EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

MATHEMATISCH-
NATURWISSENSCHAFTLICHE FAKULTÄT

InCroMAP

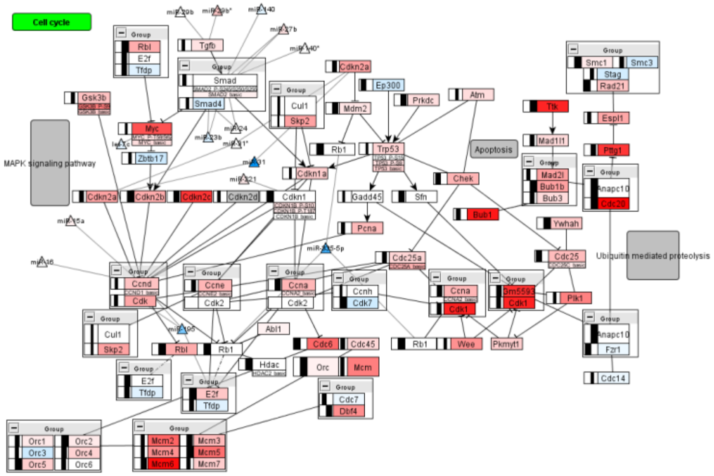
Integrated analysis of **Cro**ss-platform **M**icro**A**rray and **P**athway data

Johannes Eichner, Clemens Wrzodek, Lars Rosenbaum and Andreas Zell

October 30, 2013

Overview

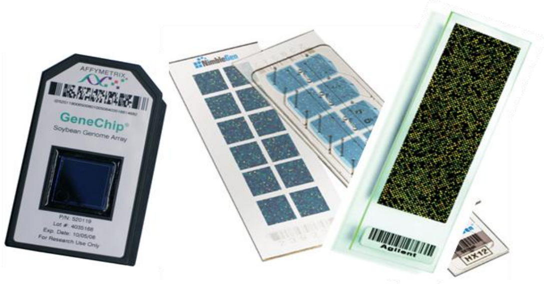
- Typical use case
- Supported platforms and required data preprocessing
- Analysis and visualization features of InCroMAP
  - Metabolic overview
  - Integrated visualization
- Pathway-based visualization
  - mRNA expression data
  - Protein abundance data
  - DNA methylation data
  - miRNA expression data
  - Metabolomics data



General use case of InCroMAP

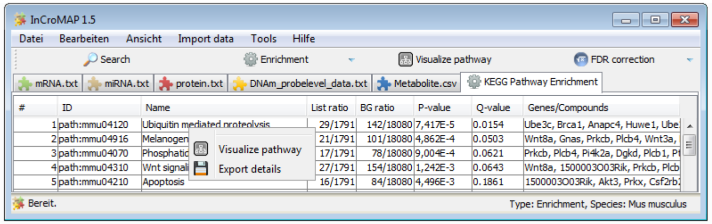
Import of processed data:

Omics data from heterogeneous platforms generated for the same set of samples



Pathway enrichment analysis:

Relevant pathways are identified based on a hypergeometric test



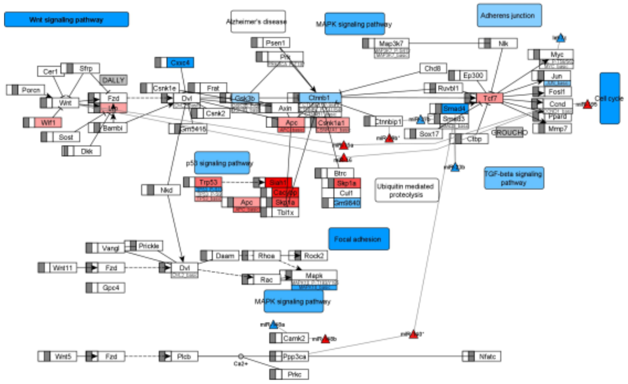
#	ID	Name	List ratio	BG ratio	P-value	Q-value	Genes/Compounds
1	path:mmu04120	Ubiquitin mediated proteolysis	29/1791	142/18080	7,417E-5	0.0154	Ube3c, Brca1, Anapc4, Huwe1, Ube
2	path:mmu04916	Melanogen	21/1791	101/18080	4,862E-4	0.0503	Wnt8a, Gnas, Prkcb, Plcb4, Wnt3a,
3	path:mmu04070	Phosphatic	17/1791	78/18080	9,004E-4	0.0621	Prkcb, Plcb4, Pi4k2a, Dgkd, Plcb1, Pt
4	path:mmu04310	Wnt signali	27/1791	154/18080	1,242E-3	0.0643	Wnt8a, 1500003O03Rik, Prkcb, Plcb
5	path:mmu04210	Apoptosis	16/1791	84/18080	4,496E-3	0.1861	1500003O03Rik, Akt3, Prkx, Csf2rb

Bereit.

