a Compared Experiment Pair.

	KEP1_5	[sample_B2] T87 cells - dark grown (10 days)	hybridized with [sample_B1]	quantitative
	KEP1_6	[sample_C1] Leaves (38 days)	hybridized with [sample_C2]	quantitative
	KEP1_7	[sample_C2] Stems (80 days)	hybridized with [sample_C1]	quantitative
	KEP1_8	[sample_D log (ratio)] T87 cells - MeJA treated vs control (2hr)	Log (ratio) data of methyljasmonate (MeJA) treated and untreated (control) T87 cells.	ratio

As the gene expression data and metabolite data must be described in log scale and linear scale, respectively, the changed values between the two quantitative data of the Compared Experiment Pair are calculated by the subtraction (for genes) and division (for metabolites) of them during the process of color representations on the maps. In the actual data processing, the "ratio" data is treated as a special case of a "quantitative" data which is paired to a control data where all the genes have value of 0 and all the metabolites have value of 1.

- ***Representation of Quantitative Data***

In addition of representations of a pair of two quantitative data (normal usage), a quantitative data from a single experiment can be represented also on the pathway maps to use this property. There are two approaches.

1) Upload the quantitative data obtained from a single experiment as a "ratio" data.

2) Upload the quantitative data obtained from a single experiment as a "quantitative" data, and upload another control data as "quantitative" which contains expression values of all the genes as 0 or values of all the metabolites as 1. Create a Compared Experiment Pair with the two data.

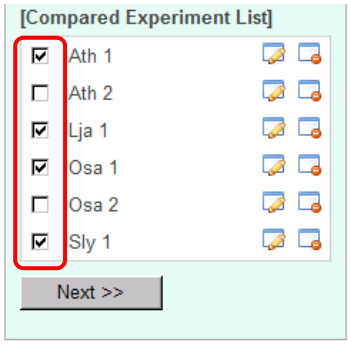
Note:

The color gradation which indicates the magnitude of the values in KaPPA-View4 is set so that no change genes and metabolites are painted in the color at the center of the gradation. As gene expression values are treated as log scale and metabolite values as linear scale, 0 and 1 as calculated ratio are the values corresponding to no change for the genes and the metabolites respectively. Therefore, to represent transcriptome and metabolome changes appropriately on the maps with the maximum range of color gradation, we recommend normalizing the data by a proper value such as global mean and median to adjust the representative to be colored in the center color of the gradation.

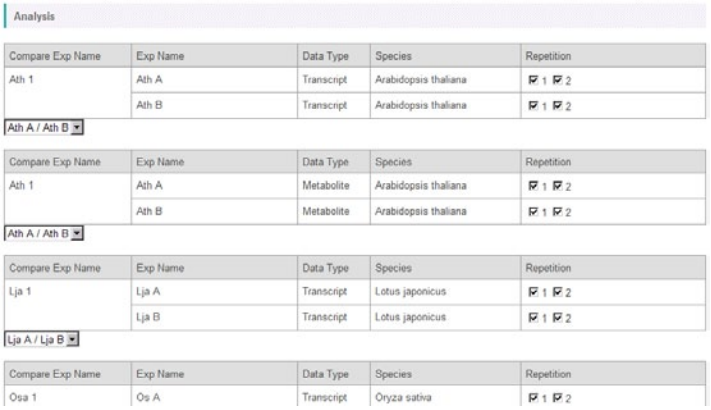
4-1-2. Selection of Compared Experiment Pairs for Browsing

Select Compared Experiment Pairs to be represented on the pathway maps. Up to 8 pairs can be selected here. The names of the selected pairs are listed at the upper part of the map browsing window, and by clicking them each data can be quickly displayed on the maps.

Check the Compared Experiment Pairs to browse at the "Compared Experiment List".

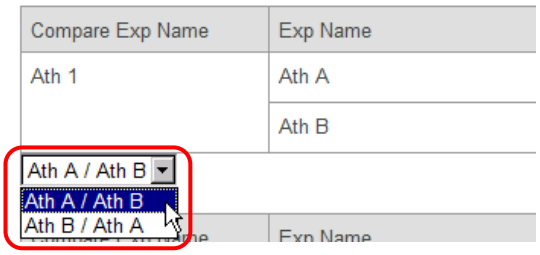


Click "Next" to transit to the next page.

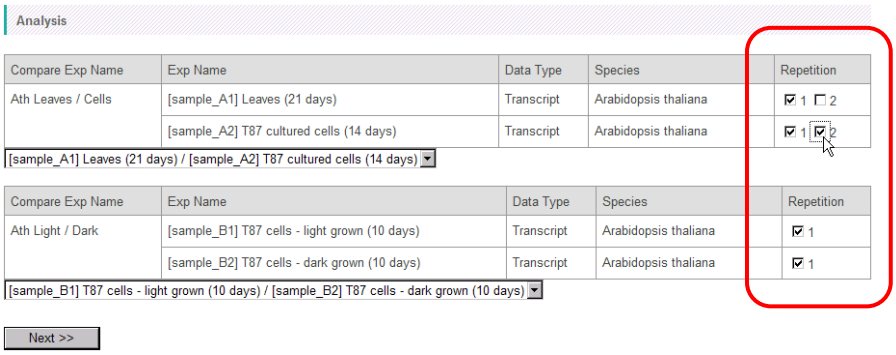


In this page appeared, users can change the directions of ratio calculations and can select the repetition data to be calculated.

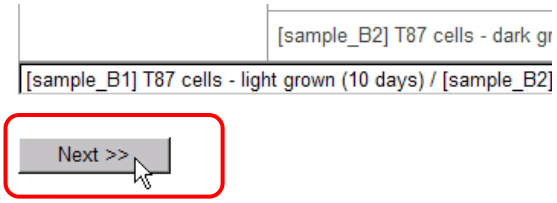
Select directions of the ratio calculations from the pull-down lists. This is disabled for "ratio" data.



An experiment data can include more than one experimental repetition and they can be distinguished by the repetition number (Refer to **9. File Format** for more details of data file format). The repetitions included in the data showed in the "Repetition" column. Uncheck the repetitions if you don't like to include them for the analysis. Representative values for the genes and metabolites are the mean values calculated only with the checked repetitions. And then, the ratios are calculated as the decided directions.

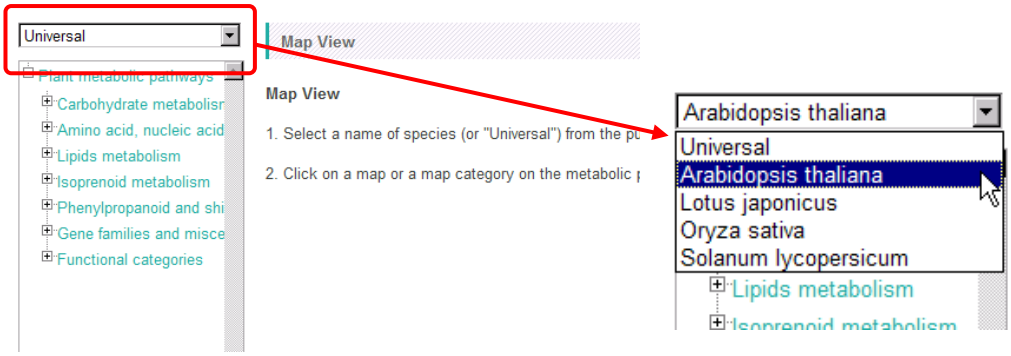


Click "Next" to proceed.

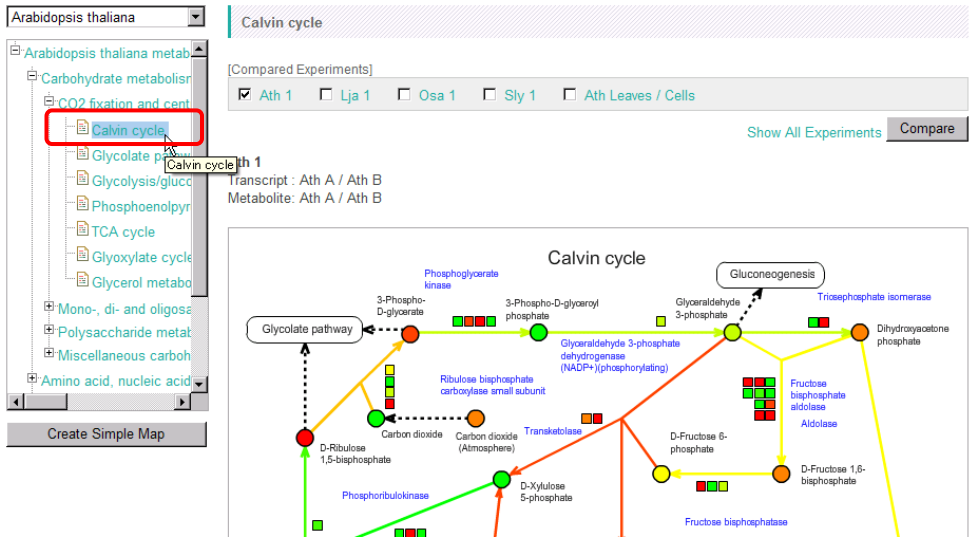


The data browsing window appears.

Select a species corresponds to the selected Compared Experiment Pair from the pull-down list.



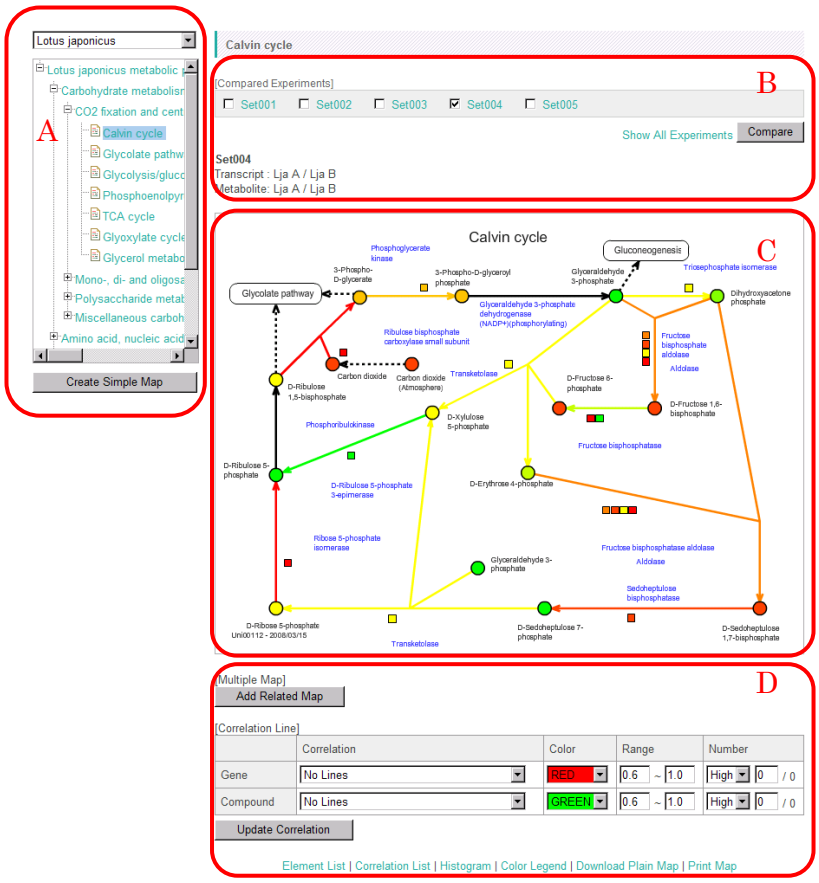
Select a pathway map from the "Metabolic Pathway Tree" to show the pathway map on the right side of the window.



4-2. Data Browsing on the Maps

4-2-1. The Data Browsing Window

The data browsing window is composed of 4 parts.



A: Pathway Tree

Metabolic pathway maps are categorized and represented as a tree structure. By selecting a species from the pull-down list, the metabolic pathway tree for the species displayed. Click a leaf of the tree to show the corresponding metabolic pathway map in the area C. When a branch of the tree is clicked a category indicator map corresponding to the category appears in the area C (Bird's Eye Maps, see **4-2-6. Bird's Eye Map**). When "Universal" is selected from the pull-down list, maps for representing all species genes will be displayed (Universal Map Mode, see **4-2-3. Map Mode**). The "Create Simple Map" button appears if a species other than "Universal" is selected, which is utilized for creating a Simple Map according to the user-input gene IDs (see **4-4-1. Simple Map**).

B: Upper Control Panel

The names of the selected Compared Experiment Pairs are displayed. By clicking one of them the corresponding data are represented on the maps in the area C. Data comparison between the Pairs can be performed through the interface in Universal Map Mode.

C: Map Area

According to the settings of the Pathway Tree and the Control Panels, metabolic pathway maps and category maps are displayed.

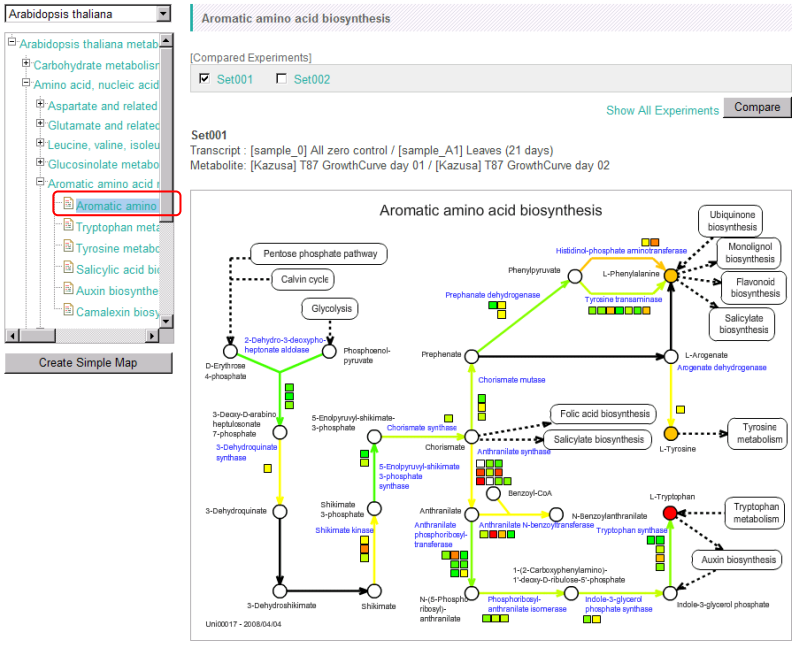
D: Lower Control Panel

There is a setting button for Multiple Map Mode where up to 4 maps are displayed in the area B (see **4-4-2. Multiple Map Mode**). There is also a control panel for correlation data overlaying (see **4-4-3. Overlay of Correlation Data**).

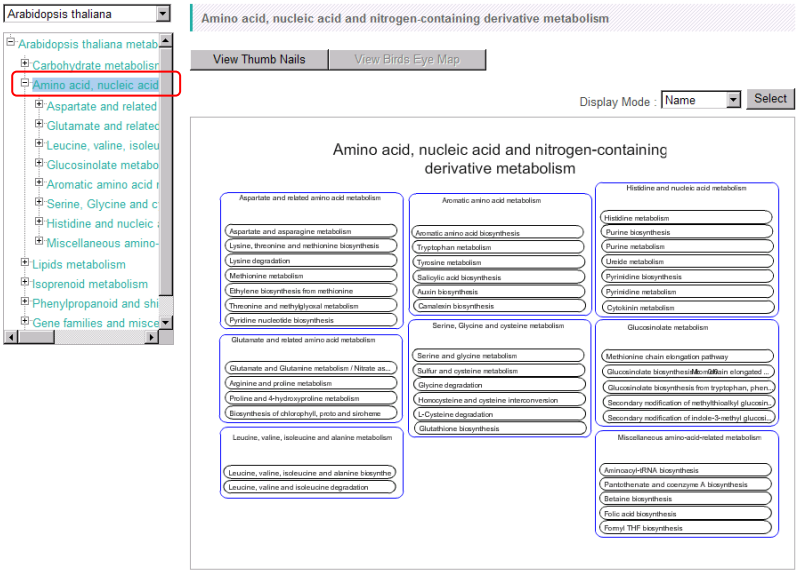
4-2-2. Metabolic Pathway Tree

The Metabolic Pathway Tree provides an interface to select species and pathway maps to view.

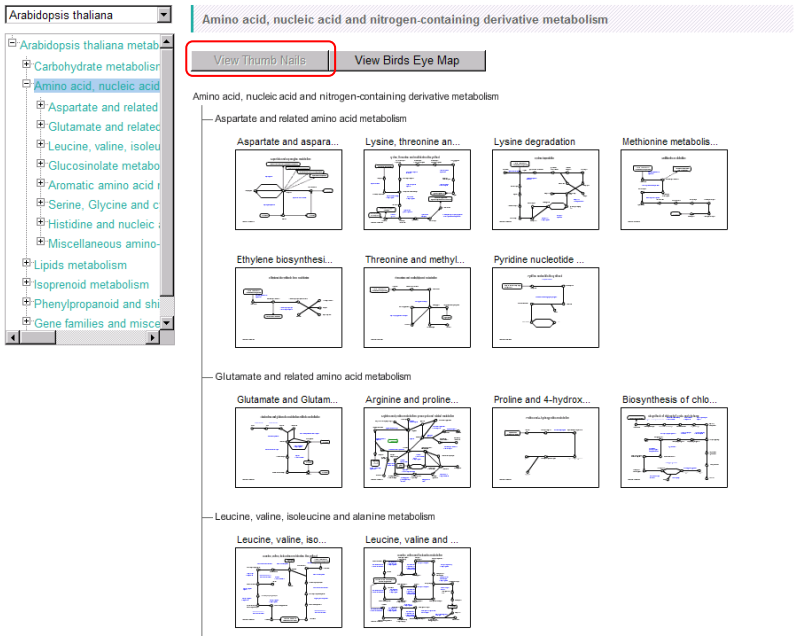
When a leaf (lower most layer of the tree) is selected, the individual pathway map is displayed.



When a branch (internal node of the tree) is selected, a map is displayed where all the pathway maps belonging to the category are schematically indicated as indicator bars (Bird's Eye Map, see **4-2-6. Bird's Eye Map**).

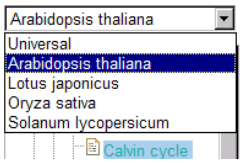


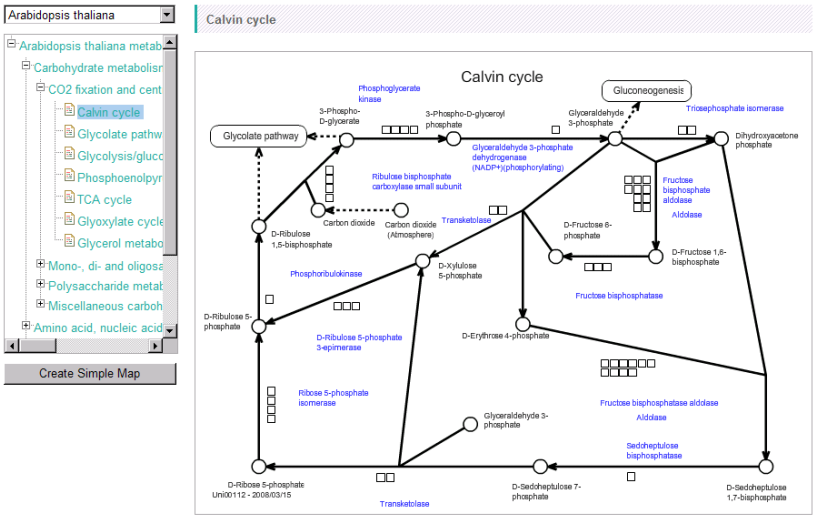
Click "View Thumbnails" button when a Bird's Eye Map is displayed to switch the Map Area to a tree-like representation of pathway map thumbnails.



