

Integration of novel data with previous biological knowledge

- Biological knowledge
- Integration methods
- Results visualization



Biological knowledge

bio::mart



intOgen



GO
the Gene Ontology



MSigDB
Molecular Signatures
Database

Biomart databases

[HOME](#)[MARTVIEW](#)[MARTSERVICE](#)[DOCS](#)[CONTACT](#)[NEWS](#)[CREDITS](#)

BioMart Project

BioMart is a query-oriented data management system developed jointly by the [Ontario Institute for Cancer Research \(OICR\)](#) and the [European Bioinformatics Institute \(EBI\)](#).

The system can be used with any type of data and is particularly suited for providing 'data mining' like searches of complex descriptive data. BioMart comes with an 'out of the box' website that can be installed, configured and customised according to user requirements. Further access is provided by graphical and text based applications or programmatically using web services or API written in Perl and Java. BioMart has built-in support for query optimisation and data federation and in addition can be configured to work as a DAS 1.5 Annotation server. The process of converting a data source into BioMart format is fully automated by the tools included in the package. Currently supported RDBMS platforms are MySQL, Oracle and Postgres.

BioMart is completely Open Source, licensed under the LGPL, and freely available to anyone without restrictions.

Powered by BioMart software:

- [BioMart Central Portal](#)
- [ICGC Data Portal](#)
- [Ensembl](#)
- [Ensembl Bacteria](#)
- [Ensembl Metazoa](#)
- [Ensembl Protists](#)
- [Ensembl Plants](#)
- [Ensembl Fungi](#)
- [Phytozome](#)
- [Gramene](#)
- [Europhenome](#)
- [UniProt](#)
- [InterPro](#)
- [HGNC](#)
- [Rice-Map](#)
- [IKMC](#)
- [Wormbase](#)
- [DroSpeGe](#)
- [ArrayExpress DW](#)
- [Eurexpress](#)
- [HapMap](#)
- [Dictybase](#)
- [COSMIC](#)
- [IntOGen](#)
- [Rat Genome Database](#)
- [GermOnLine](#)
- [PRIDE](#)
- [PepSeeker](#)
- [VectorBase](#)
- [HTGT](#)
- [Clddb](#)
- [European Mouse Mutant Archive](#)
- [Pancreatic Expression Database](#)
- [Reactome](#)
- [EU Rat Mart](#)
- [Paramecium DB](#)
- [International Potato Center \(CIP\)](#)
- [Mouse Genome Informatics \(MGI\)](#)
- [CyanoBase](#)

Third party software with BioMart Plugin:

[Bioclipse](#) [biomaRt-BioConductor](#) [Cytoscape](#) [Galaxy](#) [Gtools](#) [Ruby API](#) [Taverna](#) [WebLab](#)

Biomart databases

[HOME](#)[MARTVIEW](#)[MARTSERVICE](#)[DOCS](#)[CONTACT](#)[NEWS](#)[CREDITS](#)

BioMart Project

BioMart is a query-oriented data management system developed jointly by the [Ontario Institute for Cancer Research \(OICR\)](#) and the [European Bioinformatics Institute \(EBI\)](#).

The system can be used with any type of data and is particularly suited for providing 'data mining' like searches of complex descriptive data. BioMart comes with an 'out of the box' website that can be installed, configured and customised according to user requirements. Further access is provided by graphical and text based applications or programmatically using web services or API written in Perl and Java. BioMart has built-in support for query optimisation and data federation and in addition can be configured to work as a DAS 1.5 Annotation server. The process of converting a data source into BioMart format is fully automated by the tools included in the package. Currently supported RDBMS platforms are MySQL, Oracle and Postgres.

BioMart is completely Open Source, licensed under the LGPL, and freely available to anyone without restrictions.

Powered by BioMart software:

- | | | | | |
|--|-------------------------------|-----------------------------------|---|---|
| • BioMart Central Portal | • Phytozome | • Wormbase | • Rat Genome Database | • Pancreatic Expression Database |
| • ICGC Data Portal | • Gramene | • DroSpeGe | • GermOnLine | • Reactome |
| • Ensembl | • Europhenome | • ArrayExpress DW | • PRIDE | • EU Rat Mart |
| • Ensembl Bacteria | • UniProt | • Eurexpress | • PepSeeker | • Paramecium DB |
| • Ensembl Metazoa | • InterPro | • HapMap | • VectorBase | • International Potato Center (CIP) |
| • Ensembl Protists | • HGNC | • Dictybase | • HTGT | • Mouse Genome Informatics (MGI) |
| • Ensembl Plants | • Rice-Map | • COSMIC | • Clddb | • CyanoBase |
| • Ensembl Fungi | • IKMC | • IntOGen | • European Mouse Mutant Archive | |

Third party software with BioMart Plugin:

[Bioclipse](#) [biomaRt-BioConductor](#) [Cytoscape](#) [Galaxy](#) [Gtools](#) [Ruby API](#) [Taverna](#) [WebLab](#)

Biomart databases

 [Login](#) | [Register](#) | [BLAST/BLAT](#) | [BioMart](#) | [Tools](#) | [Downloads](#) | [Help](#) | [More...](#)

[New](#) [Count](#) [Results](#) [★ URL](#) [XML](#) [Perl](#) [Help](#)

Dataset

Homo sapiens genes (GRCh37)

Filters

Transmembrane domains :
Only

Attributes

Ensembl Gene ID
Drosophila Ensembl Gene ID
% Identity

Dataset

[None Selected]

Export all results to ☒ Unique results only

Email notification to

View rows as ☒ Unique results only

Ensembl Gene ID	Drosophila Ensembl Gene ID	% Identity
ENSG00000215764		
ENSG00000233094	FBgn0027556	19
ENSG00000233094	FBgn0033289	11
ENSG00000215608		
ENSG00000215611		
ENSG00000226460		
ENSG00000099715		
ENSG00000183704		
ENSG00000183878		
ENSG00000165246	FBgn0001987	20
ENSG00000165246	FBgn0034076	14
ENSG00000165246	FBgn0039084	17
ENSG00000165246	FBgn0033943	13
ENSG00000165246	FBgn0032132	15
ENSG00000165246	FBgn0032131	15
ENSG00000165246	FBgn0000592	14
ENSG00000165246	FBgn0010052	16

KEGG pathways



KEGG PATHWAY Database

Wiring diagrams of molecular interactions, reactions, and relations

KEGG2 **PATHWAY** **BRITE** **DISEASE** **DRUG** **KO** **GENES** **GENOME** **LIGAND** **DBGET**

Select prefix

map

Organism

Enter keywords

Go

Help

Pathway Maps

KEGG PATHWAY is a collection of manually drawn pathway maps (see [new maps](#), [change history](#), and [last updates](#)) representing our knowledge on the molecular interaction and reaction networks for:

0. Global Map

1. Metabolism

[Carbohydrate](#) [Energy](#) [Lipid](#) [Nucleotide](#) [Amino acid](#) [Other amino acid](#) [Glycan](#)
[Cofactor/vitamin](#) [Terpenoid/PK](#) [Other secondary metabolite](#) [Xenobiotics](#) [Overview](#)

2. Genetic Information Processing

3. Environmental Information Processing

4. Cellular Processes

5. Organismal Systems

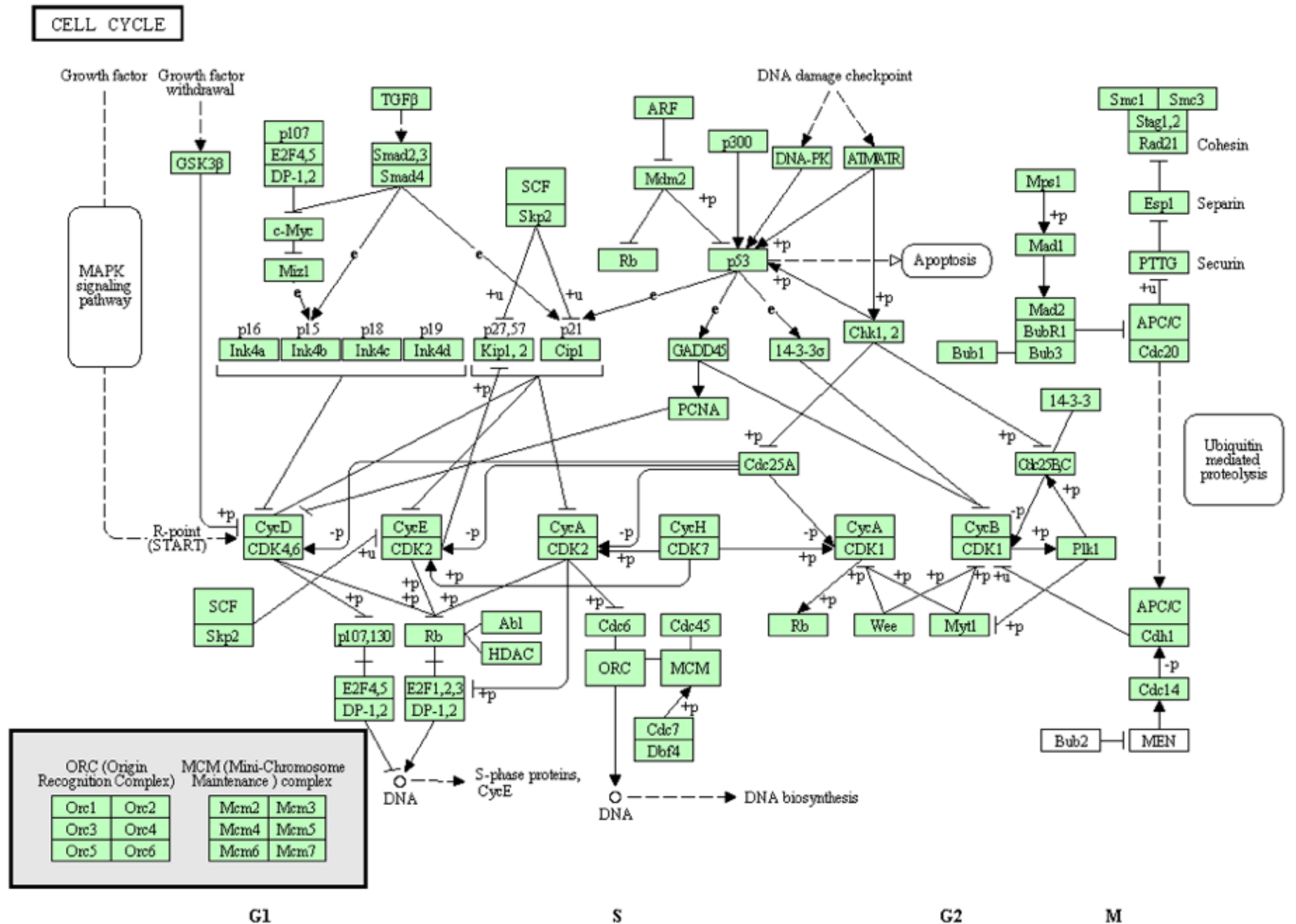
6. Human Diseases

and also on the structure relationships (KEGG drug structure maps) in:

7. Drug Development

KEGG Atlas may now be used to examine any of the KEGG pathway maps.

KEGG pathways





[Home](#) [Browser](#) [Biomart](#) [About](#)

release v02

Integrative OncoGenomics (**IntOGen**), a discovery tool for cancer researchers, is a resource that integrates multidimensional OncoGenomics Data for the identification of genes and groups of genes (biological modules) involved in cancer development.



[Intogen Newsletter](#)



[Frequently Asked Questions](#)



[Suggest datasets](#)

Latest news

2010/09/15: [IntOGen presented at CSHL Meeting on Personal Genomes](#)

2010/09/13: [ICG10 course to learn how to use IntOGen](#)

2010/07/26: [Download IntOGen data from Biomart](#)

2010/06/10: [IntOGen updates to release v02](#)

[\(more news...\)](#)

Tutorials



**GENE
SEARCH**



**EXPERIMENT
SEARCH**



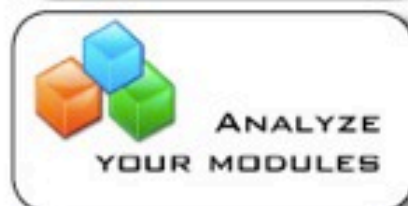
**TUMOR TYPE
SEARCH**



**CANCER
GENE
PRIORITIZATION**



**COMBINE
EXPERIMENTS**



**ANALYZE
YOUR MODULES**



**INTOGEN
IN GITOOLS**

Molecular Signature DB

c1 **positional gene sets** for each human chromosome and each cytogenetic band.

c2 **curated gene sets** from online pathway databases, publications in PubMed, and knowledge of domain experts.

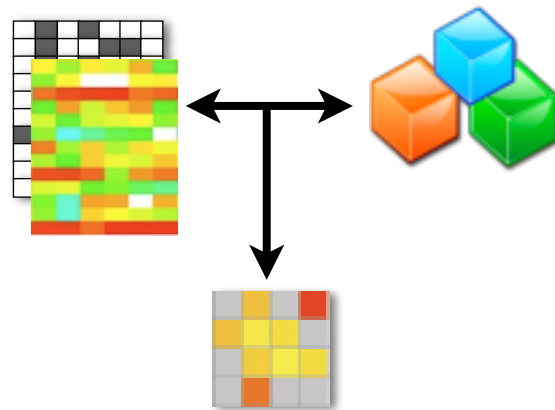
c3 **motif gene sets** based on conserved *cis*-regulatory motifs from a comparative analysis of the human, mouse, rat and dog genomes.

c4 **computational gene sets** defined by expression neighborhoods centered on 380 cancer-associated genes.

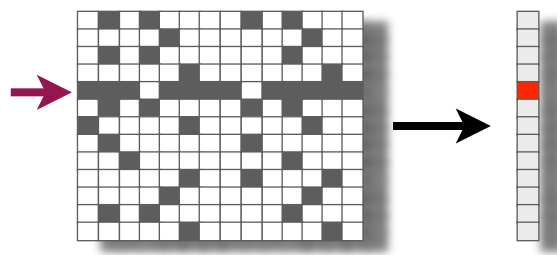
c5 **GO gene sets** consist of genes annotated by the same GO terms.