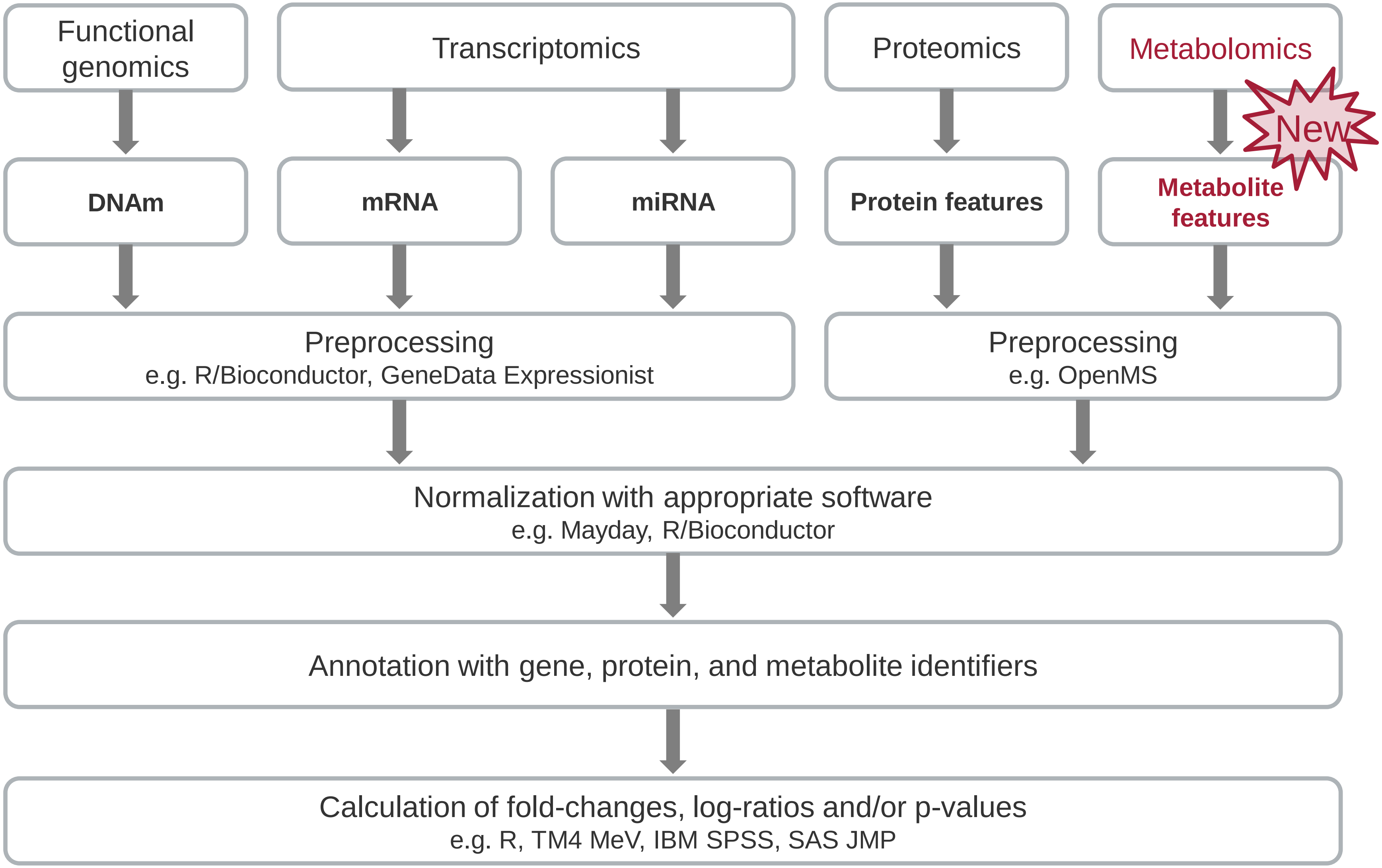
Type: Enrichment, Species: Mus musculus

Pathway visualization:

Multi-level omics data can be overlaid on specific pathways of interest



Supported platforms and required preprocessing steps



Functional genomics

Transcriptomics

Proteomics

Metabolomics

New

DNAm

mRNA

miRNA

Protein features

Metabolite features

Preprocessing
e.g. R/Bioconductor, GeneData Expressionist

Preprocessing
e.g. OpenMS

Normalization with appropriate software
e.g. Mayday, R/Bioconductor

Annotation with gene, protein, and metabolite identifiers

Calculation of fold-changes, log-ratios and/or p-values
e.g. R, TM4 MeV, IBM SPSS, SAS JMP

Omics data analysis with InCroMAP

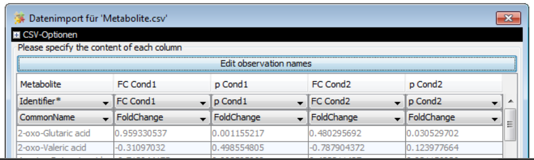
Data import

Enrichment analysis

Visualization

Run InCroMAP webstart application

Import data from heterogeneous sources


Datenimport für 'Metabolite.csv'

⊞ CSV-Optionen

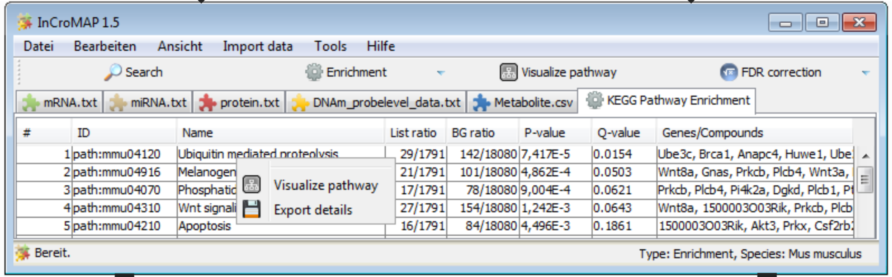
Please specify the content of each column

Edit observation names

Metabolite	FC Cond1	p Cond1	FC Cond2	p Cond2
Identifier*	FC Cond1	p Cond1	FC Cond2	p Cond2
CommonName	FoldChange	FoldChange	FoldChange	FoldChange
2-oxo-Glutaric acid	0.959330537	0.001155217	0.480295692	0.030529702
2-oxo-Valeric acid	-0.31097032	0.498554805	-0.787904372	0.123977664

Single platform enrichment

Integrated multi-omics enrichment

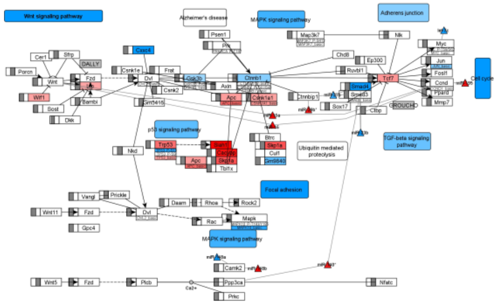


#	ID	Name	List ratio	BG ratio	P-value	Q-value	Genes/Compounds
1	path:mmu04120	Ubiquitin mediated proteolysis	29/1791	142/18080	7,417E-5	0.0154	Ube3c, Brca1, Anapc4, Huwe1, Ube
2	path:mmu04916	Melanogen	21/1791	101/18080	4,862E-4	0.0503	Wnt8a, Gnas, Prkcb, Plcb4, Wnt3a,
3	path:mmu04070	Phosphatic	17/1791	78/18080	9,004E-4	0.0621	Prkcb, Plcb4, Pi4k2a, Dgkd, Plcb1,
4	path:mmu04310	Wnt signali	27/1791	154/18080	1,242E-3	0.0643	Wnt8a, 1500003O03Rik, Prkcb, Plcb
5	path:mmu04210	Apoptosis	16/1791	84/18080	4,496E-3	0.1861	1500003O03Rik, Akt3, Prkx, Csf2rb

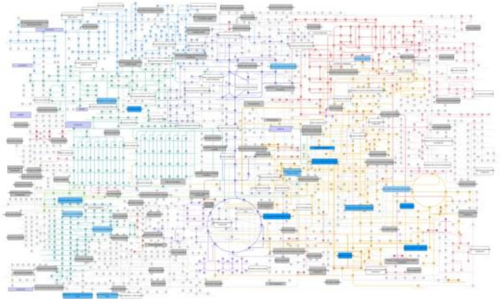
Bereit.