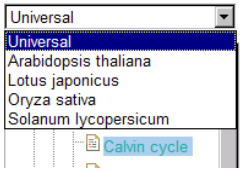
4-2-3. Map Mode

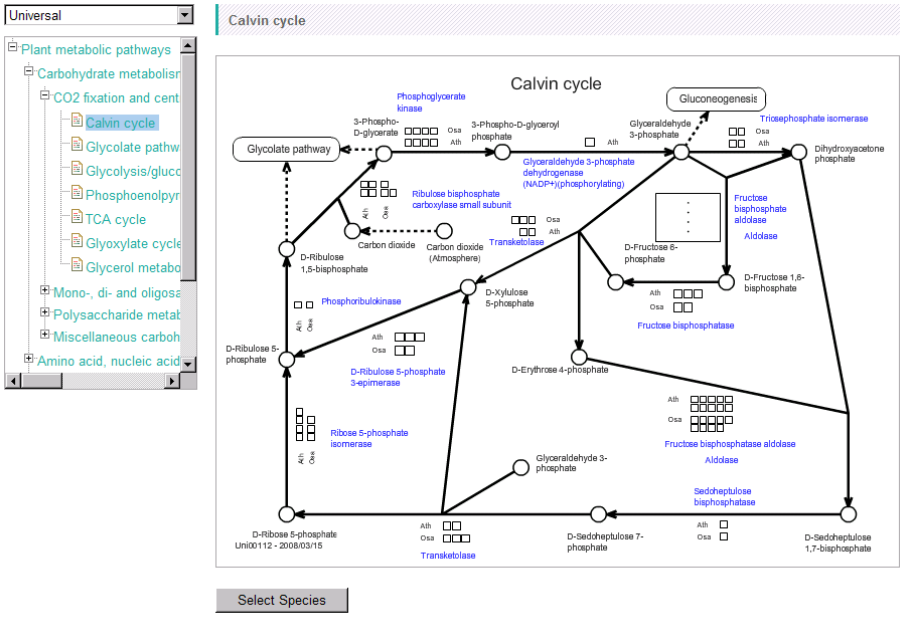
The pull-down list above the Pathway Tree is an interface to switch the species to view. When a species name is selected in the pull-down list, the Pathway Tree corresponding to the species appears, and the pathway maps with the gene symbols for the species are displayed. We call this mode as "Species Map Mode".



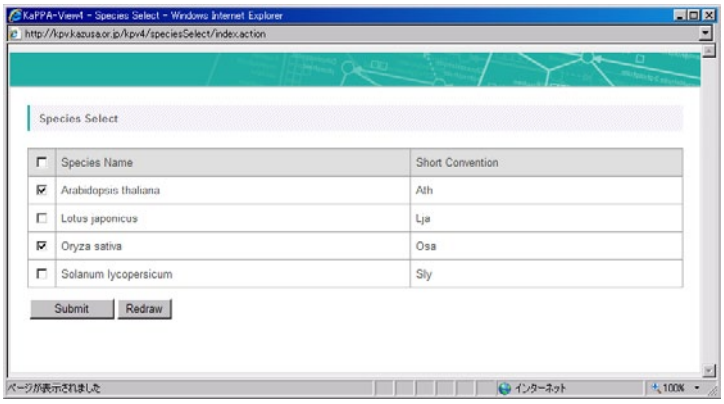


When an item "Universal" is selected in the pull-down list, a Pathway Tree for all species appears and the pathway maps with the genes from all the species are displayed. We call this as "Universal Map Mode".





Users can select the kind of the species that are represented on the pathway maps in Universal Map Mode. Click "Select Species" button below the Map Area to show a pop up window for species settings. Check the species, click "Submit" to fix the change, and click "Redraw" to refresh the Map Area.



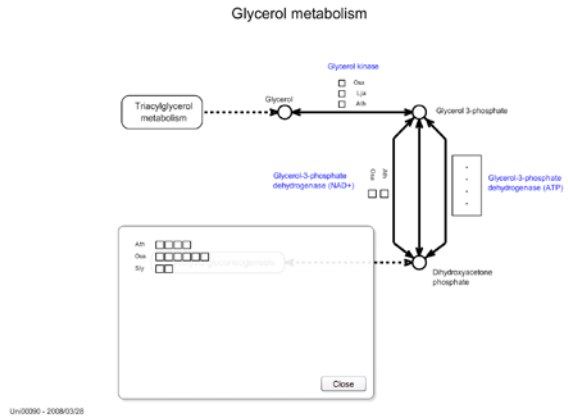
4-2-4. Metabolic Pathway Map

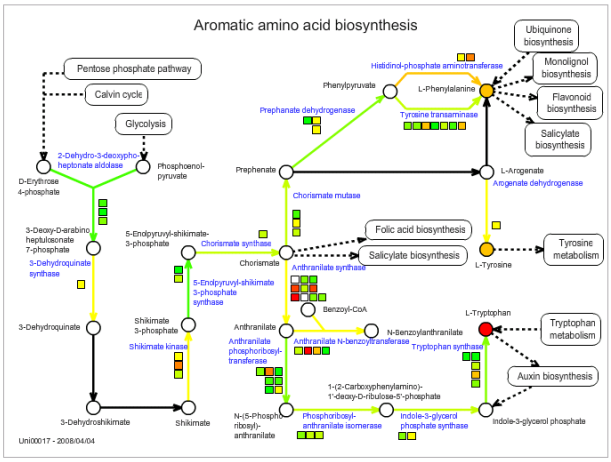
- ***Map Symbols***

The elements on the pathway maps such as genes, compounds, and enzymatic reactions are represented as the following symbols.

Element	Symbol	Note
genes	☐ (Squares)	

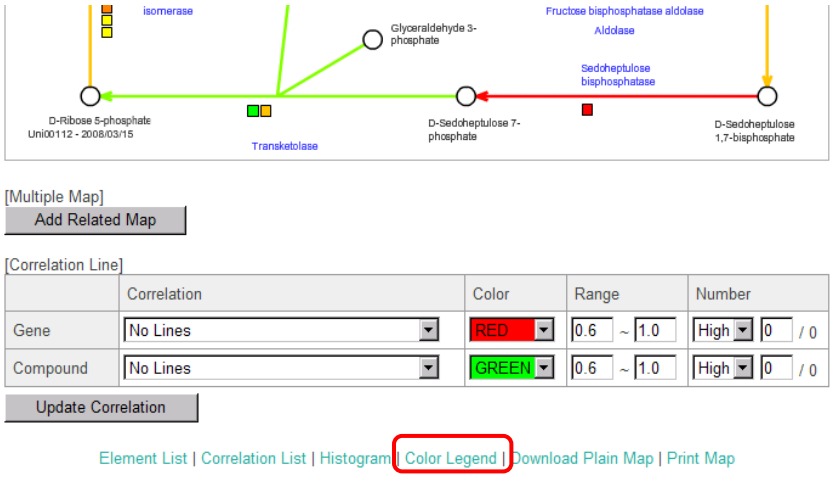
metabolites	◯ (Circles)	
enzyme reactions	⟶ (arrows)	The color of the arrows correspond to the mean value of the genes assigned to the reactions.
links to the other maps	pathway name (round rectangles with a text)	By clicking, the corresponding pathway maps is displayed.
genes	(Squares with)	When there is not an enough space to draw all the genes near by the enzyme reactions, this symbol is displayed. By clicking this, the symbols of the genes are shown in a pop-up window.

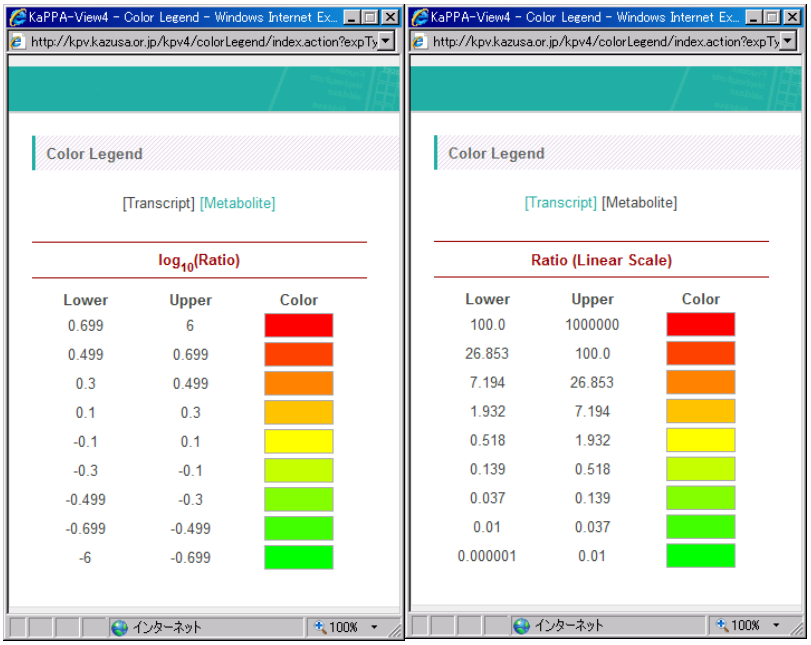




- ***Color Painting on the Symbols***

In the "Analysis" function, the symbols are painted as colors according to the experiment data. Click "Color Legend" to check the color settings.

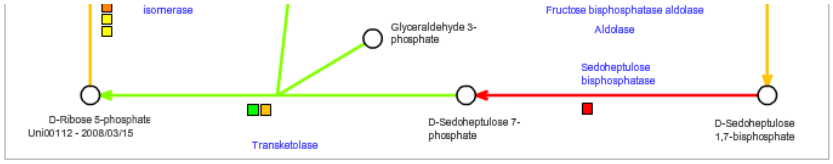




The pop up shows the colors and the corresponding ranges of the values of genes and metabolites that are calculated as ratio between the experiment data paired in the Compared Experiment Pair. The value ranges are shown in the scale of logarithm based on 10 for transcripts and linear scale for metabolites.

A color of an arrow representing enzymatic reaction is based on the average value calculated with the values of the genes assigned to the reaction. The arrow never painted in the Universal Map Mode.

Users can change the range of the values for the color paintings. Click "Histogram" to show a setting window.



[Multiple Map]

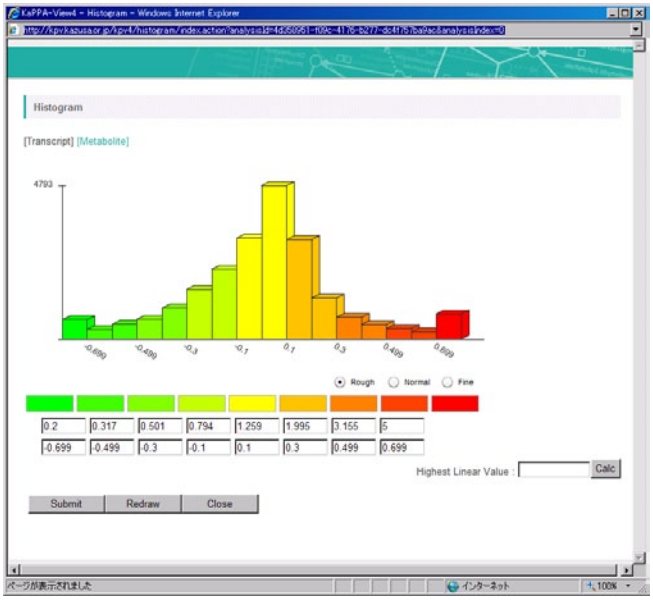
Add Related Map

[Correlation Line]

	Correlation	Color	Range	Number
Gene	No Lines	RED	0.6 ~ 1.0	High 0 / 0
Compound	No Lines	GREEN	0.6 ~ 1.0	High 0 / 0

Update Correlation

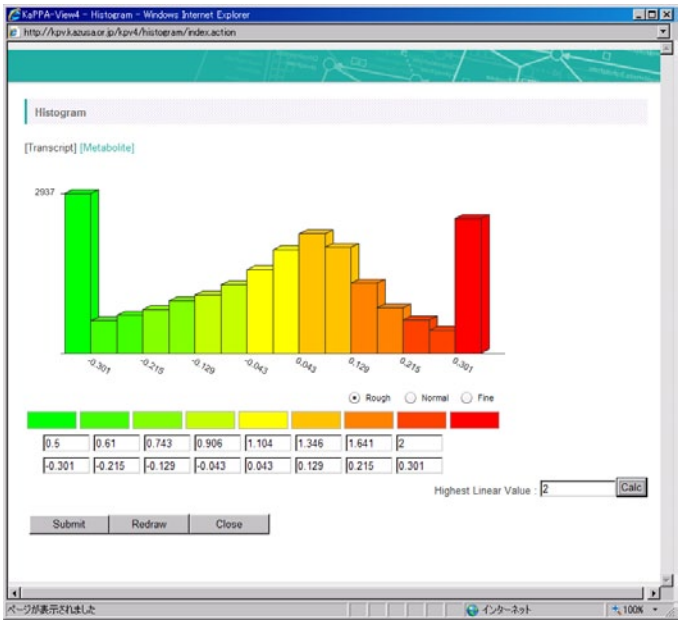
Element List | Correlation List | Histogram | Color Legend | Download Plain Map | Print Map



A histogram of changes of gene expression values in the current data is represented. The boundary values of the ranges are displayed in both log and linear scale. Click "Metabolite" to refer a similar histogram for changes of metabolite values.

In default, 5 for transcripts and 100 for metabolites (as linear scale) are set as a boundary value to express the strongest color (red). In other words, genes that expressions were changed more than 5 fold is represented in red.

Users can change the boundary value for the strongest color. Input a value into the "Highest Linear Value" field, and click "Calc" button. The histogram is refreshed.

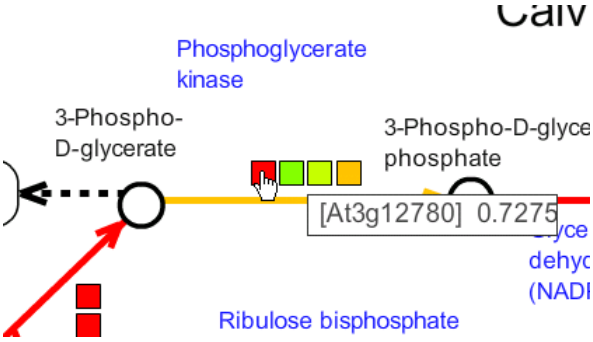


Click "Submit" to fix the alteration. Press "Redraw" button to refresh the Map Area.

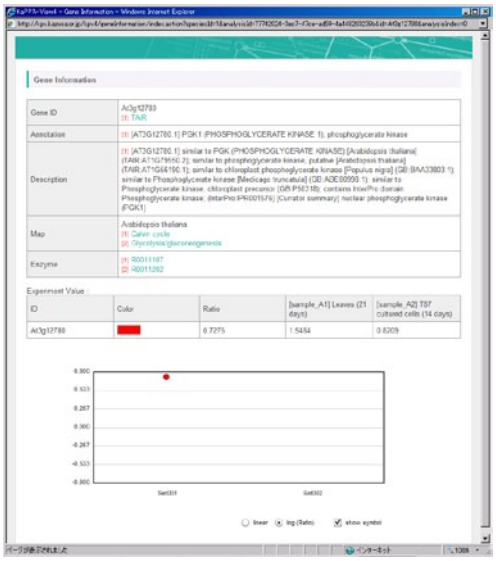
- ***Experiment Values***

The value of each symbol can be seen in several ways.

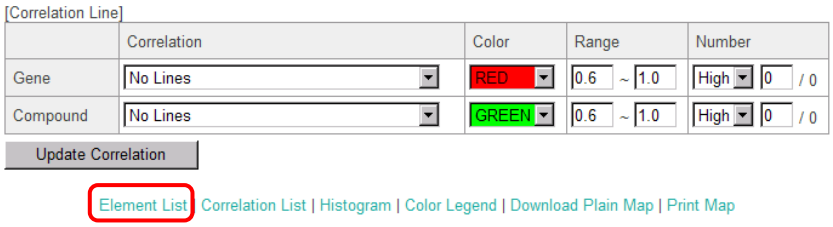
- Place the mouse cursor on the symbols. The change value appears in a tool chip.

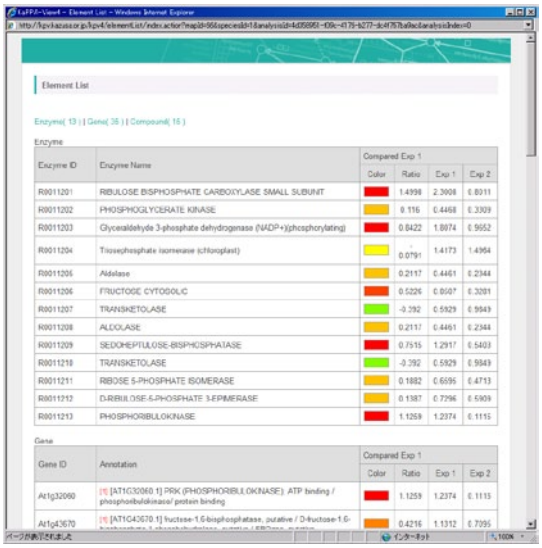


- Click a symbol to open the Pop-up Information Window for more details (see **4-2-5. Pop-up Information Window**).

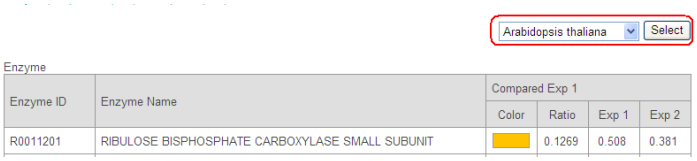


- Click "Element List" to show a list of the genes, compounds and reactions on the current pathway map.



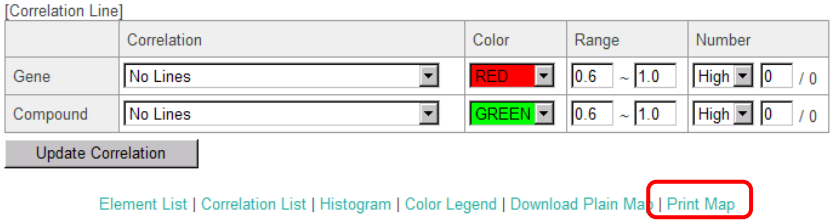


In Universal Map Mode, users can select the species to list up.



- ***Printing Maps***

Click "Print Map" to print the current map. A pop up window opens to show the printout image. Right click on the image and select "Print" to print.