Methods for data integration

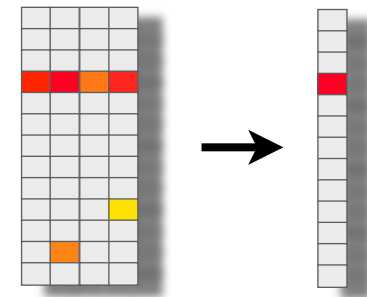
ENRICHMENT ANALYSIS



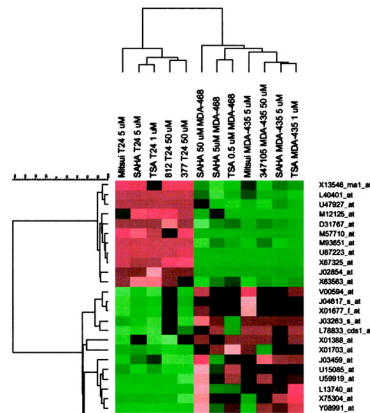
DRIVER ALTERATIONS



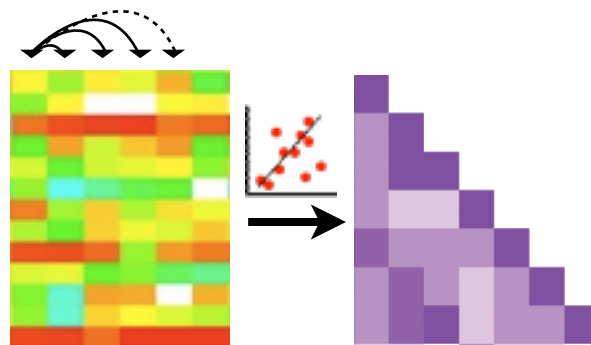
RESULTS COMBINATION



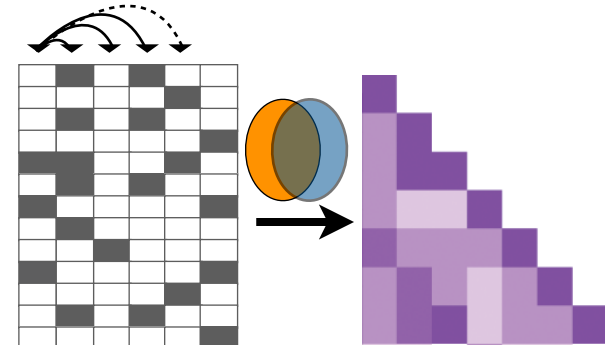
CLUSTERING



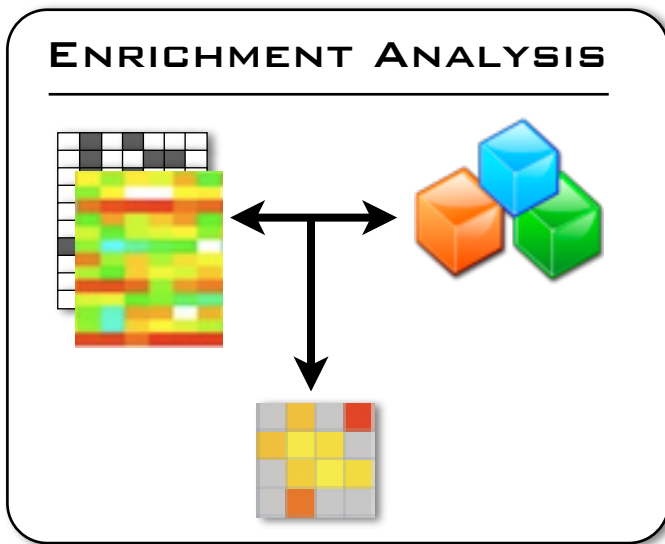
CORRELATIONS



OVERLAPS



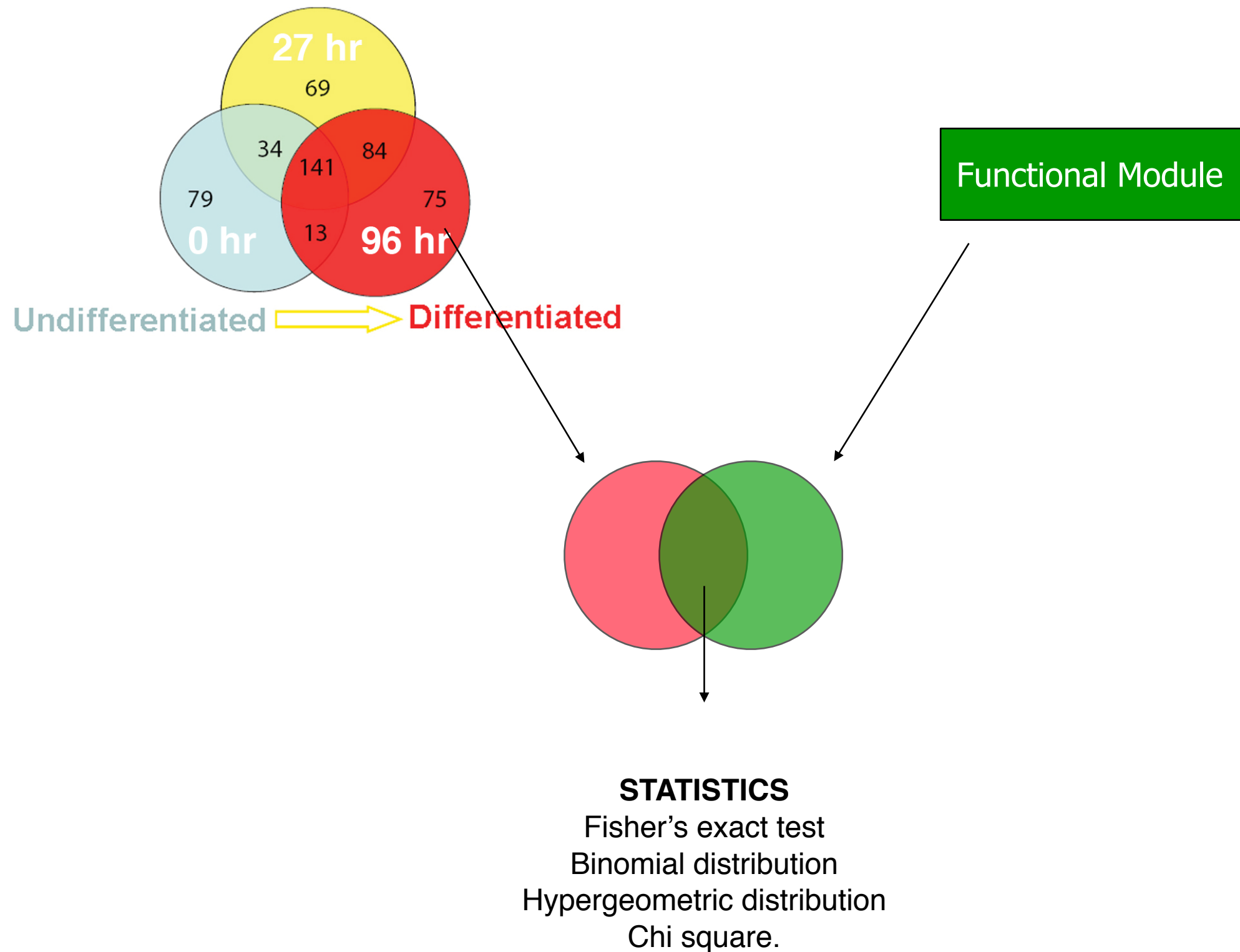
Enrichment Analysis



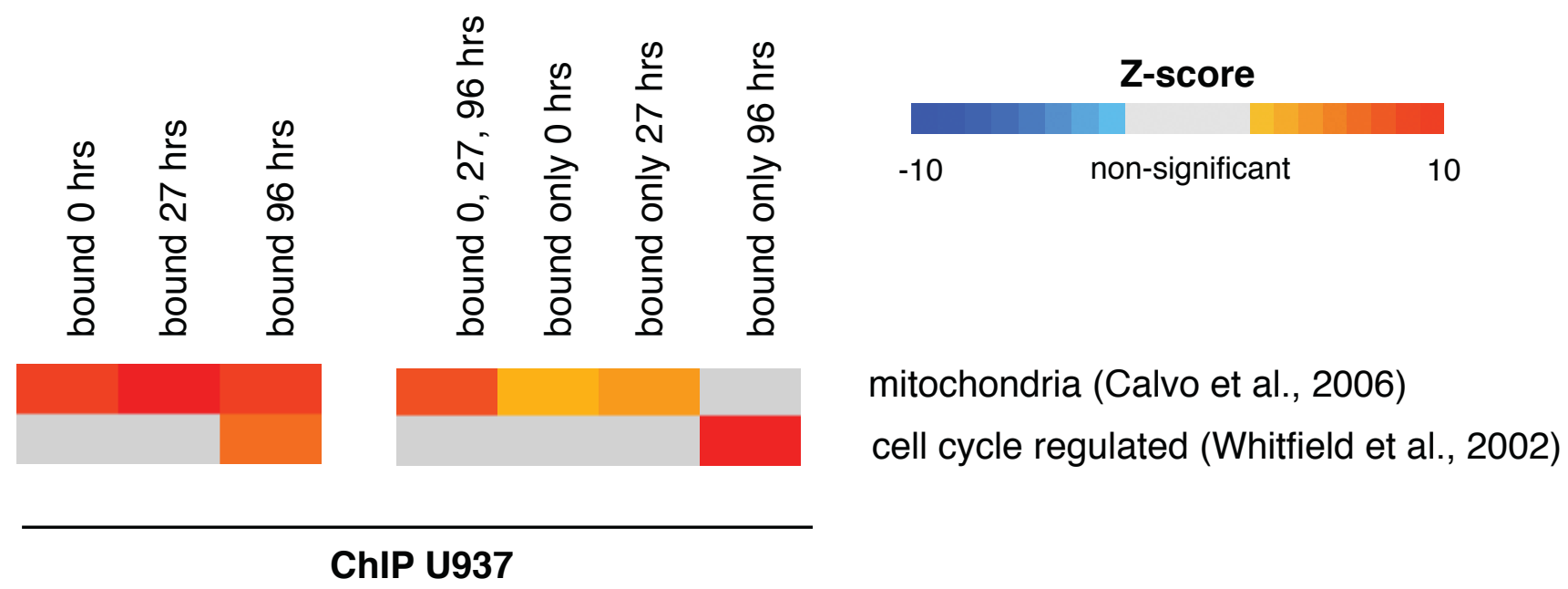
EAL: Enrichment analysis for a list of genes

EAC: Enrichment analysis for continuous values

Enrichment Analysis for a list of genes

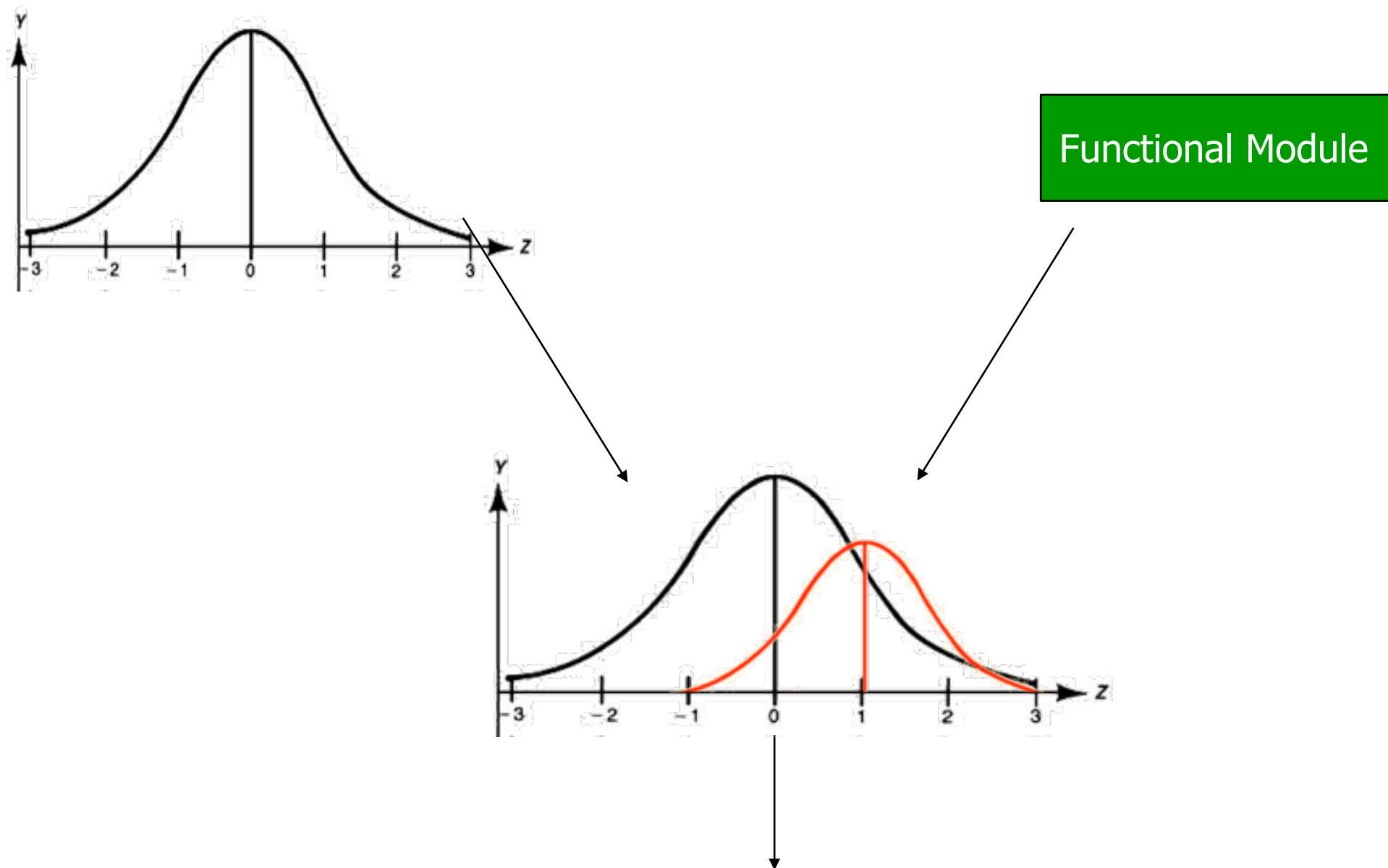


Functional Enrichment Analysis



	cell cycle regulated			mitochondrial proteins		
targets	p (n) targets	p (n) all	p-value*	p (n) targets	p (n) all	p-value*
0 hrs	0.05 (12)	0.04 (449)	0.82	0.16 (39)	0.05 (512)	<1E-16
27 hrs	0.05 (15)	0.05 (474)	0.83	0.18 (54)	0.05 (543)	<1E-16
96 hrs	0.12 (33)	0.05 (474)	9.8E-09	0.16 (44)	0.05 (543)	<1E-16
bound 0, 27, 96 hrs	0.05 (7)	0.05 (474)	0.88	0.19 (25)	0.05 (543)	<1E-16
bound only 0hrs	0.03 (2)	0.05 (474)	0.62	0.14 (9)	0.05 (543)	2.0E-03
bound only 27 hrs	0 (0)	0.05 (474)	0.08	0.16 (11)	0.05 (543)	7.7E-05
bound only 96 hrs	0.28 (19)	0.05 (474)	<1E-16	0.06 (4)	0.05 (543)	0.81

Enrichment Analysis for continuous values



STATISTICS

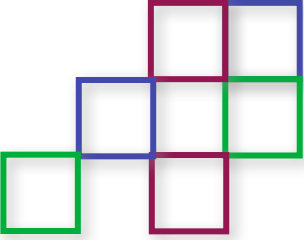
Kolmogorov–Smirnov rank statistics

Fisher's exact

Permutations and zscore

Mutual information

::: Introduction.



Gene Set Enrichment Analysis - GSEA



(Subramanian et al. PNAS. 2005.)

GSEA

MIT
Broad Institute

v 2.0 available since Jan 2007

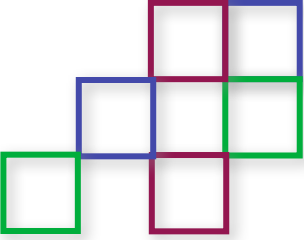
v 2.0.1 available since Feb 16th 2007

Version 2.0 includes Biocarta, Broad Institute,
GeneMAPP, KEGG annotations and more...

Platforms: Affymetrix, Agilent, CodeLink, custom...



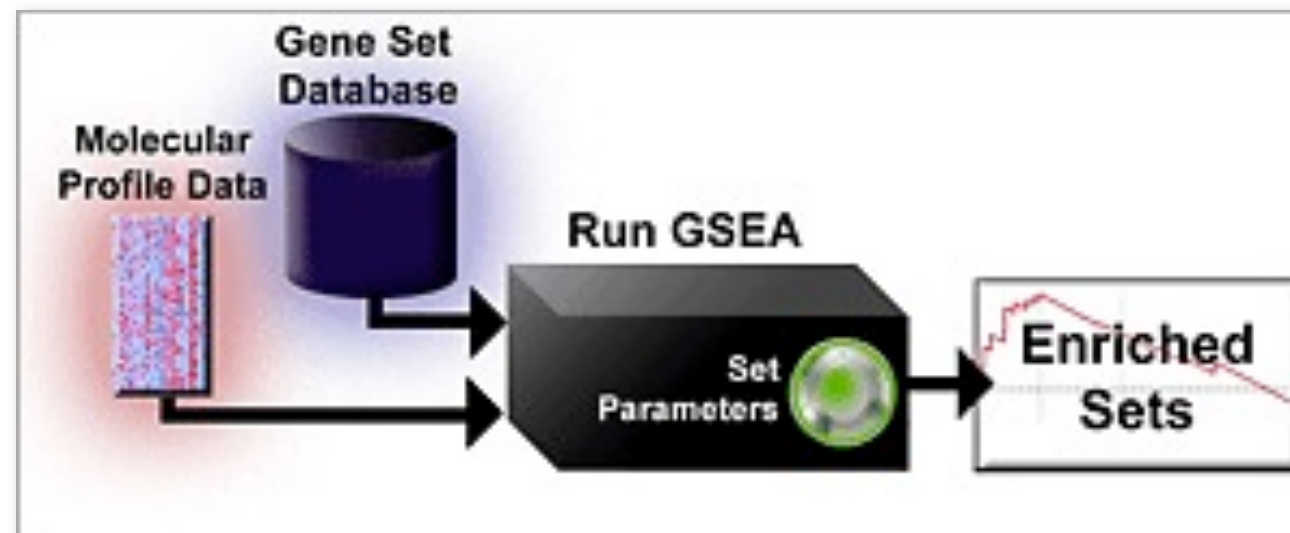
::: Introduction.



Gene Set Enrichment Analysis - GSEA -

::: How works GSEA?

GSEA applies Kolmogorov-Smirnov test to find asymmetrical distributions for defined blocks of genes in datasets whole distribution.

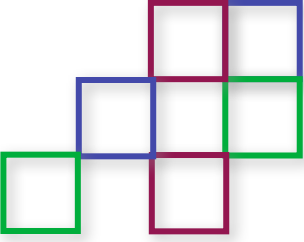


Is this particular **Gene Set** enriched in my experiment?



Genes selected by researcher, Biocarta pathways, GeneMAPP sets,
genes sharing cyto band, genes targeted by common miRNAs
...up to you...

::: Introduction.



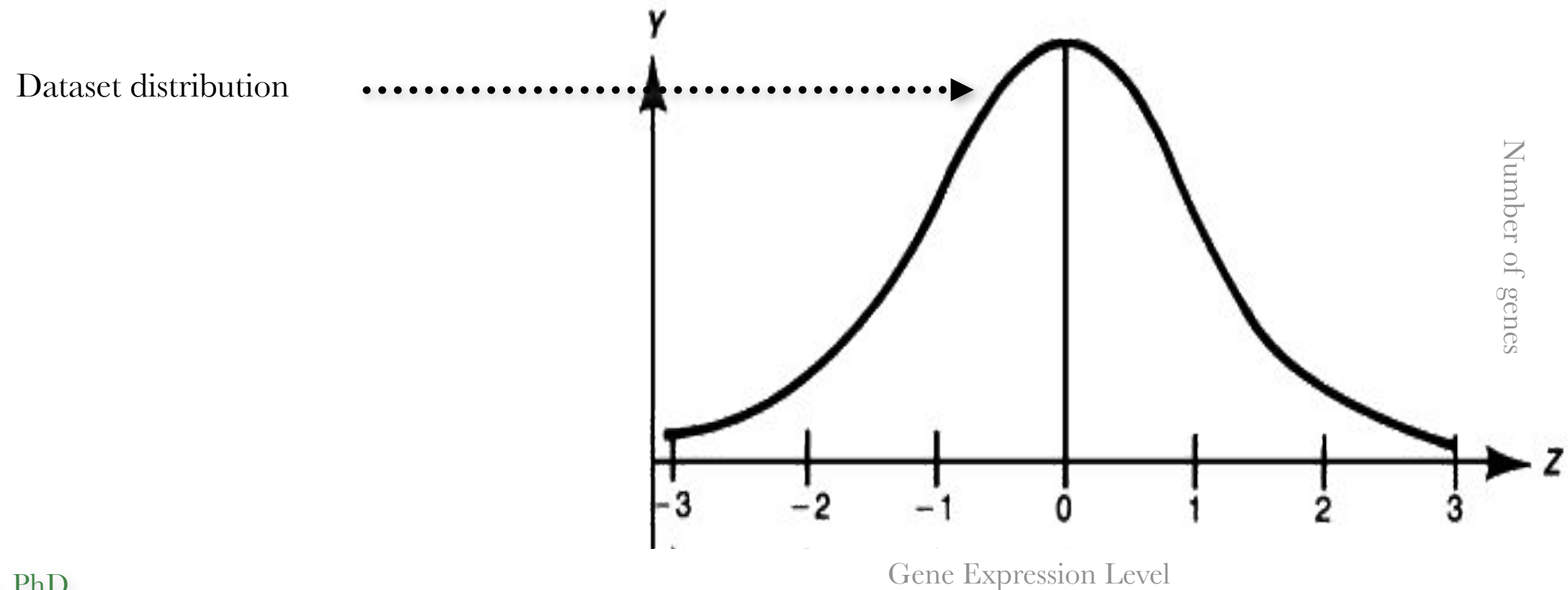
Gene Set Enrichment Analysis - GSEA -

::: K-S test

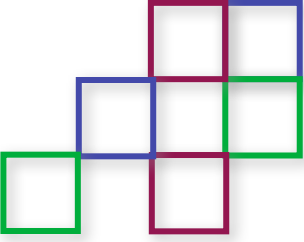
The Kolmogorov–Smirnov test is used to determine whether two underlying one-dimensional probability distributions differ, or whether an underlying probability distribution differs from a hypothesized distribution, in either case based on finite samples.

The one-sample KS test compares the empirical distribution function with the cumulative distribution functions specified by the null hypothesis. The main applications are testing goodness of fit with the normal and uniform distributions.

The two-sample KS test is one of the most useful and general nonparametric methods for comparing two samples, as it is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples.



::: Introduction.