Type: Enrichment, Species: Mus musculus

Select pathway of interest for integrated visualization

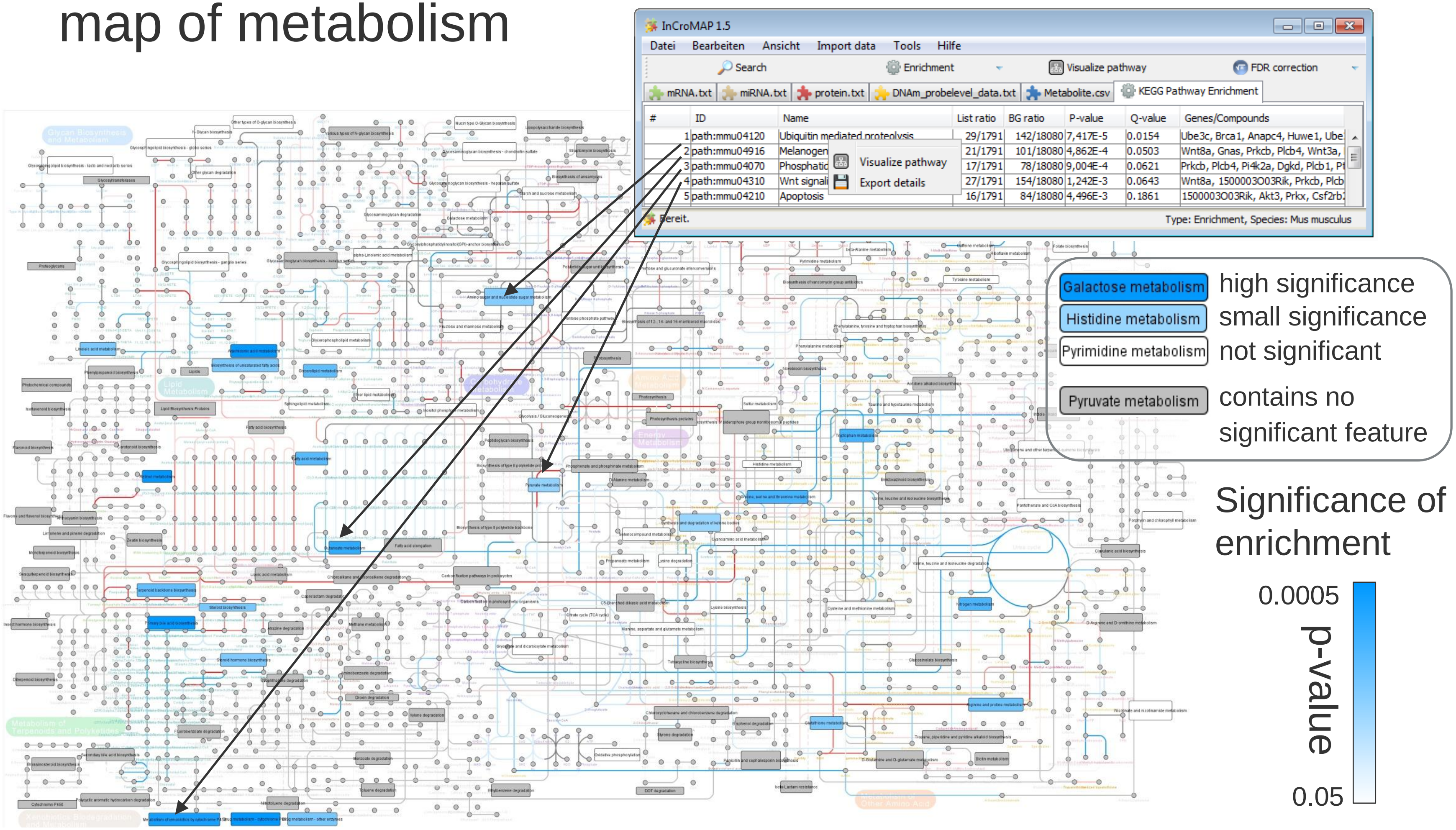


Integrated visualization of enrichment in metabolic overview



Metabolic overview visualization

- Visualization of pathway enrichment results in global map of metabolism



InCroMAP 1.5

#	ID	Name	List ratio	BG ratio	P-value	Q-value	Genes/Compounds
1	path:mmu04120	Ubiquitin mediated proteolysis	29/1791	142/18080	7,417E-5	0.0154	Ube3c, Brca1, Anapc4, Huwe1, Ube
2	path:mmu04916	Melanogen	21/1791	101/18080	4,862E-4	0.0503	Wnt8a, Gnas, Prkcb, Plcb4, Wnt3a,
3	path:mmu04070	Phosphatid	17/1791	78/18080	9,004E-4	0.0621	Prkcb, Plcb4, Pi4k2a, Dgkd, Plcb1, P
4	path:mmu04310	Wnt signali	27/1791	154/18080	1,242E-3	0.0643	Wnt8a, 1500003O03Rik, Prkcb, Plcb
5	path:mmu04210	Apoptosis	16/1791	84/18080	4,496E-3	0.1861	1500003O03Rik, Akt3, Prkx, Csf2rb

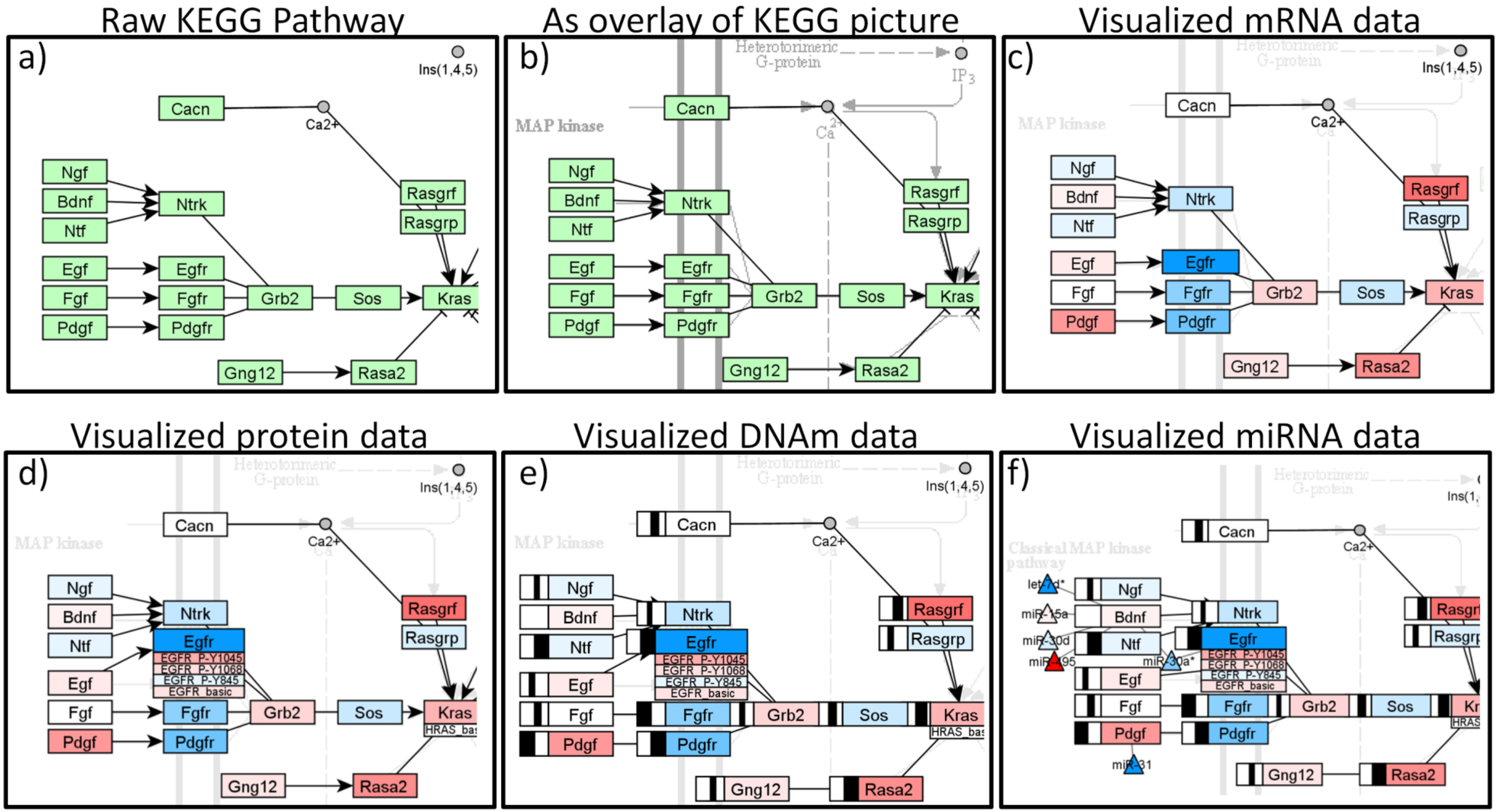
Bereit.