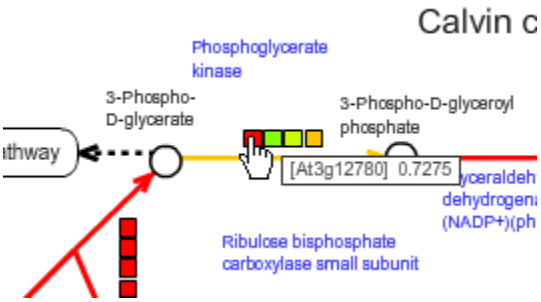


- ***Downloading Maps***

The pathway maps on the KaPPA-View4 are prepared in scalable vector graphics (SVG) format. Click "Print Pain Map" to download the SVG file for the current map. The downloaded files can be used as templates for creating User Maps (Refer to the **Manual on User Map Creation**).

4-2-5. Pop-up Information Window

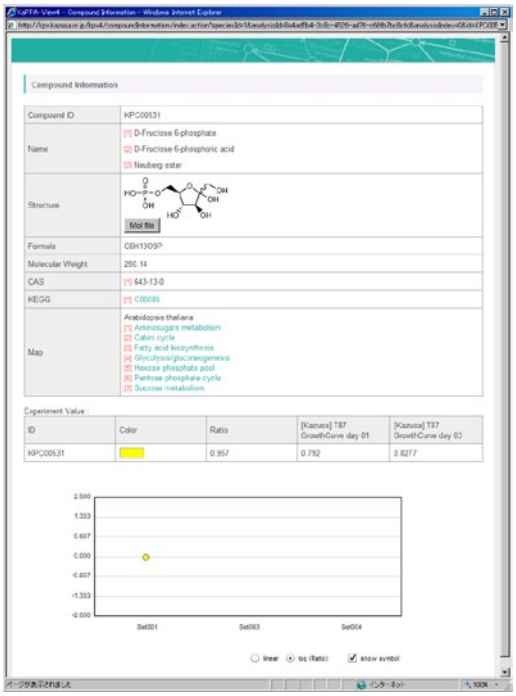
By clicking on the symbols for genes, compounds, and enzymatic reactions, the Pop-up Information Window opens to show detailed information for the element. Users can see the experiment values, the other maps which also have the element on them, links to the related databases and so on.



An example of a Pop-up Information Window for a gene is shown below. When several Compared Experiment Pairs are analyzed, the values are represented in a line graph.

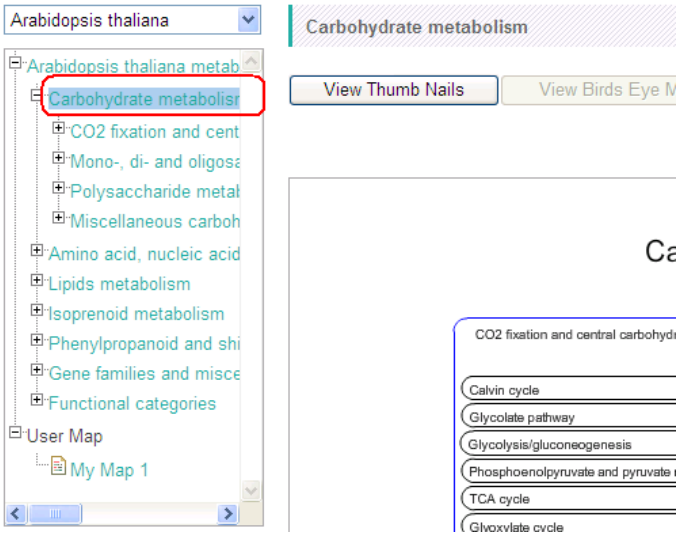
On the information window for an enzymatic reaction, the values of all the genes assigned to the reaction are represented. By clicking on a line in the graph, users can jump to the information page for the clicked gene.

An example of compound information page is shown below.



4-2-6. Bird's Eye Map

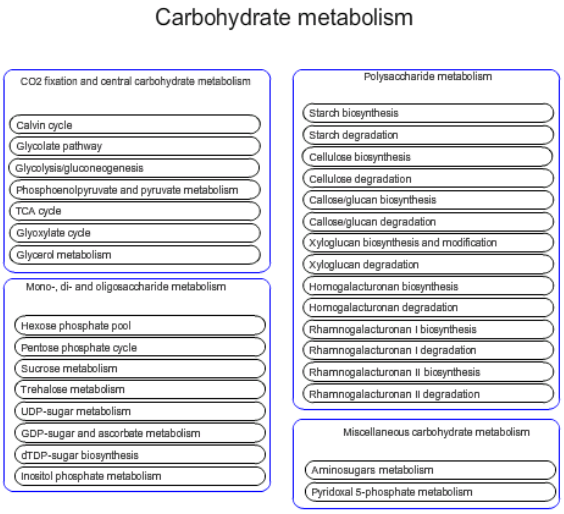
By clicking on an inter nodes of the Pathway Tree, the corresponding categories are represented as a Bird's Eye Map where all the pathway maps belong to the category are indicated as indicator bars. Users can jump to the individual pathway maps and other Bird's Eye Maps for child categories by clicking on the indicator bars and the area surrounding the bars, respectively.



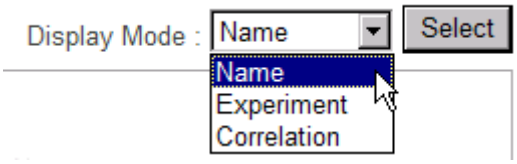
As described below, Bird's Eye Map can be used to summarize the experiment data for each pathway map and overview all of them.

- ***Map Names***

Just after clicking on the internal nodes on the Pathway Tree, the names of the pathway maps are represented in the indicator bars.

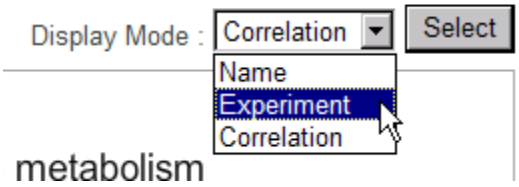


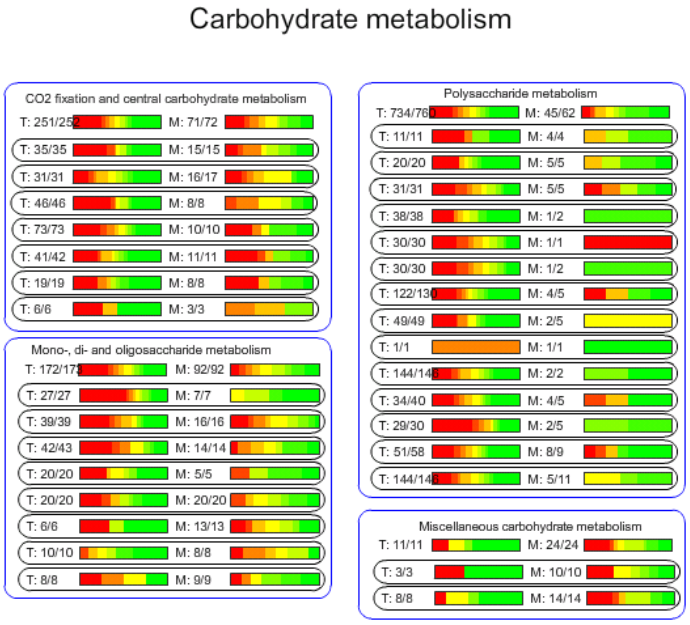
When experiment data or correlation data are represented on the Bird's Eye Map in Species Map Mode, select "Name" from the "Display Mode" pull-down list at the top-right of the Map Area to switch to the map name representation.



- ***Experiment Data***

Select "Experiment" from the "Display Mode" pull-down list in the Species Map Mode to represent summaries of experiment data.





Each indicator bar shows following information.

ex.)

T: Transcripts

The denominator shows the number of genes drawn on the pathway map (43), and the numerator shows the number of genes having valid values in the current experiment data (42).

M: Metabolites

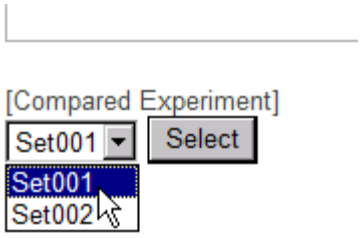
The denominator shows the number of compounds drawn on the pathway map (14), and the numerator shows the number of compounds having valid values in the current experiment data (8).

Colors on the bar:

It shows proportions of the genes (left) and metabolites (right) painted in the color on the pathway map.

Therefore, if a large proportion of the color bar for genes on a pathway map was painted heavily in red, we can recognize that expressions of most of the genes on the pathway were enhanced in the current data comparison, implies the pathway might be activated.

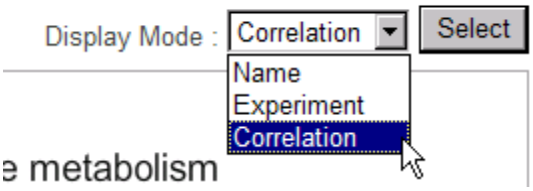
When multiple Compare Experiment Pairs were analyzed, a data selection pull-down list appears at the bottom-left of the Map Area. Select an item and click "Select" button to switch the data.



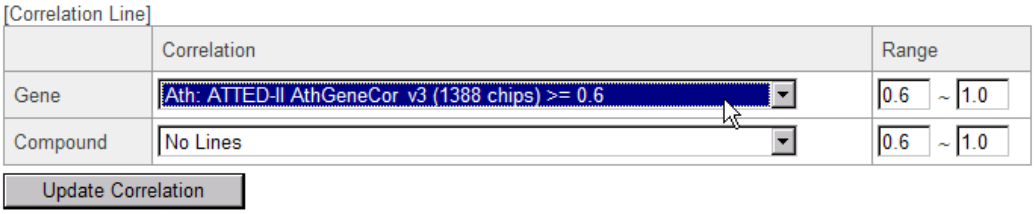
- ***Correlation Data***

Number of gene-to-gene and metabolite-to-metabolite relationships on the pathway maps can be represented on the Bird's Eye Maps. Refer to **4-4-3. Overlay of Correlation Data** for more details on representation of correlation data.

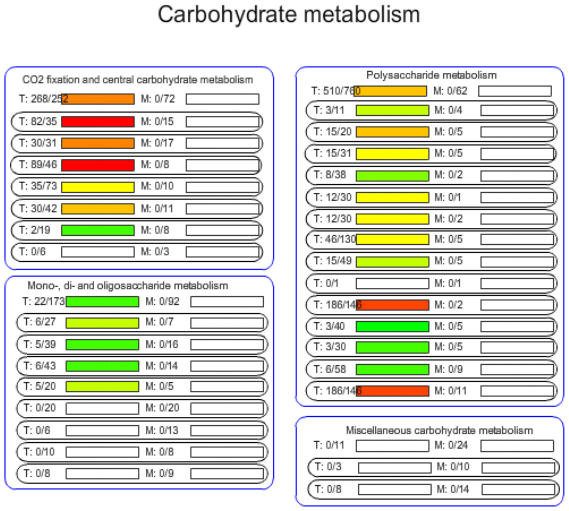
Select "Correlation" for "Display Mode" and click "Select" button in Species Map Mode.



The correlation control panel appears in the Lower Control Panel. Select a data and set ranges to compile and view data.



The relationship numbers are represented in the Bird's Eye Map as shown below.



The numbers written beside "T:" or "M:" (numerators) show the numbers of the correlation lines on the pathway maps. The numbers after the slashes ("/") (denominators) indicate the numbers of the genes or metabolites drawn on the maps. For the indicators of the middle tiers, both of the numerators and the denominators are the sum of the line and element numbers included under the tier.

The color of the bar is decided as follows.

When defined,

D: $\log_{10}$(line number / element number) of a map,

Dmax: The maximum value of D among the maps under the current tier, and

Dmin: The minimum value of D among the maps under the current tier,

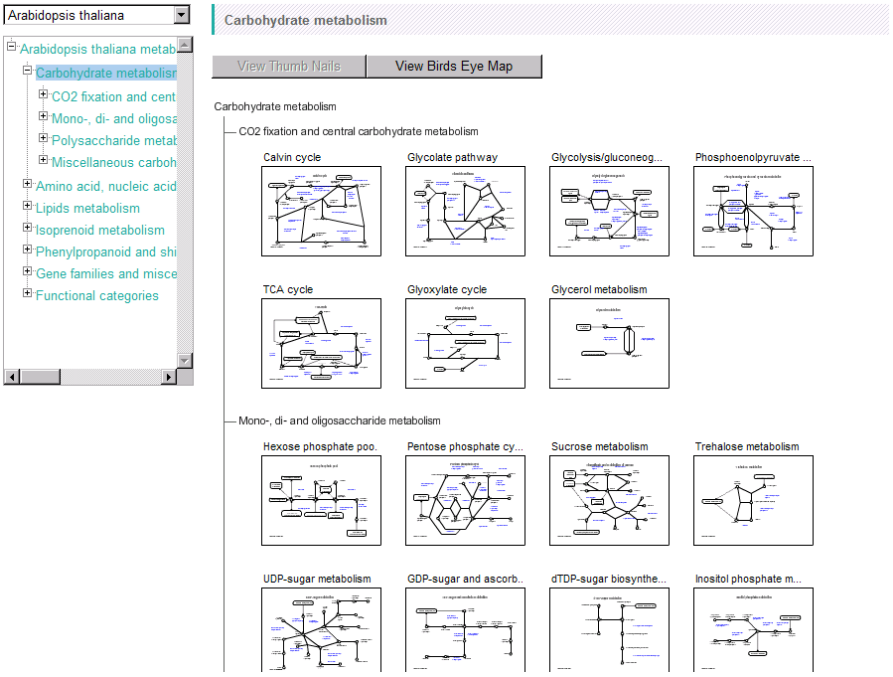
the bar of the map having Dmax is painted in the strongest color (red), and the map of Dmin is painted in the weakest color (green).

Therefore, the maps having dense relationships are painted stronger colors.

*Correlations calculated in an element (self correlations) are excluded from the line number counting, even if they are included in the uploaded files.

4-2-7. Thumbnails of the Maps

Click on a internal node of Pathway Tree and then on "View Thumbnails" button to view the thumbnails of the pathway maps arranged as a tree-like structure. Click on the thumbnails to jump to the corresponding pathway maps.



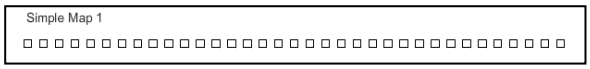
4-3. Data Analysis Functions

This section covers various data analysis functions available on KaPPA-View4. Combinations of these functions will accelerate the studies for gene functions.

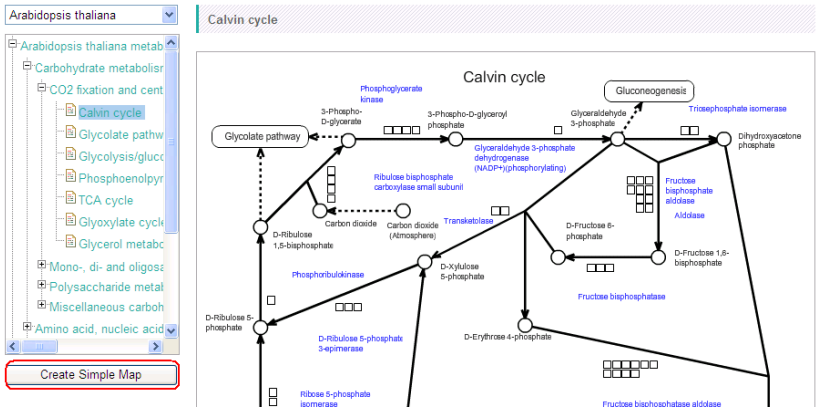
4-3-1. Simple Map

As KaPPA-View4 provides data viewing functions based on the pathway maps, the genes not drawn on the maps cannot be analyzed even if the values of such genes are included in the experiment data. To represent the values of the genes not drawn on the default pathway maps, we provide the Simple Map function.

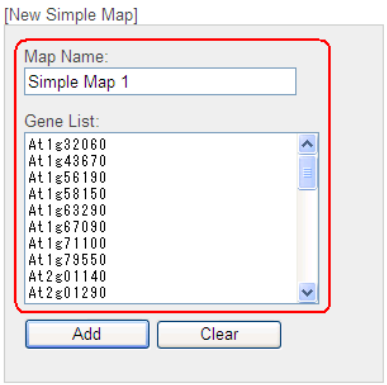
Any genes registered in the system with proper IDs can be automatically aligned on a Simple Map (as shown below) according to the user input of IDs, and the resulted Simple Map can be used in the data analysis as same as the default pathway maps. The gene IDs registered in the KaPPA-View4 are available through "Download" function (see **7. Download**).



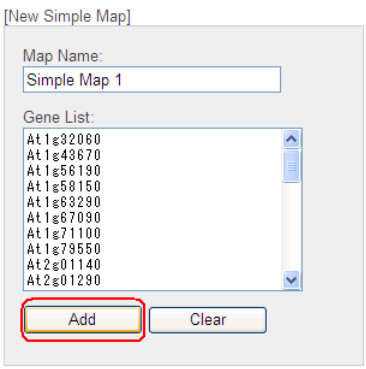
Click "Create Simple Map" at the bottom of the Pathway Tree in the Species Map Mode. This button is disabled in the Universal Map Mode, Bird's Eye view mode, and thumbnail view mode. A pop up window appears.

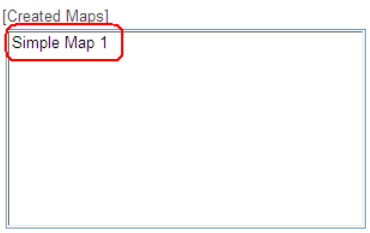


Input the IDs into the "Gene List", give a name of the Simple Map.

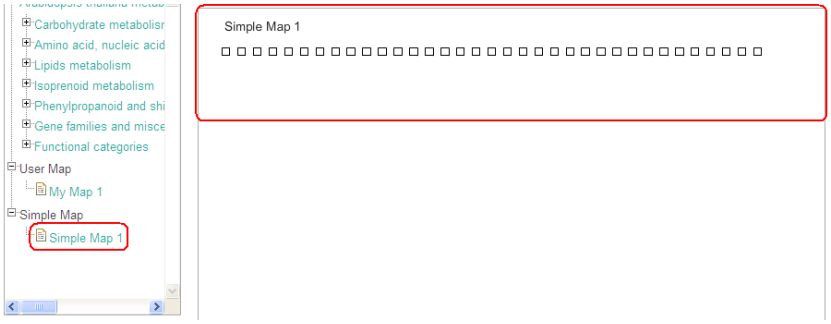


Click "Add" button to fix the input.





Click "Redraw" to refresh the main window. The created Simple Map appears in the Pathway Tree.



There are three ways to input the gene IDs into the "Gene List" area.

- Direct input by manual

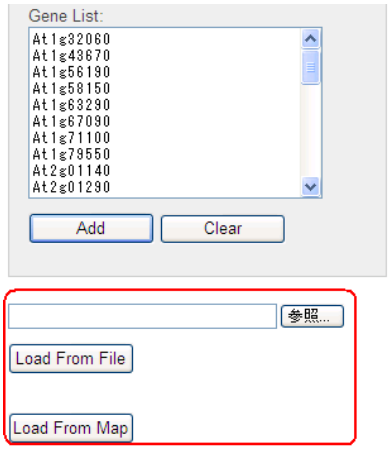
Input the gene IDs directly into the area. The IDs must be separated by newlines (returns).

- Load from File

Click "Browse" button to select a text file where gene IDs are written in each row. Click "Load From File" to load and input the IDs into the Gene List area.

- Load from current Map

Gene IDs are retrieved from the pathway map currently displayed. Click "Load From Map" button.