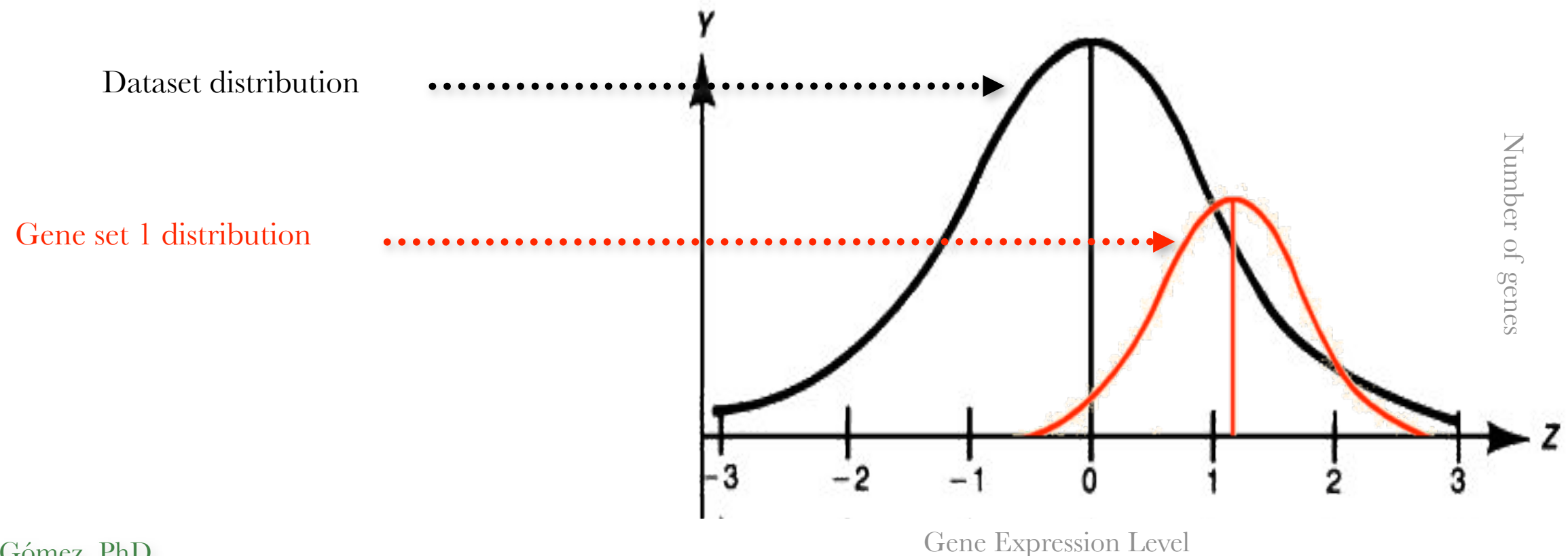
Gene Set Enrichment Analysis - GSEA -

::: K-S test

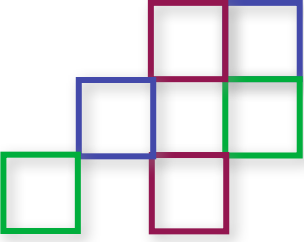
The Kolmogorov–Smirnov test is used to determine whether two underlying one-dimensional probability distributions differ, or whether an underlying probability distribution differs from a hypothesized distribution, in either case based on finite samples.

The one-sample KS test compares the empirical distribution function with the cumulative distribution functions specified by the null hypothesis. The main applications are testing goodness of fit with the normal and uniform distributions.

The two-sample KS test is one of the most useful and general nonparametric methods for comparing two samples, as it is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples.



::: Introduction.



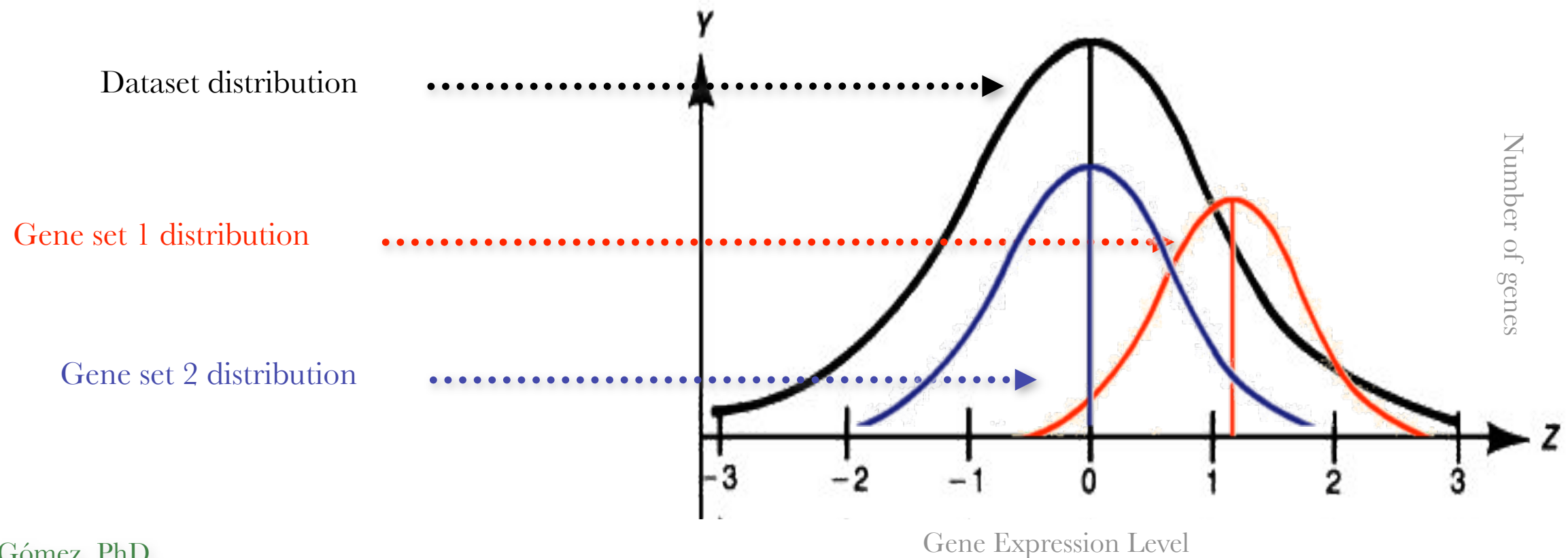
Gene Set Enrichment Analysis - GSEA -

::: K-S test

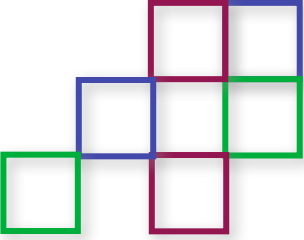
The Kolmogorov–Smirnov test is used to determine whether two underlying one-dimensional probability distributions differ, or whether an underlying probability distribution differs from a hypothesized distribution, in either case based on finite samples.

The one-sample KS test compares the empirical distribution function with the cumulative distribution functions specified by the null hypothesis. The main applications are testing goodness of fit with the normal and uniform distributions.

The two-sample KS test is one of the most useful and general nonparametric methods for comparing two samples, as it is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples.

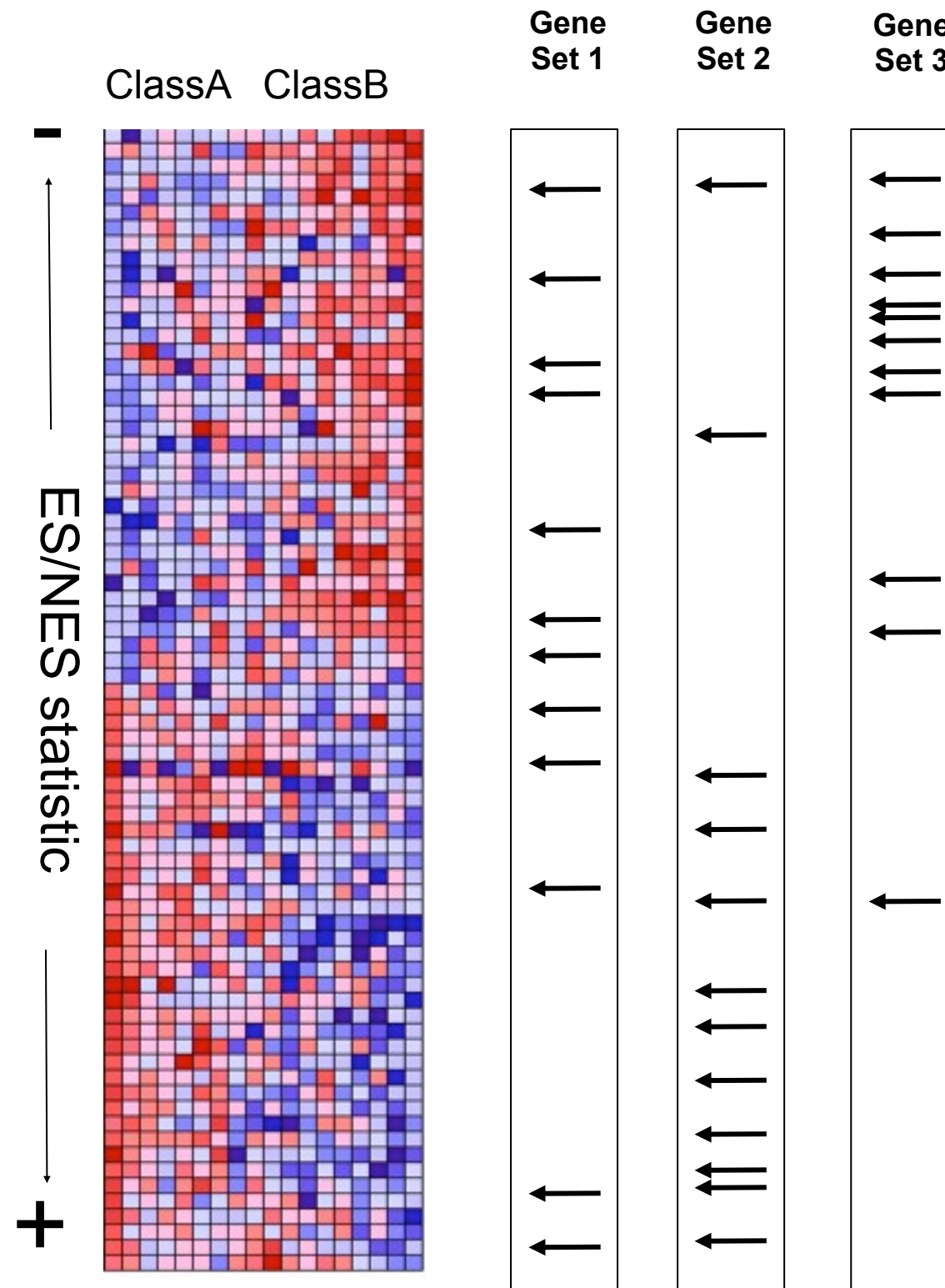


::: Introduction.

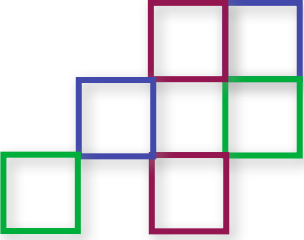


Gene Set Enrichment Analysis - GSEA -

::: How works GSEA?

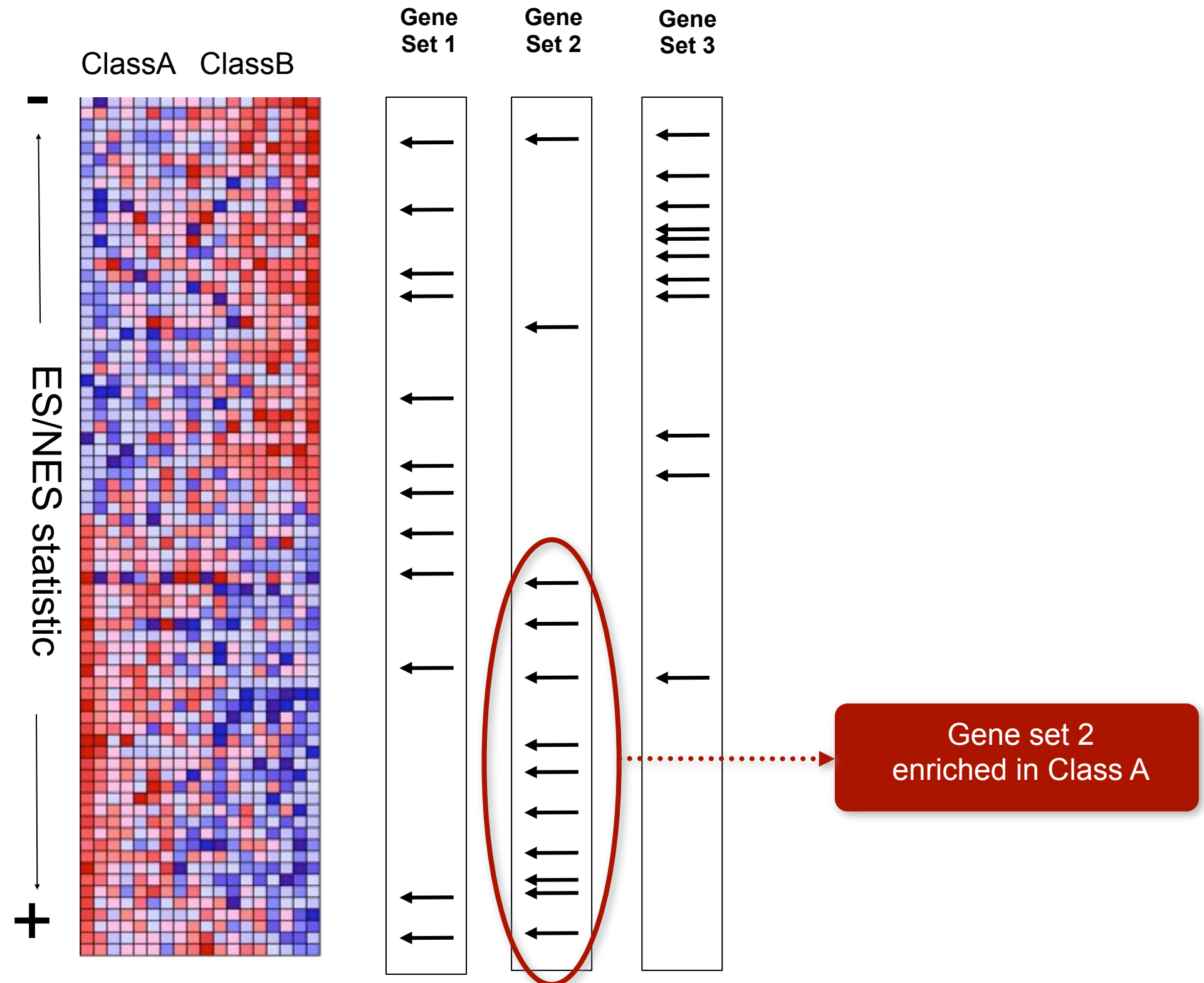


::: Introduction.

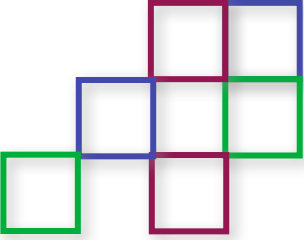


Gene Set Enrichment Analysis - GSEA -

::: How works GSEA?

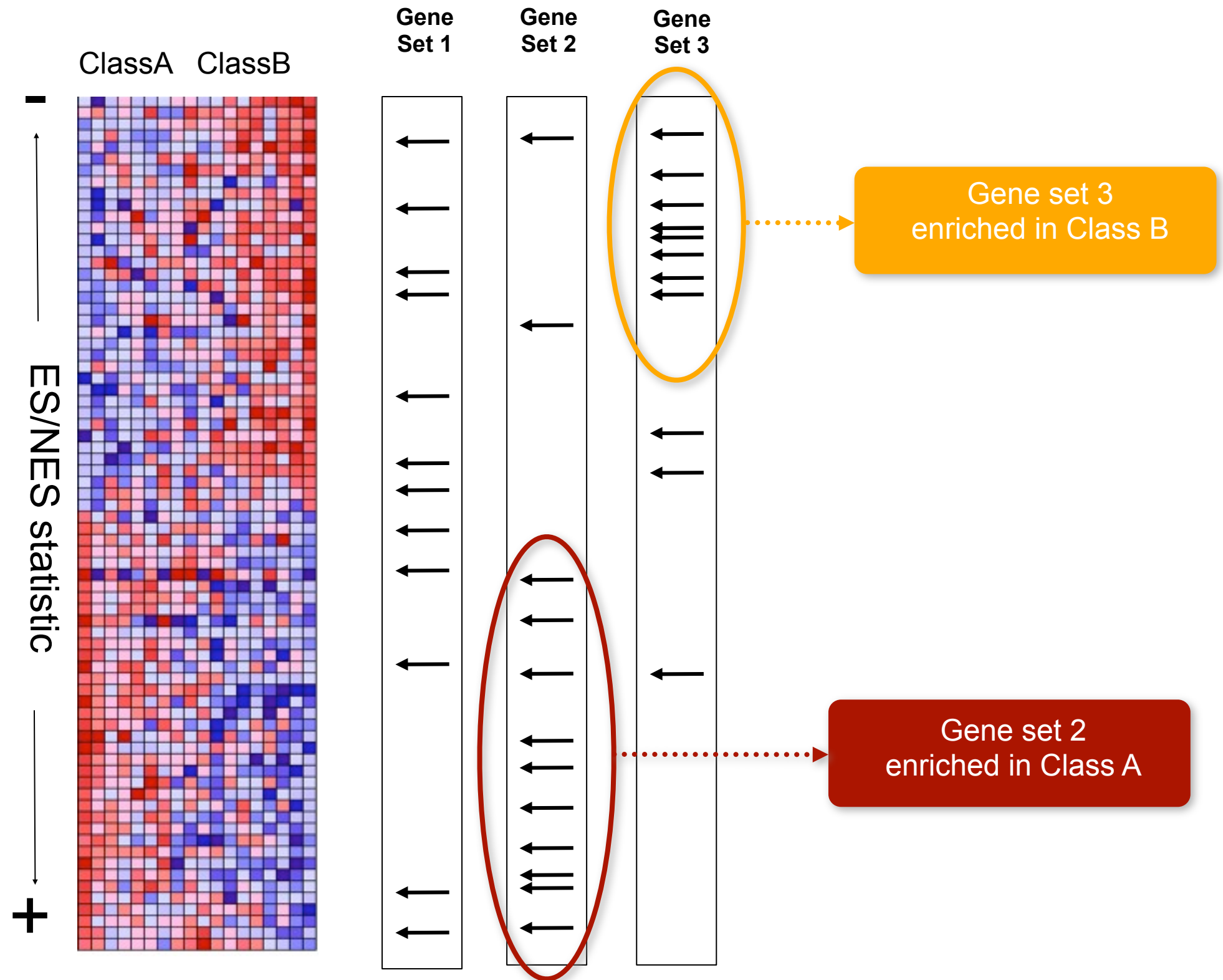


::: Introduction.

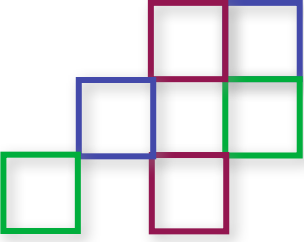


Gene Set Enrichment Analysis - GSEA -

::: How works GSEA?

