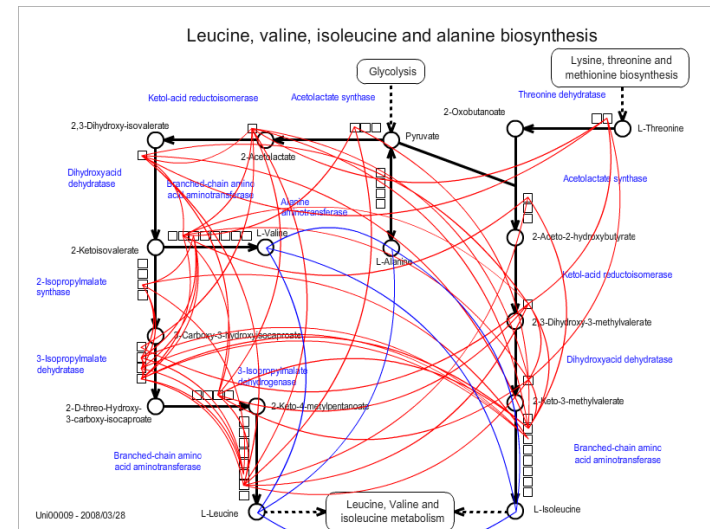


Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010

# **KaPPA-View4:** **A Pathway Database for** **Gene Co-expression and** **Metabolite Co-accumulation** **Analysis**



2010/10/13 Biocuration 2010 (Tokyo International Exchange Center)

**Nozomu Sakurai**

Kazusa DNA Research Institute  
 Lab. Genome Biotechnology

# Outline

1. About KaPPA-View4
2. Advanced Data Analysis
3. Perspectives

# KaPPA-View4



"Development of Fundamental Technologies for Controlling the Material Production Process of Plants" (2002-2010)

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010

Species

KaPPA - View 4  
Kazusa Plant Pathway Viewer

Main Temporary Upload Analysis **Map View** Search Download

Arabidopsis thaliana

Calvin cycle

Arabidopsis thaliana metab

- Carbohydrate metabolism
  - CO<sub>2</sub> fixation and cent
    - Calvin cycle
    - Glycolate pathw
    - Glycolysis/glucc
    - Phosphoenolpyr
    - TCA cycle
    - Glyoxylate cycle
    - Glycerol metabo
  - Mono-, di- and oligos
  - Polysaccharide metat
  - Miscellaneous carboh
  - Amino acid, nucleic acid

Create Simple Map

Calvin cycle

Glycolate pathway

3-Phospho-D-glycerate

3-Phospho-D-glyceroyl phosphate

Glyceraldehyde 3-phosphate

Gluconeogenesis

Triosephosphate isomerase

Dihydroxyacetone phosphate

Fructose biphosphate aldolase

Aldolase

D-Fructose 6-phosphate

D-Fructose 1,6-biphosphate

Fructose biphosphatase

Fructose biphosphatase aldolase

Aldolase

Sedoheptulose biphosphatase

D-Erythrose 4-phosphate

D-Xylulose 5-phosphate

Fructose biphosphatase

D-Ribulose 5-phosphate

D-Ribulose 5-phosphate 3-epimerase

Phosphoribulokinase

Transketolase

Carbon dioxide (Atmosphere)

Carbon dioxide

Ribulose biphosphate carboxylase small subunit

Glyceraldehyde 3-phosphate

D-Ribulose 1,5-biphosphate

Phosphoglycerate kinase

D-Ribose 5-phosphate

Ribose 5-phosphate isomerase

Glyceraldehyde 3-phosphate

D-Sedoheptulose 7-phosphate

D-Sedoheptulose 1,7-biphosphate

Transketolase

Uni0112 - 2008/03/15

\* To go full screen and to magnify the map, please right-click on the map.

Pathway Map

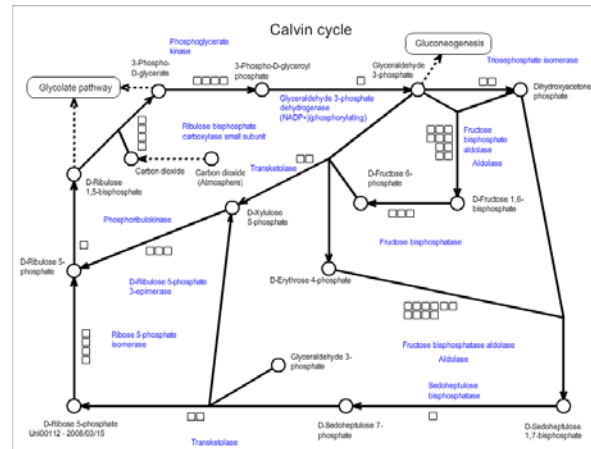
# History and the Pathway Maps

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010



2005 ver.1

*~150 original maps  
for Arabidopsis*



Dr. Tokimatsu T.  
(post-doc)

*moved Mar. 2005*



2006 ver.2



2008 ver.3

KaPPA - View 4  
Kazusa Plant Pathway Viewer

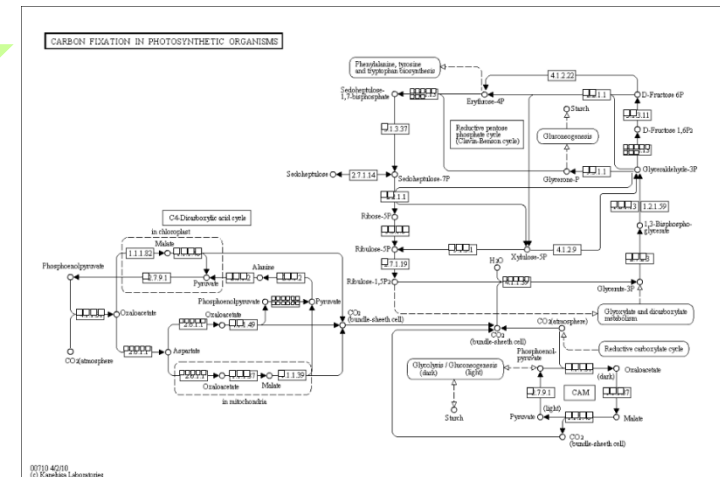
2009 ver.4

KaPPA - View 4  
Kazusa Plant Pathway Viewer

2010 ver.4 KEGG



*~300 maps for various species*



# Viewer Functions

## Relationship Data

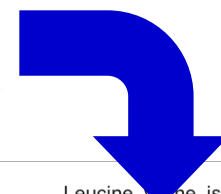
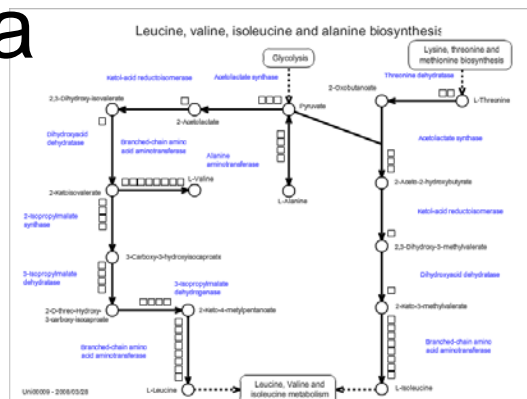
Gene co-expression