Type: Enrichment, Species: Mus musculus

Galactose metabolism	high significance
Histidine metabolism	small significance
Pyrimidine metabolism	not significant
Pyruvate metabolism	contains no significant feature

Significance of enrichment

Integrated pathway-based visualization


Raw KEGG Pathway

As overlay of KEGG picture

Visualized mRNA data

Visualized protein data

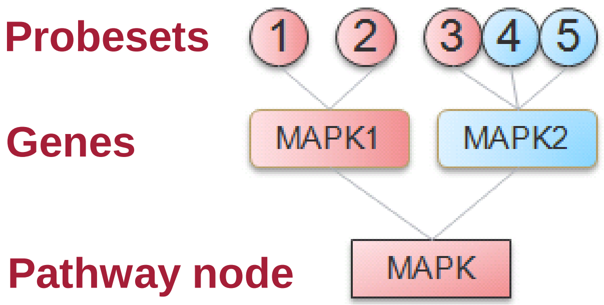
Visualized DNAm data

Visualized miRNA data

http://www.cogsys.cs.uni-tuebingen.de/software/InCroMAP/

Visualization of mRNA data

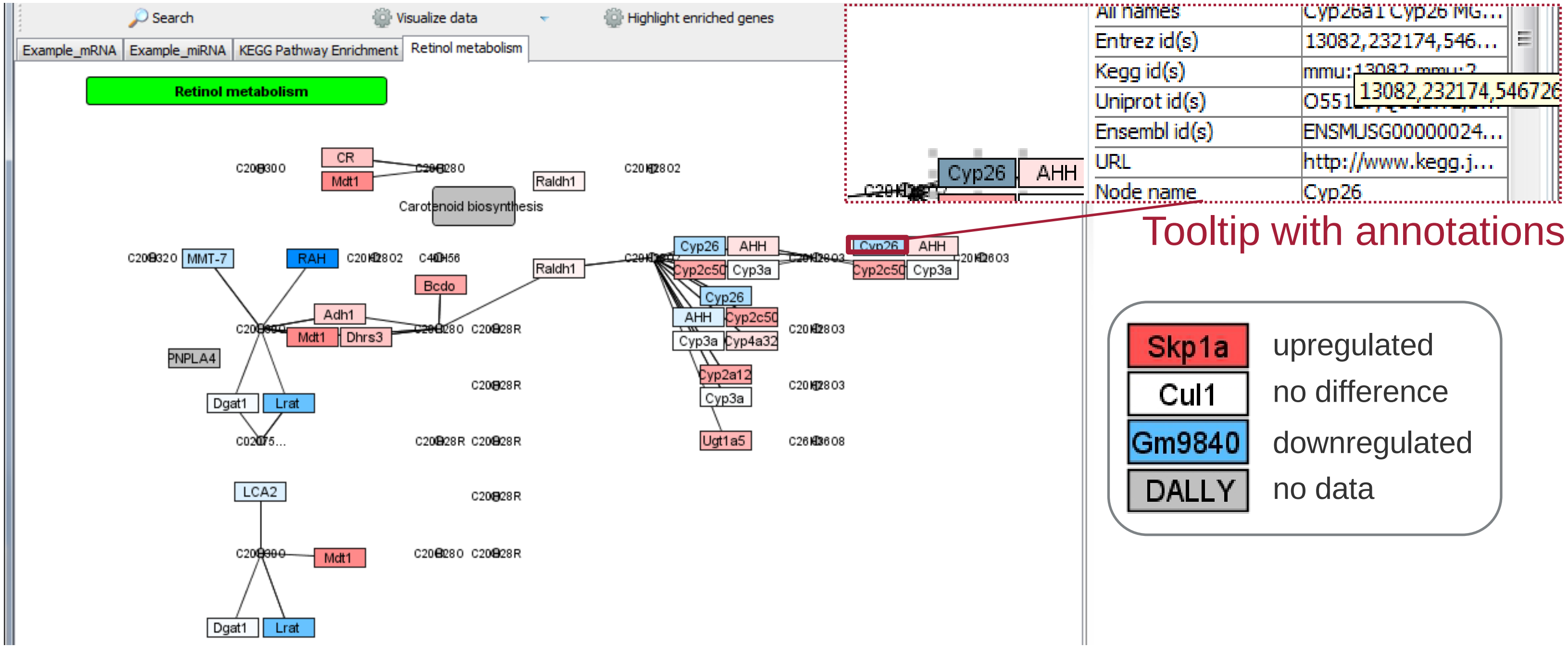
- Import of normalized data (e.g., fold-changes, p-values) in tabular format
- Summarization of values from multiple probesets and genes



Probesets 1 2 3 4 5

Genes MAPK1 MAPK2

Pathway node MAPK



All names	Cyp26a1 Cyp26 MG...
Entrez id(s)	13082,232174,546...
Kegg id(s)	mmu:13082 mmu:2
Uniprot id(s)	O551
Ensembl id(s)	ENSMUSG00000024...
URL	http://www.kegg.j...
Node name	Cyp26

13082,232174,546726

Tooltip with annotations

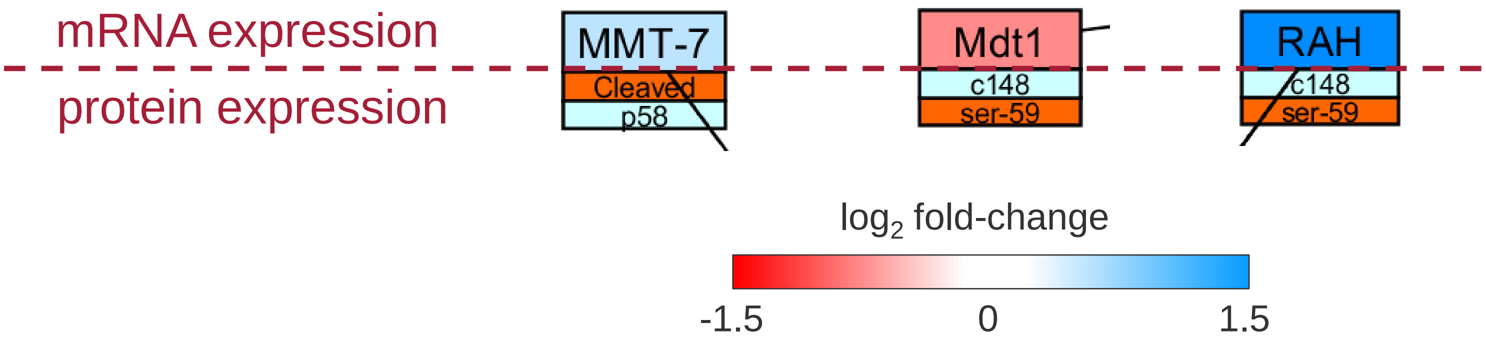
Skp1a	upregulated
Cul1	no difference
Gm9840	downregulated
DALLY	no data