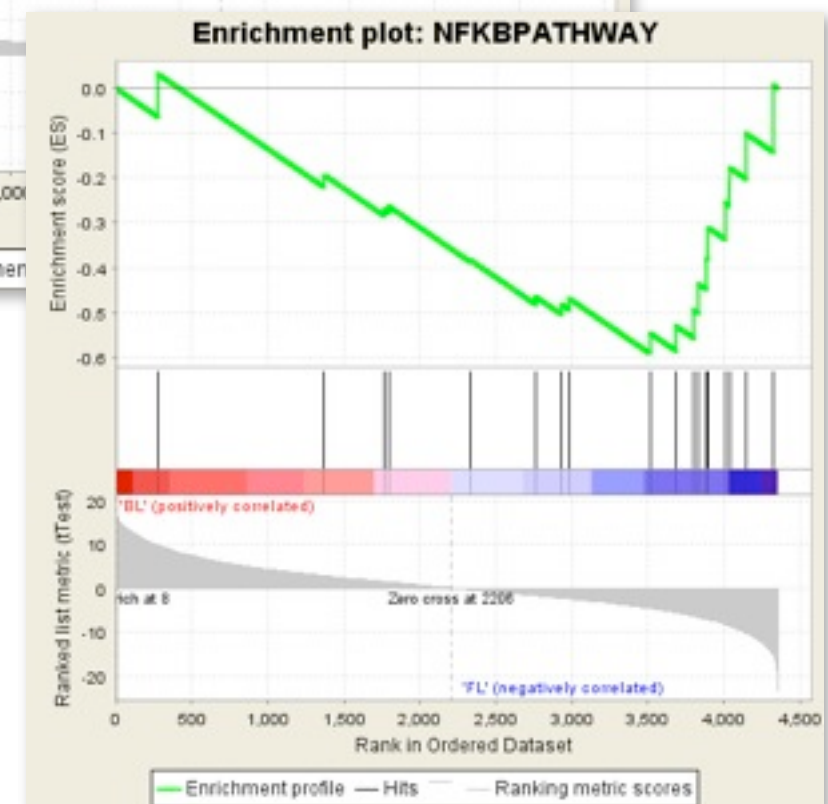
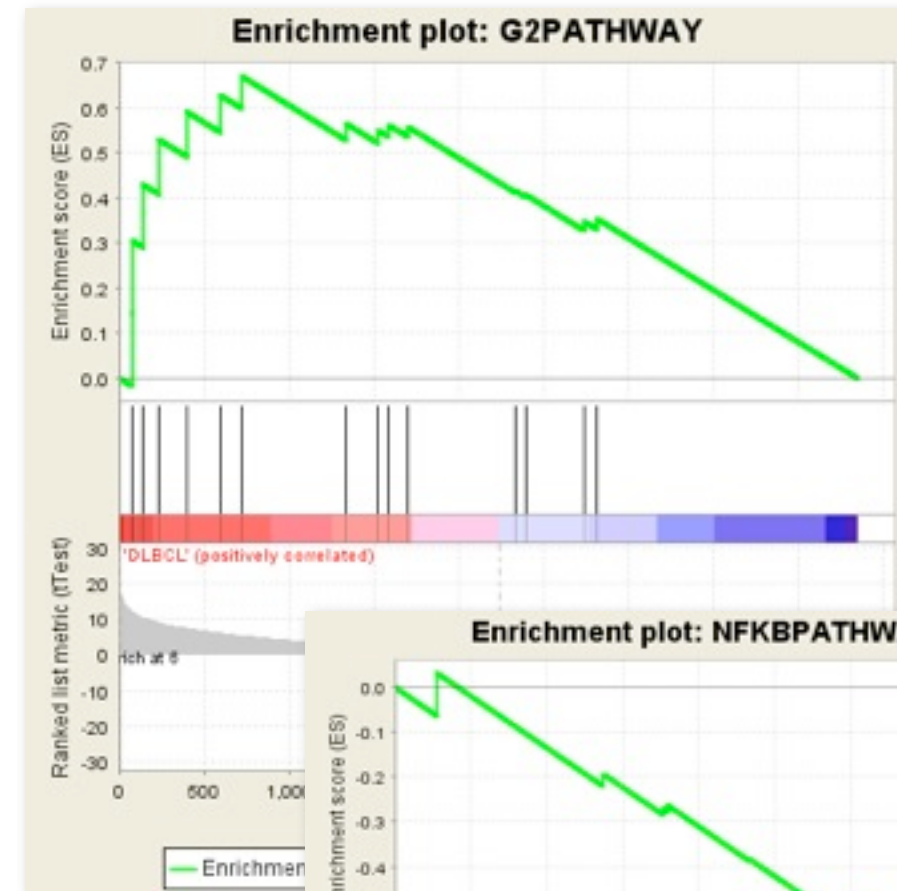
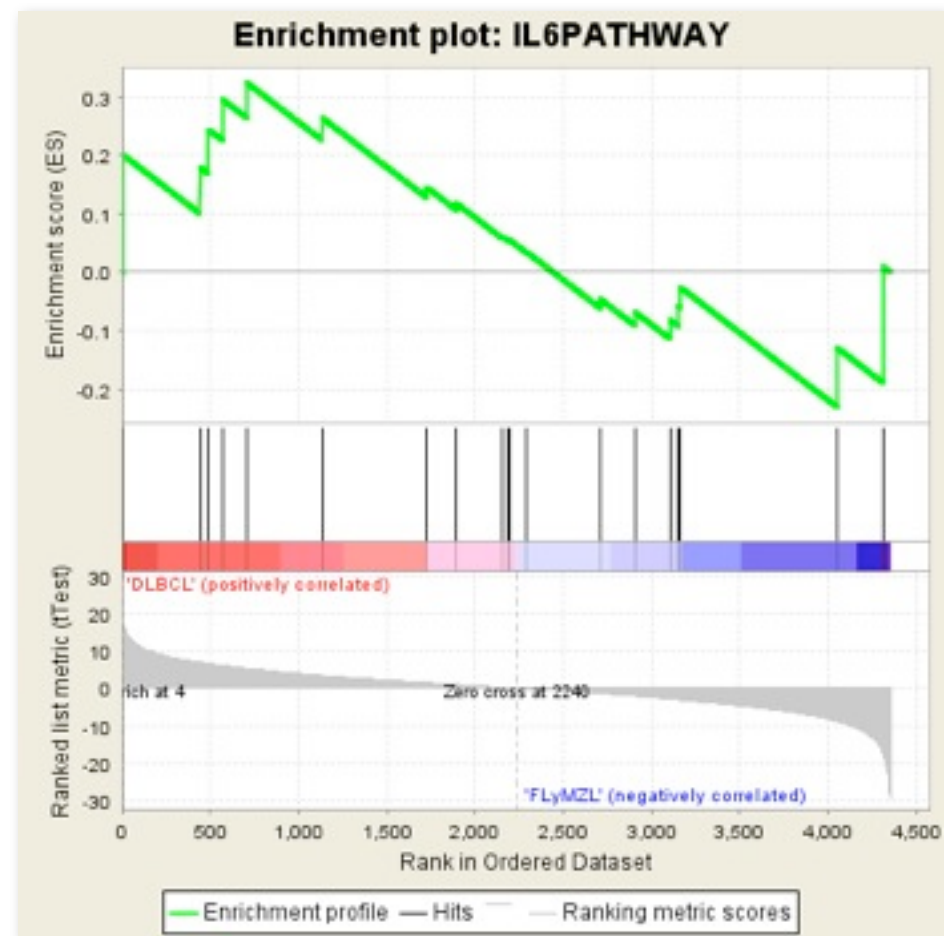


::: Introduction.



Gene Set Enrichment Analysis - GSEA -

ES examples :::



Correction for multiple testing

When detecting differentially expressed genes, we want to detect ONLY the differentially expressed, with no false positives !

Statistical Corrections

In many analysis of biological experiments, a great number of false positives are found among the results. When making multiple comparisons, we need to apply a statistical correction to our threshold, to remove the maximum of false positives.

Commonly available statistical corrections:

Method	Complexity	Time	Method	Results	Drawback
Bonferroni correction	simplest	fastest	Most conservative	keeping only the most significant results, removing every possible noise, or putative results.	a lot of significant information is removed along with the noise
False Discovery Rate (FDR)			Less conservative	a good compromise between keeping only really significant hits, and having too much false positives.	Some false positives...

http://cbi.labri.fr/outils/BlastSets/BlastSets_web_manual/principles.html

Some Enrichment Analysis Tools

Method	URL	Platform Type	EA type
DAVID	http://david.abcc.ncifcrf.gov	WT	EAL (modified Fisher's exact - EASE)
Babelomics	http://www.babelomics.org	WT	EAL, EAC (Fisher's exact)
GOMINER	http://discover.nci.nih.gov/gominer	WT	EAL (Fisher's exact)
ConceptGene	http://conceptgen.ncibi.org	WT	EAL (modified Fisher's exact - EASE)
Ingenuity	http://www.ingenuity.com/	WT	-
GSEA	http://www.broadinstitute.org/gsea/	DS, CL, RP	EAC (Kolmogorov–Smirnov rank statistics)
Gitools	http://www.gitools.org	DS, CL	EAL (Binomial distribution, Fisher's exact), EAC (z-test)
BiNGO	http://www.psb.ugent.be/cbd/papers/BiNGO	CP	EAL (Hypergeometric distribution)
GOstats	http://www.bioconductor.org	RP	EAL (conditional hypergeometric test)
iPAGE	http://tavazoielab.princeton.edu/iPAGE/	DS, CL, WT	EAC (mutual information)
ToppGene	http://toppgene.cchmc.org	WT	EAL (Hypergeometric distribution)

Code for platform type:

WT: Web tool

DS: Downloadable software

CL: Command line options

RP: R package (Bioconductor)

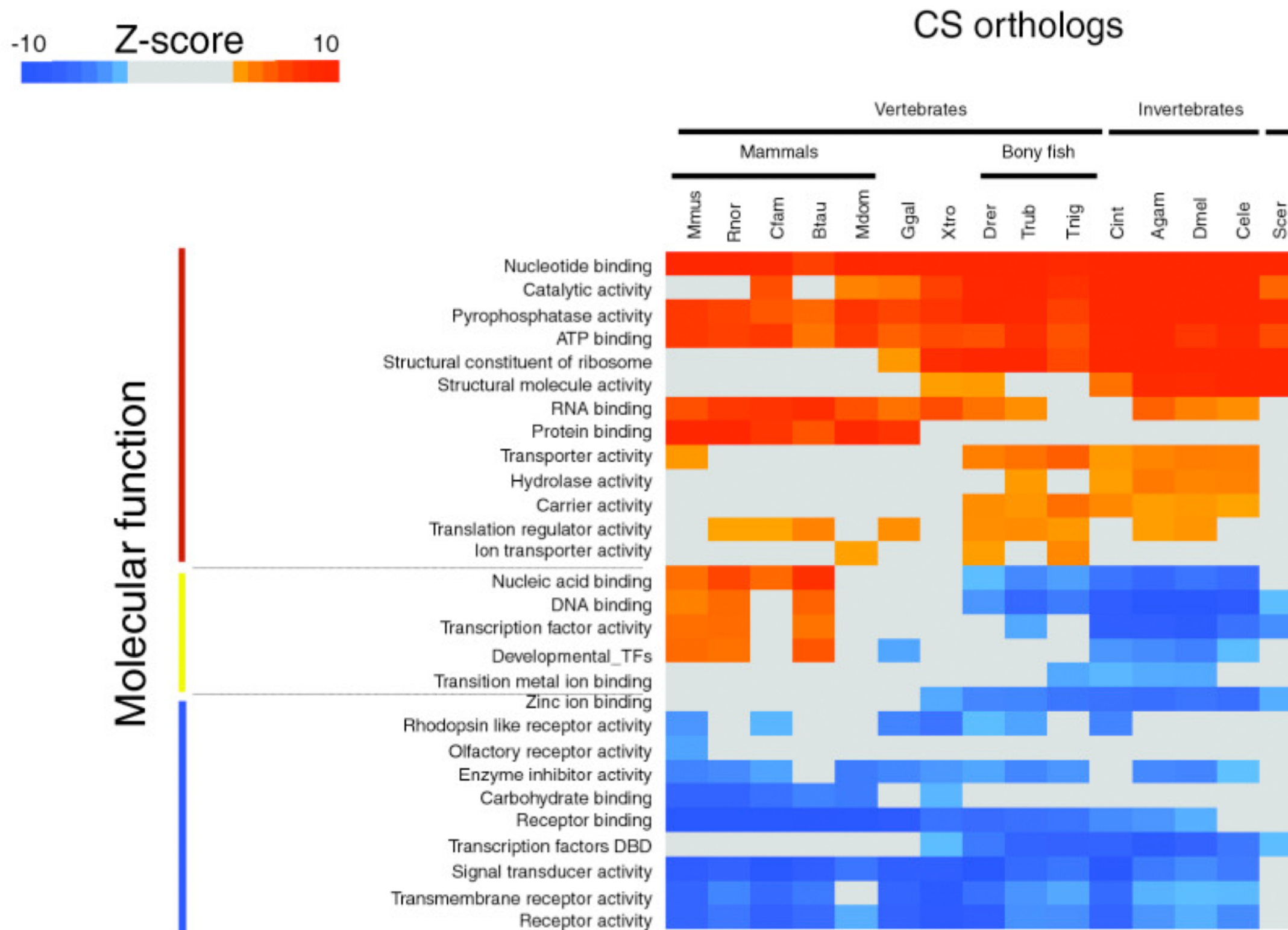
CP: Cytoscape plugin

Code for EA type:

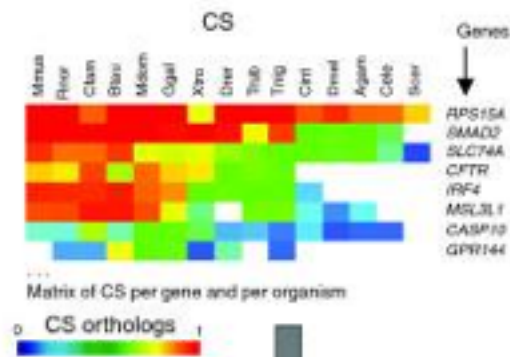
EAL: Enrichment analysis for a list of genes

EAC: Enrichment analysis for continuous values

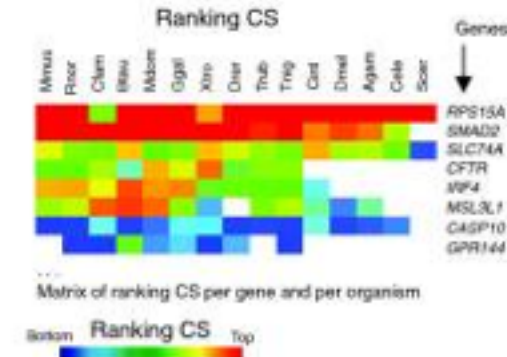
Example of use of EA for continuous values



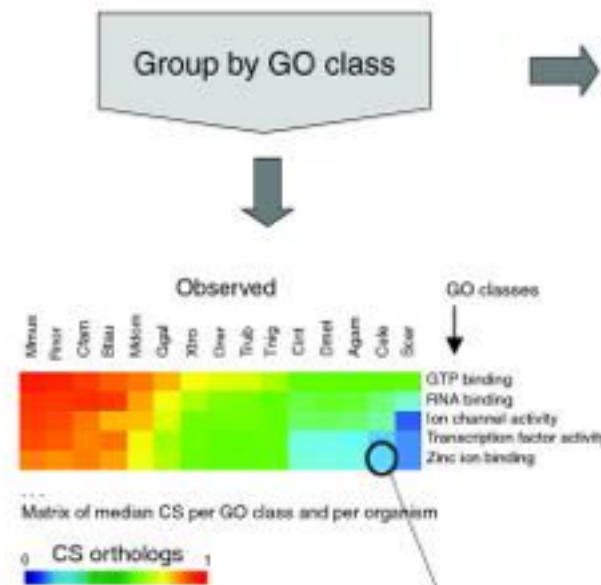
(a) Conservation score (CS) per gene and per organism



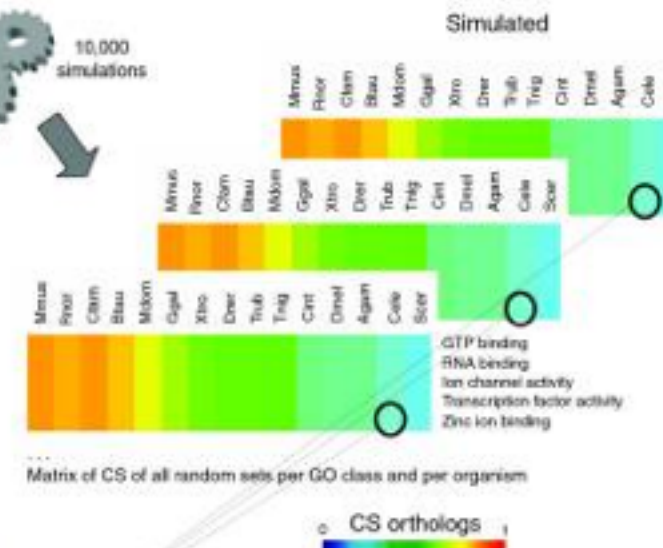
(b) Conservation relative to each genome



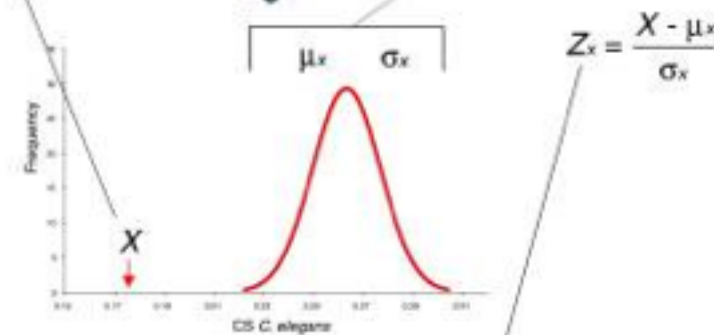
(c) Conservation of functional categories



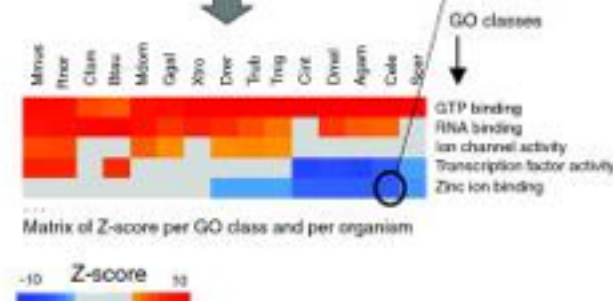
(d) Expected conservation for the group



(e) Statistical analysis (Z-score)