Met. co-accumulation

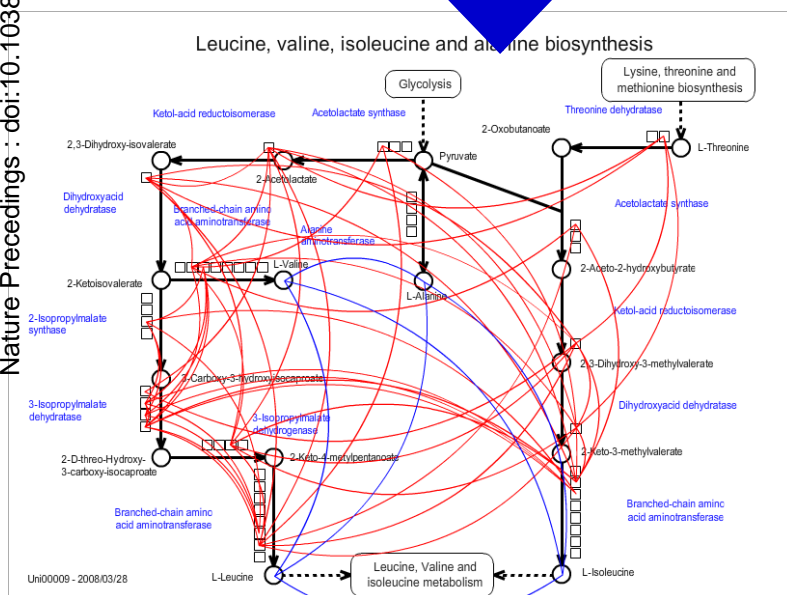
## Omics Data

Transcriptome

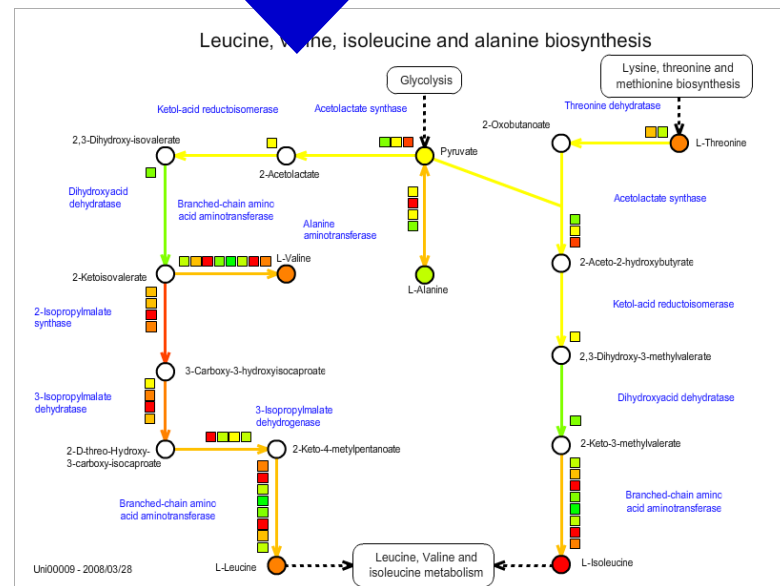
Metabolome



Leucine, valine, isoleucine and alanine biosynthesis



Leucine, valine, isoleucine and alanine biosynthesis



# Representation of Correlations



Arabidopsis, Rice



Human, Mouse, Rat

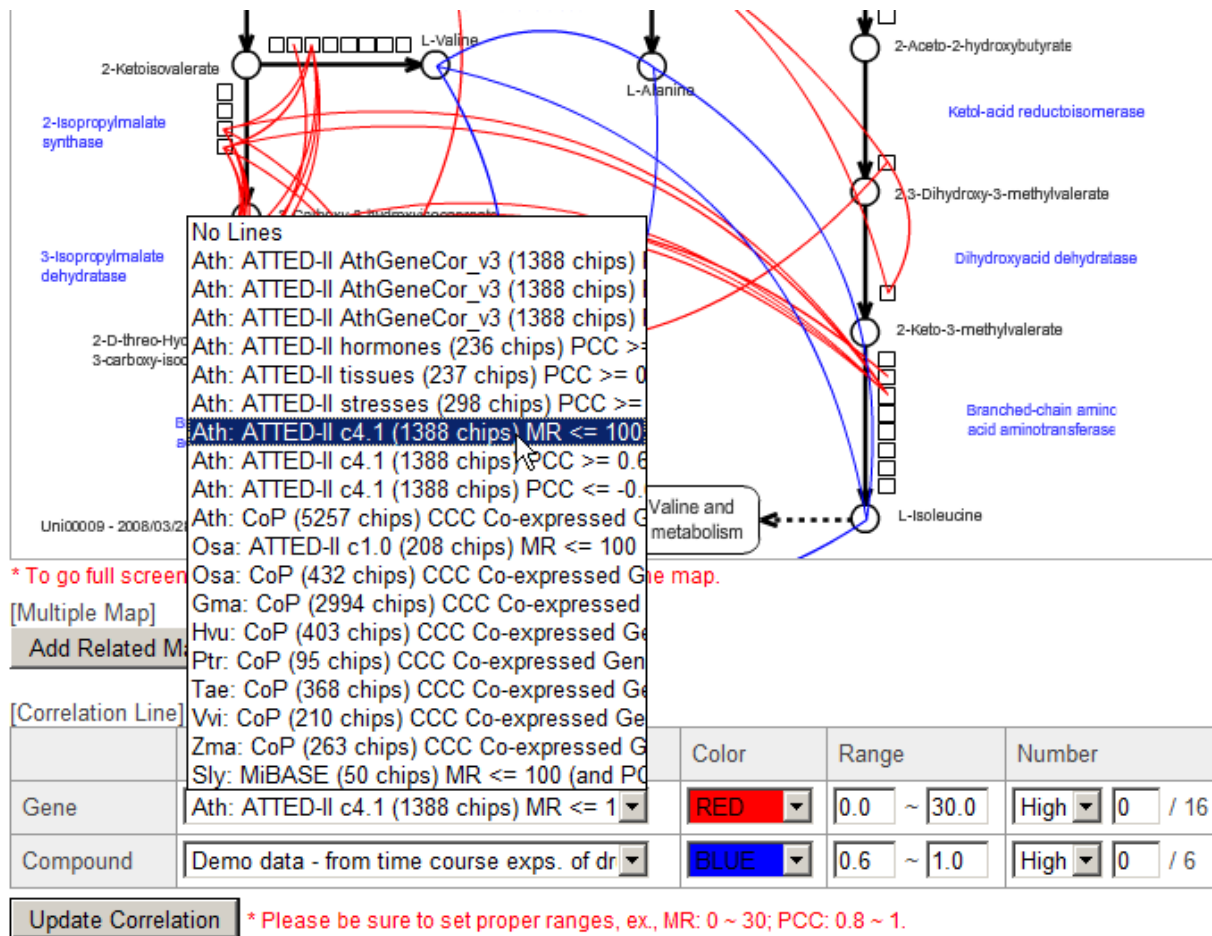


Arabidopsis, poplar, soy bean, barley, rice, wheat, grape, maize



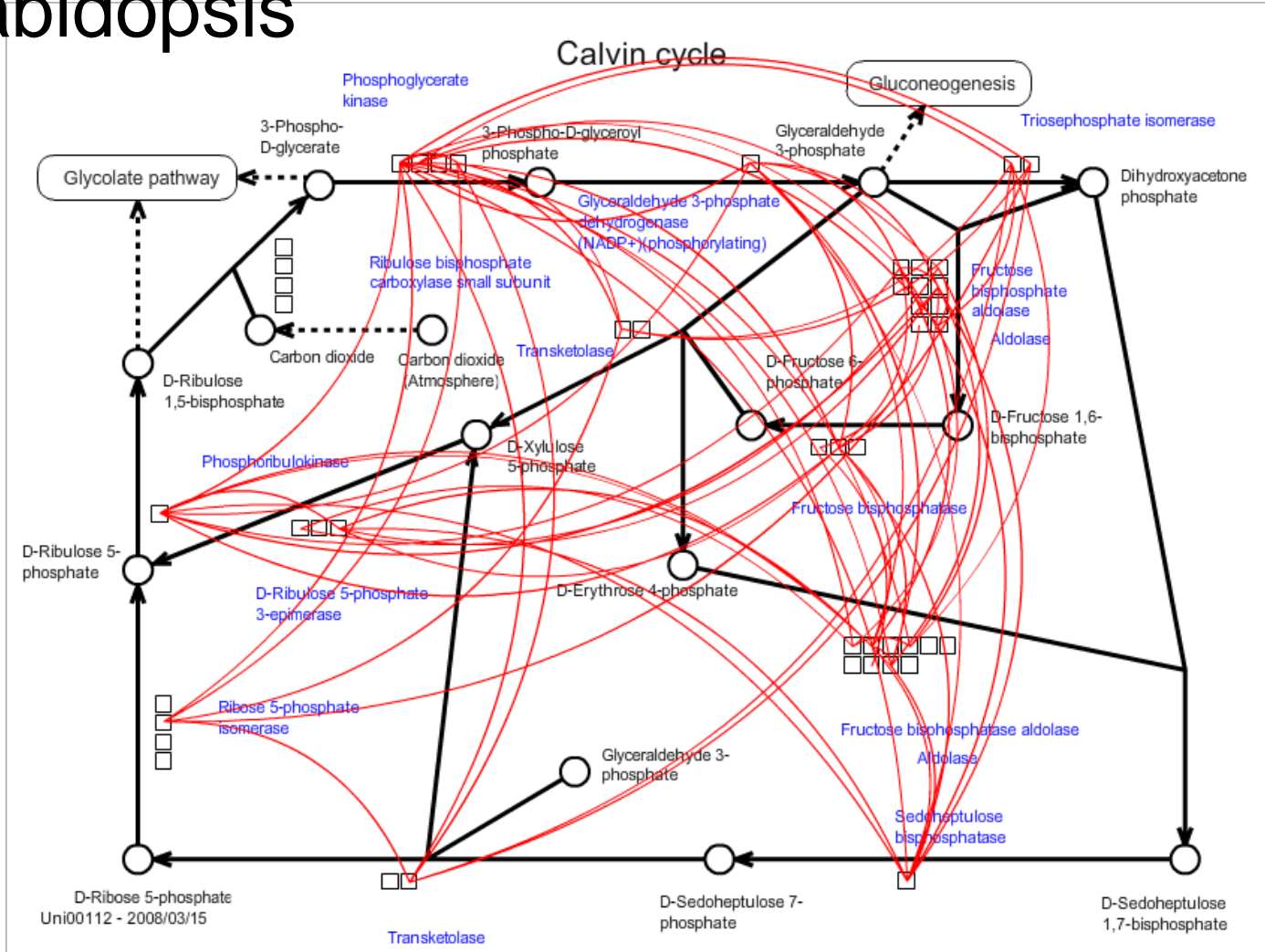