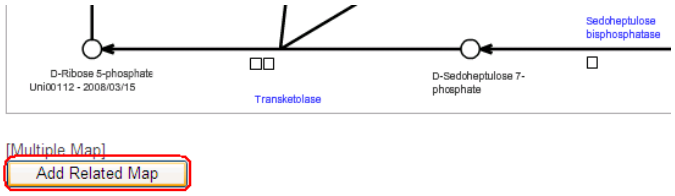


4-3-2. Multiple Map Mode

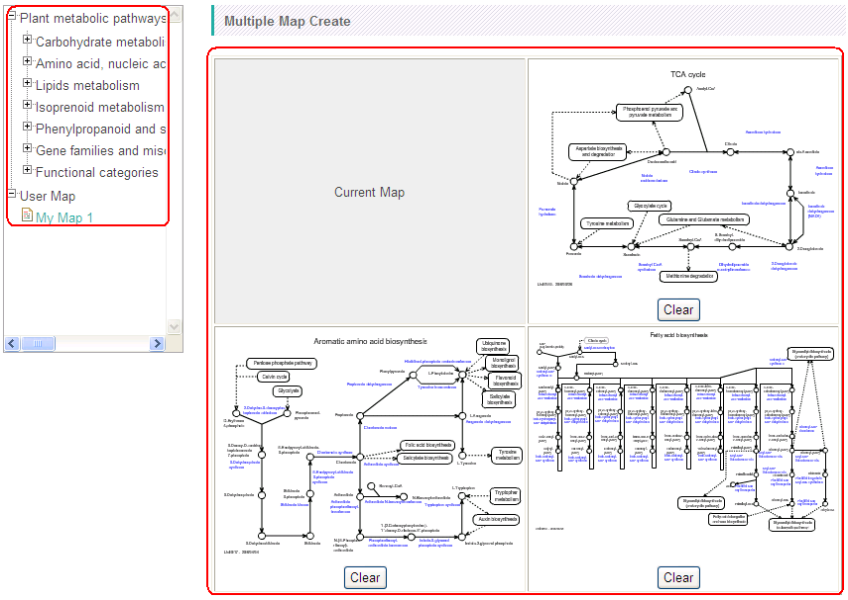
Up to 4 maps are displayed on a single browser window and they can be used simultaneously for color representations and correlation overlaying. The mode referred as "Multiple Map Mode" helps users to

compare omics values between several maps, investigating regulatory relationships between genes on a pathway map and maps for gene families such as transcription factors, and so on.

Click "Add Relation Map" button at the bottom-left of the Map Area in Species Map Mode. A setting window pops up.



Select a pathway map from the pathway tree on the pop up window to set the map on the preview panel on the right. Up to three maps can be selected here. The upper-left position of the resulted multiple map is interlocked to the Pathway Tree on the data browsing window, therefore users cannot select the map for this position (indicated as "Current Map" as shown below).



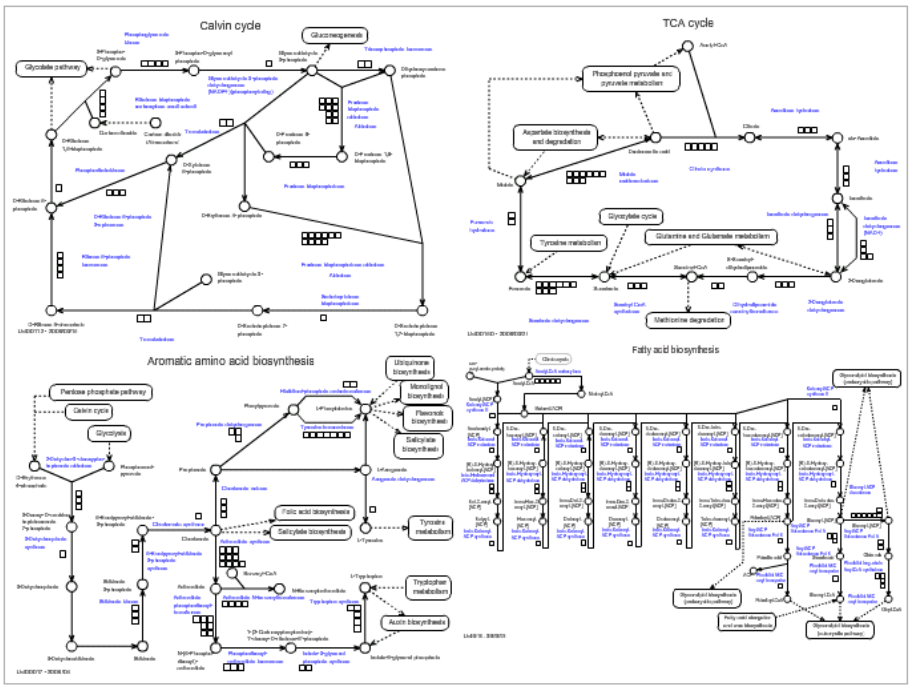
Click "Clear" button at the bottom of each map preview if you want to remove the selection.

After selecting the maps, enter the map set name and click "Add" button to register the combination of the maps.

Click "Redraw" to refresh the main window.

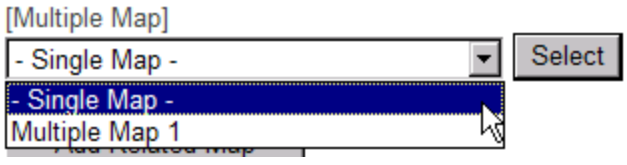
Close the pop up window.

A pull-down list named "Multiple Map" appears at the bottom of the Map Area. Select the registered name and click "Select" to show the map set (Multiple Map Mode).



The upper-left panel of the Multiple Map is interlocked to the Pathway Tree. By selecting a pathway map from the Pathway Tree, the upper-left map on the Multiple Map is replaced to the selected map.

Select "- Single Map-" from the pull-down list and click "Select" button to return to the normal display of the maps (Single Map Mode).



4-3-3. Overlay of Correlation Data

In the recent years, a huge number of microarray data are available on public, and it contributes to generate co-expression data as a novel data resource. A group of genes which are involved in a certain biological system could be expressed in coordinate manner throughout various conditions. Therefore, focusing on the unknown genes which co-expressing with well known genes could give a hint to uncover the functions of

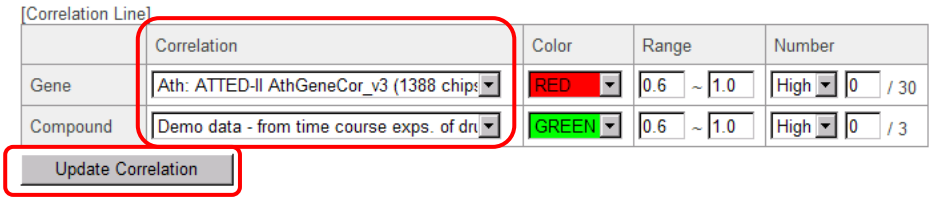
the unknown genes. ATTED-II (http://atted.jp/), for example, is one of vanguards of such approaches, and it can list up co-expressing genes of Arabidopsis for a query gene of researcher's interest.

KaPPA-View4 provides a function to overlay gene co-expression data onto the pathway maps. Data representation in this manner helps to grasp the gene-to-gene relationships on the aspect of metabolisms. KaPPA-View4 can represent metabolite-to-metabolite correlations too.

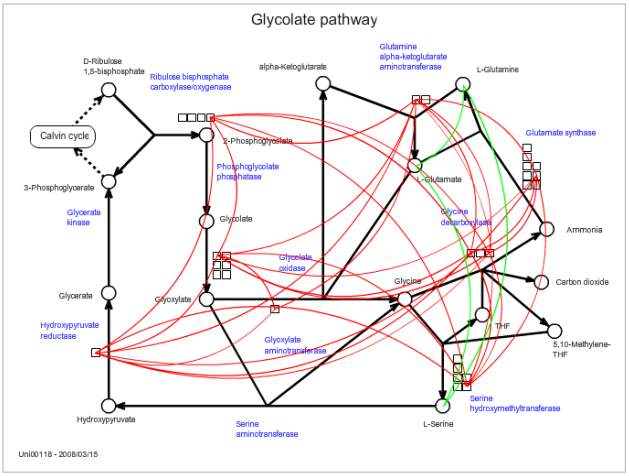
As an index of co-expression between the genes, correlation coefficients have been typically used. Hence, the functions of KaPPA-View4 concerning to the co-expression of the genes or co-accumulation of the metabolites are referred like "correlation functions". However, the data which KaPPA-View4 accepts is not restricted in the correlations. Any data which represents gene-to-gene or metabolite-to-metabolite relationships as numerical values can be utilized. For example, protein to protein interaction data described by 0 or 1 is acceptable. Please try to project your own ideas onto the pathway maps with KaPPA-View4.

- ***Data Selection and Display***

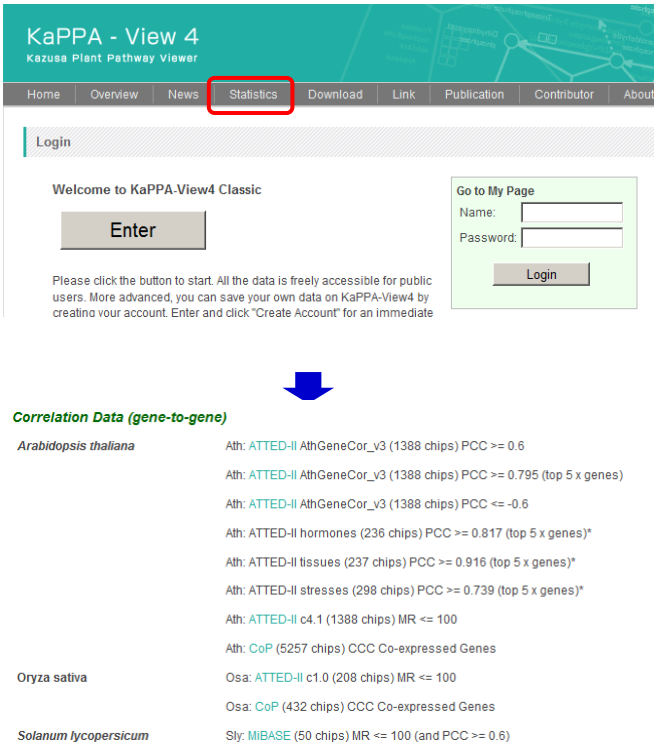
A control panel entitled "Correlation Line" appears at the Lower Control Panel in Species Map Mode (Correlation Control Panel). Select correlation data for genes and/or compounds from the pull-down lists and click "Update Correlation" button to overlay the data on the pathway map.



Gene-to-gene and/or metabolite-to-metabolite relationships are represented as curves. The color of the curves can be set on the Correlation Control Panel too.



Information about the default correlation data are available on the "Statistics" page of the KaPPA-View4. Refer to the supplemental file of KaPPA-View4 paper too for more details.

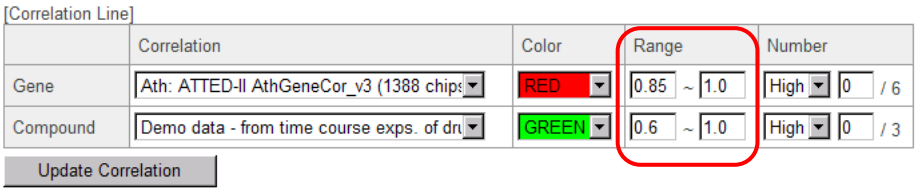


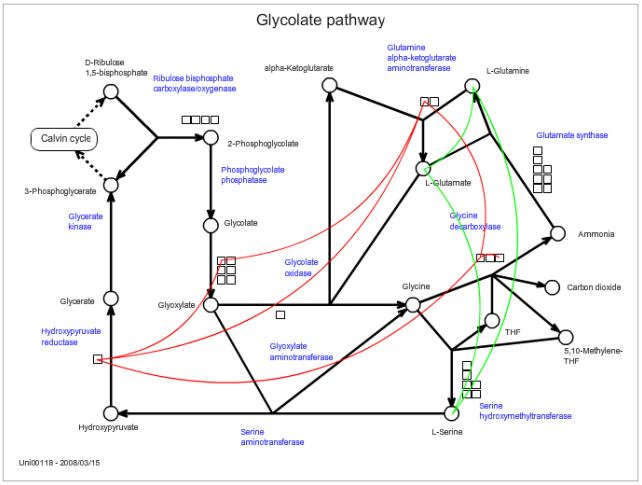
Users can upload their own correlation data too (see **3-4. Uploading Correlation Data**). The uploaded data can be selected from the pull-down lists.

Users can filter the correlation data to display by the following ways.

- ***Data Filtering by Value Ranges***

Set the ranges of the correlation data at the "Range" column on the Control Panel. In default, the ranges are set as 0.6 - 1.0.





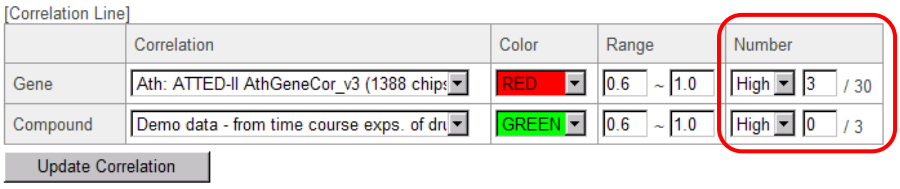
Note:

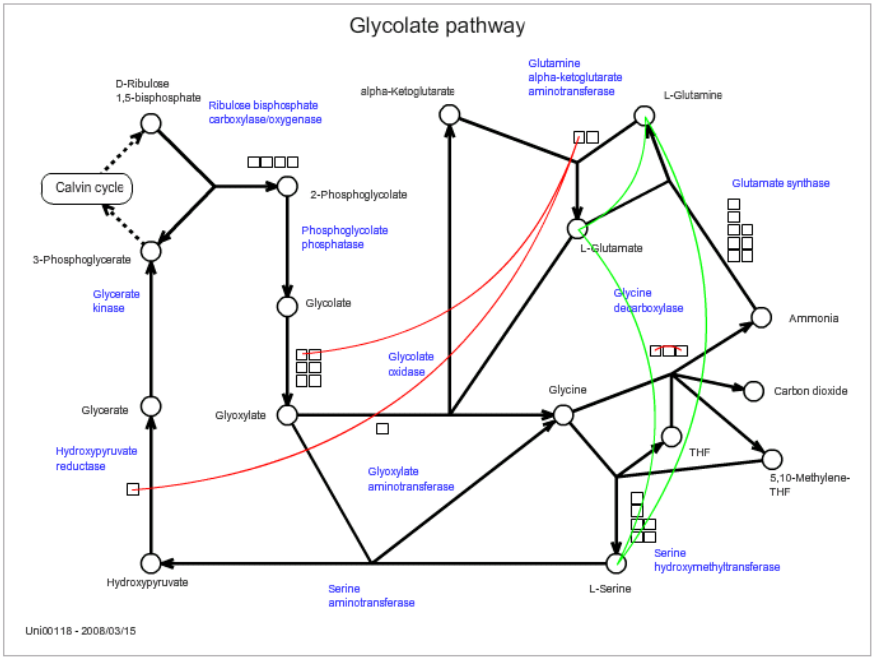
A correlation coefficient takes a value in a range from -1.0 to 1.0. However, in some default data, similarity values are not represented in correlation coefficient but in mutual rank (MR) which can take more than zero and up to several tens of thousands. Please be sure to set ranges properly.

- ***Data Filtering by Curve Numbers***

Users can set the number of curves to be represented on the pathways in the "Number" column. Select "High" or "Low" to determine which of the correlation values to be displayed, the higher most ones or the lower most ones, and enter the number of the curves. Click "Update Correlation" to refresh the representation. The number setting is applied to the curves that passed through the range filter, and the number of filtered curves by range settings is shown as a number after the slash "/".

An example is shown below. In total 30 curves for gene correlations are filtered by the range settings (0.6 - 1.0). By the number settings of "High" and "3", only the curves having the higher most 3 correlation values within the range are to be represented.

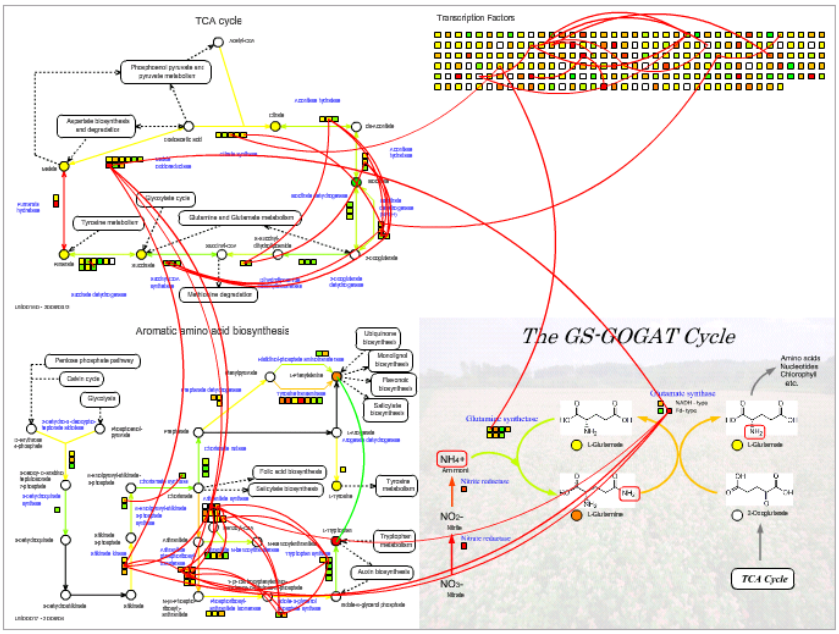




Enter "0" and click "Update Correlation" button to unset the number settings.

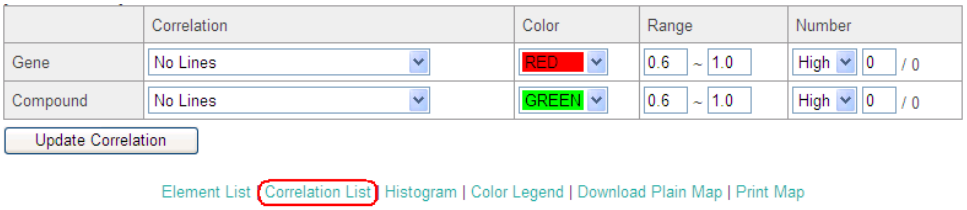
- ***Correlation Curves in Multiple Map Mode***

The overlaying function of correlation data is available in the Multiple Map Mode too. The figure below shows an example of correlation representations between a default pathway maps, a Simple Map, and a User Map.



- ***Correlation Data List***

Click "Correlation List" at the bottom of the window to view a list of the correlation curves currently displayed.



A pop up opens.