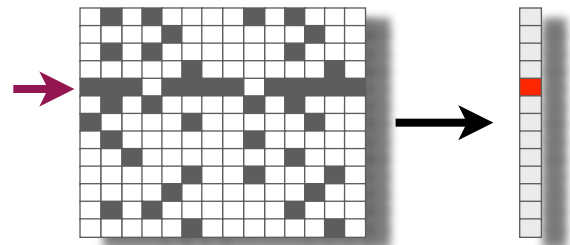


(f) Resulting matrix of Z-scores

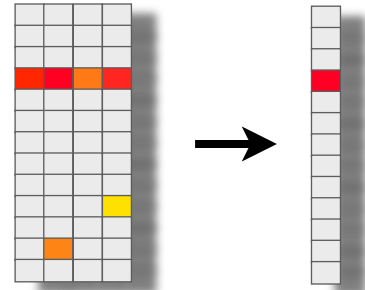


Driver alterations and results combination

DRIVER ALTERATIONS

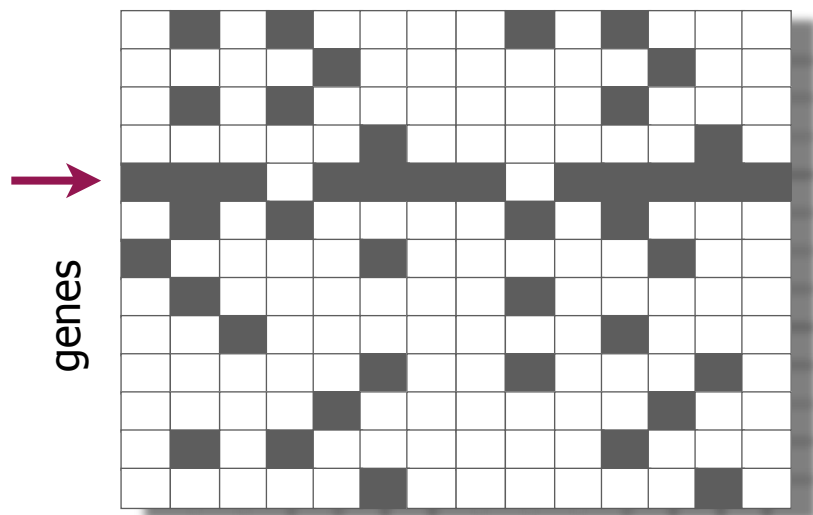


RESULTS COMBINATION



experiment 1

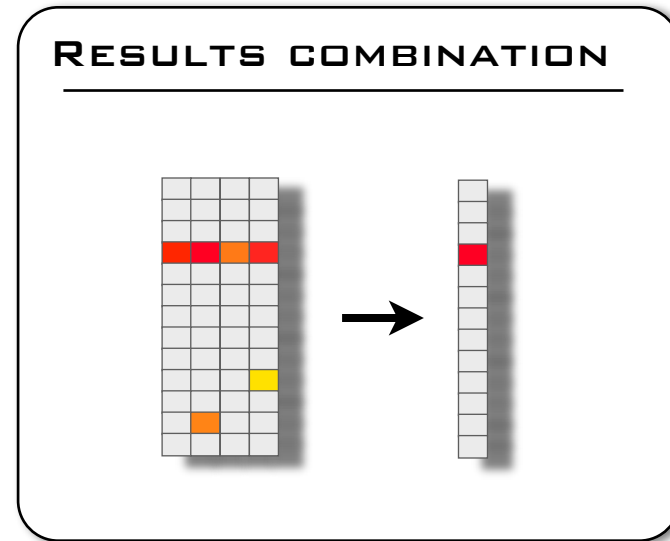
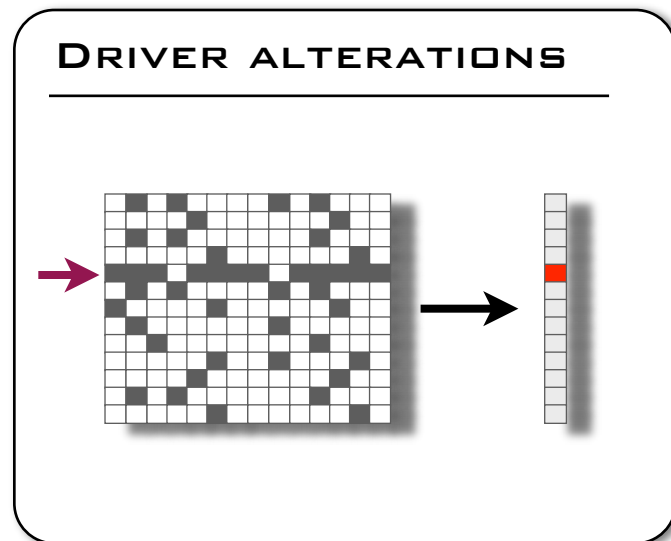
samples



■ altered

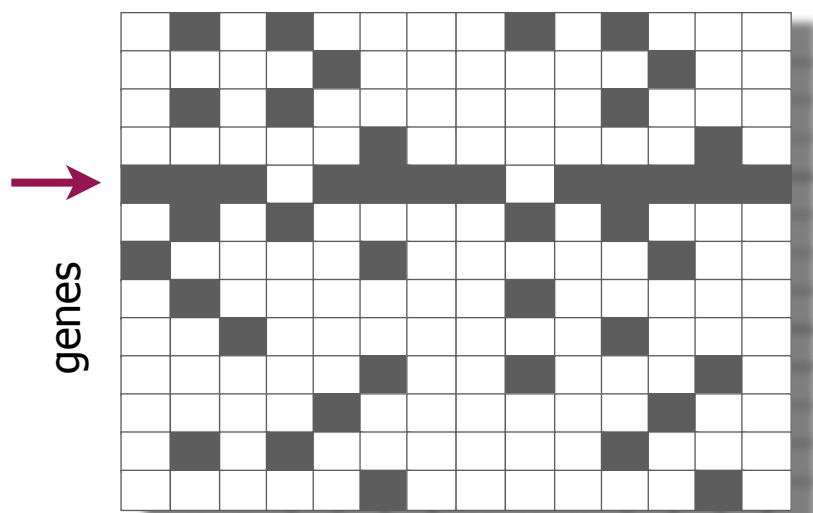
□ not altered

Driver alterations and results combination



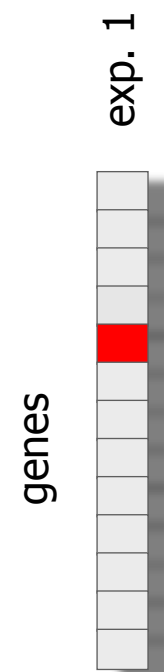
experiment 1

samples



STEP 1

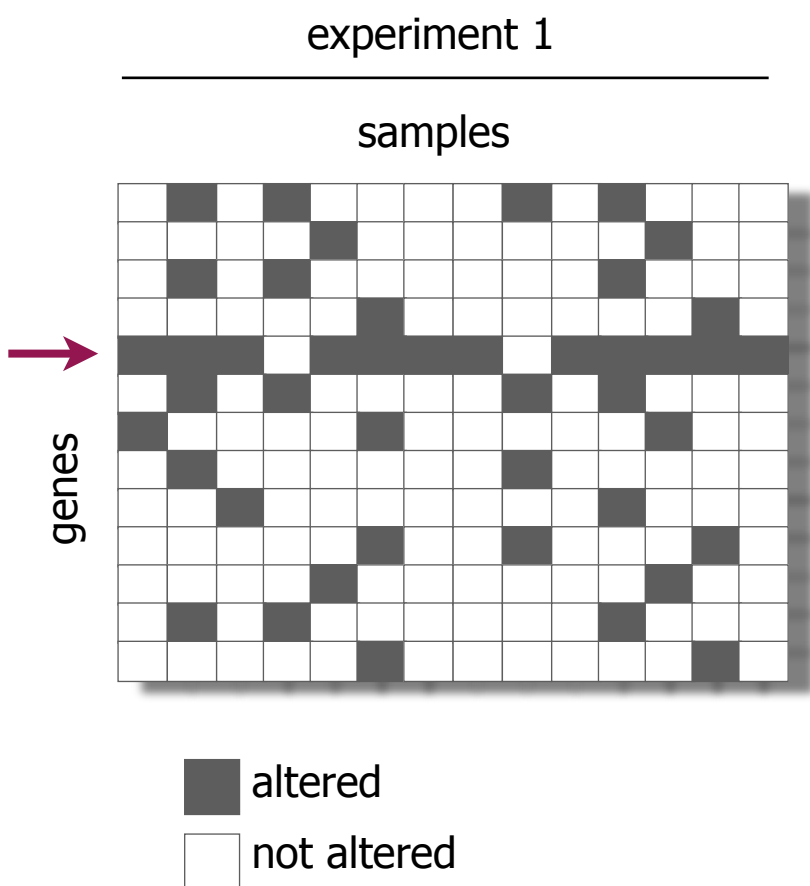
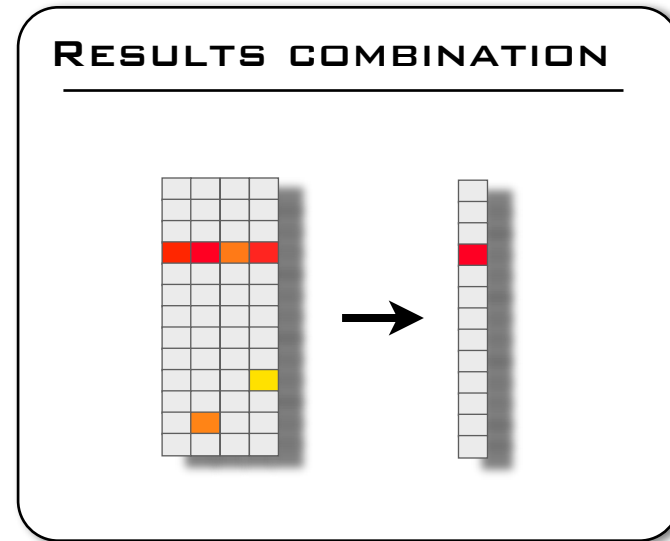
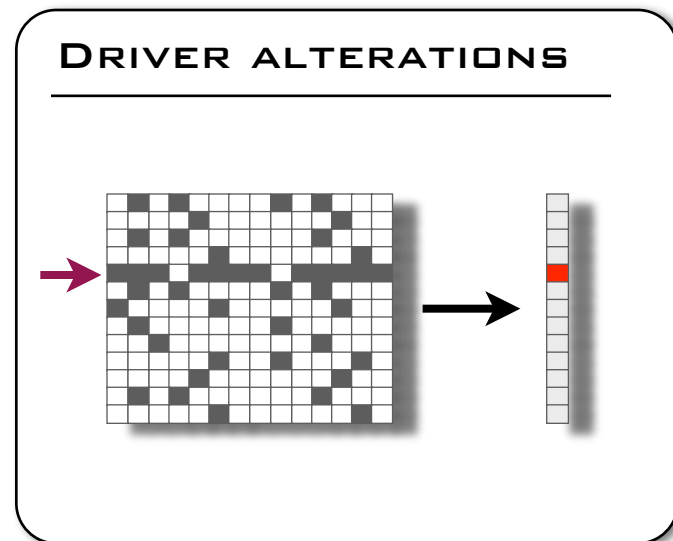
identification of
driver alterations



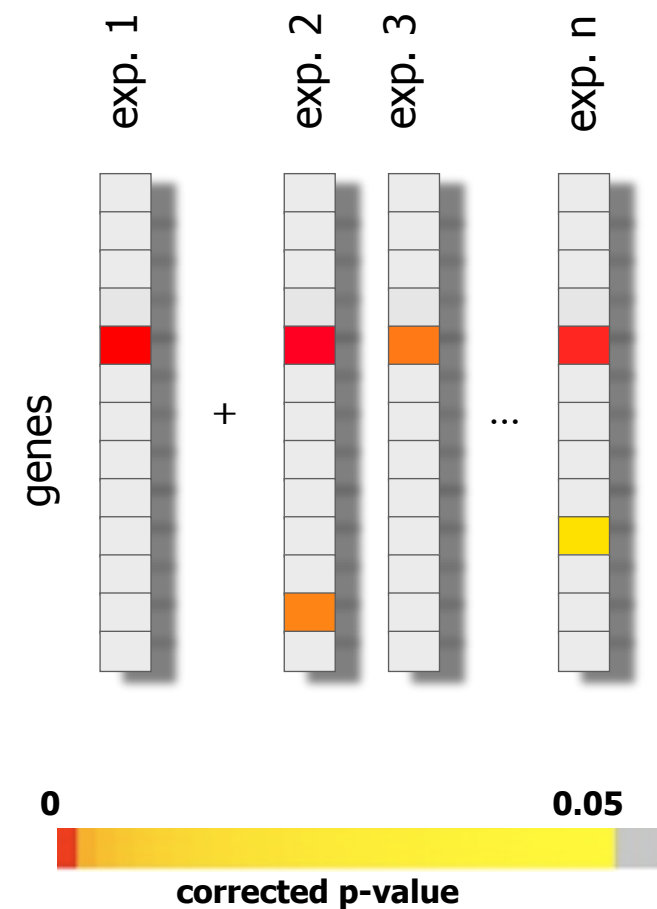
■ altered
□ not altered



Driver alterations and results combination



STEP 1
identification of
driver alterations



STEP 2
combination of
experiments

$$Z_{comb,g} = \frac{\sum_{i=1}^n \omega_i Z_i}{\sqrt{\sum_{i=1}^n \omega_i^2}}$$

Cancer type A

Driver alterations and results combination

CDKN2A (p16) Up-regulation in Breast Cancer

	N samples	Observed	Expected	p-value	corrected p-value	
Yu K et al 2008	183	85	31	6.04E-20	6.64E-19	
Finak G et al 2008	52	46	15	6.74E-19	1.47E-17	
Richardson A et al 2007	43	25	5	2.65E-13	1.01E-11	
Richardson AL et al 2006	40	23	5	2.54E-12	9.29E-11	
Weigelt B et al 2005	64	39	14	1.70E-11	3.37E-10	
Mostafavi ZB et al 2008	53	28	9	1.24E-08	2.80E-07	
Casey T et al 2008	28	10	4	8.93E-03	0.071	
Davis SR et al 2008	54	2	4	0.90	1	
Radvanyi L et al 2005	51	0	0	1	1	

Driver alterations and results combination

CDKN2A (p16) Up-regulation in Breast Cancer

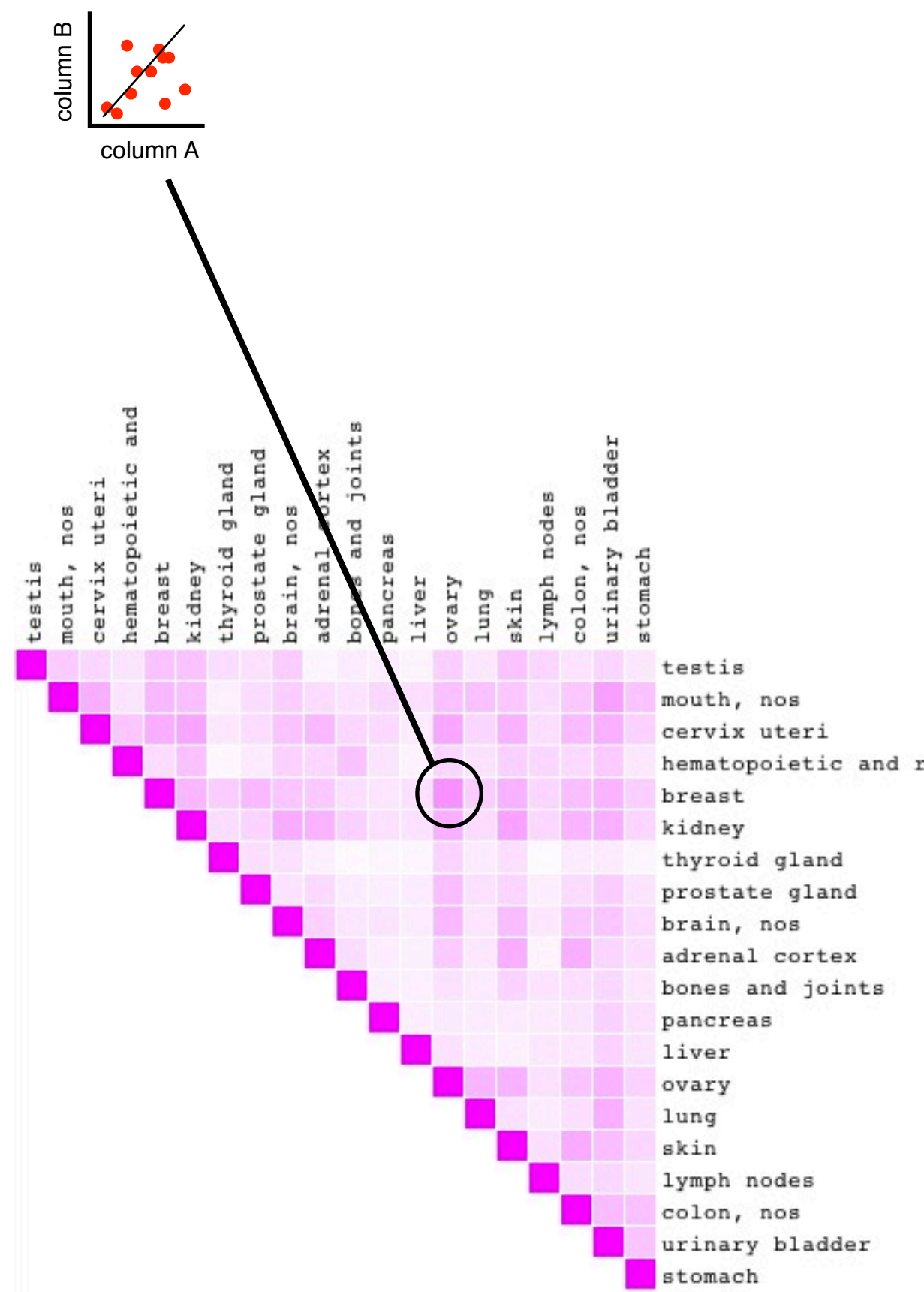
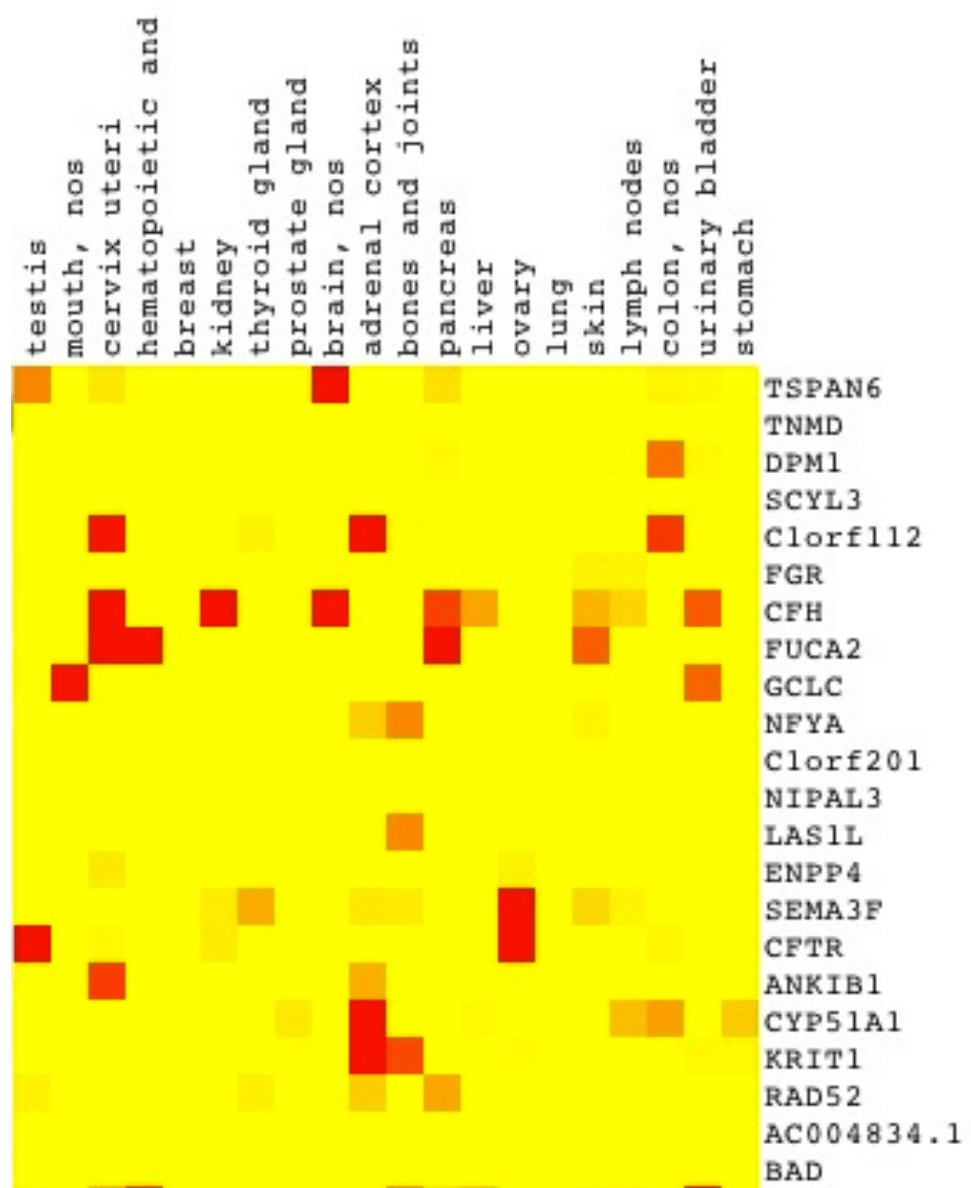
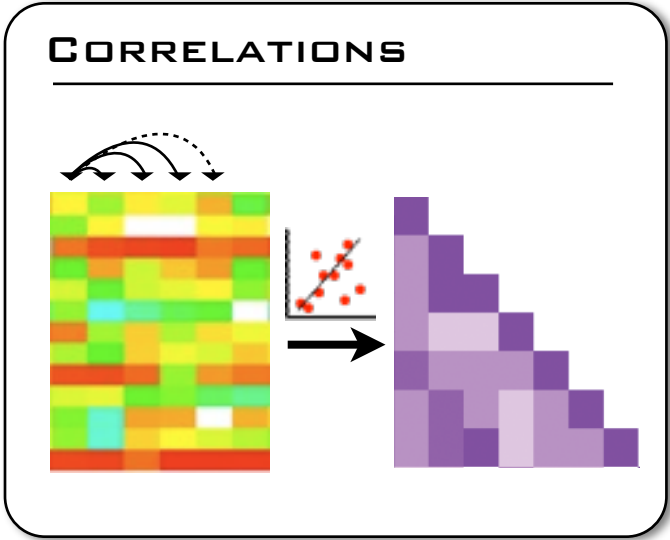
	N samples	Observed	Expected	p-value	corrected p-value	
Yu K et al 2008	183	85	31	6.04E-20	6.64E-19	
Finak G et al 2008	52	46	15	6.74E-19	1.47E-17	
Richardson A et al 2007	43	25	5	2.65E-13	1.01E-11	
Richardson AL et al 2006	40	23	5	2.54E-12	9.29E-11	
Weigelt B et al 2005	64	39	14	1.70E-11	3.37E-10	
Mostafavi ZB et al 2008	53	28	9	1.24E-08	2.80E-07	
Casey T et al 2008	28	10	4	8.93E-03	0.071	
Davis SR et al 2008	54	2	4	0.90	1	
Radvanyi L et al 2005	51	0	0	1	1	