Protein modification expression data

- Data has to be provided in special format

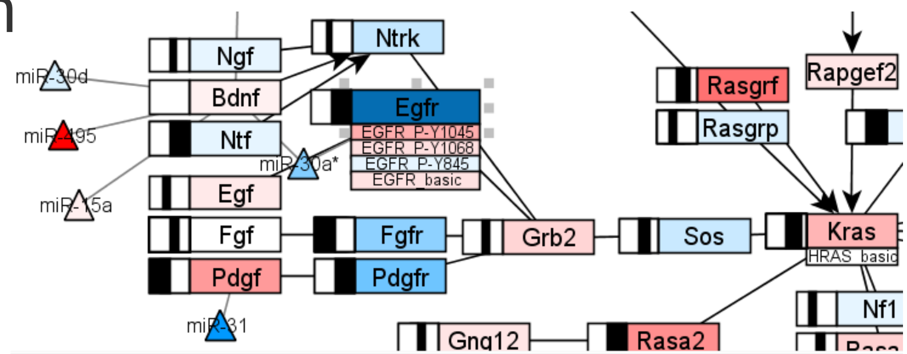
Analyte Short Name	HGNC Symbol	Modification	Treatment 1 (fold-change)	Treatment 2 (fold-change)	Treatment 3 (fold-change)
MMT-7	n/a	cleaved	0.044	1.544	-0.220
MMT-7	n/a	phospho-Ser58	-0.209	-0.509	-0.124
Mdt1	RDH11	cleaved-148	-0.151	-0.451	-0.199
Mdt1	RDH11	phospho-Ser59	-0.1308	1.130	-0.085
RAH	RAB34	Cleaved-148	-0.155	-0.455	-0.197
RAH	RAB34	phospho-Ser59	-0.287	2.287	0.006

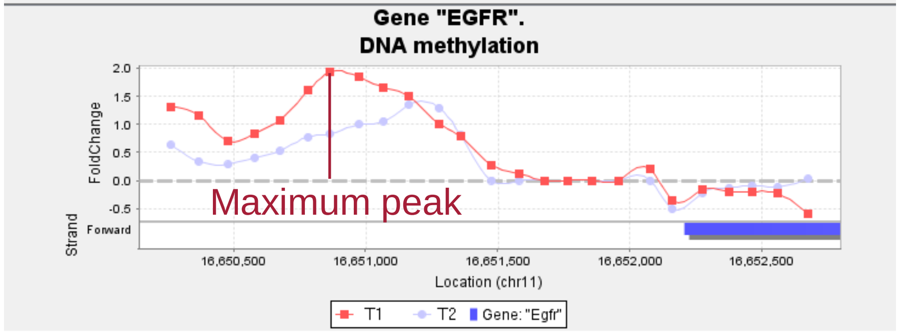
- Proteins and their modifications are added as small colored boxes below gene nodes



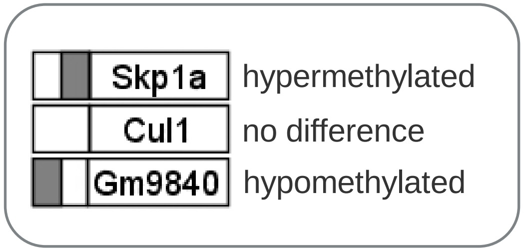
Visualization of DNAm data

- Consider probes in region between -2000bp and +500bp relative to TSS
- Compute summary value based on maximum peak