Gene 1	Gene 2	Coefficient
At1g32060	At3g55800	0.926
At1g32060	At3g26650	0.921
At1g32060	At3g54050	0.909
At2g21330	At3g55800	0.903
At3g55800	At4g38970	0.9
At1g32060	At4g38970	0.9
At3g26650	At4g38970	0.893
At3g54050	At3g55800	0.893
At3g12780	At3g54050	0.887
At2g21330	At3g54050	0.879

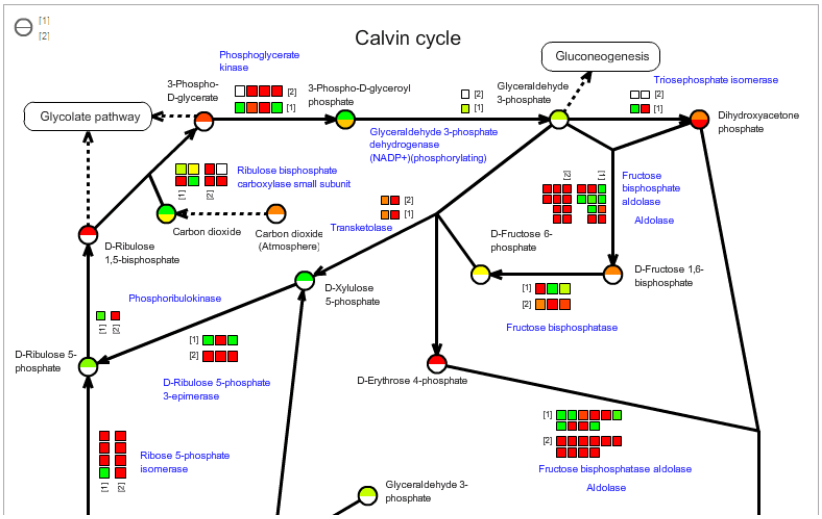
Compound 1	Compound 2	Coefficient
KPC00486	KPC00182	0.86572265625
KPC01099	KPC00570	0.8447265625

4-3-4. Comparison of Two Experiments in One Species

When multiple Compared Experiment Pairs for one species are included in the current analysis, two of them can be selected and comparatively displayed on the maps (Compare Mode). The Compare Mode is not available in Universal Map Mode, on User Maps, and on Simple Maps.

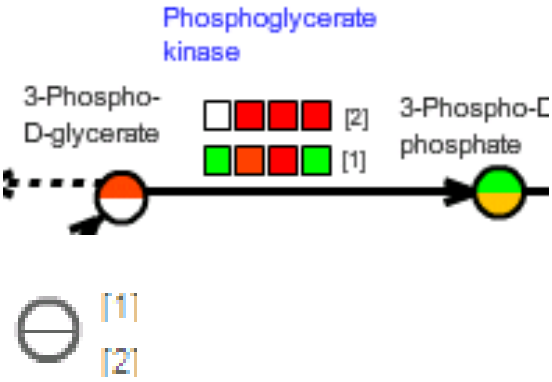
Check two of the Compared Experiment Pairs on the Upper Control Panel, and click "Compare" button.

[1] **Set001**
Transcript : Ath A / Ath B
Metabolite: Ath A / Ath B
[2] **Set002**
Transcript : dark grown T878233
Metabolite: Ath A / comp test1



The selected data are numbered as [1] and [2] and the details are shown on top of the Map Area. Two set of gene symbols are created on the pathway maps to show the values derived from each data.

The compound symbols (circles) are divided into upper and lower parts and they represent values from each data as abbreviated on the top-left corner of the map.

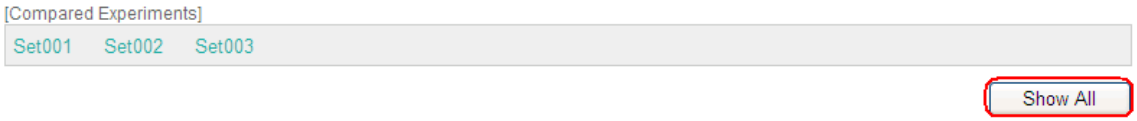


The arrows for the enzymatic reactions are not colored in the Compare Mode.

4-3-5. Data Comparison between the Species

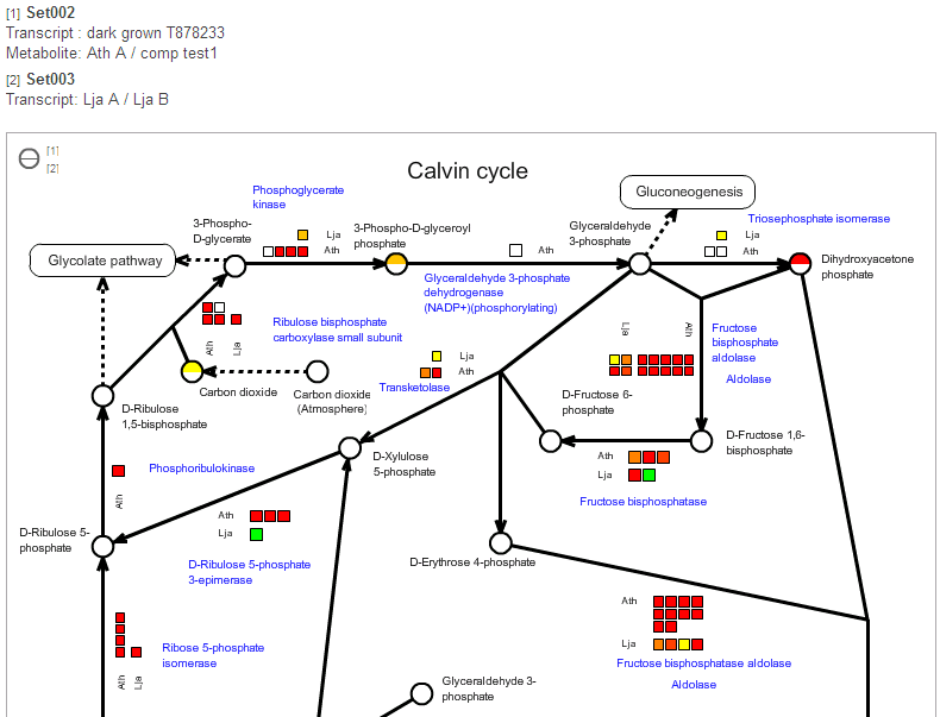
When Compared Experiment Pairs derived from multiple species are included in the current analysis, one of the data for each species can be compared and displayed on the pathway maps. This function is only available in Universal Map Mode.

Click the "Show All" button in the Upper Control Panel in the Universal Map Mode. An extra menu appears.



Select one of data for each species, and then click "Submit". The figure below is an example of the case two Arabidopsis data and one *Lotus japonicus* data exist.

A representation similar to the Compare Mode appears, except the caption of the genes are the abbreviations of the species.



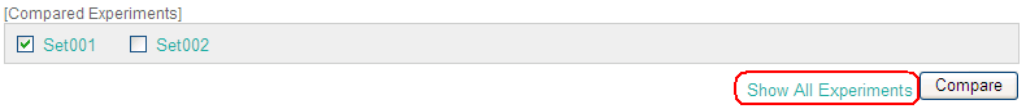
Note:

The species compared by this function must be the same as the species that allowed to display on the Universal Maps. If you got unexpected results, try to set the species again according to the section **4-2-3. Map Mode**, and then perform the Show All settings.

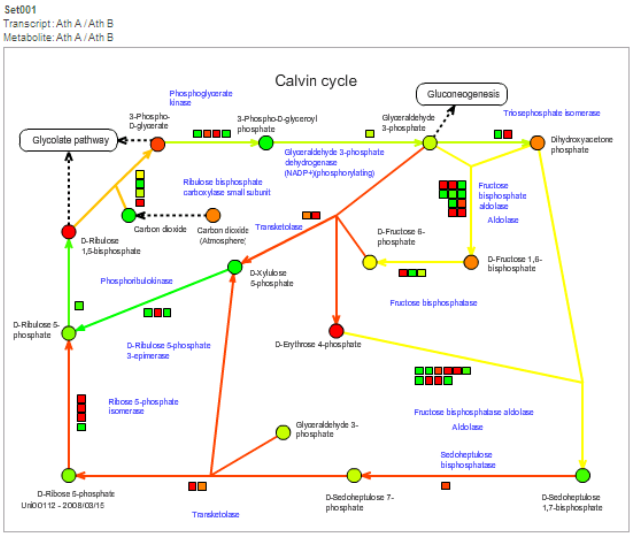
4-3-6. Displaying All Experiment Data

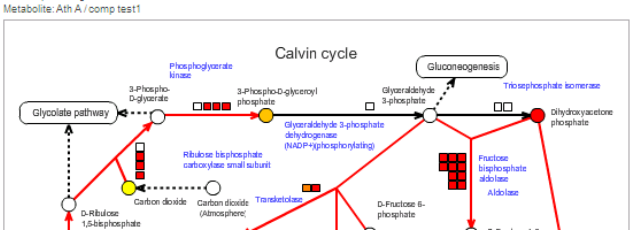
When multiple Compared Experiment Pairs are included in the current analysis, all of them can be displayed in a single browser window. This function is not available in Universal Map Mode.

Click "Show All Experiments" on the Upper Control Panel.



A window pops up that shows a list of the current maps representing each data.





5. Map Browsing Functions

The map representations without color painting by the experiment data are available too. Click "Map View" on the Main Menu.

The user interfaces for the functions depending on the omics data, such as Compare Mode, are not shown in the Map View function, but similar representations of the maps are available.

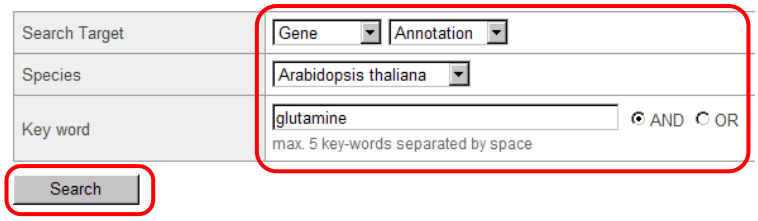
6. Searching

The "Search" on the Main Menu provides search functions for the genes, compounds, enzymatic reaction. A gene search functions by the BLAST program based on the sequence similarity is also available. Users can investigate the maps where your favorite genes are drawn.

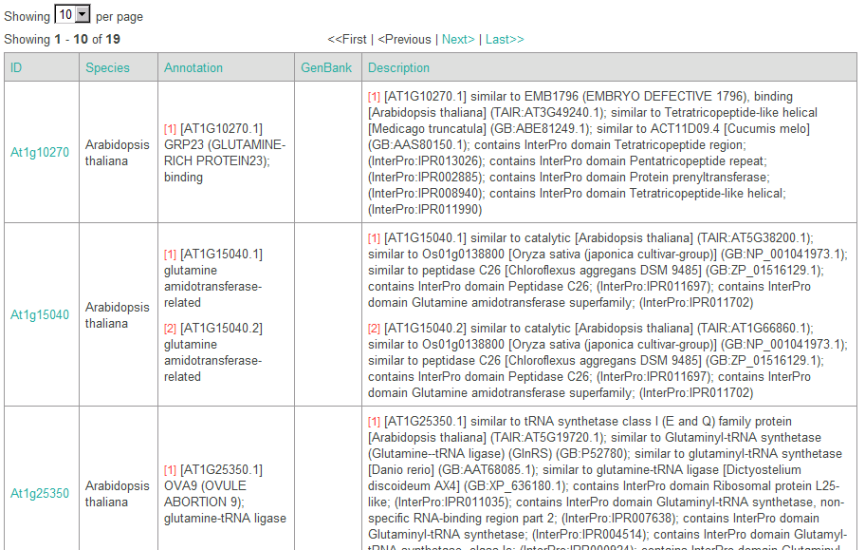
6-1. Gene, Compound and Enzymatic Reactions

Click the "Search" on the Main Menu. The searching user interface appears.

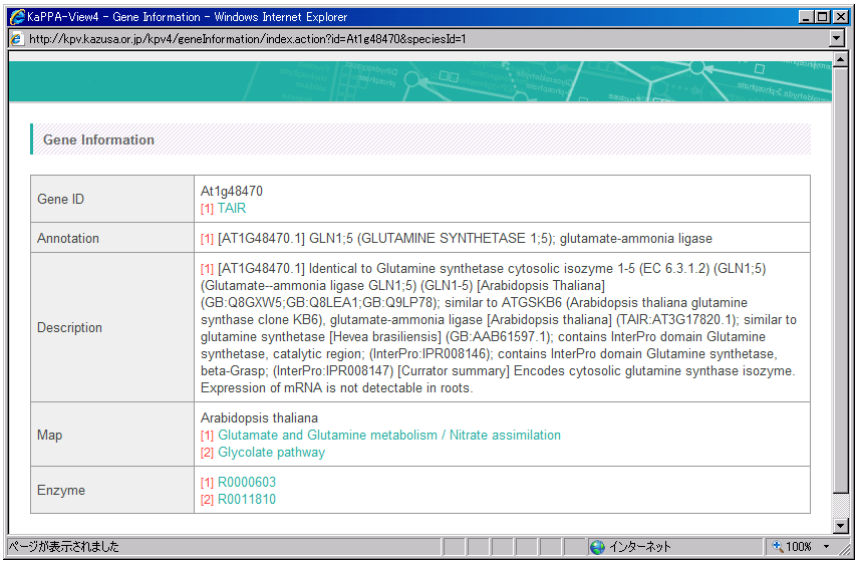
Enter the conditions into the form and click "Search" button. The figure below is an example of a searching condition for searching gene annotations matched to a keyword "glutamine" in Arabidopsis thaliana.



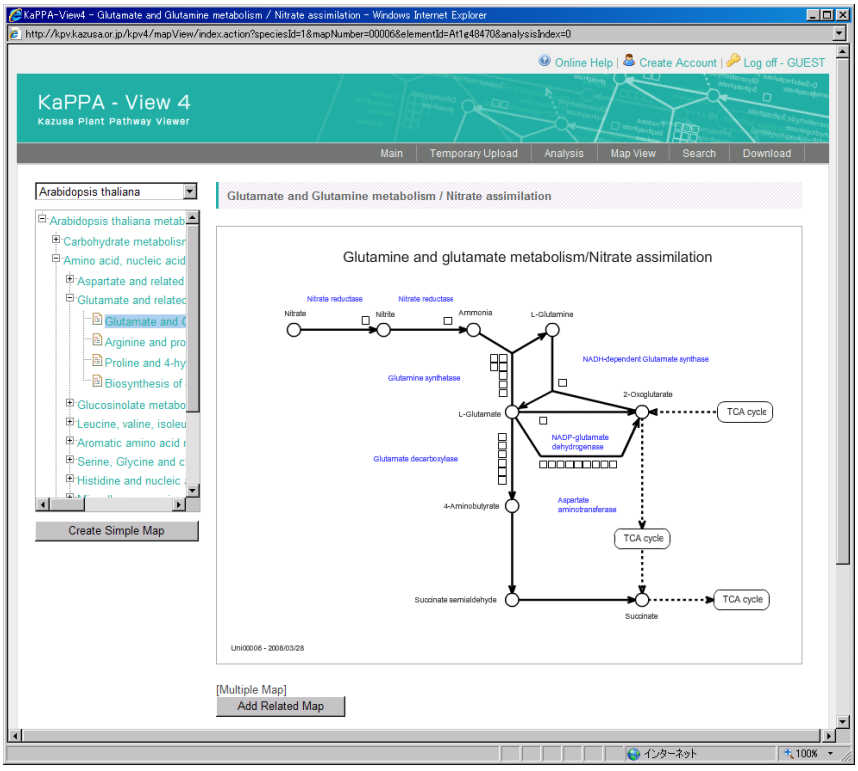
The results are listed.



Click the ID to view more details. A window pops up.



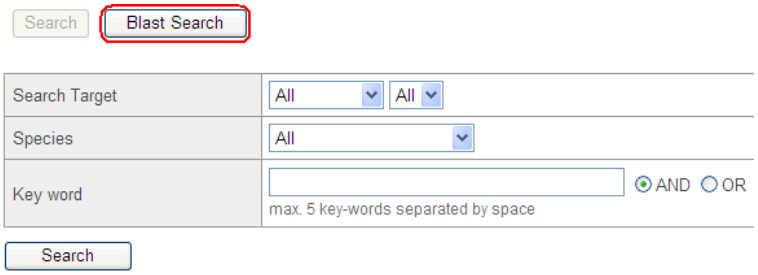
Click on the links written in the "Map" item to jump to the maps where the gene is drawn. The gene is highlighted by blinking on the map.



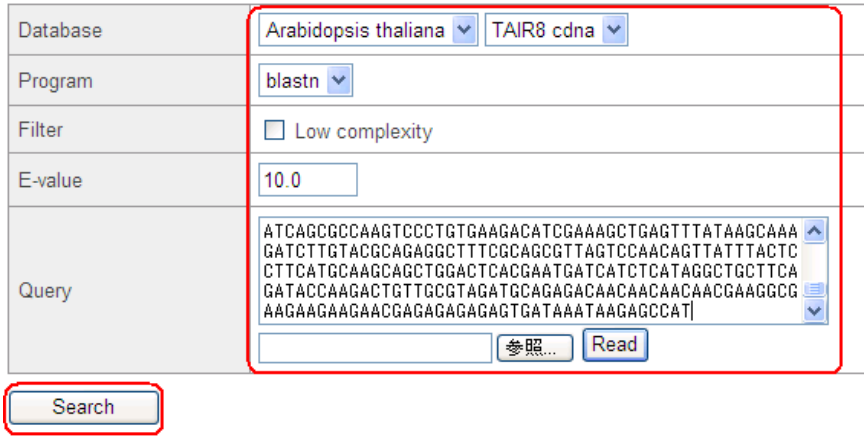
6-2. Blast Searching

This serves a gene search function based on the similarity of the nucleotide- or amino acid- sequences between the queried sequence and the genes registered in the KaPPA-View4 system using the Blast program.

Click "Blast Search" button.



The Blast searching interface appears. Enter searching conditions and click "Search".



Only a single sequence must be entered in "Query" input area. Multi-FASTA format is not accepted. The query can be loaded from a file.

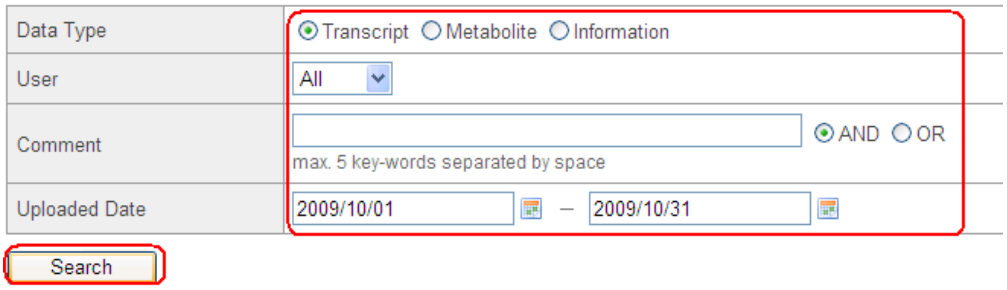
The results are listed. Click a gene ID to refer the details.