$$Z_{comb,g} = \frac{\sum_{i=1}^n \omega_i Z_i}{\sqrt{\sum_{i=1}^n \omega_i^2}}$$

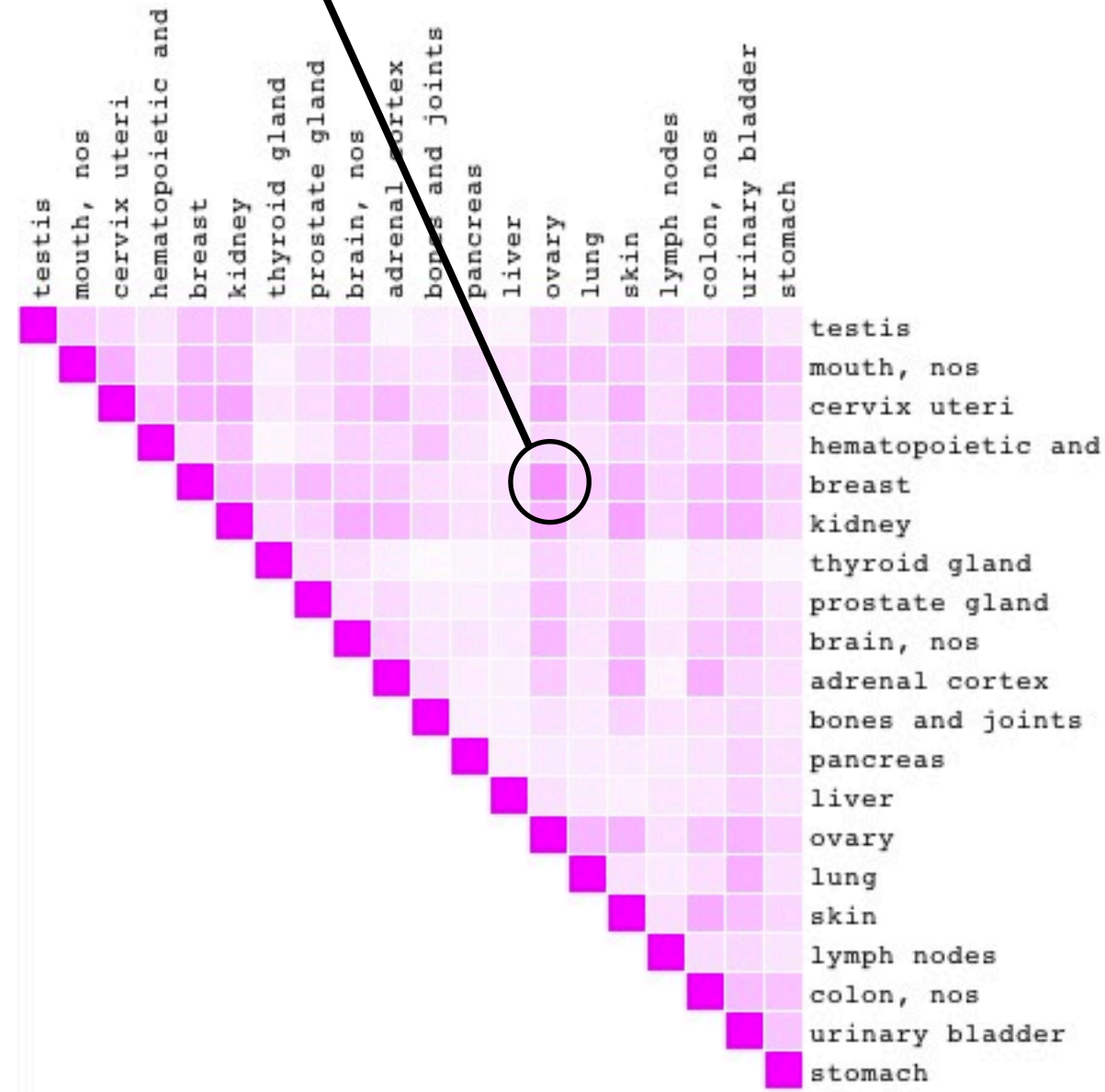
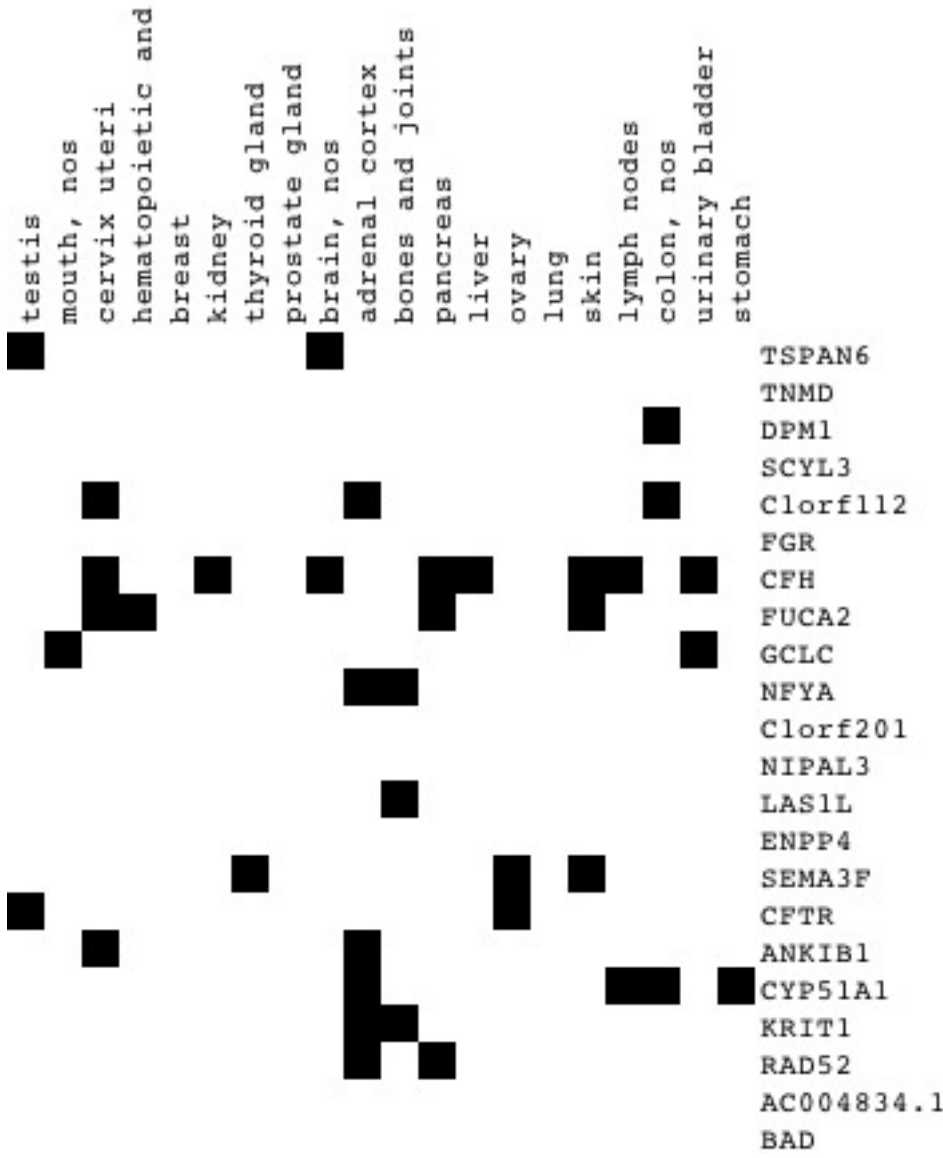
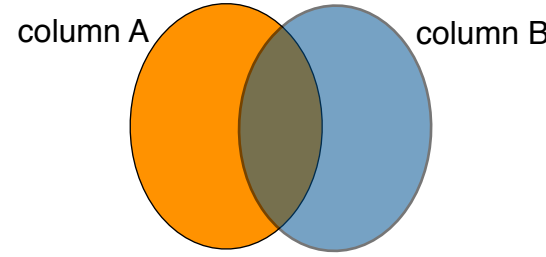
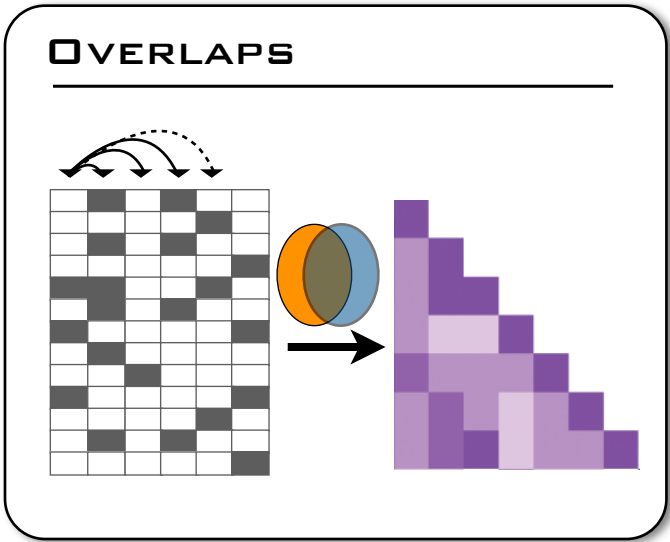


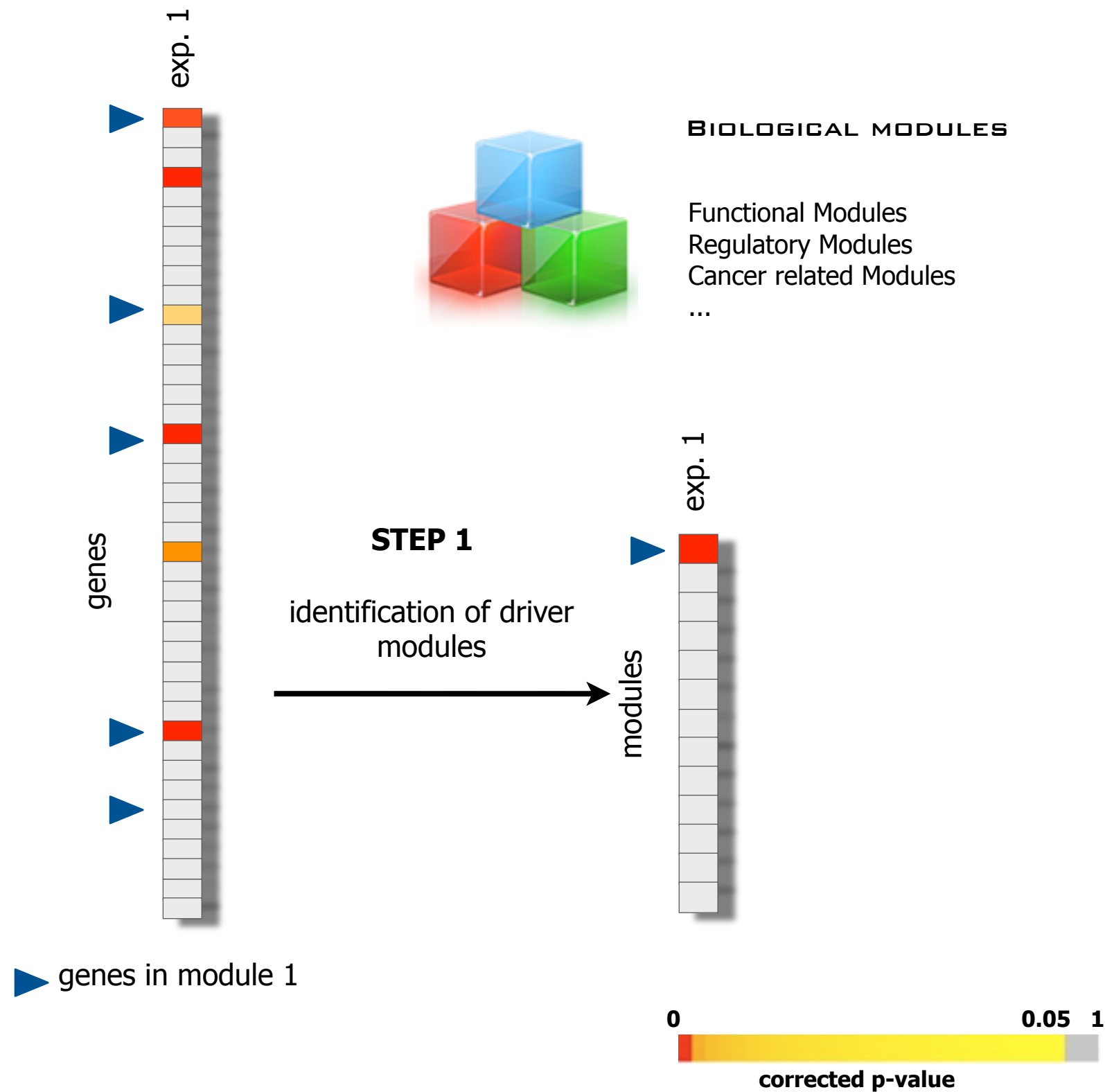
	N experiments	p-value	
Breast Cancer	9	9.99E-16	