MiBASE

Tomato



# Correlations on the maps

## Arabidopsis

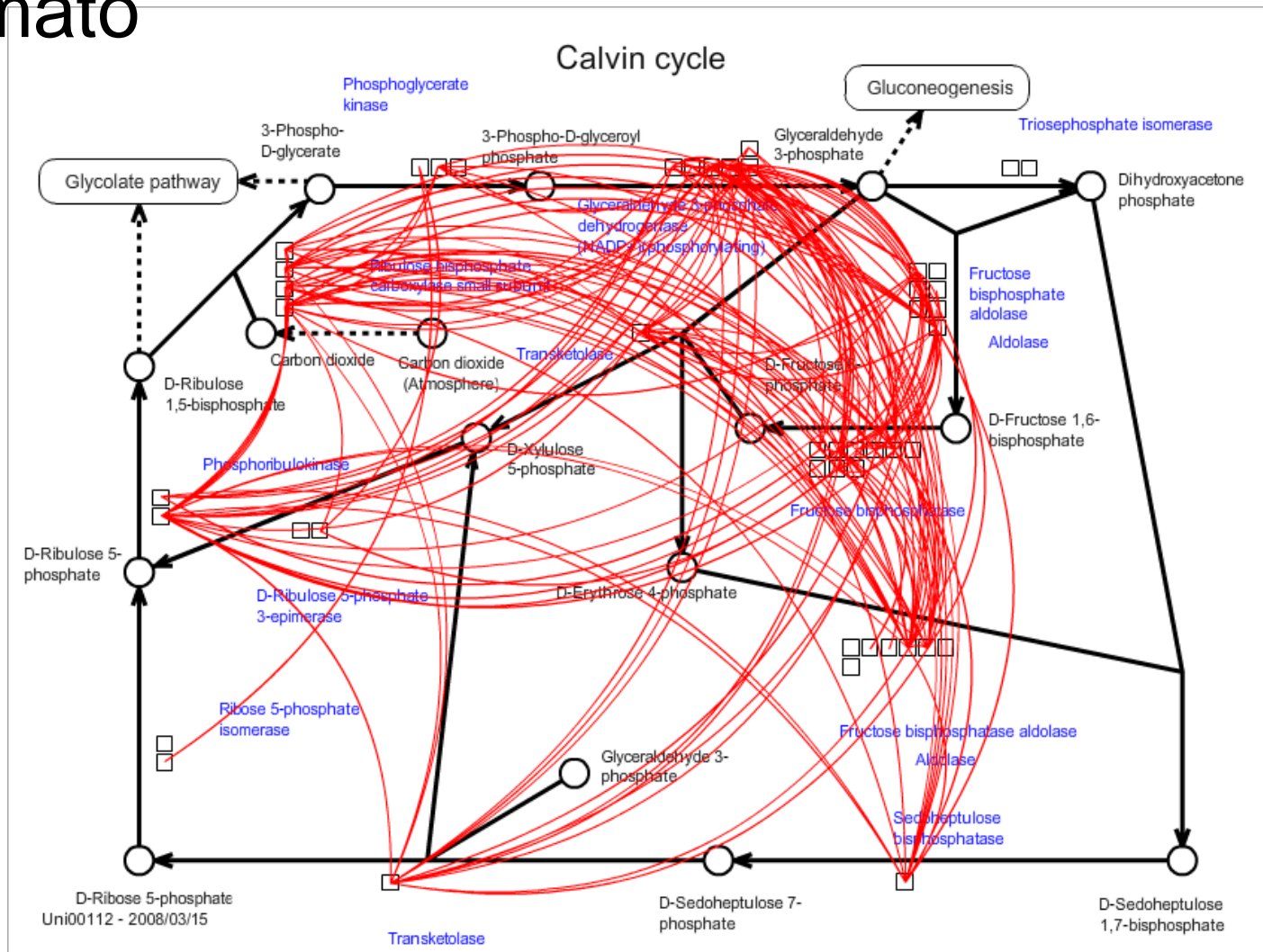


Ath: ATTED-II c4.1 (1388 chips) MR <= 100

# Correlations on the maps

## Tomato

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010



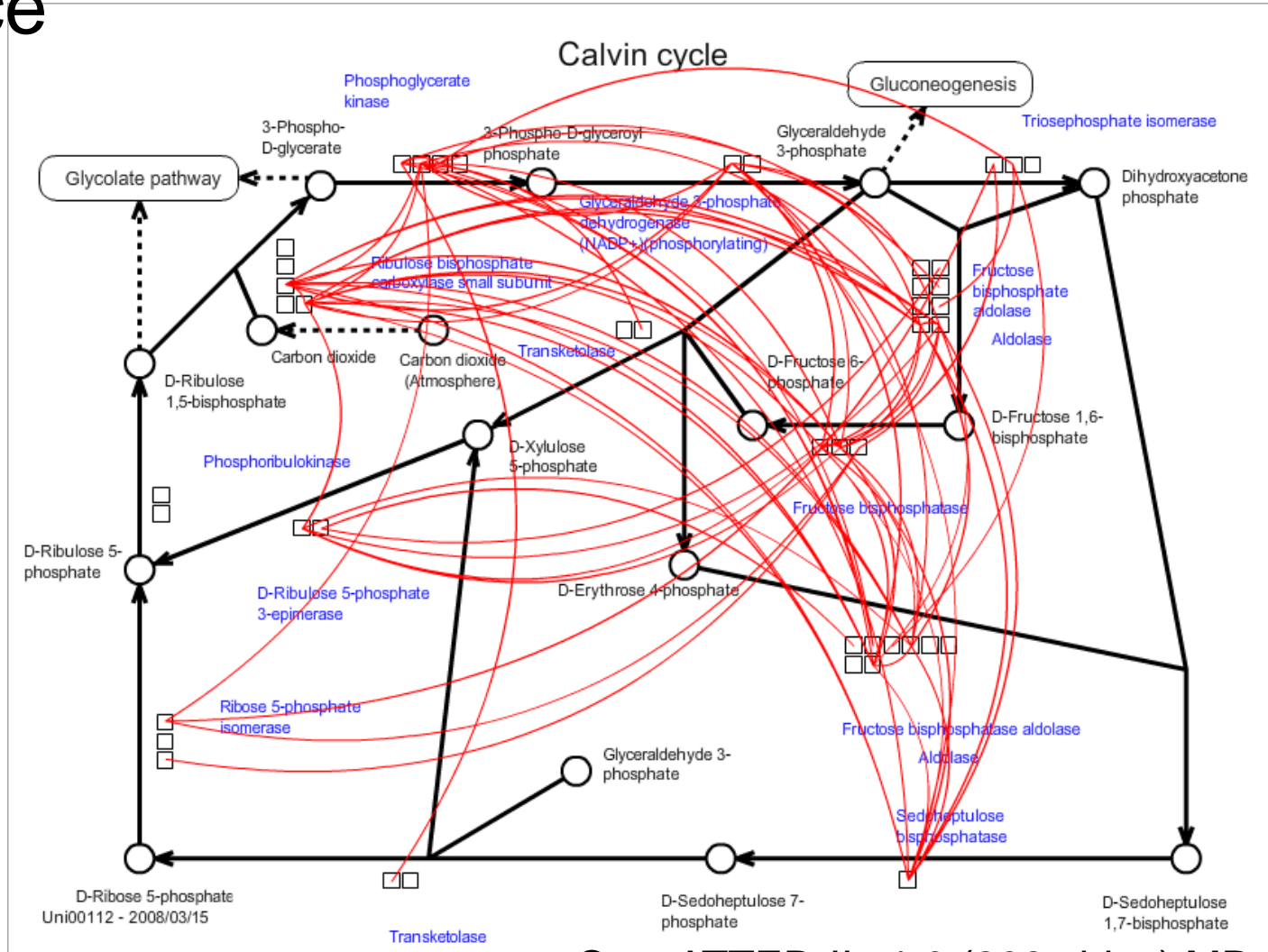
Sly: MiBASE (50 chips) MR  $\leq 100$  (and PCC  $\geq 0.6$ )



# Correlations on the maps

## Rice

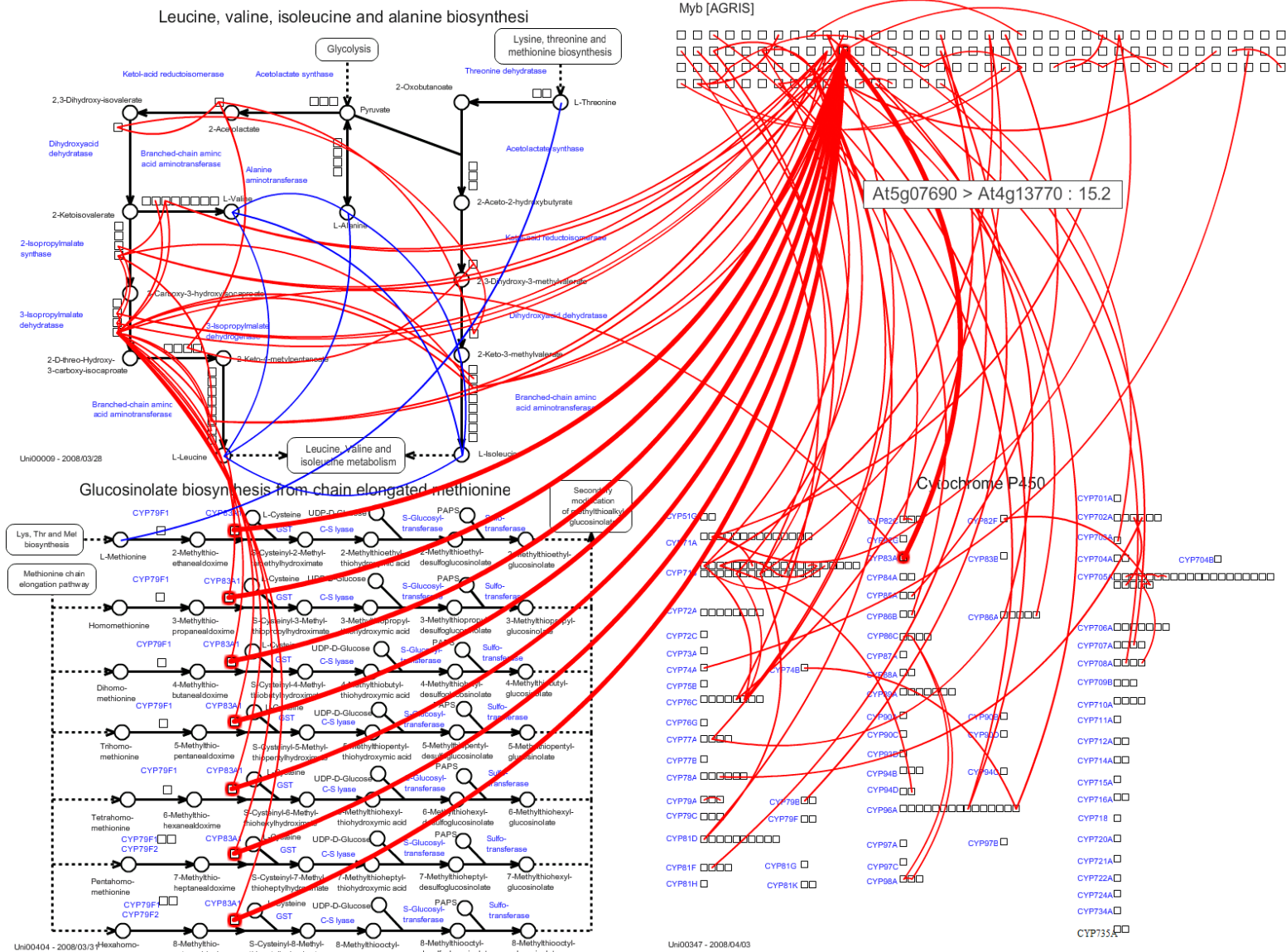
Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010