Gene ID	E-Value	Enzyme ID	Ec No	Map
AT1G15970	0.00	R0006404 R0006411 R0006423 R0006429 R0006433	1.2.1.44 1.2.1.44 1.2.1.44 1.2.1.44 1.2.1.44	Cinnamate-monolignol pathway / sinapoyl ester biosynthesis
AT1G80850	7E-76			
AT5G44680	2E-5			
AT3G12710	7E-5			
AT5G27230	0.068			
AT1G75090	0.068			

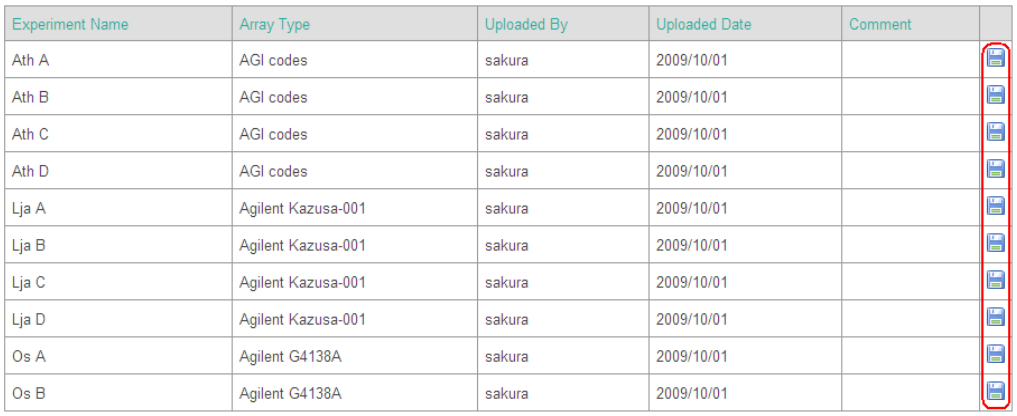
7. Download

Click "Download" on the Main Menu to download experiment data and information files used in the KaPPA-View4 system.

Enter conditions to search files and click "Search".



The results are listed. To start downloading, click a download icon ( ) or right-click on the icon and then select "Save as...".



Experiment Name	Array Type	Uploaded By	Uploaded Date	Comment	
Ath A	AGI codes	sakura	2009/10/01		
Ath B	AGI codes	sakura	2009/10/01		
Ath C	AGI codes	sakura	2009/10/01		
Ath D	AGI codes	sakura	2009/10/01		
Lja A	Agilent Kazusa-001	sakura	2009/10/01		
Lja B	Agilent Kazusa-001	sakura	2009/10/01		
Lja C	Agilent Kazusa-001	sakura	2009/10/01		
Lja D	Agilent Kazusa-001	sakura	2009/10/01		
Os A	Agilent G4138A	sakura	2009/10/01		
Os B	Agilent G4138A	sakura	2009/10/01		

Select "Information" for the "Data Type" to download the basic information file used in the system. The basic information files are entitled as follows. "[date]" represents the released date. "[Pre]" is the abbreviation of a species or "Uni" used in Universal Map Mode.

- ***The files commonly used for all species.***

Compounds: Uni_compoundInfo_[date].csv

Enzymatic reactions: Uni_enzymeInfo_[date].csv

- ***Species specific files***

Genes: [Pre]_geneInfo_[date].csv

Gene assignments to reactions:

[Pre]_geneBoxInfo_[date].csv or [Pre]_geneGroupInfo_[date].csv

Relationships between the DNA microarray probes and the Genes:

[Pre]_featureGene_[date].csv

Map assignment to the Pathway Tree: [Pre]_mapRelation_[date].csv

Tips:

Use the compound information file to investigate the compound IDs in the KaPPA-View4.

We don't provide files for gene-to-map, compound-to-map, or reaction-to-map information. These relationships are calculated dynamically by the system from the pathway map files provided in SVG format and the other information files. To investigate the elements drawn on the map, refer to the Element List (see **Experiment Values in the section 4-2-2**).

Note:

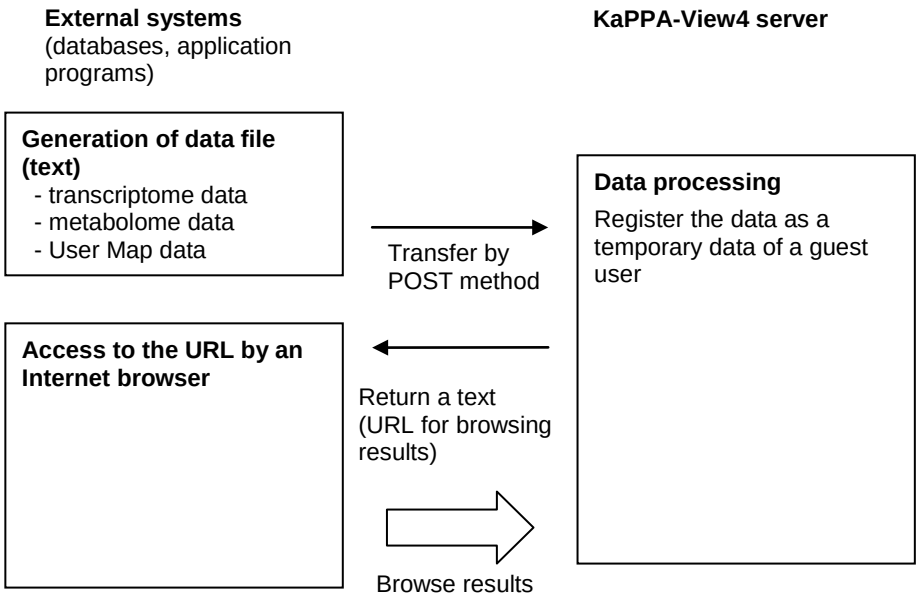
The Chemical Abstract Service (CAS) numbers are written in the compound information file. A CAS number consists of three number parts separated by hyphens. Some of them might be miss-recognized by Microsoft Excel as date data and be replaced to other unexpected values. Be careful to edit and utilize the compound information file.

8. Utilization from External Systems

This chapter covers the ways for utilization of KaPPA-View4 from external systems. The accessibility from the external systems is greatly improved in KaPPA-View4. We newly opened APIs (Application Program Interface) that allows uploading the experiment data from the external systems without logging-in to the KaPPA-View4 system through the graphical user interfaces (GUIs) of the website. Therefore, the developers of other databases and application programs incorporate functions to view their omics data directly on KaPPA-View4.

8-1. Data Uploading APIs

KaPPA-View4 provides new APIs for uploading experiment data and user created map data from external systems. The outline of the communications between the clients and the KaPPA-View4 server are shown in the schematic figure below.



The clients must generate a text file to send the data to the KaPPA-View4 server to ensure stable transfer. Transcriptome data, metabolome data, User created Map data, and some meta data for them can be included in the file. Refer to **9-5. Data for POST Transferring** for details of the file format.

The POST method of HTTP is used to transfer the data to the KaPPA-View4 server. Additional libraries might be required to execute POST transfer in some programming languages.

8-1-1. Data Transferring Procedure

Send the generated file to the URL below by POST method. Access to the URL received as response by an Internet browser to view the results.

URL	http://kpv.kazusa.or.jp/kpv4/upload.action
Method	POST
Content-Type	multipart/form-data
Parameter	uploaded (*Data File*)
Response	*Success*: an URL starts with "http" *Error*: error messages. It might be written in multiple lines.

8-1-2. Sample Codes

- ***HTML Form***

```
<form action="http://kpv.kazusa.or.jp/kpv4/upload.action" method="post"
enctype="multipart/form-data">
<input type="file" name="uploaded" />
<input type="submit">
</form>
```

Open the html file by an Internet Browser. Select a data file from the web form, and click "Submit" button. If the data file accepted, an URL is displayed on the browser window. Access to the URL by inputting it into the address field of the browser.

This is not a practical example as a web application, but good for understanding the behavior and usage of the API.

- ***PHP***

This is a sample code using a PEAR library.

```
require_once 'HTTP/Client.php';

$client = new HTTP_Client();

$url = "http://kpv.kazusa.or.jp/kpv4/upload.action";
$file = "./uploadfiles/postdata.txt";

$postdata = array();
$postfile = array('uploaded', $file);
$capsule = array($postfile);

$client->post($url, $postdata, false, $capsule);
$response = $client->currentResponse();
```

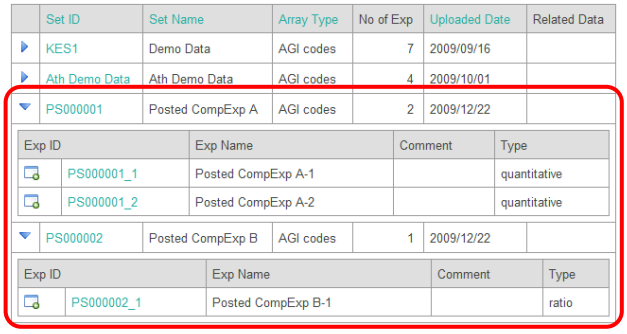
Redirect to the URL received as the response to browse the results.

8-1-3. Behavior of KaPPA-View4 after the Transferring

The users are recognizes as Guest Users by the KaPPA-View4 system after accessing through the URL received as the response. The transfered data are registered as temporary files and Compared Experiment Pairs are automatically created. Therefore, users can start analysis with the KaPPA-View pathway maps just after accessing to the URL. The individual experiment data appears in the data list of "Analysis" function, allows users to compare them with other arbitral data.

The data are expired as the same as those of Guest Users', i.e., the data are deleted from the server when the user logoff, all the browser windows closed or no operation have done in 60 minutes.

The transferred data appears in the "Analysis" window as shown below.



The transferred data is designated as Experiment Set ID of PS[*six-digit serial number*]. If the transfered data is a pair of two experiment data, they are registered as quantitative data and the Experiment IDs are assigned as [Exp. Set ID]_1 and [Exp. Set ID]_2. If the transfered data includes only one experiment data, it is registered as ratio data and the Experiment ID is assigned as [Exp. Set ID]_1.

8-2. Direct Access by IDs of Gene, Compound, Reaction, and Map

The genes, compounds, enzymatic reactions and the pathway maps used in the KaPPA-View4 system have their own ID. The information pages for the elements and pathway maps can be directly accessed from the external systems by describing URLs in defined formats with the IDs and by accessing to the URL.

If the user is not logged-in to the system, the user is automatically recognized as a Guest User.

8-2-1. URL Format

- ***Gene***

`http://kpv.kazusa.or.jp/kpv4/geneInformation/view.action?id=`***Gene ID***

ex)

http://kpv.kazusa.or.jp/kpv4/geneInformation/view.action?id=At1g58150

A gene information page appears.

Gene Information

Gene ID	At1g58150 [1] TAIR
Annotation	[1] [AT1G58150.1] unknown protein
Description	[1] [AT1G58150.1]
Map	[1] Calvin cycle [2] Glycolysis/gluconeogenesis
Enzyme	[1] R0011107 [2] R0011202

- ***Compound***

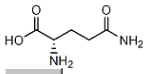
http://kpv.kazusa.or.jp/kpv4/compoundInformation/view.action?id=*Compound ID*

Ex)

http://kpv.kazusa.or.jp/kpv4/compoundInformation/view.action?id=KPC00697

A compound information page appears.

Compound Information

Compound ID	KPC00697
Name	[1] L-Glutamine [2] L-2-Aminoglutaramic acid
Structure	 Mol file
Formula	C5H10N2O3
Molecular Weight	146.14
CAS	[1] 56-85-9
KEGG	[1] C00064
Map	[1] Aminoacyl-tRNA biosynthesis [2] Glutamate and Glutamine metabolism / Nitrate assimilation [3] Glycolate pathway

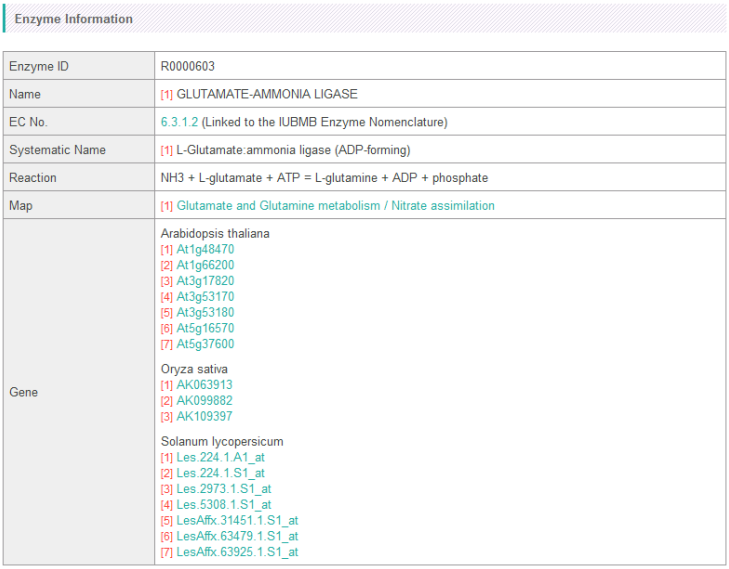
- ***Enzymatic Reaction***

http://kpv.kazusa.or.jp/kpv4/enzymeInformation/view.action?id=*Reaction ID*

Ex)

http://kpv.kazusa.or.jp/kpv4/enzymeInformation/view.action?id=R0000603

An information page for the enzymatic reaction appears.



Enzyme Information

Enzyme ID	R0000603
Name	[1] GLUTAMATE-AMMONIA LIGASE
EC No.	6.3.1.2 (Linked to the IUBMB Enzyme Nomenclature)
Systematic Name	[1] L-Glutamate:ammonia ligase (ADP-forming)
Reaction	NH3 + L-glutamate + ATP = L-glutamine + ADP + phosphate
Map	[1] Glutamate and Glutamine metabolism / Nitrate assimilation
Gene	Arabidopsis thaliana [1] At1g48470 [2] At1g66200 [3] At3g17820 [4] At3g53170 [5] At3g53180 [6] At5g16570 [7] At5g37600 Oryza sativa [1] AK063913 [2] AK099882 [3] AK109397 Solanum lycopersicum [1] Les.224.1.A1_at [2] Les.224.1.S1_at [3] Les.2973.1.S1_at [4] Les.5308.1.S1_at [5] LesAffx.31451.1.S1_at [6] LesAffx.63479.1.S1_at [7] LesAffx.63925.1.S1_at

- ***Pathway Map***

http://kpv.kazusa.or.jp/kpv4/mapView/view.action?mapNumber=*Map Number*

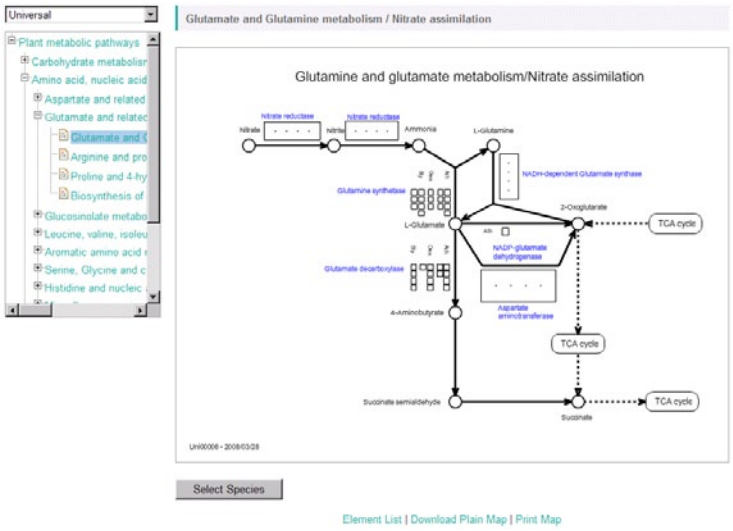
Trim initial three alphabets from the Map ID to get "Map Number". The map is shown in Universal Map Mode with this description of URL. Specify a species by attaching another parameter to the description to show the species specific maps (see next section).

Ex)

http://kpv.kazusa.or.jp/kpv4/mapView/view.action?mapNumber=00006

http://kpv.kazusa.or.jp/kpv4/mapView/view.action?mapNumber=00028f

The corresponding pathway map appears.



8-2-2. Specifying Species

By adding a parameter below in the URL, users can specify the species.

&speciesName=*Species Name*

"Species Name" is the name of species registered in KaPPA-View4. You can describe it case-insensitively. The text must be URL encoded, therefore, try to replace the space between the family name and the genus name to "%20" if it's not accepted properly. When an unrecognizable Species Name were written, the system returns information in Universal Map Mode.

Ex)

http://kpv.kazusa.or.jp/kpv4/geneInformation/view.action?id=At1g58150&speciesName=Arabidopsis%20thaliana

http://kpv.kazusa.or.jp/kpv4/compoundInformation/view.action?id=KPC00697&speciesName=Lotus%20japonicus

http://kpv.kazusa.or.jp/kpv4/enzymeInformation/view.action?id=R0000603&speciesName=Oryza%20sativa

http://kpv.kazusa.or.jp/kpv4/mapView/view.action?mapNumber=00006&speciesName=Solanum%20lycopersicum

```
http://kpv.kazusa.or.jp/kpv4/mapView/view.action?mapNumber=00415&species
Name=Lotus%20japonicus
```

9. File Format

This chapter covers the file formats of the following data.

Experiment Data:

The omics data such as transcriptome data obtained DNA microarray and metabolome data detected by metabolomics approaches. It is used for the analysis on the pathway maps.

Correlation Data:

Gene-to-gene or metabolite-to-metabolite relationship data such as gene co-expression data and metabolite co-accumulation data. It is used for the overlaying of relationship data on the pathway maps.

User Map Data:

The map data originally created by the users.

Data for POST transferring:

The data created by the client systems when the data is represented directly on KaPPA-View4 via the POST transferring API.

9-1. General Notices

Don't include any multibyte letters in the data files. Don't attach the byte order mark (BOM) when the file has to be encoded in UTF-8.

9-2. Experiment Data for Uploading

The omics data such as transcriptome data obtained DNA microarray and metabolome data detected by metabolomics approaches. By uploading the data file, users can analyze the data on KaPPA-View4.

In addition to the file format for the previous version, the format for KaPPA-View4 can include some headers for describing sample information and so on. The previous format is still acceptable, but we recommend attaching header information to avoid confusions if you upload a lot of data.

9-2-1. Data Part

Same as the previous format as described below. The data files for the previous version can be accepted as they are.

- ***Extension from the Previous Version***

- If the compound data includes 0 as a value, the system alerts it. In the previous version, such compound data was ignored.

- A tab delimited text (tab separated vector: TSV) format becomes to be acceptable as well as the comma separated vector (CSV) format.

- Arbitral gene and compound IDs (with prefix of TMG and TMC, respectively) that are not registered in the system can be included in the data file. Those IDs enables to paint colors for the gene and compound symbols that have corresponding IDs on the User Maps.

- ***Format***

The data of transcripts and metabolites have to be prepared in independent files. Prepare data following below with proper tool such as Microsoft Excel, and save as a CSV or TSV file. When you add headers (described later), save the data in TSV format.

Line	Column	Description
1st	1st	Define the type of data. Enter "(arrayexp)" for transcript data or "(compexp)" for metabolite data with the parentheses.
	2nd ~	Enter the experiment names. Exactly the same experiment names have to be written in the columns that hold the experimental repetition data as defined in the 2nd line. The system recognizes the data of different names as independent experiments. The experiment names written here also have to match exactly to those in the header information to keep correct correspondence.
2nd	1st	Enter "(rep)" with the parentheses. "rep" stands for "repetition".
	2nd ~	Enter the repetition numbers like 1, 2, 3... for the data of experimental repetition. Don't assign the same repetition number for the columns having the same experiment name (written in line 1).
3rd ~	1st	Enter IDs for the elements, the feature IDs for transcript data and compound IDs for metabolite data. *Download the basic information file to investigate these IDs from the "Download" function. Users can input other IDs for the representation of values of arbitral genes and compounds. In such case, enter IDs in the following formats. "TMG[*number*]" for genes. ex) TMG1, TMG02 "TMC[*number*]" for compounds. ex) TMC003, TMC045 *To represent the data, the corresponding symbols that have the same IDs have to be drawn on the User Maps.
	2nd ~	Enter the values. The values for genes have to be log-transferred (based on 10 is recommended). The values for compounds have to be linear scale and 0 (zero) and negative values cannot be included. The values are treated by the system as log and linear scale for genes and compounds.