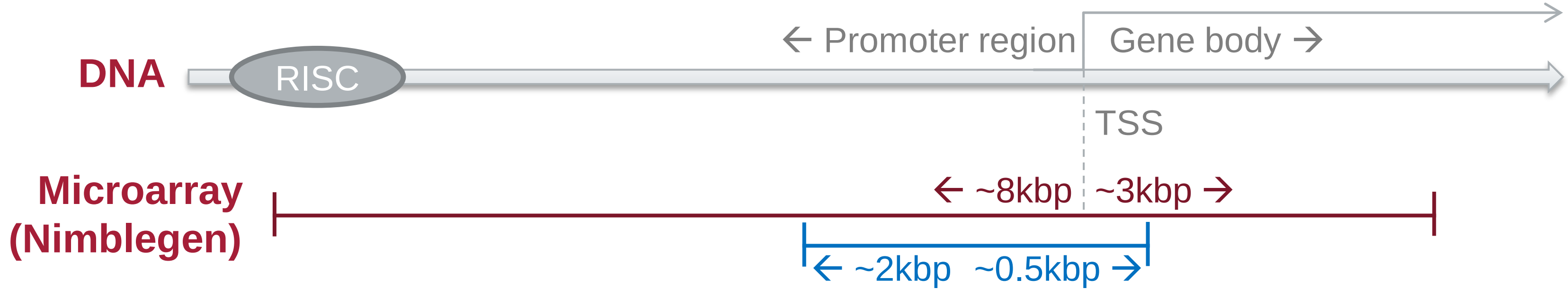




Maximum peak

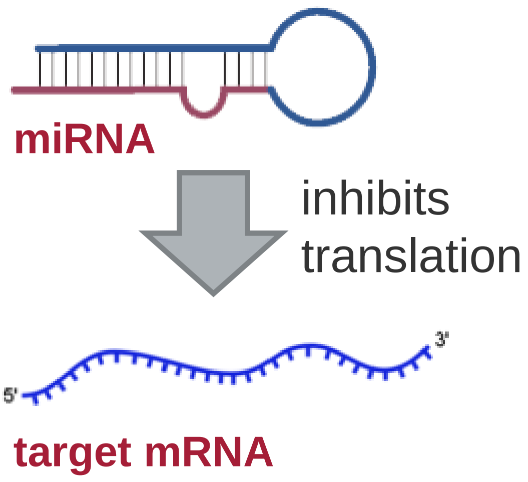


Skp1a	hypermethylated
Cul1	no difference
Gm9840	hypomethylated



Visualization of miRNA data

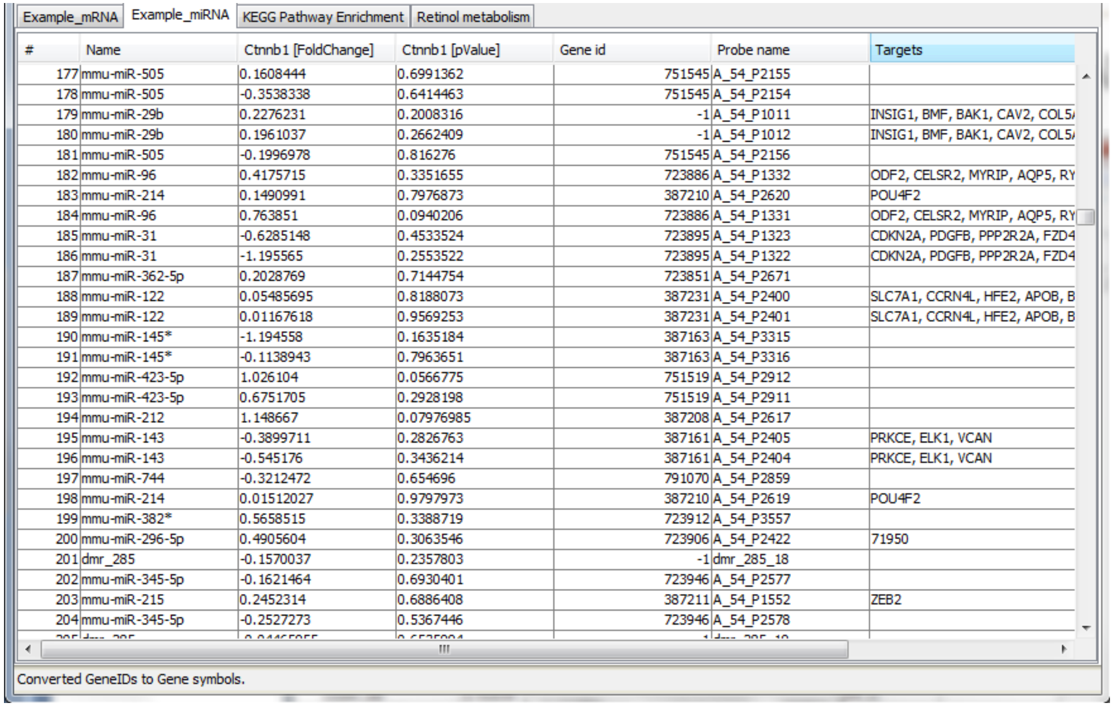
- Import of normalized data (e.g., fold-changes, p-values) into InCroMAP
- Mapping of miRNAs to experimentally confirmed and predicted mRNA targets