Osa: ATTED-II c1.0 (208 chips) MR <= 100

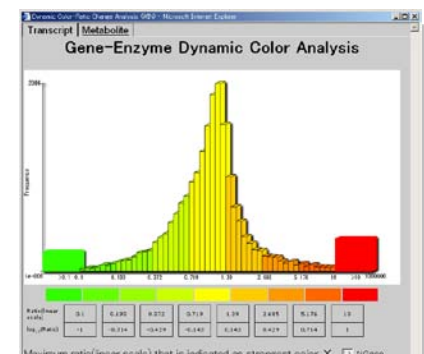
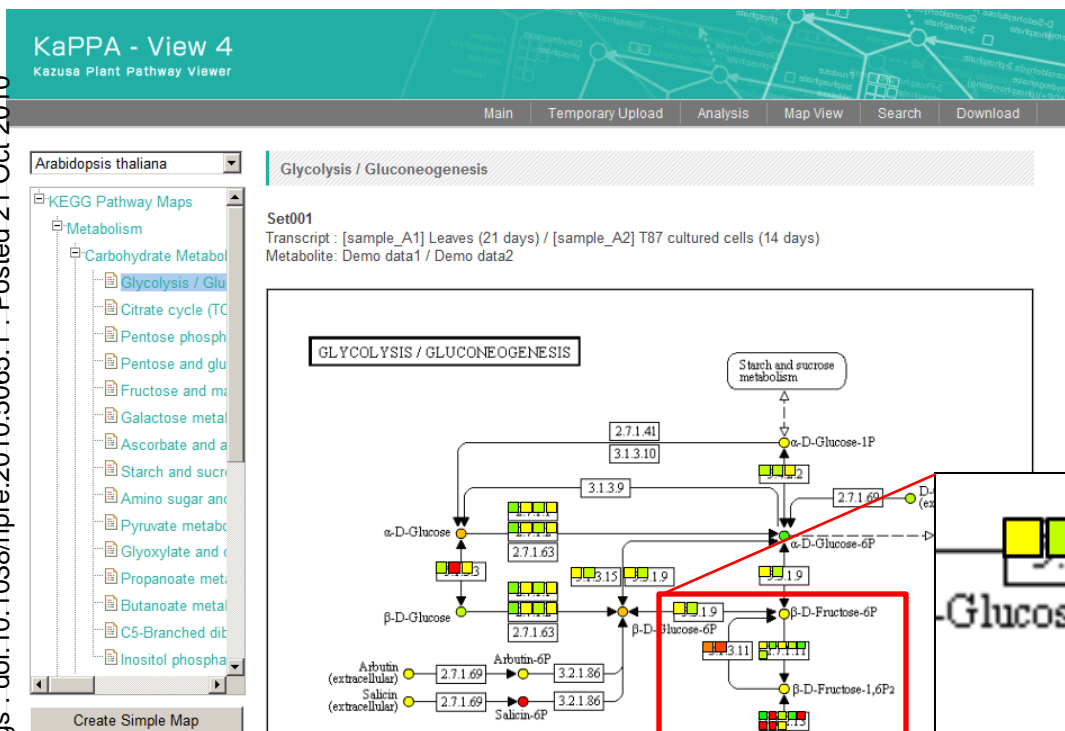
# Correlation curves across the maps

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010



# Representation of Omics Data

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010



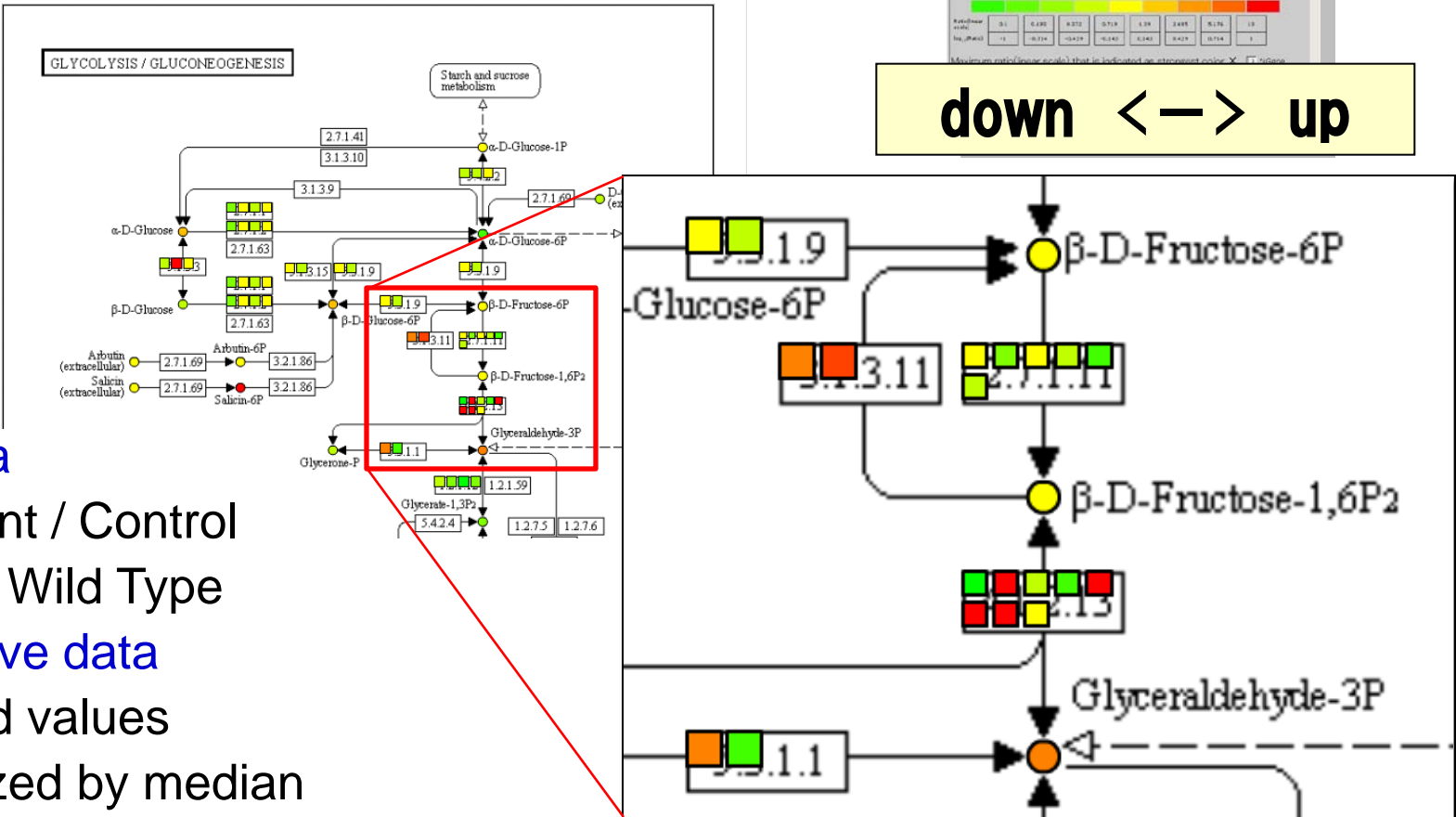
down  $\leftarrow$   $\rightarrow$  up

## Ratio data

- Treatment / Control
- Mutant / Wild Type

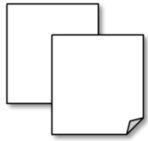
## Quantitative data

- Detected values normalized by median



# Procedure for Omics data representation

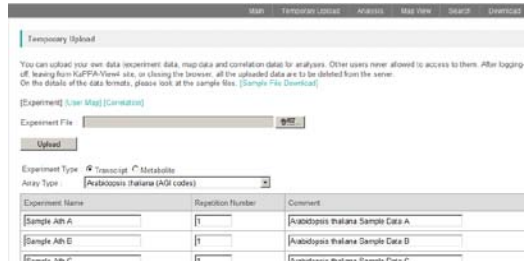
## Preparation of data files



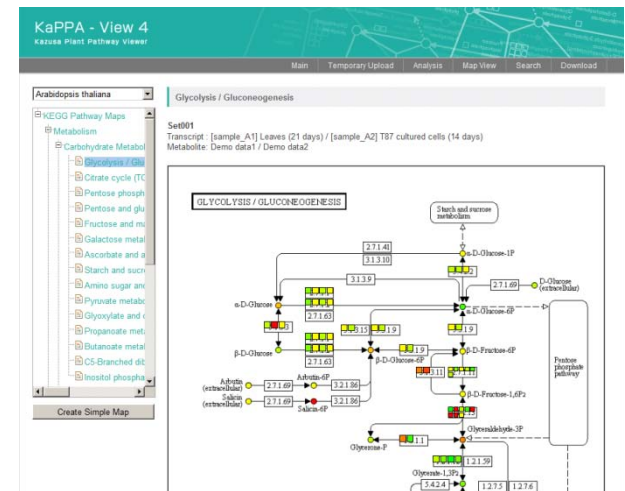
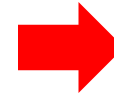
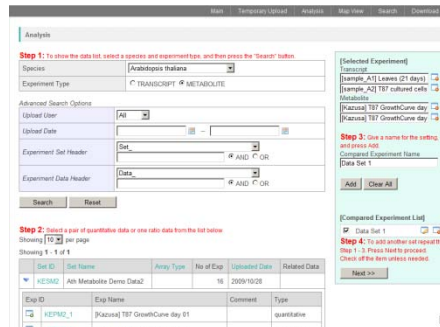
	A	B	C	D
1	(arrayexp)	Sample Ath A	Sample Ath B	Sample Ath B
2	(rep)	1	1	
3	At1_g01010	-0.82847516	0.646901298	-1.1904
4	At1_g01020	-1.32244998	-0.52435562	-0.1669
5	At1_g01030	1.118873492	1.41118869	-1.2666
6	At1_g01040	-1.21811202	-0.96145002	0.0746
7	At1_g01050	-0.66493581	0.126151787	1.17505
8	At1_g01060	0.809261597	1.566740565	0.57903