- ***Sample***

The figures below show examples data preparation with Microsoft Excel.

Transcript data.

	A	B	C	D	E	F	G
1	(arrayexp)	Treatment A	Treatment A	Treatment B	Treatment B	Control	Control
2	(rep)	1	2	1	2	1	2
3	At1g01010	-0.8285	-0.9025	0.6469	0.5278	-1.1905	-1.0772
4	At1g01020	-1.3224	-1.2137	-0.5244	-0.5612	-0.1669	-0.2220
5	At1g01030	1.1189	1.1208	1.4112	1.3381	-1.2666	-1.2963
6	At1g01040	-1.2181	-1.2986	-0.9615	-0.8966	0.0747	0.1132
7	At1g01050	-0.6649	-0.7699	0.1262	0.2110	1.1751	1.1702
8	At1g01060	0.8093	0.7658	1.5667	1.5299	0.5790	0.6099
9	At1g01070	-1.2044	-1.2744	-0.8895	-0.8087	0.6100	0.6117
10	At1g01080	0.9042	0.9041	1.0593	0.9300	0.0773	0.0601
11	At1g01090	1.4873	1.6288	-0.3224	-0.3106	0.6158	0.5189
12	At1g01100	1.2858	1.2653	-0.0542	-0.0030	-0.3353	-0.3996
13	At1g01110	-0.5739	-0.4931	0.3831	0.2371	-1.0614	-1.2403
14	At1g01120	0.6451	0.4880	1.4538	1.4422	-0.8261	-0.9685
15	At1g01130	0.2656	0.4062	0.0312	0.0338	1.1975	1.1718
16	At1g01140	0.8378	0.7601	0.1744	0.1655	-0.0671	-0.0313
17	At1g01150	0.4551	0.5159	0.9629	0.9596	0.3979	0.4504
18	At1g01160	-0.0261	-0.0262	-0.6591	-0.5196	1.3509	1.2571
19	At1g01170	-0.8704	-0.9254	0.0169	0.0314	-0.6277	-0.6292
20	At1g01180	-0.7845	-0.7282	-0.0663	-0.1267	-0.0682	-0.1228
21	At1g01190	-0.3691	-0.3608	-0.1055	-0.0091	0.0961	0.0572

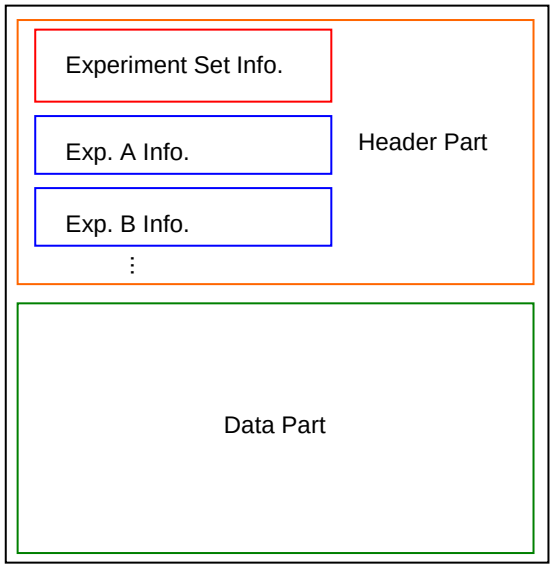
Metabolite data.

	A	B	C	D	E	F	G
1	(compexp)	Treatment C	Treatment C	Treatment D	Treatment D	Control	Control
2	(rep)	1	2	1	2	1	2
3	KPC00001	657	663	124	125	5	5
4	KPC00002	9	9	9967	10274	9	9
5	KPC00003	241184	238124	864	864	22	22
6	KPC00004	122	120	18	18	180	177
7	KPC00005	5140	5026	372	377	407378	393058
8	KPC00006	463	479	51	51	7	7
9	KPC00007	458	459	428122	438831	10175	10204
10	KPC00008	50484	50152	752	772	86	85
11	KPC00009	53	52	235794	238191	43	42
12	KPC00010	64323	64567	890	896	24	25
13	KPC00011	91	93	1299	1294	1323	1342
14	KPC00012	78	81	21	21	709	701
15	KPC00013	13949	13348	6455	6697	4135	4038
16	KPC00014	850	871	147273	145685	72457	71906
17	KPC00015	15519	14881	5104	4996	7	7
18	KPC00016	18904	18830	4	4	55	57
19	KPC00017	2011	2042	553280	553895	618	632
20	KPC00018	269	270	331455	320680	49	50
21	KPC00019	280	282	8606	8886	52	53
22	KPC00020	670745	672289	68416	66720	2853	2855
23	KPC00021	15492	15524	1926	1878	2081	2107
24	KPC00022	7633	7860	95970	91740	300	304
25	KPC00023	1427	1406	202	207	91220	93044
26	KPC00024	5	5	74107	74234	19	19

9-2-2. Header Part

The data files prepared following the directions above can include header information at the head of the files to describe the metadata of the experiments. The data files without headers can be accepted by the KaPPA-View4 system, but we recommend attaching header information so that the data are recognized unambiguously on KaPPA-View4 by the user.

Several related experiment data can be grouped in a Experiment Set, the header includes parts for experiment set and each experiment data. The structure of the whole data file is schematically represented below.



The Data File (CSV or TSV)

- ***Format***

Each item of header information should be described in a line according to the following formats.

>[Item Name] <*tab*> [Description]

The lines in the header must start with ">", therefore descriptions must not contain newlines (returns). Blank lines are ignored.

There are two types of the Item Name, the Reserved Headers and User Defined Headers.

- ***Reserved Headers***

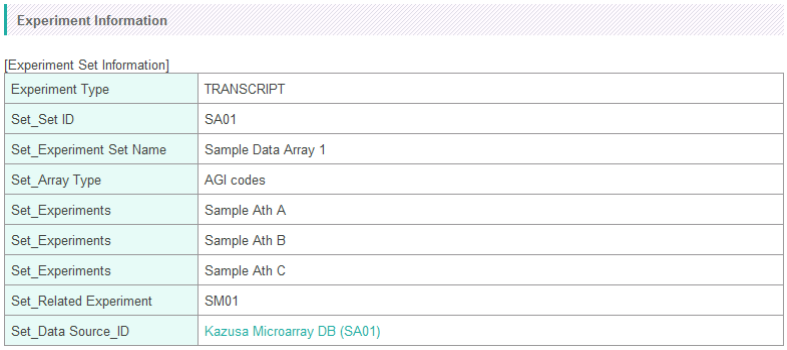
The Reserved Headers are recognized and the descriptions are processed properly by the system.

The Reserved Headers for the Experiment Set.

Reserved Headers	Explanation	Required
>datatype	Define the type of the data, "array" or "compound". It is	yes

>Set_Set ID	Define the ID of the Experiment Set. If it's omitted, the system automatically assigns an ID like "TempSet_[*number*]".	no
>Set_Experiment Set Name	Define the Name of the Experiment Set. If it's omitted, the system automatically attaches a name like "TempSet_[*number*]".	no
>Set_Species	Define the species name that the data derived. The name must match exactly to the one used in the system. If it's omitted, it has to be defined during the uploading process.	no
>Set_Array Type	Define the appropriate Array Type name corresponding to the feature IDs described in the data file. The array type name must match exactly to the one registered to the system. If it's omitted, it has to be defined during the uploading process.	no
> Set_Related Experiment	Define the Experiment Set ID that related to this data. The ID appears as a link to another Exp. Set on the data selection list of "Analysis" function. For example, a transcriptome Exp. Set can be related a metabolome Exp.Set that are obtained from the same samples. In this case, link to the metabolome Exp. Set will help a smooth data selection in the "Analysis".	no
>Set_***_ID *and* >Set_***_link *(*** is an arbitral text string)*	On the Experiment Information window (see below), users can make arbitral links. Define the caption of the link as ">Set_***_ID" header, and define the URL corresponding to the caption as ">Set_***_link".	no

The figure below is an example of the Experiment Information window. The window appears when the Exp. Set ID is clicked on the experiment data list of the "Analysis" function.



Experiment Information

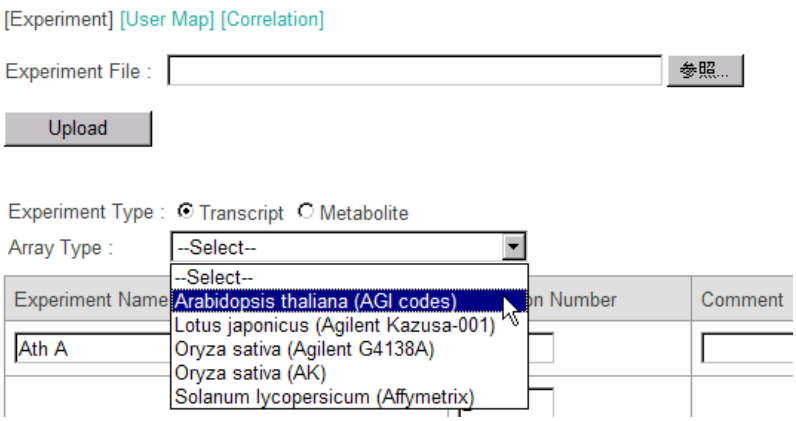
[Experiment Set Information]

Experiment Type	TRANSCRIPT
Set_Set ID	SA01
Set_Experiment Set Name	Sample Data Array 1
Set_Array Type	AGI codes
Set_Experiments	Sample Ath A
Set_Experiments	Sample Ath B
Set_Experiments	Sample Ath C
Set_Related Experiment	SM01
Set_Data Source_ID	Kazusa Microarray DB (SA01)

The Array Type names registered to the system can be investigated as followings.

Click "Temporary Upload" on the Main Menu after logging-in to KaPPA-View4. Click "Browse" to select a data file of transcripts and then press "Upload". The system checks the data file and the summary of the file appears. Click on the "Array Type" pull-down list to check the names of the Array Type (written in the

parentheses) registered and the corresponding species names. Download the basic information file for feature-gene relationships to check the Feature IDs (see **7. Download**).



The Reserved Headers for experiment data.

Reserved Header	Explanation	Required
>Data_Experiment_ID	Define the Experiment D. If it's omitted, the system assigns automatically an ID like "TempExp_[*number*]".	no
>Data_Experiment_Name	Defines the name of the Experiment. The name has to be described exactly the same as the corresponding data name written in the Data part.	yes
>Data_Value Type	Define the type of the experiment data, "ratio" or "quantitative" (see **Type of Experiment Data**). If it's omitted, it's set automatically to "quantitative".	no
>Data_Comments	Enter comments.	no
>start	The header doesn't have the description. It is a marker indicating the beginning of one experiment data.	yes
>end	The header doesn't have the description. It is a marker indicating the end of one experiment data.	yes
>Data_***_ID *and* >Data_***_link *(*** is an arbitral text string)*	On the Experiment Information window (see below), users can make arbitral links. Define the caption of the link as ">Data_***_ID" header, and define the URL corresponding to the caption as ">Data_***_link".	no

The figure below is an example of the Experiment Information window. The window appears when the Exp. ID is clicked on the experiment data list of the "Analysis" function.

[Experiments]

Data_Experiment_ID	SEA01
Data_Experiment_Name	Sample Ath A
Data_Value Type	quantitative
Data_Comments	Arabidopsis thaliana Sample Data A
Data_Data Source_ID	Kazusa Microarray DB (SEA01)

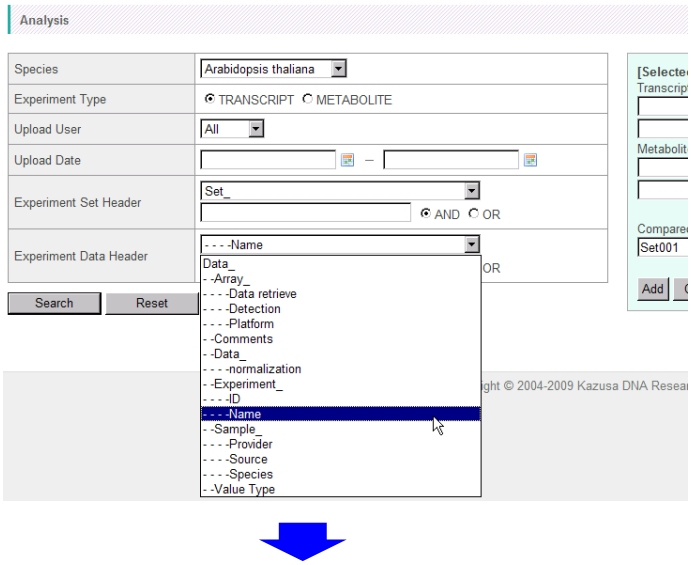
Data_Experiment_ID	SEA02
Data_Experiment_Name	Sample Ath B
Data_Value Type	quantitative
Data_Comments	Arabidopsis thaliana Sample Data B

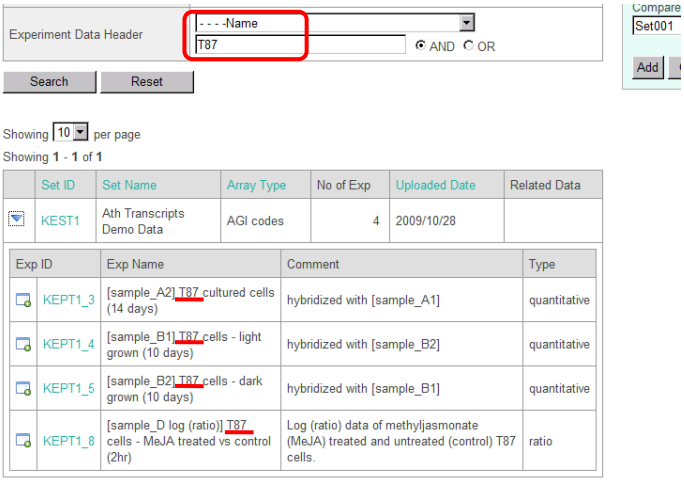
Data_Experiment_ID	SEA03
Data_Experiment_Name	Sample Ath C
Data_Value Type	ratio
Data_Comments	Arabidopsis thaliana Sample Data C

- ***User Defined Headers***

To describe other metadata, users can define their original headers (User Defined Headers). The headers concerning to the Experiment Set and Experiment Data must starts with ">Set_" and ">Data_" respectively. The User Defined Headers must not have the same name as the Reserved Headers.

As the under bar "_" represents a separator and the name of header can be separated into several keywords, the set of headers are treated by the system as a tree structure where the keywords represent each node. Users can search the experiment data based on the descriptions of the headers through the "Analysis" interface which can specify the nodes (see below).





- ***Sample***

	A	B	C	D	E	F	G
1	>datatype	array					
2	>Set_Set ID	SA01					
3	>Set_Experiment Set Name	Sample Data Array 1					
4	>Set_Species	Arabidopsis thaliana					
5	>Set_Array Type	AGI codes					
6	>Set_Experiments	Sample Ath A					
7	>Set_Experiments	Sample Ath B					
8	>Set_Experiments	Sample Ath C					
9	>Set_Related Experiment	SM01					
10	>Set_Data Source_ID	Kazusa Microarray DB (SA01)					
11	>Set_Data Source_link	http://www.kazusa.or.jp/microarray_db/search?setid=SA01					
12							
13	>start						
14	>Data_Experiment_ID	SEA01					
15	>Data_Experiment_Name	Sample Ath A					
16	>Data_Value Type	quantitative					
17	>Data_Comments	Arabidopsis thaliana Sample Data A					
18	>Data_Data Source_ID	Kazusa Microarray DB (SEA01)					
19	>Data_Data Source_link	http://www.kazusa.or.jp/microarray_db/search?expid=SEA01					
20	>end						
21	>start						
22	>Data_Experiment_ID	SEA02					
23	>Data_Experiment_Name	Sample Ath B					
24	>Data_Value Type	quantitative					
25	>Data_Comments	Arabidopsis thaliana Sample Data B					
26	>end						
27	>start						
28	>Data_Experiment_ID	SEA03					
29	>Data_Experiment_Name	Sample Ath C					
30	>Data_Value Type	ratio					
31	>Data_Comments	Arabidopsis thaliana Sample Data C					
32	>end						
33							
34	(arrayexp)	Sample Ath A	Sample Ath A	Sample Ath B	Sample Ath B	Sample Ath C	Sample Ath C
35	(rep)	1	2	1	2	1	2
36	At1g01010	-0.828475158	-0.902463025	0.646901298	0.527831687	-1.190483883	-1.077241722
37	At1g01020	-1.322449978	-1.213715999	-0.524355622	-0.561203464	-0.166943424	-0.222037326
38	At1g01030	1.118873492	1.1207812	1.41118869	1.338139967	-1.266626461	-1.29633268
39	At1g01040	-1.218112023	-1.298647166	-0.961450022	-0.896647384	0.07469906	0.11324618
40	At1g01050	-0.664935805	-0.769907958	0.126151787	0.211021608	1.175051608	1.170206475
41	At1g01060	0.809261597	0.765810307	1.566740565	1.529944679	0.579037425	0.609907787
42	At1g01070	-1.204413205	-1.274388803	-0.889537352	-0.808698789	0.61000164	0.611701704

Refer to the sample data too which are available from the KaPPA-View4 main page after logging-in.

9-3. Correlation Data

The gene-to-gene and metabolite-to-metabolite correlation data are used for representing on the pathway maps as curves (see **4-4-3. Overlay of Correlation Data**).

9-3-1. Format

Make data describing following information in each line, and save it as a CSV file.

Column	Description	Data type	Required
1st	ID1	Text string (max. 100 letters)	yes
2nd	ID2	Text string (max. 100 letters)	yes
3rd	A value	Real number	yes

The gene IDs and compound IDs registered to KaPPA-View4 can be entered into the first and second column. Gene-to-gene relationships and compound-to-compound relationships should be described in separate files, because they are separately uploaded. Gene-to-metabolite relationships are not accepted by the system.

If there are multiple lines that having the same combination of ID1 and ID2, an alert message appears in the uploading process when they have different values, but they are accepted as one record when they have the same values.

9-3-2. Sample

```
1  At1g01060,At2g46830,0.855
2  At1g01060,At3g09600,0.813
3  At1g01060,At4g38960,0.800
4  At1g01080,At3g48730,0.900
5  At1g01080,At5g55220,0.893
6  At1g01080,At3g29185,0.891
7  At1g01080,At4g29060,0.890
8  At1g01080,At1g32990,0.888
9  At1g01080,At1g05190,0.886
10 At1g01080,At1g79850,0.882
11 At1g01080,At2g37660,0.881
12 At1g01080,At1g64510,0.880
13 At1g01080,At5g13510,0.880
14 At1g01080,At5g47190,0.877
15 At1g01080,At3g55330,0.876
16 At1g01080,At5g45930,0.876
17 At1g01080,At3g01480,0.875
18 At1g01080,At1g74970,0.874
19 At1g01080,At1g62780,0.873
20 At1g01080,At2g35500,0.872
21 At1g01080,At5g42765,0.872
22 At1g01080,At1g48350,0.871
23 At1g01080,At3g47650,0.871
```

Note:

File size of correlation data tend to become large. If the file was not accepted by the server due to the restriction of the file size, an alert message appears to inform you it. Consider to reduce the file size by cutting off the entries based on a certain threshold of correlation values, or by rounding the values at proper digit to reduce the letter numbers.