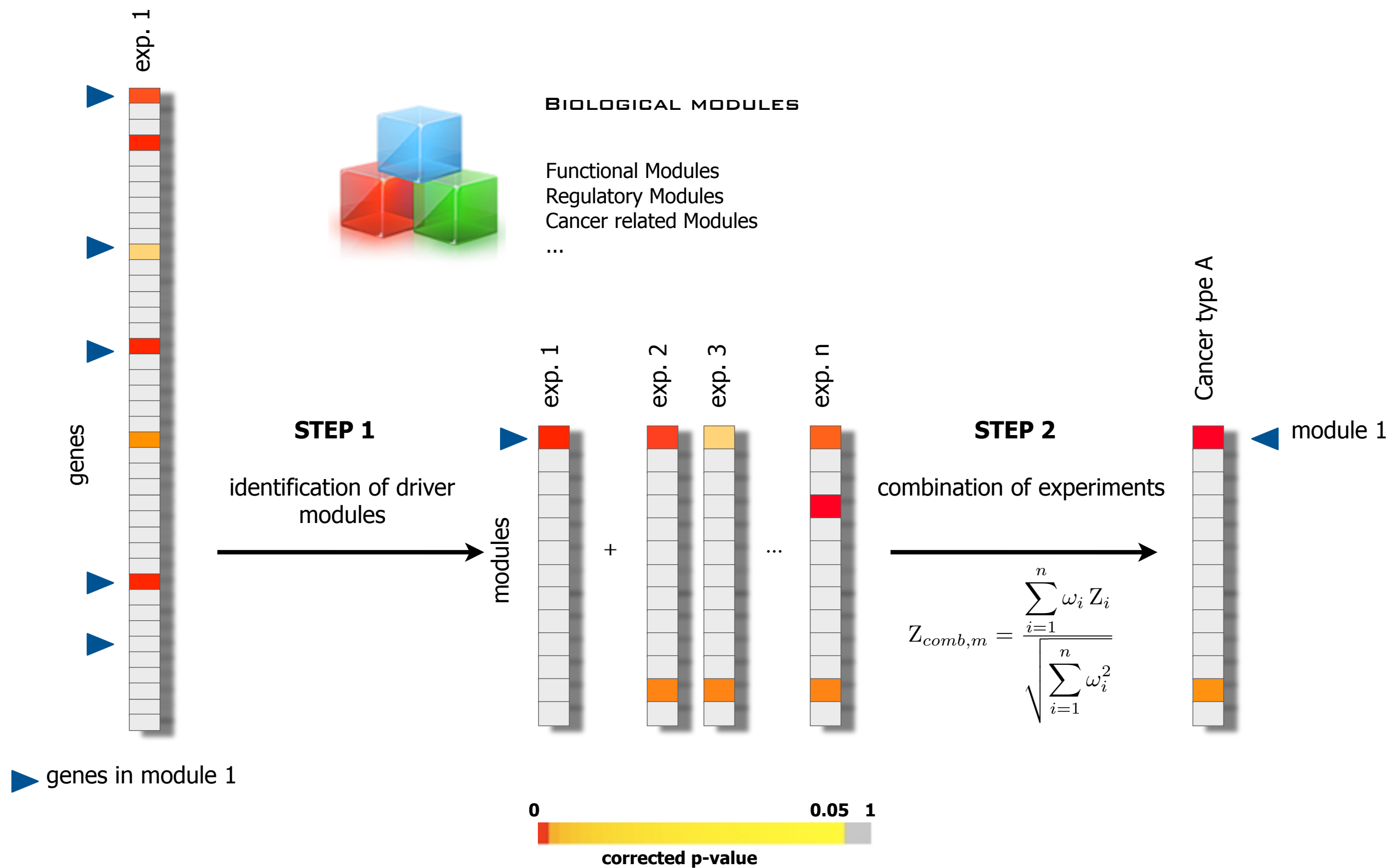
Correlations



Overlaps

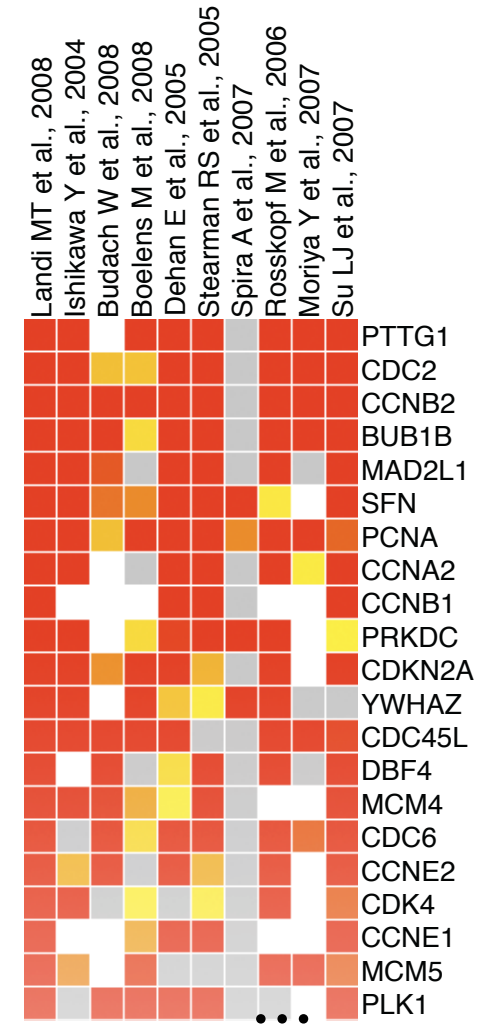






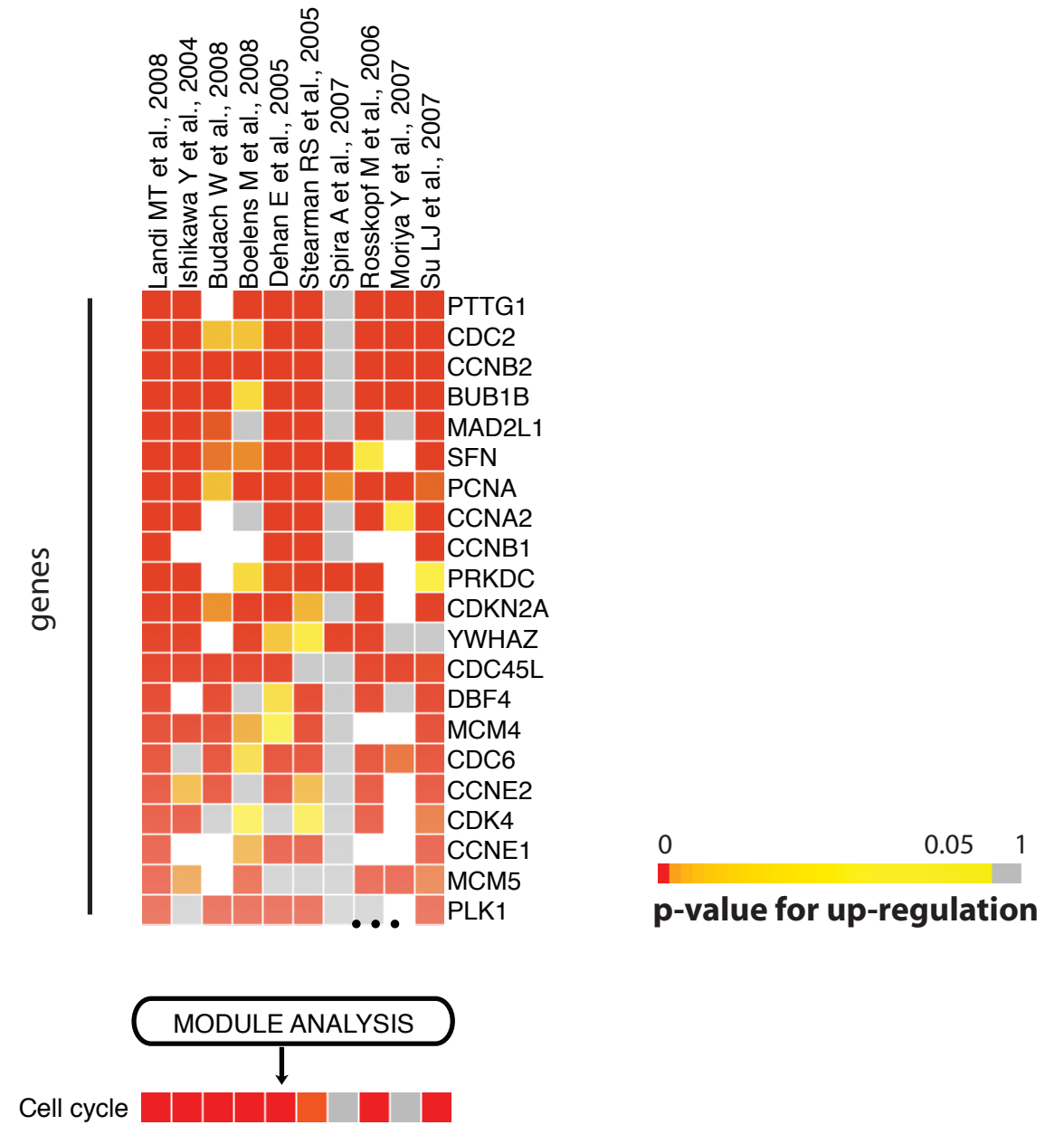
Experiments for lung tissue

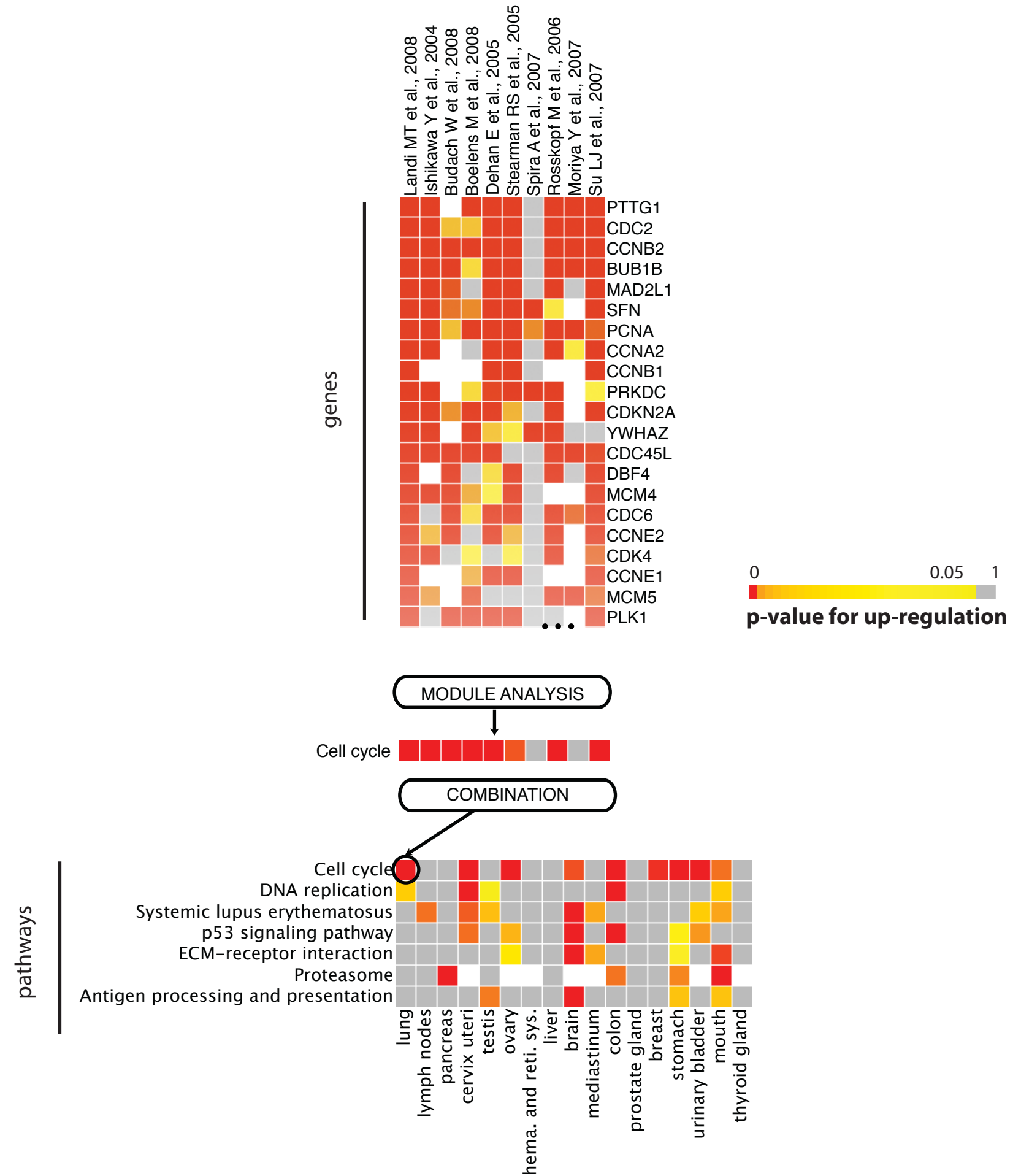
genes



0 0.05 1
p-value for up-regulation

Experiments for lung tissue







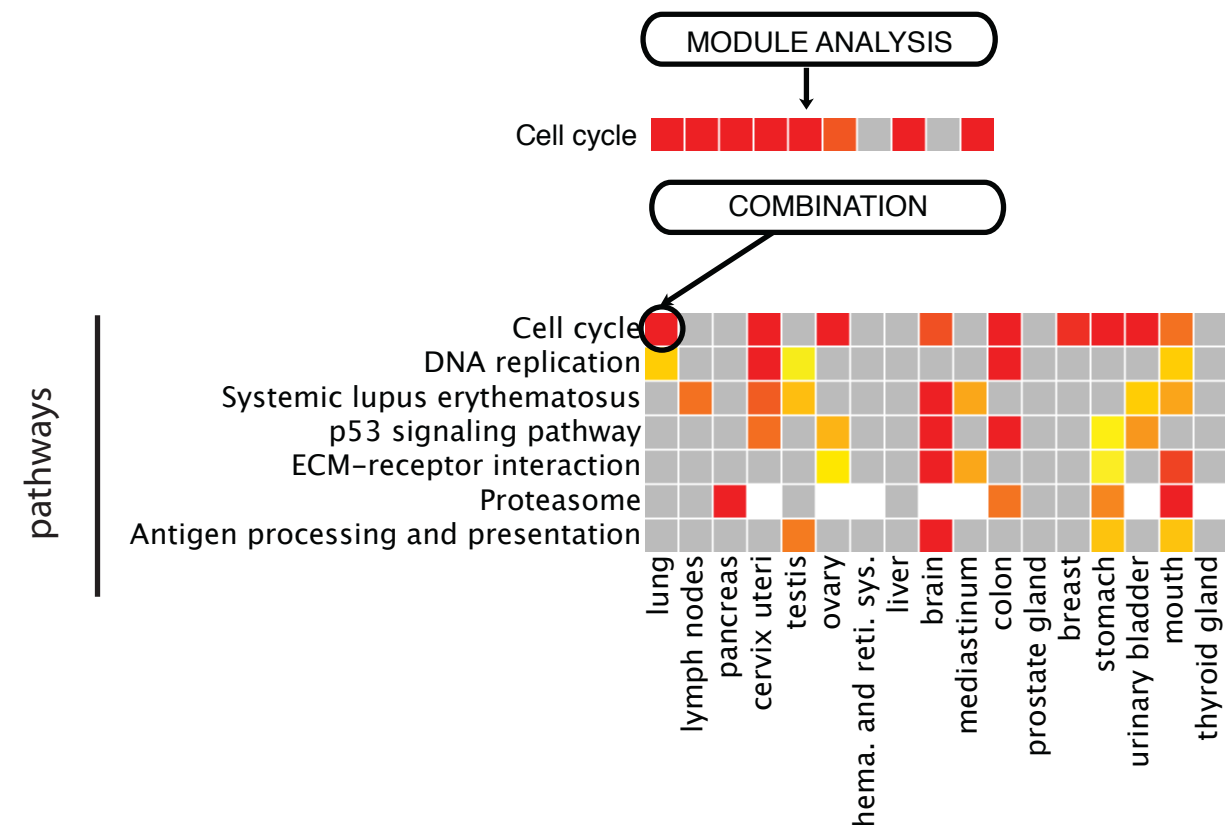
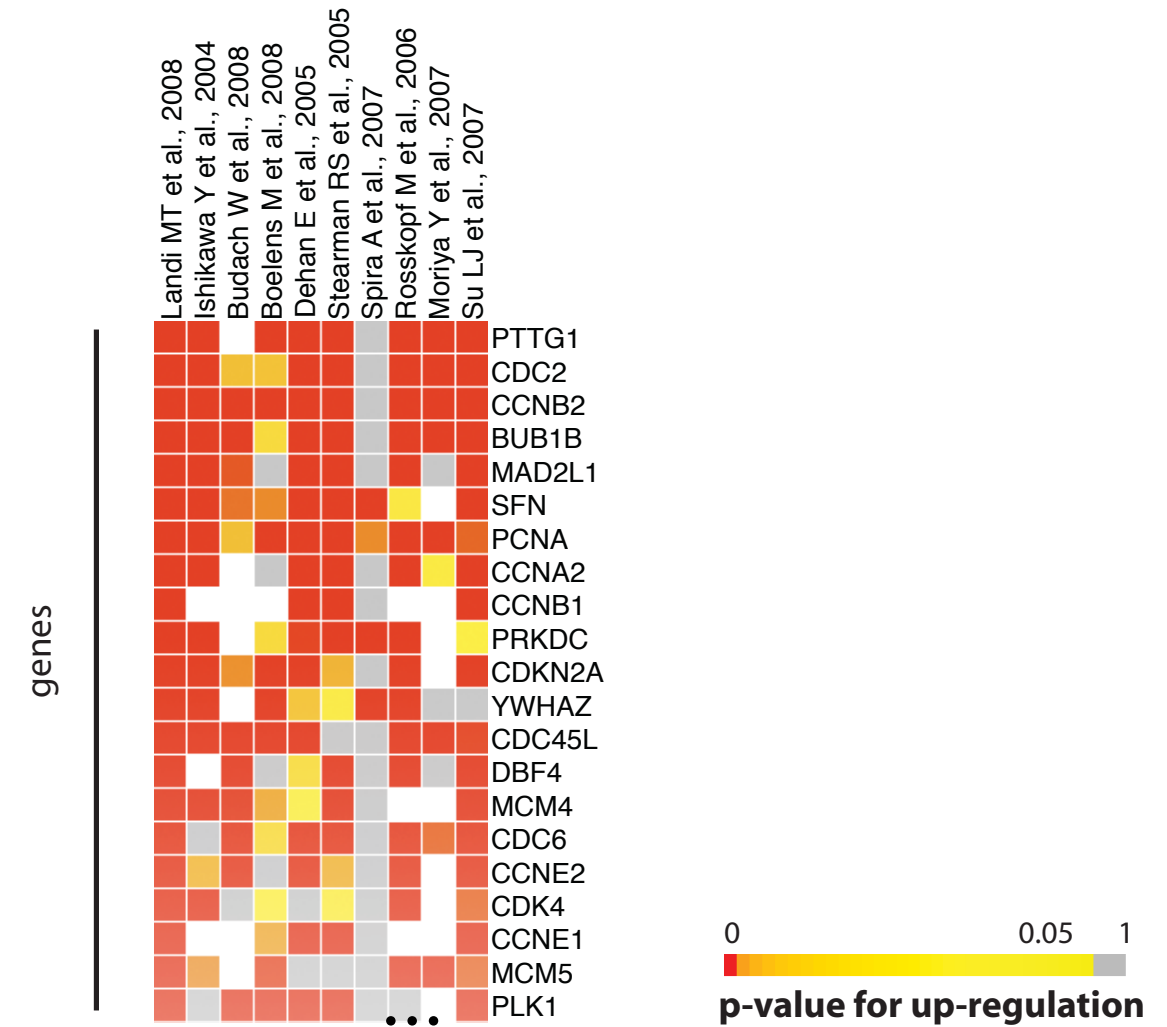
BIOLOGICAL MODULES