	A	B	C	D	S
1	(compexp)	Sample Ath A	Sample Ath B	Sample Ath B	
2	(rep)	1	1	2	
3	KPD00001	657	124	125	
4	KPC00002	9	9967	10274	
5	KPC00003	241184	864	864	
6	KPC00004	122	18	18	
7	KPC00005	5140	372	377	
8	KPC00006	463	51	51	

## Logging-in, and Uploading



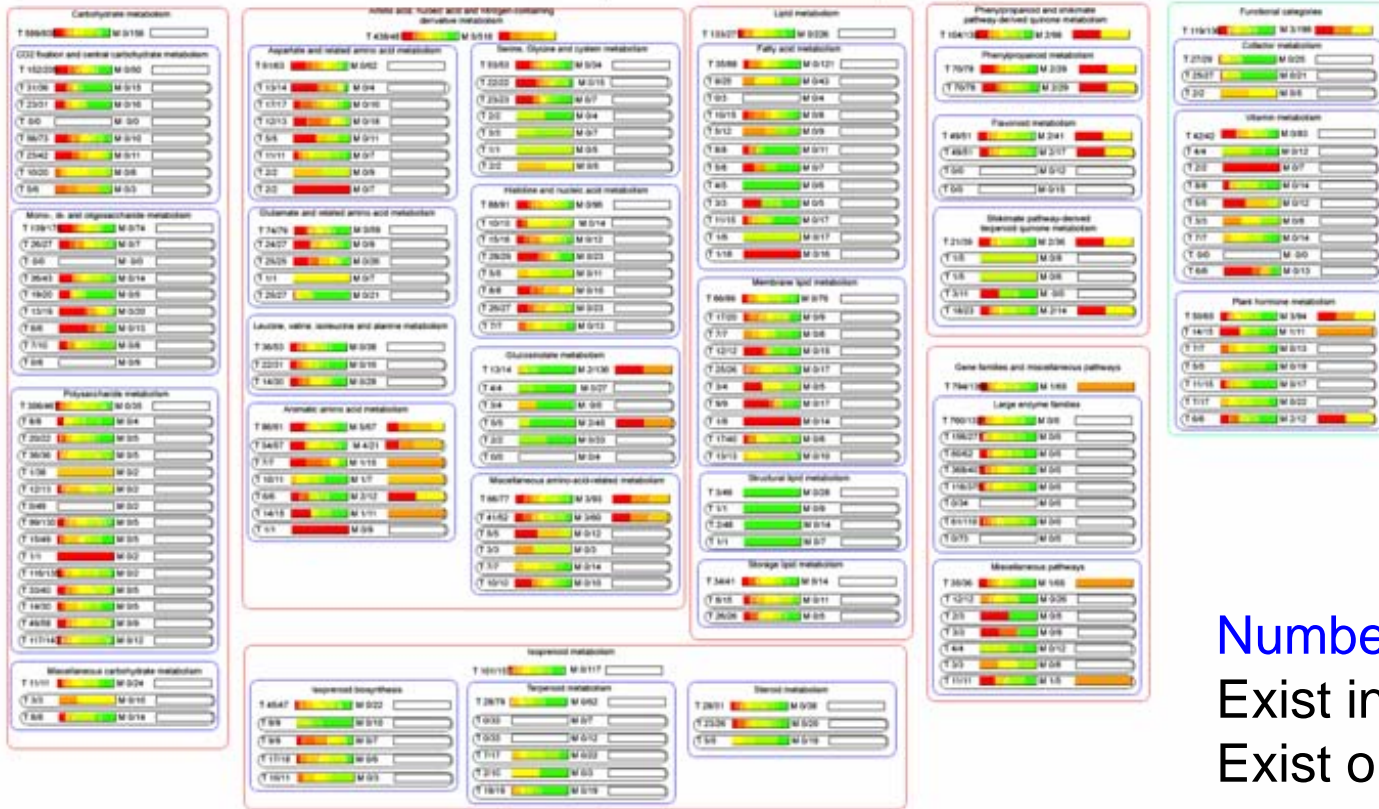
## Selecting data sets for representation



## Getting representations

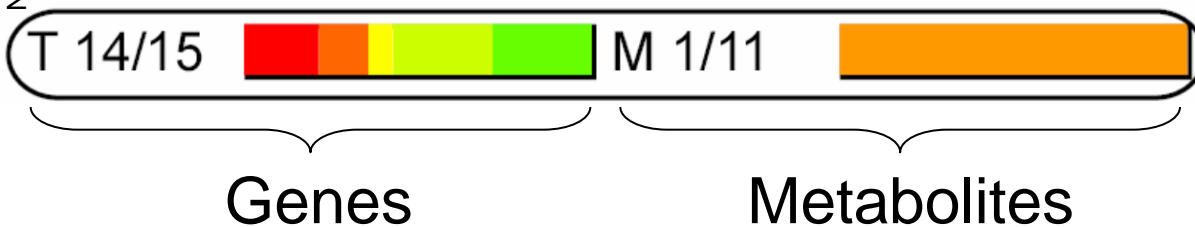
# Bird's Eye View

Arabidopsis thaliana metabolic pathway



Numbers:  
Exist in Input Data /  
Exist on the Map

Color indicator:  
Proportion of Genes or  
Metabolites painted with  
the color



# Outline

1. About KaPPA-View4

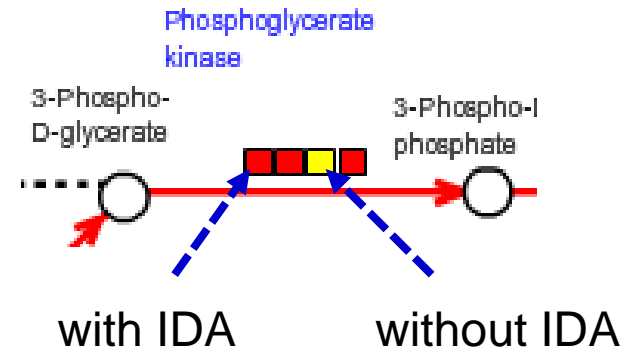
2. Advanced Data Analysis

3. Perspectives

# Non-omics data representation

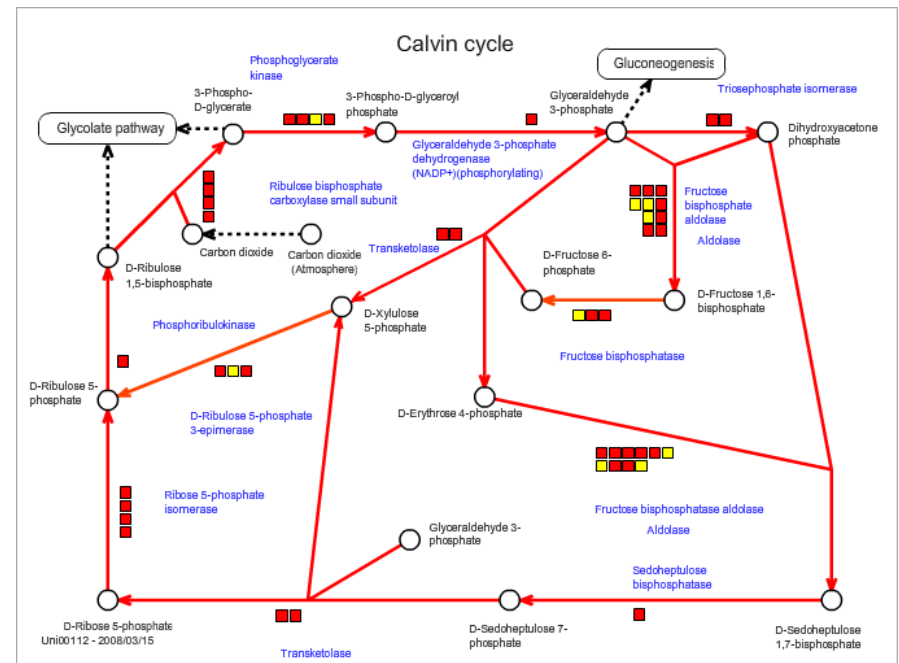
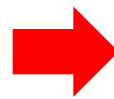
Experimental value of gene expression