9-4. User Map

This section introduces briefly about the User Maps and their format restrictions. Refer to the "**Manual on User Map Creation**" for more details.

Users can create their own pathway maps (User Maps) and utilize them for their analysis. By creating pathway maps that are not provided as defaults, are edited and curated more carefully, are better designed for presentations, and include genes and compounds not registered to the system, the possibility of KaPPA-View4 analysis can be expanded.

The User Map files can be utilized by being uploaded by the users through KaPPA-View4 GUI, or by transferring in POST methods through the KaPPA-View4 API. The Power Users can send their own User Maps to the KaPPA-View administrator with simple steps to discuss about improvement of default maps.

The User Maps should be created in a Scalable Vector Graphics (SVG) format. To paint colors on the symbols for genes, compounds, and reactions dynamically by the system, a proper ID should be assigned for each symbol. We recommend using the free drawing software Inkscape (http://www.inkscape.org/).

A summary of ID formats for the SVG objects on the User Maps.

Element	Explanation	ID format
Gene	Represent genes.	text+(integer number)_g Gene IDs used in KaPPA-View4 system or user-defined ID starting with "TMG". Use different integer numbers to draw one gene on more than one place of the same map. Ex) At1g12340(1)_g, TMG001(1)_g
Gene box	Used for defining a region to represent a group of genes corresponding to an enzymatic reaction.	B+number The number corresponds to the numeric part of enzymatic reaction ID starting with "R" Ex) B00001
Compound	Represent compounds.	KPC+number, C+number, G+number, D+number, OR TMC+number KPC+number is compound ID format used in KaPPA-View4 Classic system. C, G or D+number are compound ID format used in KaPPA-View4 KEGG. TMC+number is a user-defined ID. Ex) KPC00005, TMC00001
Enzymatic reaction	Represent enzymatic reactions.	R+number Enzymatic reaction IDs used in the system. Ex) R0000101

Link to a related map	Enables to jump to another map	3 letters+integer number(+1 letter) Map IDs used in the system. Ex) Uni00001, Uni00034f, Lja00017

The SVG objects that are allowed for drawing the element symbols.

Element	SVG object	Note
Gene	rect	Be sure to set color to the fill, not to set "No paint".
Gene Box	rect	Be sure to set color to the fill, not to set "No paint".
Compound	rect or circle	Be sure to set color to the fill, not to set "No paint". rect を使用する場合、rx, ry 属性で角のカーブを作成することにより、丸い化合物シンボルを作ることができます。
Enzymatic reaction	line or path	Dotted lines can only be represented properly with "line" object. Stylized arrow heads of path objects are not properly displayed.
Link to a related map	rect or path	

About Other SVG objects.

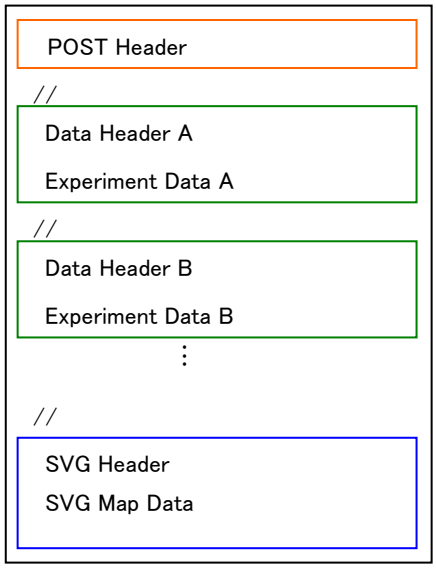
Text objects	Use Arial font. Conversion to outlines is recommended.
Images	By embedding image files to the SVG canvas, they can be displayed on KaPPA-View4.

9-5. Data for POST Transferring

The data transferred by POST transferring APIs (**8-1. Data Uploading APIs**) is called "POST data" here.

9-5-1. Structure of POST Data

The figure below schematically represents the structure of a POST data file.



POST Data File (text)

General information such as the species name are described in the POST Header part.

The following Experiment Data part are separated by a delimiter "//", and can include more than one experiment data.

The SVG Map Data part contains SVG data of User Map. It can be omitted.

9-5-2. Format

- ***Header Part of POST Data***

POST Header contains general information on the POST data. Each item of header information should be described in a line according to the format below.

>[Item Name] <*tab*> [Description]

The lines in the header must start with ">" and the description must not contain newlines (returns). Blank lines are ignored.

Write following items.

Item Name	Explanation	Required
>species	Define the species name. It must be exactly the same as the one registered to the system.	yes

>array	Define the Array Type name. It must be exactly match to the name registered to the system. See **p.72** for investigating valid Array Type names.	yes (when the file included any transcriptome data)
>default_map	Define the map which appears first after the POST transferring. Trim first three alphabets and write it as [Description]. When an SVG data contained in the file, the map ID defined in the ">svg_name" header (see below) can be described. When it is omitted or invalid ID was described, no default map appears.	no

- ***Experiment Data Part***

Be careful that the Experiment Data Part for POST data is different from those of the data files for GUI uploading (described in **9-2. Experiment Data for Uploading**). The Experiment Data in the POST data corresponds to the Compared Experiment Pairs which are created during the "Analysis" steps.

More than one Experiment Data can be included and they must be separated by a delimiter "//". Each Experiment Data contains Header part and Data part.

The format of the header is the same as POST headers. Describe following information.

Item Name	Explanation
>set_name	Define the name of the data which corresponds to the Compared Experiment Pair name.
>data_type	Define the type of the experiment, "array" or "compound".

If there is more than one Experiment Data that have the same ">set_name" and ">data_type" among the Experiment Data separated by "//", only the former data is registered.

When there are pair of Experiment Data that have "array" and "compound" for ">data_type" and shares the same ">set_name", the data is recognized as a Compared Experiment Pair which includes both of transcript and metabolite data.

The Data part follows to the Header Part. Describe the items below. Be careful that they are different from the ones of data for GUI-uploading (described in **9-2. Experiment Data for Uploading**).

Column	Explanation
1st	Enter Feature IDs for transcripts data and Compound IDs for metabolites. To investigate the Feature IDs and Compound IDs used in the system, download the basic information files (see **7. Download**). The user defined IDs, TMG[*number*] for genes and TMC[*number*] for compounds can be entered too to paint colors on the SVG objects that having these IDs on the User Maps.
2nd	Enter experiment values. The values should be transformed to logarithm scale for

	genes (based on 10 is recommended), and linear scale for metabolites. The compound values must not negative nor 0. If the 3rd column have a value, values in the 2nd and 3rd column recognized by the system as quantitative data of the numerator and the denominator, respectively, for the Compared Experiment Pair. If there is no data in the 3rd column, the value in the 2nd column recognized as ratio data.
3rd	Enter experiment values. It can be omitted. The values should be transformed to logarithm scale for genes (based on 10 is recommended), and linear scale for metabolites. The compound values must not negative nor 0. The value is used as a denominator for the ratio calculation with the value in 2nd column. If the 3rd column have no data, the value in the 2nd column is recognized as the ratio data.

- ***Map Data Part***

The header item below is required. The format of the header is the same as the POST header. Following to the header, the SVG data (text) of the User Map (**9-4. User Map**) should be written. Open the SVG file of the User Map with a text editor such as the Notepad for Windows to get the SVG data in text format.

Item	Explanation
>svg_name	Define the name of the map. The name appears on the Pathway Tree. When the name is written in the ">default_map" item in the POST header, the map appears first on the browser after transferring the POST data.

9-5-3. Sample

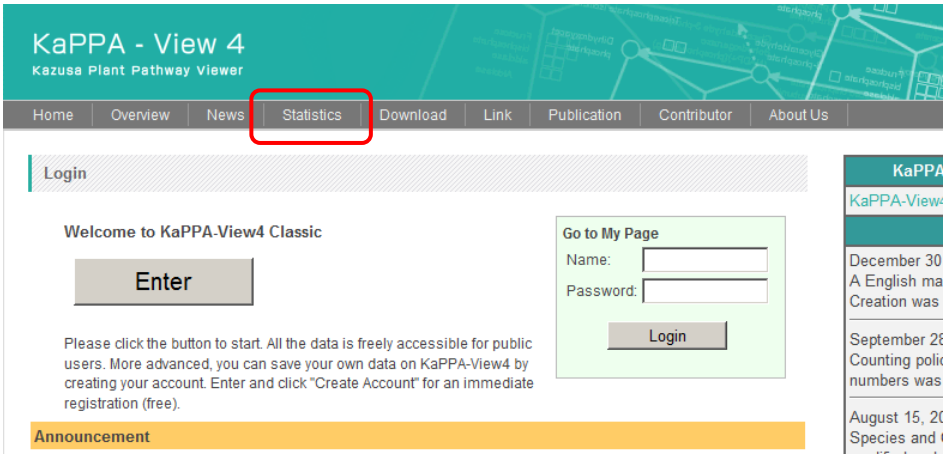
```
>species^   Arabidopsis thaliana↵
>array^ AGI Codes↵
>default_map^   AthTMAP02↵
//↵
>set_name^   Posted CompExp A↵
>data_type^ array↵
TMG001    ^   0.07652^    0.08213↵
At1g01010^  -0.58741^    0.10697↵
At1g01030^  -0.40190^   -0.62194↵
At1g01040^  -0.61347^    0.19020↵
At1g01050^   0.08326^    1.59364↵
  .
  .
  .
//↵
>set_name^  Posted CompExp B↵
>data_type^ array↵
TMG001    ^   0.00150↵
At1g01010^  -0.98685↵
At1g01030^  -0.16034↵
At1g01040^   0.07234↵
At1g01050^   2.03661↵
  .
  .
  .
//↵
>set_name^  Posted CompExp B↵
>data_type^ compound↵
TMC001    ^  1008^     13↵
TMC002    ^    98^   1287↵
KPC00001^    123^    143↵
KPC00002^   1056^    978↵
KPC00003^    348^    417↵
  .
  .
  .
//↵
>svg_name^  AthTMAP02↵
<?xml version="1.0" encoding="UTF-8" standalone="no"?>↵
<!-- Created with Inkscape (http://www.inkscape.org/) -->↵
<svg↵
   xmlns:svg="http://www.w3.org/2000/svg"↵
      xmlns="http://www.w3.org/2000/svg"↵
  .↵
  .↵
  .↵
</svg>↵
```

10. Default Data

10-1. System Data, Species, and Correlation Data

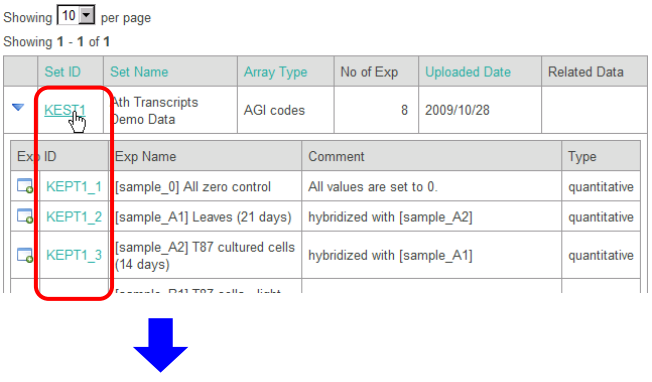
Refer to the KaPPA-View4 paper and the supplement file of it (Sakurai N, et al. (2011) Nucleic Acids Research 39: D677-684) for details of the default data (at August 15th 2010).

The "Statistics" page of KaPPA-View4 website shows the latest data installed in the system.



10-2. Experiment Data

We provide some sample data for demonstration. Click on the Experiment Set ID or Experiment ID on the experiment data list of the "Analysis" function. A window pops up to show the metadata.



Experiment Information

[Experiment Set Information]

Experiment Type	TRANSCRIPT
Set_Set ID	KEST1
Set_Experiment Set Name	Ath Transcripts Demo Data
Set_Array Type	AGI codes
Set_Description	Default data for demonstration
Set_Depositor Name	sakurai

[Experiments]

Data_Experiment_ID	KEPT1_1
Data_Experiment_Name	[sample_0] All zero control
Data_Value Type	quantitative
Data_Comments	All values are set to 0.

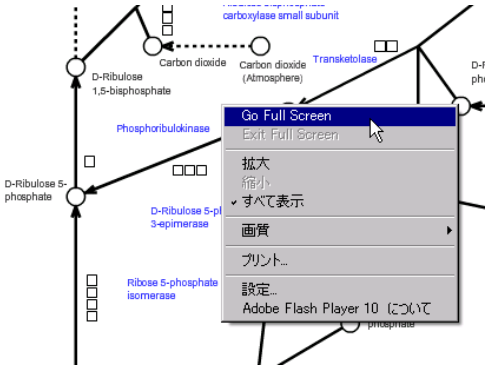
Data_Experiment_ID	KEPT1_2
Data_Experiment_Name	[sample_A1] Leaves (21 days)
Data_Value Type	quantitative
Data_Comments	hybridized with [sample_A2]
Data_Sample_Source	Arabidopsis leaves, 21 days
Data_Sample_Species	Arabidopsis thaliana
Data_Sample_Provider	Kazusa DNA Research Institute
Data_Array_Platform	Agilent 22 K
Data_Array_Detection	2 color
Data_Array_Data retrieve	Feature Extraction 9.0
Data_Data_normalization	normalized to median

Data_Experiment_ID	KEPT1_3

11. Hints and Trouble Shootings

11-1. Getting a Screen Shot

Right-click on a pathway map, and select "Go Full Screen". Press "PrtScr" button (Windows) to get a screen shot in a fixed pixel sizes.



11-2. Two-color Microarray Data

The 2-color DNA microarray data can be properly represented on the pathway maps, by following to the ways below.

1)

- Prepare the experiment data file for KaPPA-View4 using the log(ratio) data generated by the 2-color microarray analysis.

- Create a control data file where all the gene expression values are set to 1.

- Upload both of the experiment data and the control data to KaPPA-VIew4.

- Pair the two to create a Compared Experiment Pair.

2)

- Prepare the experiment data for KaPPA-View4 using the log(ratio) data generated by the 2-color microarray analysis.

- Attach header information to the data file (see **9-2. Experiment Data for Uploading**). Write "ratio" for the ">Data_Value" so that the data is recognized as "ratio" data by the system.

- Upload the data file.

- Create a Compared Experiment Pair with the data alone.

11-3. One-color Microarray Data

The log-transformed expression data can be normalized by a global mean (or median) so that the gene symbols are to be painted with a wide color ranges.

Alternatively, create the log-transformed expression data (not normalized) and a control data where all values are set to a global mean (or median). Upload the two and create a Compared Experiment Pair with them. An advantage of the later is that the experiment data can be directly used for other analyses by comparing to other 1-color data which are uploaded in the same way.

11-4. Investigating Compound IDs

In order to create a metabolite expression data for the analyses, users have to know the Compound IDs uses in the KaPPA-View4. There are several ways to investigate the Compound IDs.

* In KaPPA-View4 KEGG, the Compound IDs are the same as those of KEGG (http://www.genome.jp/kegg/).

1)

Click on the compound symbols on the pathway maps.

2)

Find the compound by "Search" function of KaPPA-View4.

3)

Download the compound information file from the "Download" menu after logging-in.

11-5. Representation of non-omics data on the maps

In usual, the function for the experimental data uploading is used to upload omics data such as gene expressions (transcriptome data) and metabolite accumulations (metabolome data). Utilizing the function, we can visualize the other data on the pathway maps as follows:

1) Significances of the data detections on the microarrays

ex) - P-value (log ratio) from the Agilent's microarray

2) Probabilities of the data changes

ex) - The results of t-test

3) Existence of the genes or compounds

ex) - Gene lists extracted from the microarray analysis

- A compound list of a standard chemical library

- Gene list for each GO evidence codes

to check the reliability of gene annotations

To prepare a data file as the same format to the experimental data files, we can upload the data and represent on the pathway maps on KaPPA-View4. In the case of 1) and 2), the significance and probability values are described in the data file instead of the detection values. Of course the values should be scaled to fit to the color gradation of KaPPA-View4. In the case of 3), existence of the genes or compounds should be described as a proper value, ex., 1 (exists) and 0 or null (not exists) for genes. By setting proper values for multiple categories of the genes or compounds, users can represent up to 9 categories in different colors simultaneously on the maps.

The functions for switching several experimental data to browse, and for representing two experiments data simultaneously on a single map, enable users to browse the real experimental data and correlation data while checking the values of user's interest.

12. Acknowledgements

KaPPA-View was developed in Kazusa DNA Research Institute with a support from the New Energy and Industrial Technology Development Organization (NEDO), Japan under the research project named "Development of Fundamental Technologies for Controlling the Material Production Process of Plants" (P02001).

13. References

Papers about KaPPA-View

Sakurai N, Ara T, Ogata Y, Sano R, Ohno T, Sugiyama K, Hiruta A, Yamazaki K, Yano K, Aoki K, Aharoni A, Hamada K, Yokoyama K, Kawamura S, Otsuka H, Tokimatsu T, Kanehisa M, Suzuki H, Saito K and Shibata D (2011) KaPPA-View4: a metabolic pathway database for representation and analysis of correlation networks of gene co-expression and metabolite co-accumulation and omics data. *Nucleic Acids Research* **39**: D677-684

Sakurai N and Shibata D (2006) KaPPA-View for integrating quantitative transcriptomic and metabolomic data on plant metabolic pathway maps. *J Pesticide Science* **31**: 293-295

Tokimatsu T, Sakurai N, Suzuki H and Shibata D (2006) KaPPA-View: A tool for Integrating Transcriptomic and Metabolomic Data on Plant Metabolic Pathway Maps. *In* Saito K, Dixon RA and Willmitzer L eds, *Biotechnology in Agriculture and Forestry*, Vol. **57**, pp. 155-163, Springer-Verlag, Berlin Heidelberg

Tokimatsu T, Sakurai N, Suzuki H, Ohta H, Nishitani K, Koyama T, Umezawa T, Misawa N, Saito K and Shibata D (2005) KaPPA-view: a web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiol* **138**: 1289-1300

Other Related Papers

Obayashi T, Hayashi S, Saeki M, Ohta H and Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Research* **37**: D987-991

14. About Us

Team KaPPA-View

E-mail: kappa-view at kazusa.or.jp (replace at to @)

The Administrators and developing team of KaPPA-View in Kazusa DNA Research Institute (http://www.kazusa.or.jp/e/). Please send your all inquiries to this e-mail address.

Axiohelix Co. Ltd.

E-mail: info at axiohelix.com (replace at to @)

The system programming was carried out by AXIOHELIX.

http://www.axiohelix.com/en/

Corporate Headquarters

Okinawa Sangyo Shien Center 502,1831-1, Oroku, Naha-shi, Okinawa 901-0152, Japan

TEL:+81-98-858-2887, FAX:+81-98-987-0288

Tokyo Office

Ando Bldg. 8F, 3-6-9, Kitashinagawa, Shinagawa-ku, Tokyo 140-0001, Japan

TEL:+81-3-6698-6731, FAX:+81-3-6698-6734

15. Manual Versions

- 1.0, January 8th 2011

The first release.