miRNA

inhibits translation

target mRNA

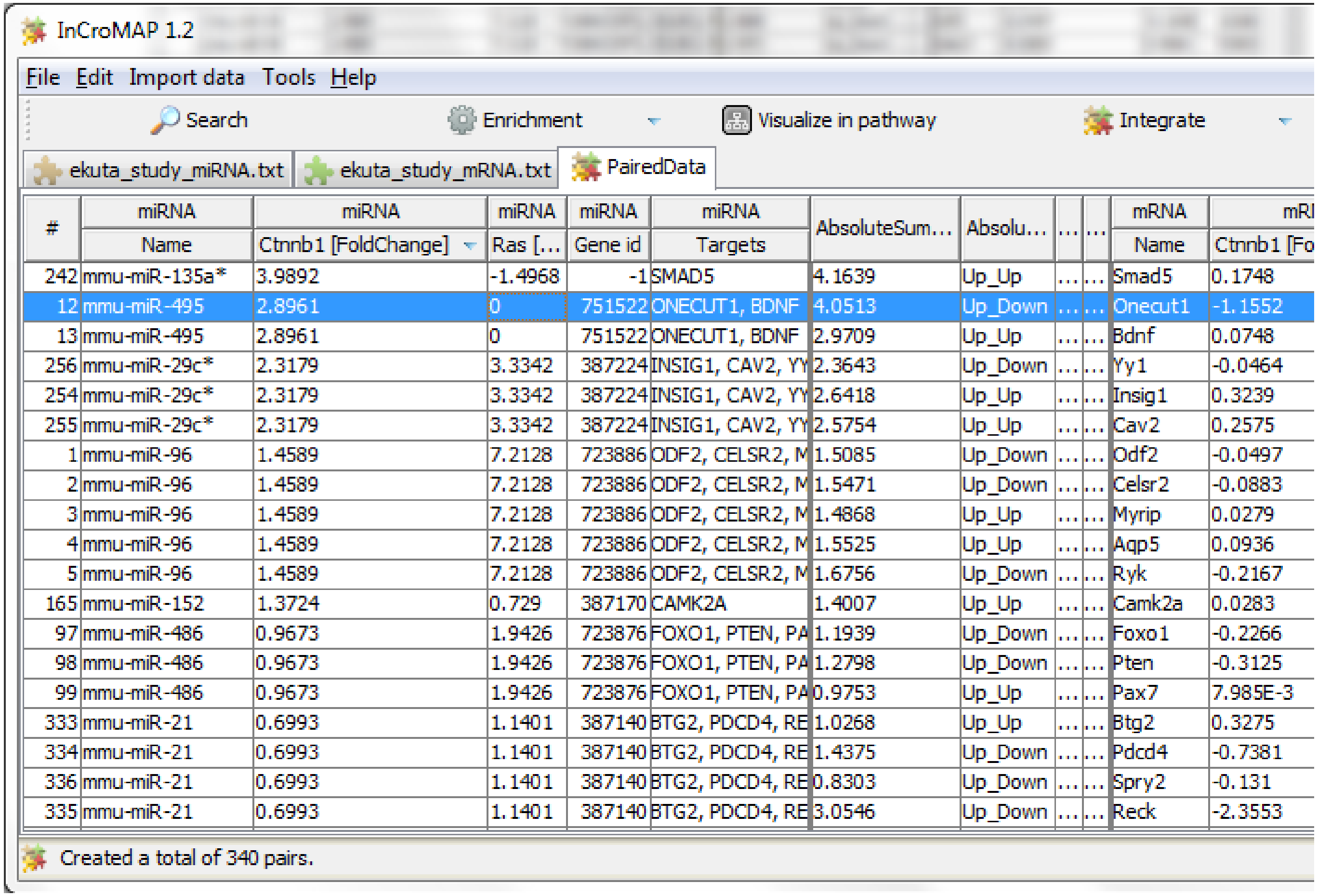


#	Name	Ctnnb1 [FoldChange]	Ctnnb1 [pValue]	Gene id	Probe name	Targets
177	mmu-miR-505	0.1608444	0.6991362	751545	A_54_P2155	
178	mmu-miR-505	-0.3538338	0.6414463	751545	A_54_P2154	
179	mmu-miR-29b	0.2276231	0.2008316	-1	A_54_P1011	INSIG1, BMF, BAK1, CAV2, COL5
180	mmu-miR-29b	0.1961037	0.2662409	-1	A_54_P1012	INSIG1, BMF, BAK1, CAV2, COL5
181	mmu-miR-505	-0.1996978	0.816276	751545	A_54_P2156	
182	mmu-miR-96	0.4175715	0.3351655	723886	A_54_P1332	ODF2, CELSR2, MYRIP, AQP5, RY
183	mmu-miR-214	0.1490991	0.7976873	387210	A_54_P2620	POU4F2
184	mmu-miR-96	0.763851	0.0940206	723886	A_54_P1331	ODF2, CELSR2, MYRIP, AQP5, RY
185	mmu-miR-31	-0.6285148	0.4533524	723895	A_54_P1323	CDKN2A, PDGFB, PPP2R2A, FZD4
186	mmu-miR-31	-1.195565	0.2553522	723895	A_54_P1322	CDKN2A, PDGFB, PPP2R2A, FZD4
187	mmu-miR-362-5p	0.2028769	0.7144754	723851	A_54_P2671	
188	mmu-miR-122	0.05485695	0.8188073	387231	A_54_P2400	SLC7A1, CCRN4L, HFE2, APOB, B
189	mmu-miR-122	0.01167618	0.9569253	387231	A_54_P2401	SLC7A1, CCRN4L, HFE2, APOB, B
190	mmu-miR-145*	-1.194558	0.1635184	387163	A_54_P3315	
191	mmu-miR-145*	-0.1138943	0.7963651	387163	A_54_P3316	
192	mmu-miR-423-5p	1.026104	0.0566775	751519	A_54_P2912	
193	mmu-miR-423-5p	0.6751705	0.2928198	751519	A_54_P2911	
194	mmu-miR-212	1.148667	0.07976985	387208	A_54_P2617	
195	mmu-miR-143	-0.3899711	0.2826763	387161	A_54_P2405	PRKCE, ELK1, VCAN
196	mmu-miR-143	-0.545176	0.3436214	387161	A_54_P2404	PRKCE, ELK1, VCAN
197	mmu-miR-744	-0.3212472	0.654696	791070	A_54_P2859	
198	mmu-miR-214	0.01512027	0.9797973	387210	A_54_P2619	POU4F2
199	mmu-miR-382*	0.5658515	0.3388719	723912	A_54_P3557	
200	mmu-miR-296-5p	0.4905604	0.3063546	723906	A_54_P2422	71950
201	dmr_285	-0.1570037	0.2357803	-1	dmr_285_18	
202	mmu-miR-345-5p	-0.1621464	0.6930401	723946	A_54_P2577	
203	mmu-miR-215	0.2452314	0.6886408	387211	A_54_P1552	ZEB2
204	mmu-miR-345-5p	-0.2527273	0.5367446	723946	A_54_P2578	

Converted GeneIDs to Gene symbols.