	A	B	C
1	(arrayexp)	Sample Ath A	Sample
2	(rep)	1	
3	At1_g01010	-0.82847516	0.6468
4	At1_g01020	-1.32244998	-0.52
5	At1_g01030	1.118873492	1.41
6	At1_g01040	-1.21811202	-0.96
7	At1_g01050	-0.66493581	0.126
8	At1_g01060	0.809261597	1.566
9	At1_g01070	-1.20441321	-0.88
10	At1_g01080	0.904213049	1.059
11	At1_g01090	1.487257461	-0.32
12	At1_g01100	1.285759095	-0.0



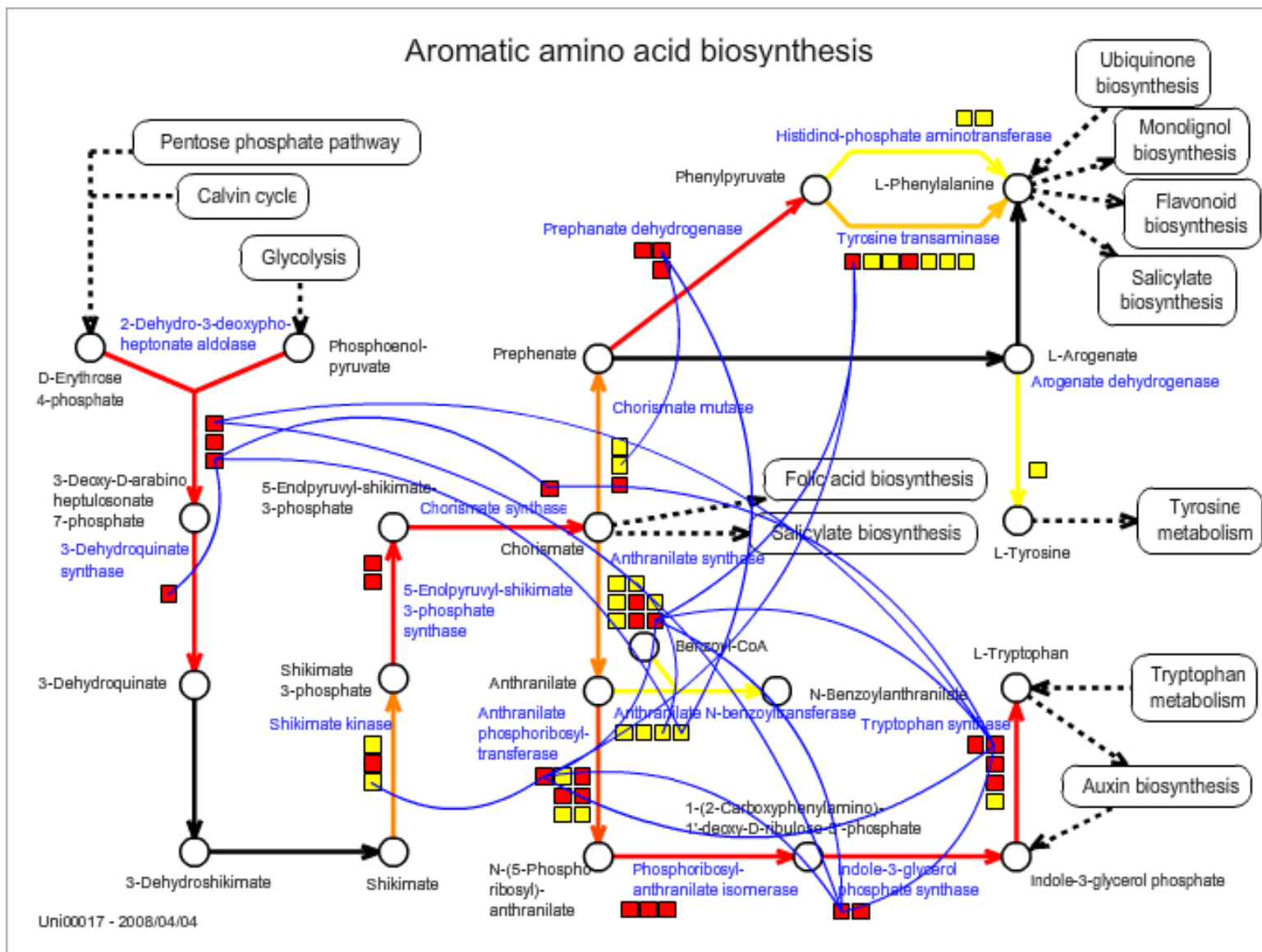
Ex) Attachment of an evidence code of Gene Ontology

	A	B	C	D
1	(arrayexp)	IDA: Inferred from Direct Assay		
2	(rep)	1		
3	At1_g01010	0		
4	At1_g01020	1.0		
5	At1_g01030	0		
6	At1_g01040	1.0		
7	At1_g01050	1.0		
8	At1_g01060	0		
9	At1_g01070	0		
10	At1_g01080	1.0		
11	At1_g01090	1.0		
12	At1_g01100	1.0		



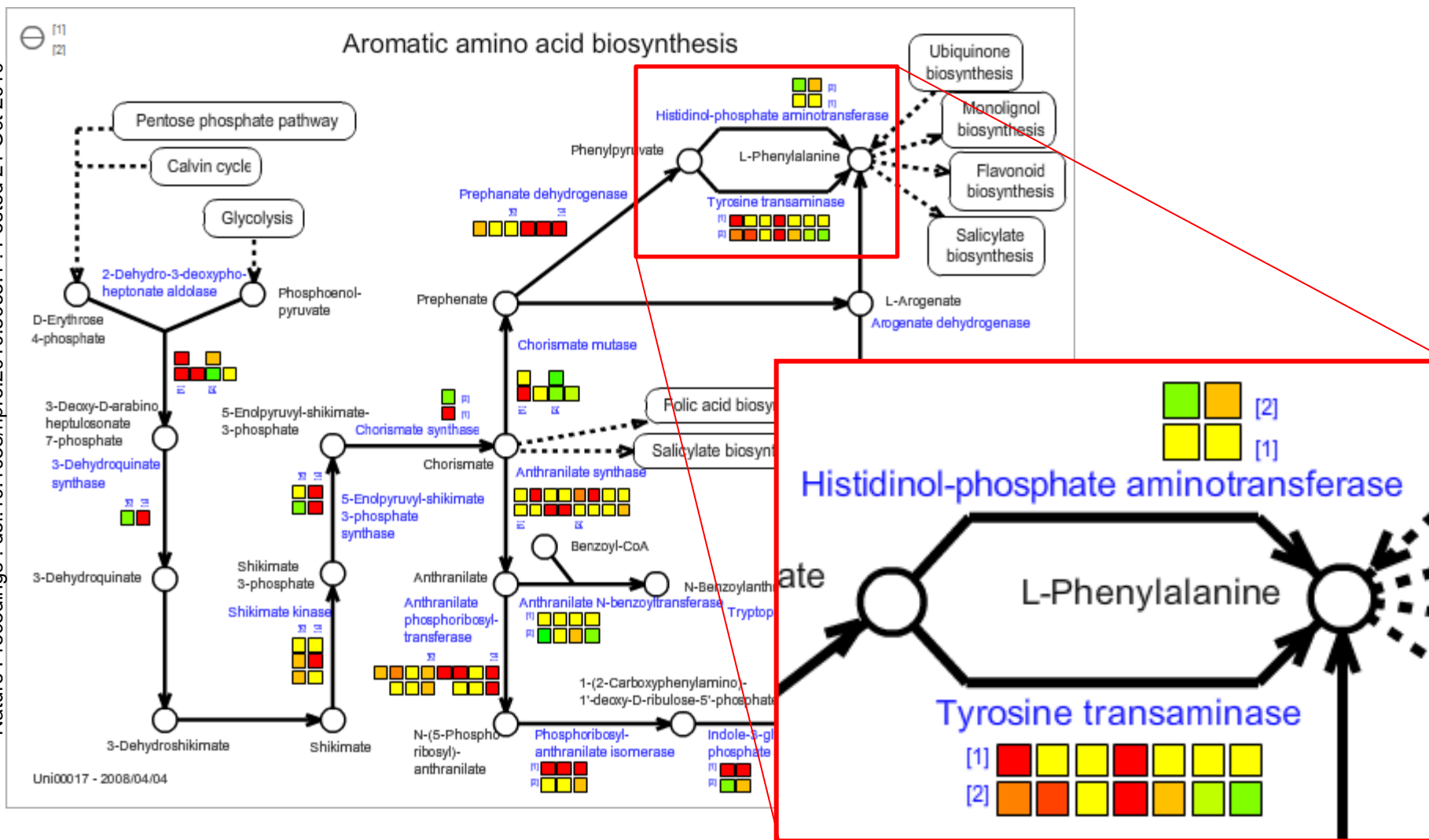
# Analysis with evidence code (1)

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010





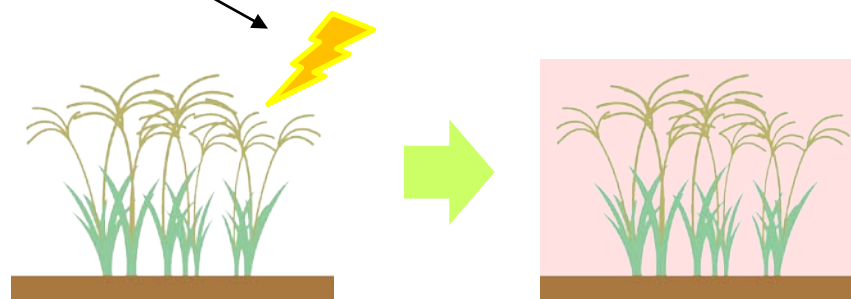
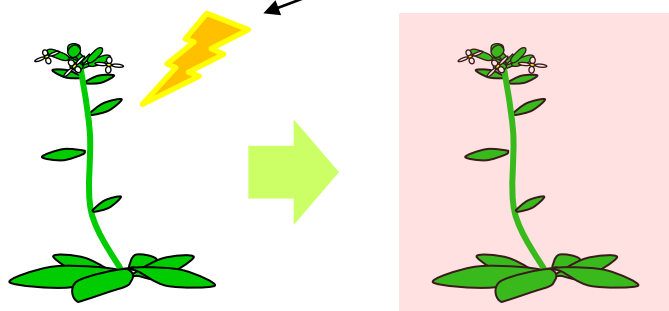
# Analysis with evidence code (2)



[1] with or without IDA, [2] Experimental Data

# Comparative Omics Analysis

Similar treatment



Arabidopsis DNA microarray

Rice DNA microarray

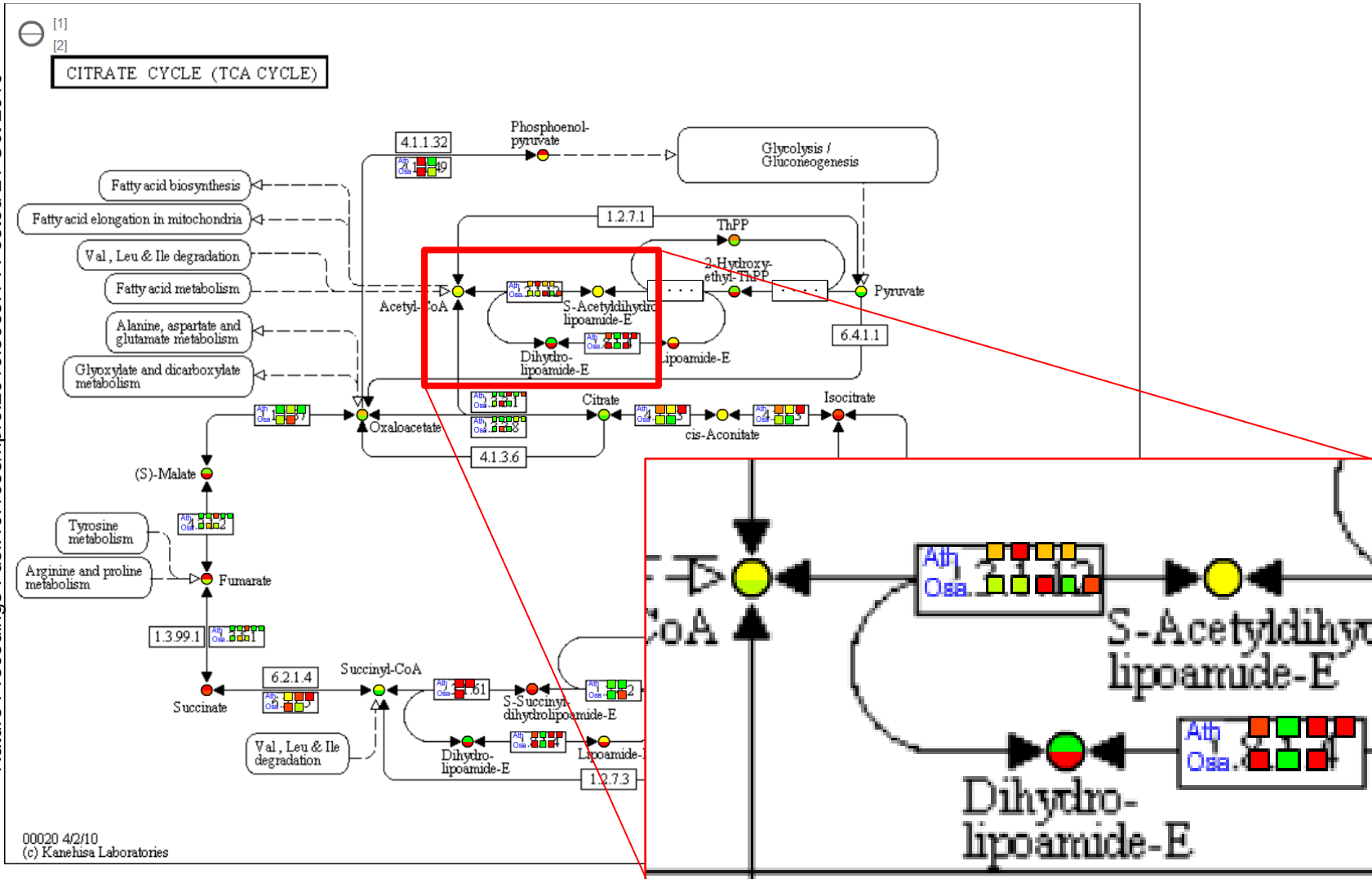
Data

Data

KaPPA - View 4  
Kazusa Plant Pathway Viewer

# Omics data comparison between the species

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010



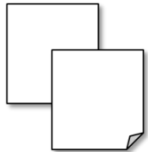
Ath: Arabidopsis, Osa: Rice

# Outline

1. About KaPPA-View4
2. Advanced Data Analysis
3. Perspectives

# Procedure for Omics data representation

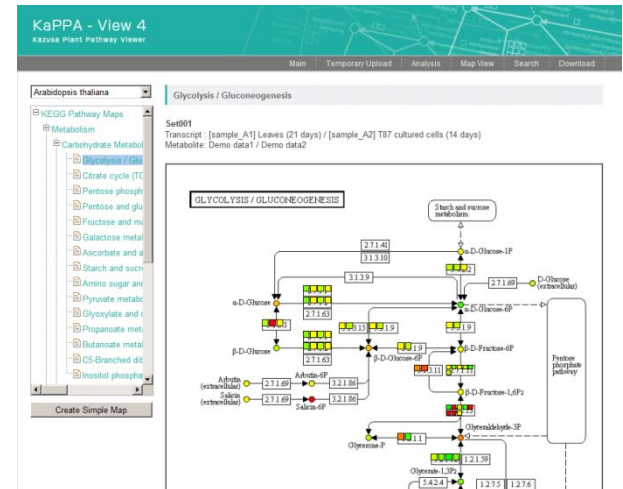
## Preparation of data files



1	A	B	C	D	E
1	(arrayexp)	Sample Ath A	Sample Ath B	Sample Ath C	Sample Ath C
2	(rep)	1	1	1	2
3	At1.g01.01	-0.82847516	0.646901298	-1.19048388	-1.07724172
4	At1.g01.020	-1.32244989	-0.62435662	-0.18684342	-0.22203793
5	At1.g01.030	1.118973492	1.41118969	-1.26662646	-1.29633266
6	At1.g01.040	-1.21811202	-0.98145002	0.07466906	0.11324818
7	At1.g01.050	-0.66493591	0.120151787	1.175051908	1.170206475
8	At1.g01.060	0.80201597	1.569740605	0.579037425	0.690907787
9	At1.g01.070	-1.20441521	-0.88953735	0.61000164	0.611701704
10	At1.g01.080	0.904213049	1.059304533	0.07728201	0.060123451
11	At1.g01.090	1.487257461	-0.32237127	0.615823494	0.518919148
12	At1.g01.100	1.2857959095	-0.0542007	-0.33025193	-0.39955302
13	At1.g01.110	-0.57385919	0.383083021	-1.00143081	-1.24032402
14	At1.g01.120	0.645001067	1.453814354	-0.82605022	-0.86852184
15	At1.g01.130	0.265570431	0.021234679	1.197526294	1.171799542
16	At1.g01.140	0.837828789	0.174417061	-0.06702859	-0.03127468
17	At1.g01.150	0.455129018	0.952882013	0.397943407	0.450393106

## Logging-in, and Uploading

## Selecting data sets for representation

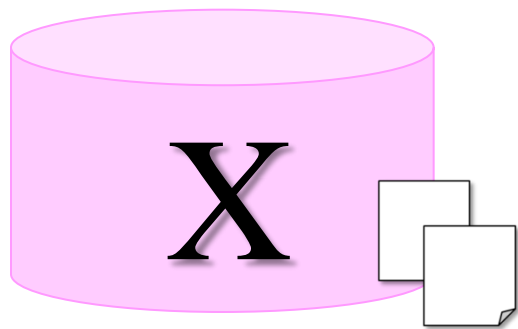


## Getting representations

# Post transferring system



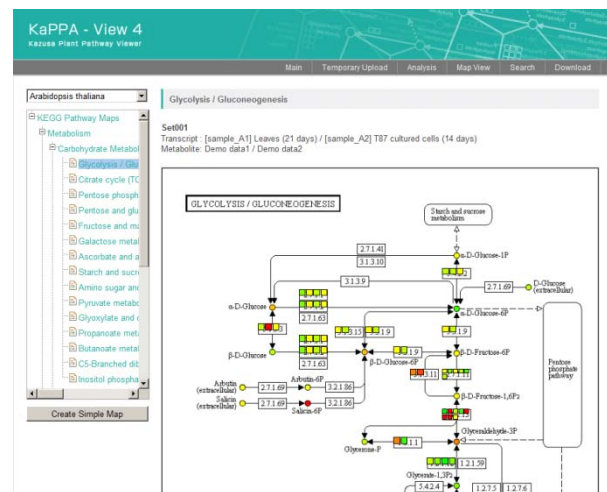
Data selection,  
and triggering "View on KaPPA-View" action