Pairing of miRNA and mRNA data

- Detection of putative miRNA-mRNA interactions

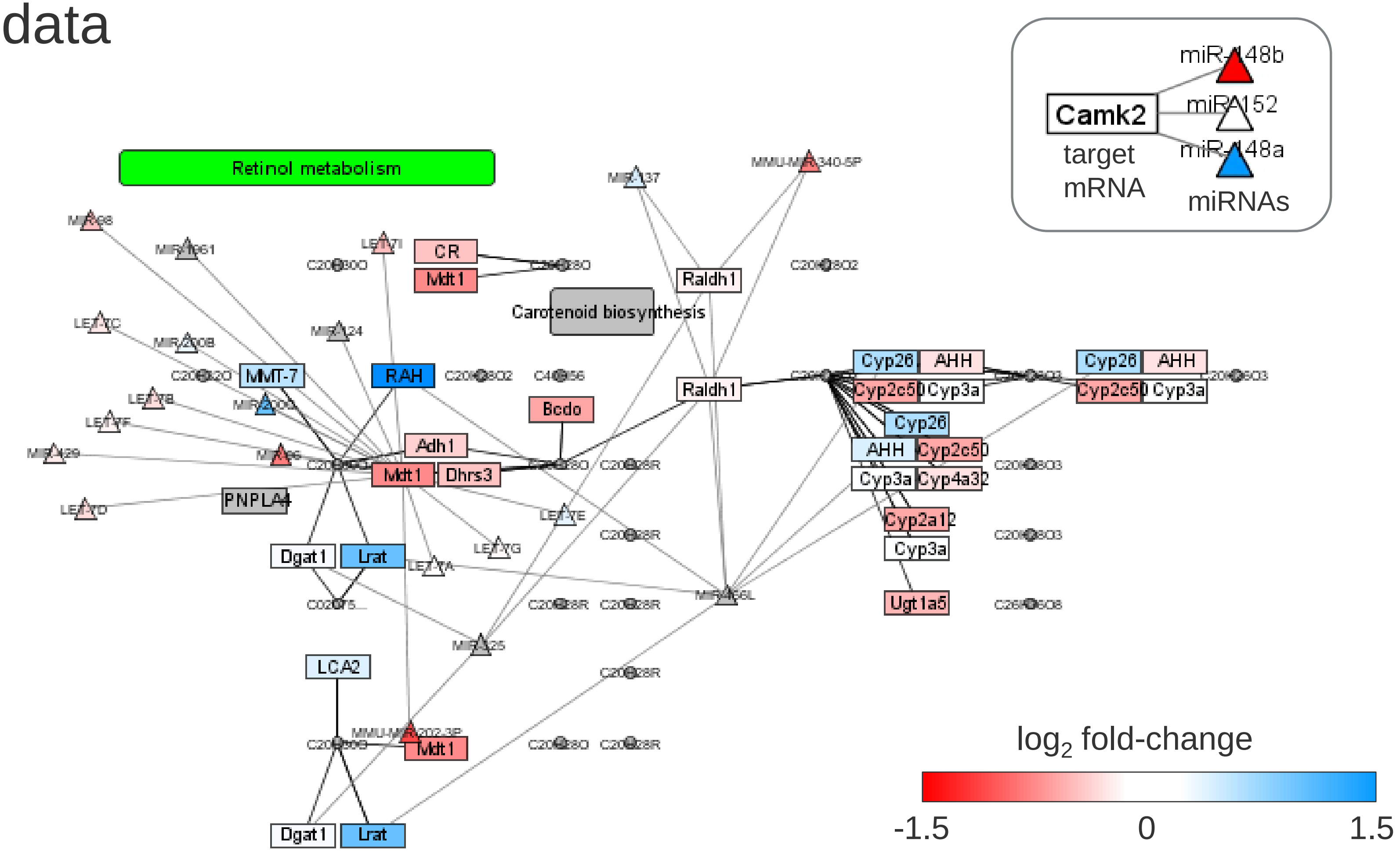


#	miRNA	miRNA	miRNA	miRNA	miRNA	AbsoluteSum...	Absolu...	...	...	mRNA	mRI
	Name	Ctnnb1 [FoldChange]	Ras [...	Gene id	Targets					Name	Ctnnb1 [Fo
242	mmu-miR-135a*	3.9892	-1.4968	-1	SMAD5	4.1639	Up_Up	...	...	Smad5	0.1748
12	mmu-miR-495	2.8961	0	751522	ONECUT1, BDNF	4.0513	Up_Down	...	...	Onecut1	-1.1552
13	mmu-miR-495	2.8961	0	751522	ONECUT1, BDNF	2.9709	Up_Up	...	...	Bdnf	0.0748
256	mmu-miR-29c*	2.3179	3.3342	387224	INSIG1, CAV2, YY	2.3643	Up_Down	...	...	Yy1	-0.0464
254	mmu-miR-29c*	2.3179	3.3342	387224	INSIG1, CAV2, YY	2.6418	Up_Up	...	...	Insig1	0.3239
255	mmu-miR-29c*	2.3179	3.3342	387224	INSIG1, CAV2, YY	2.5754	Up_Up	...	...	Cav2	0.2575
1	mmu-miR-96	1.4589	7.2128	723886	ODF2, CELSR2, M	1.5085	Up_Down	...	...	Odf2	-0.0497
2	mmu-miR-96	1.4589	7.2128	723886	ODF2, CELSR2, M	1.5471	Up_Down	...	...	Celsr2	-0.0883
3	mmu-miR-96	1.4589	7.2128	723886	ODF2, CELSR2, M	1.4868	Up_Up	...	...	Myrip	0.0279
4	mmu-miR-96	1.4589	7.2128	723886	ODF2, CELSR2, M	1.5525	Up_Up	...	...	Aqp5	0.0936
5	mmu-miR-96	1.4589	7.2128	723886	ODF2, CELSR2, M	1.6756	Up_Down	...	...	Ryk	-0.2167
165	mmu-miR-152	1.3724	0.729	387170	CAMK2A	1.4007	Up_Up	...	...	Camk2a	0.0283
97	mmu-miR-486	0.9673	1.9426	723876	FOXO1, PTEN, PA	1.1939	Up_Down	...	...	Foxo1	-0.2266
98	mmu-miR-486	0.9673	1.9426	723876	FOXO1, PTEN, PA	1.2798	Up_Down	...	...	Pten	-0.3125
99	mmu-miR-486	0.9673	1.9426	723876	FOXO1, PTEN, PA	0.9753	Up_Up	...	...	Pax7	7.985E-3
333	mmu-miR-21	0.6993	1.1401	387140	BTG2, PDCD4, RE	1.0268	Up_Up	...	...	Btg2	0.3275
334	mmu-miR-21	0.6993	1.1401	387140	BTG2, PDCD4, RE	1.4375	Up_Down	...	...	Pdcd4	-0.7381
336	mmu-miR-21	0.6993	1.1401	387140	BTG2, PDCD4, RE	0.8303	Up_Down	...	...	Spry2	-0.131
335	mmu-miR-21	0.6993	1.1401	387140	BTG2, PDCD4, RE	3.0546	Up_Down	...	...	Reck	-2.3553

Created a total of 340 pairs.

Visualization of miRNA data

- Color nodes according to expression as done for mRNA data



Visualization of metabolomics data

- Small circles are colored according to measured metabolite concentrations

New

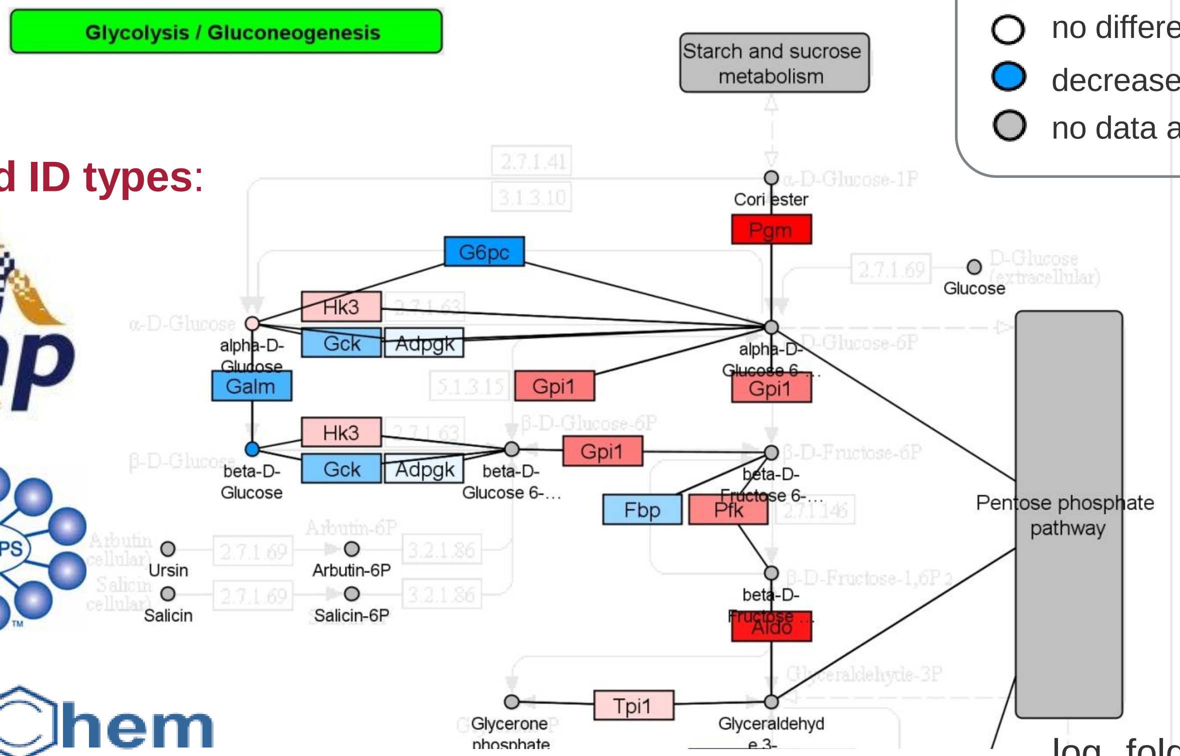
Supported ID types:

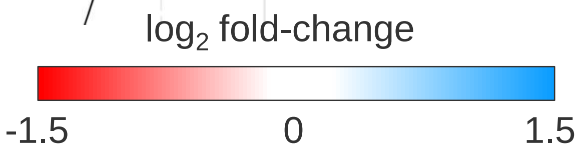
hmp
human metabolome project

LIPID MAPS

PubChem

🔴 increased abundance
○ no difference
🔵 decreased abundance
● no data available





Summary

- InCroMAP enables the visualization of cross-omics datasets in a pathway-centered manner
- Metabolic alterations can be displayed in an interactive global pathway map
- Linkage of miRNA transcripts to pathways based on confirmed/predicted interactions
- Consistent graphical representations are used to visualize data from specific biological layers
- New in InCroMAP version 1.6
  - Enhanced support for metabolomics data
  - Uses KEGG data from 2013

Thanks for your attention.

Contact:

Johannes Eichner
University of Tübingen
Wilhelm-Schickard-Institute for Computer Science
Dept. Cognitive Systems

Sand 1, 72076 Tübingen
Telefon: +49 7071 29-78970
E-Mail: johannes.eichner@uni-tuebingen.de