Database storing "Omics" data



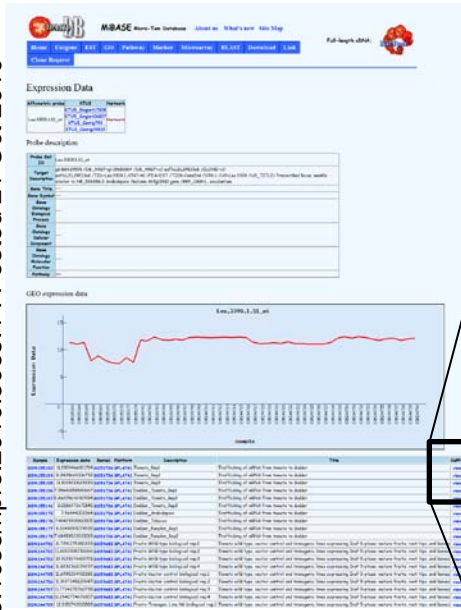
*Post Transfer*



Getting representations

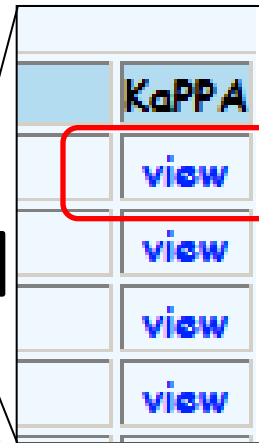
# Omics data browsing from outside systems

Nature Precedings : doi:10.1038/npre.2010.5065.1 : Posted 21 Oct 2010

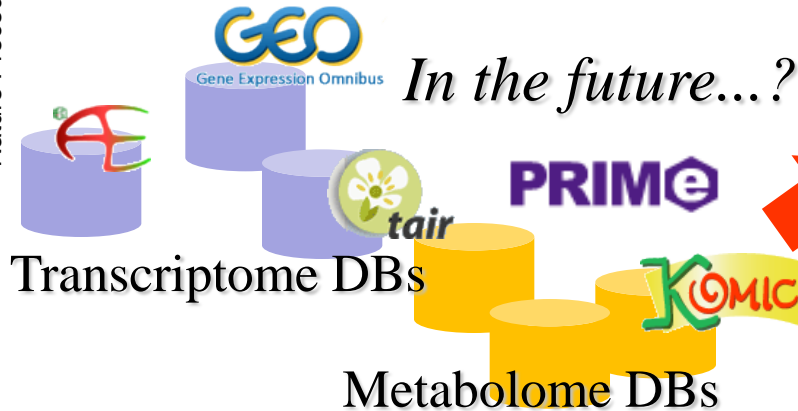
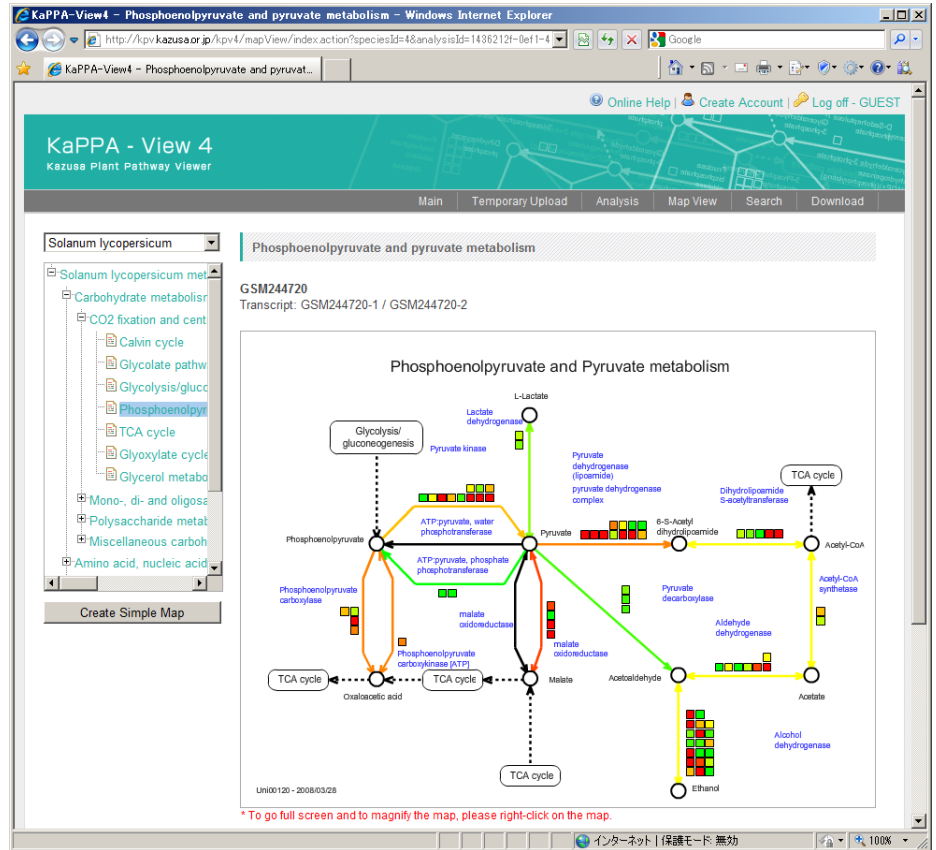


MiBASE (Micro-Tom Database)

<http://www.pgb.kazusa.or.jp/mibase/>

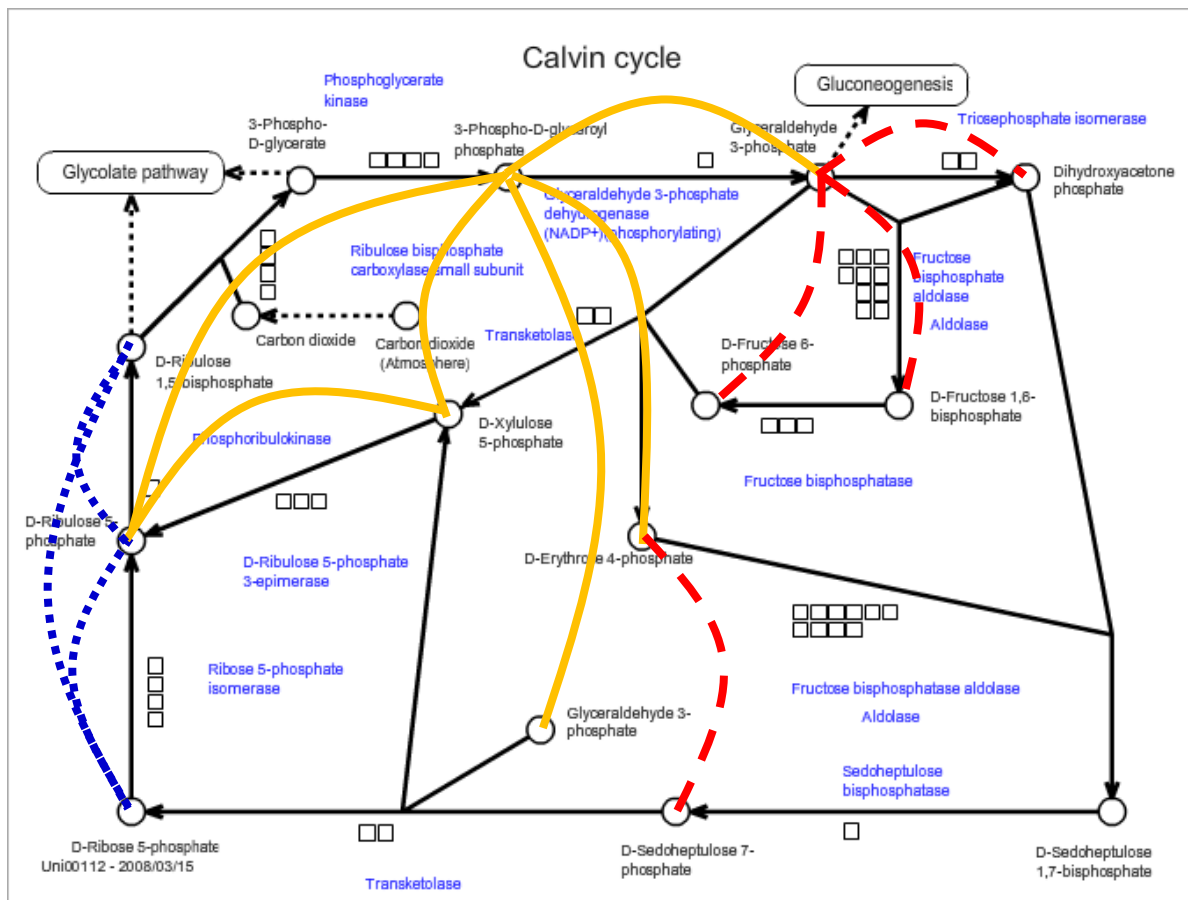


Expression Data View

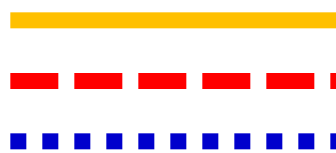


\* To go full screen and to magnify the map, please right-click on the map.

# Comparative Correlation Analysis



Comparison between  
Correlation Data A  
Correlation Data B



Common  
Data A specific  
Data B specific



# Summary

- KaPPA-View4 can represent gene co-expressions, metabolite co-accumulations, and omics data on the pathway maps.
- Non-omics data such as evidence codes of GO can be represented with correlation data and omics data.
- Omics data for different species can be represented on the same pathway map.

# Acknowledgement

Kazusa DNA Research Institute

Daisuke Shibata

Yoshiyuki Ogata

Takeshi Ara

Ryosuke Sano

Takashi Ohno

Kenjiro Sugiyama

Kiyoshi Yamazaki

Koh Aoki

Hideyuki Suzuki

Programming AXIOHELIX Co. Ltd.

New Energy and Industrial Technology Development Organization (NEDO, Japan)

- "Development of Fundamental Technologies for Controlling the Material Production Process of Plants" [P02001]
- "New Energy Technology Development" [P07015]

Kyoto Univ.

Minoru Kanehisa

Toshiaki Tokimatsu

RIKEN, Chiba Univ.

Kazuki Saito

Tohoku Univ.

Takeshi Obayashi

Meiji Univ.

Kentaro Yano