Functional Modules

Regulatory Modules

OncoModules

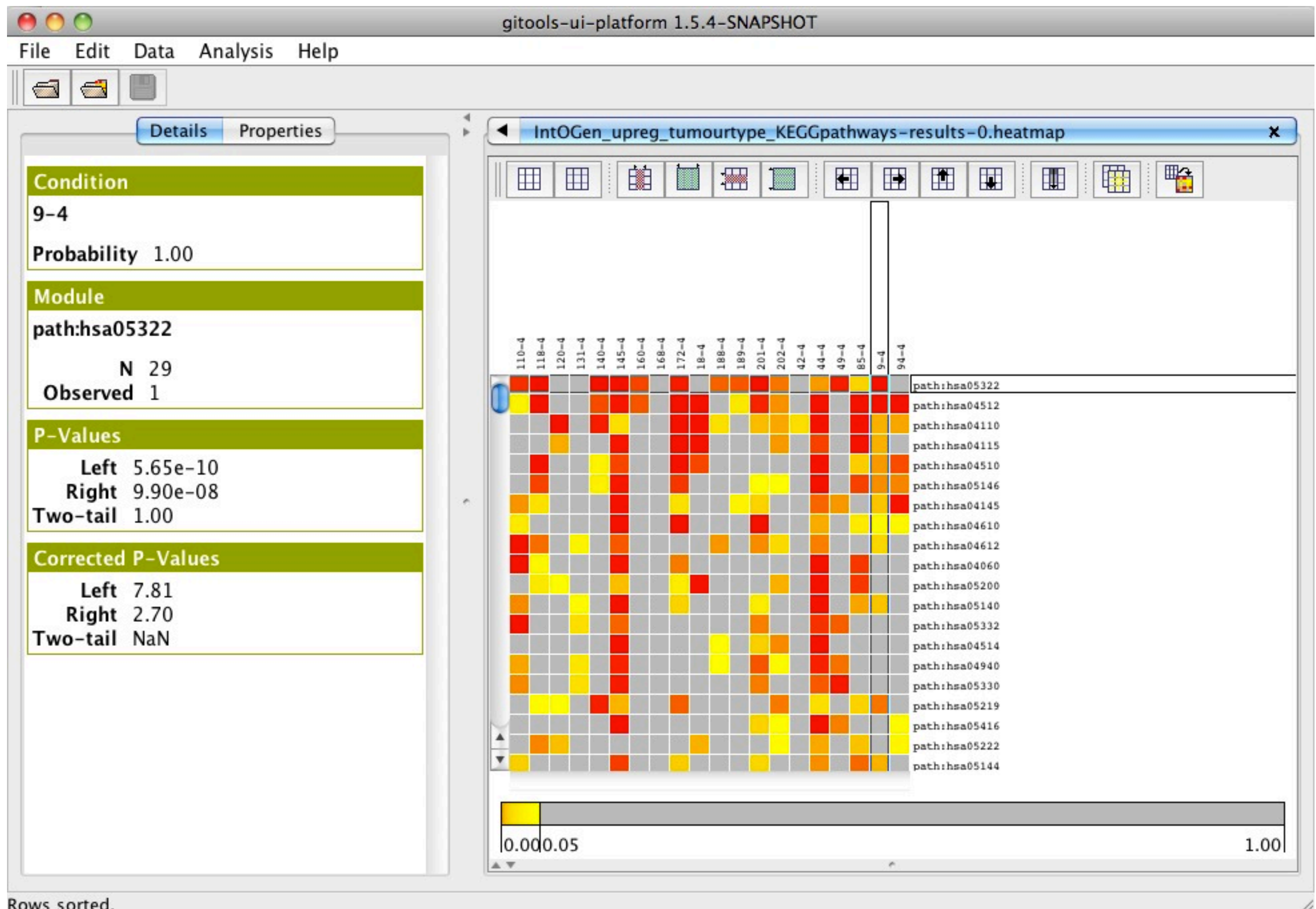
■ ■ ■

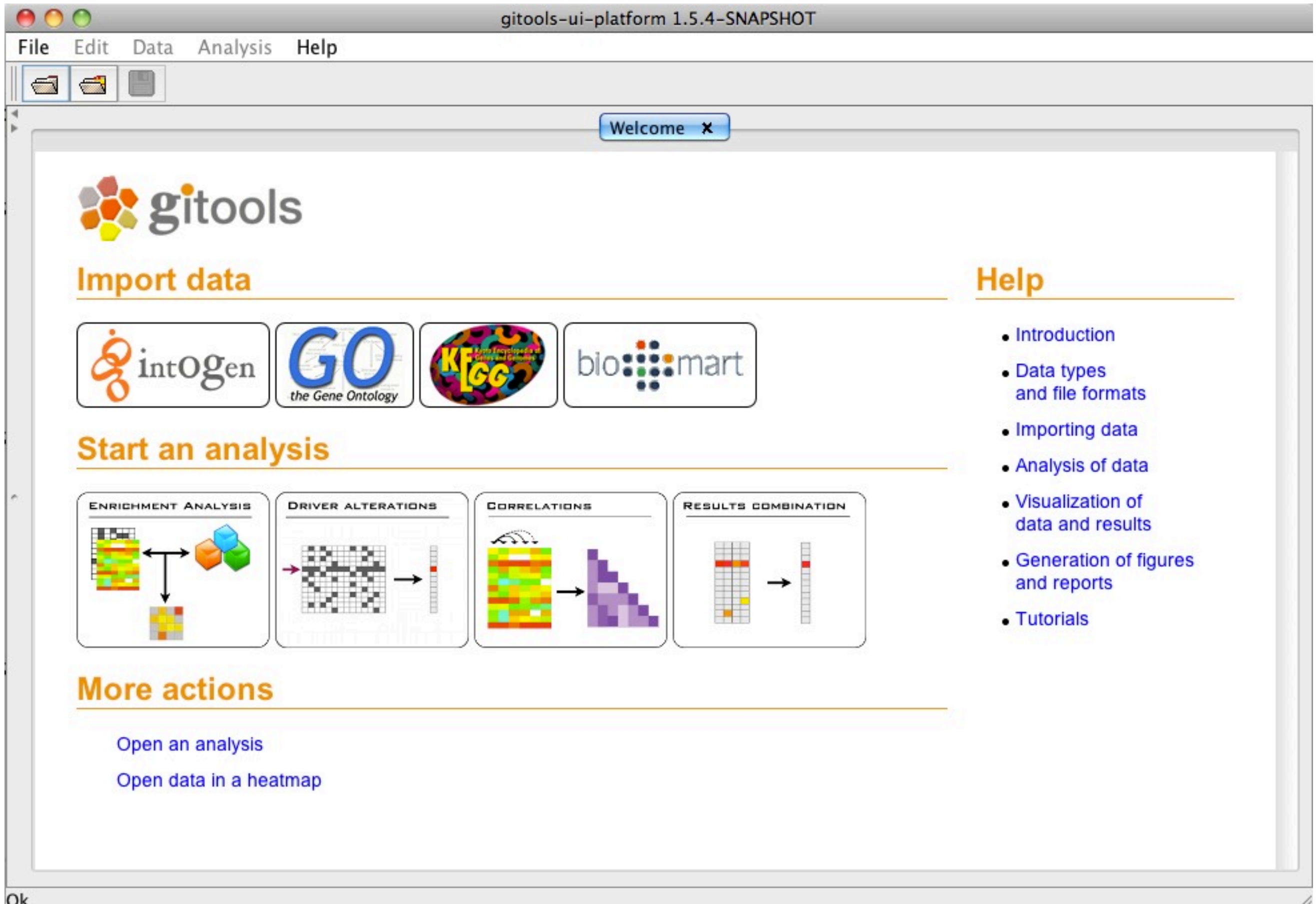




Analysis and visualization of genomic data using interactive heat-maps

Interactive heatmaps







Data



Analysis



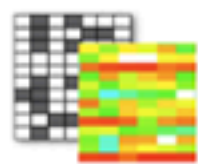
Browse



Export

Three Data Types

Data Matrices



Data Matrices provide values for a number of items (i.e. genes) in several conditions. They can represent continuous or binary values.

Modules



Modules (gene-sets) are lists of genes or other items that share a common biological property. For example, genes in the same pathway.

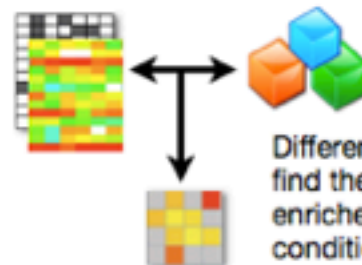
Tables



Tables are files that provide mappings to synonymous ids, descriptions or other information about the items used in the analysis.

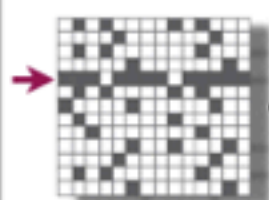
Perform different types of analysis over matrices and modules

Enrichment Analysis (EA)



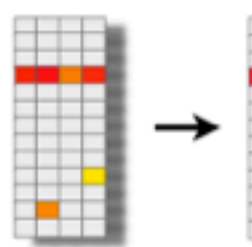
Different EA statistics to find the gene-sets enriched in each of the conditions of your matrix

Oncodrive



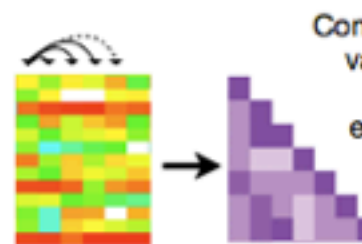
Method to find genes or items more altered than expected by chance taking into account the whole matrix

Results Combination



Z-transform method to combine p-values from different experiments. For example to combine the results for different independent experiments that analyze the same tumor type

Correlations



Compare the pattern of values of all columns against all. For example to compare the log2 ratios of several samples.

Results in Gitools are represented in the form of heat-maps. This allows easy comparison of results for different columns (e.g. conditions) and rows (e.g. genes or genesets)

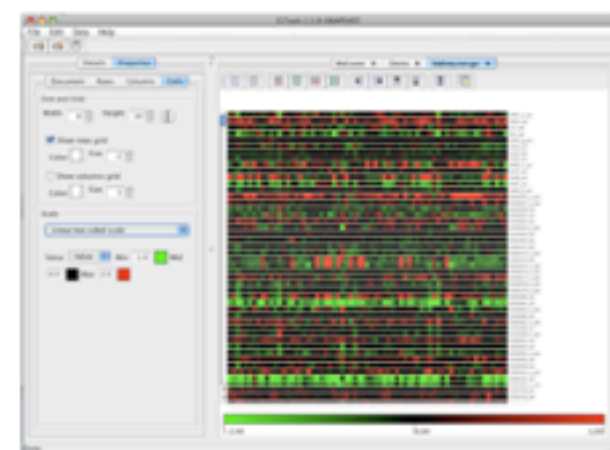
Browse results

Move, search, filter and sort rows and columns. This helps you to explore matrices and results easily



Edit heat-maps

Change colors, grid, font, sizes, labels etc. to display the heat-maps in the more convenient way for your type of data.



Gitools results can be exported in different formats

Export data and results

Export matrices and analysis results in tab separated or excel format.



Export heat-maps as images

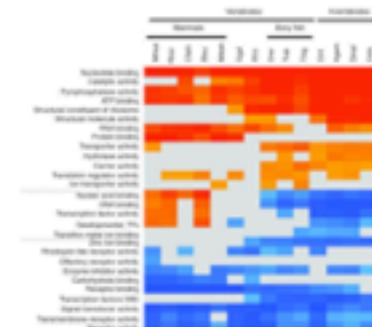


Fig2 in Lopez-Bigas et al Genome Biology 2008

After selecting rows and columns with the browser and choosing the labels, colors and other properties of your results heat-map you can export it as an image for presentations or manuscripts.

Export as HTML

Static html where you can browse the genes, modules and statistics. It is useful to share the analyses with collaborators.



Import data from Biomart, IntOGen and KEGG



Easy to prepare data formats

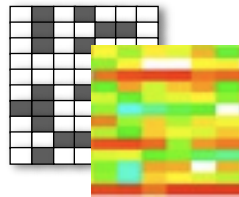
	A	B	C	D
1	exp	exp	exp	exp
2	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
3	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
4	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
5	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
6	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
7	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
8	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
9	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000
10	ENSG00000100000	ENSG00000100000	ENSG00000100000	ENSG00000100000



Data

Three Data Types

Data Matrices



Data Matrices provide values for a number of items (i.e. genes) in several conditions. They can represent continuous or binary values.

Modules



Modules (gene-sets) are lists of genes or other items that share a common biological property. For example, genes in the same pathway.

Tables



Tables are files that provide mappings to synonymous ids, descriptions or other information about the items used in the analysis.

Import data from Biomart, IntOGen and KEGG



Integrative
Onco
Genomics



Easy to prepare data formats

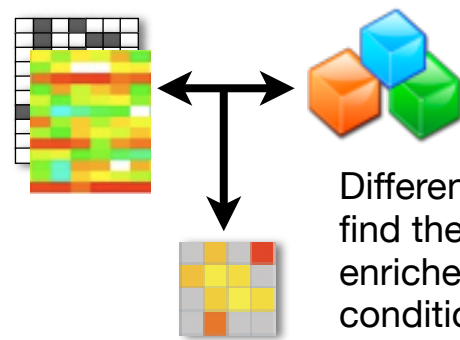
	A	B	C	D
1	enag	exprRBP2_fc_0vs27	exprRBP2_fc_0vs96	exprRBP2_fc_27vs96
2	ENSG00000103326	1.03804564	1.037601202	-1.007718185
3	ENSG00000126773	-1.058802415	1.01532742	1.075031125
4	ENSG00000170619	1.176885084	1.222509964	1.147081081
5	ENSG00000126775	1.102293892	-1.035618755	-1.141556227
6	ENSG00000126777	1.152791949	1.375878556	1.193518533
7	ENSG00000170613	1.041398515	-1.003648134	-1.045197676
8	ENSG00000126778	1.032234841	-1.003728351	-1.036083375
9	ENSG00000170615	-1.079261284	1.03497675	1.117010335
10	ENSG00000156973	1.103467112	1.314476077	1.191223611



Analysis

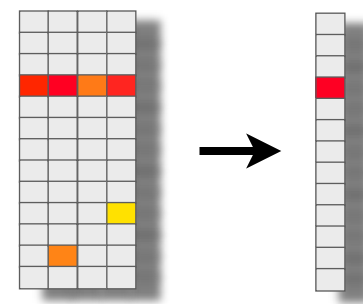
Perform different types of analysis over matrices and modules

Enrichment Analysis (EA)



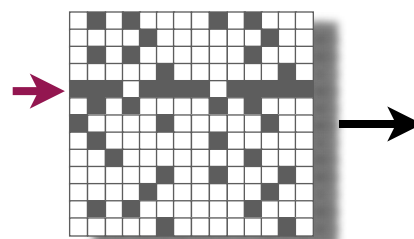
Different EA statistics to find the gene-sets enriched in each of the conditions of your matrix

Results Combination



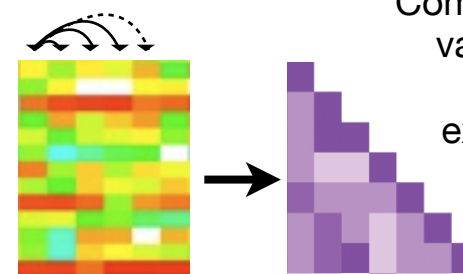
Z-transform method to combine p-values from different experiments. For example to combine the results for different independent experiments that analyze the same tumor type

Oncodrive



Method to find genes or items more altered than expected by chance taking into account the whole matrix

Correlations



Compare the pattern of values of all columns against all. For example to compare the log2 ratios of several samples.

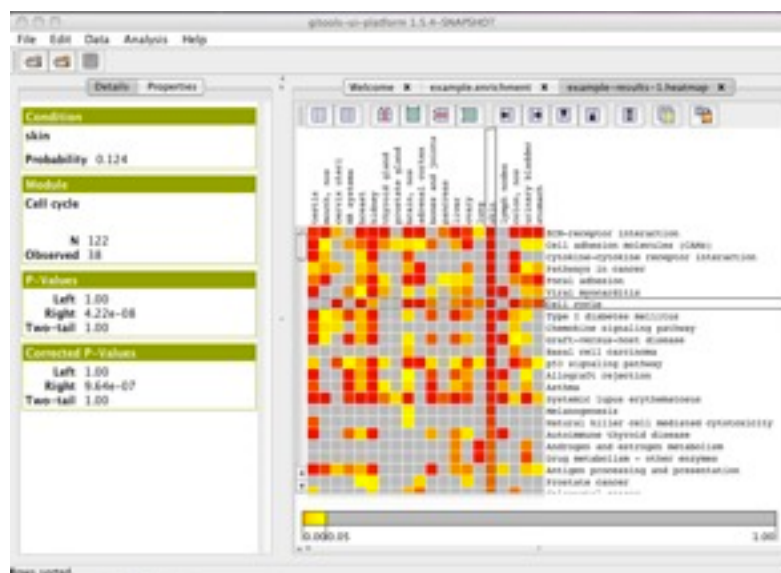


Browse

Results in Gtools are represented in the form of heat-maps. This allows easy comparison of results for different columns (e.g. conditions) and rows (e.g. genes or genesets)

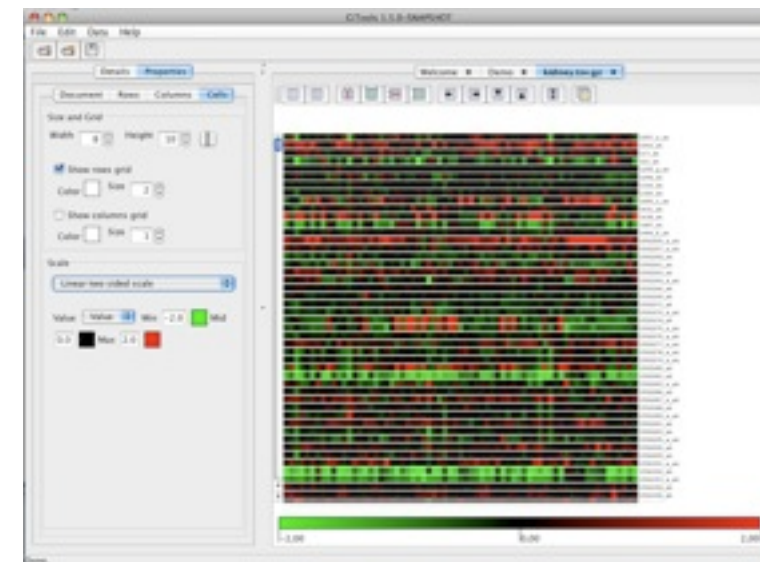
Browse results

Move, search, filter and sort rows and columns. This helps you to explore matrices and results easily



Edit heat-maps

Change colors, grid, font, sizes, labels etc. to display the heat-maps in the more convenient way for your type of data.





Export

Gitools results can be exported in different formats

Export data and results

Export matrices and analysis results in tab separated or excel format.



Export as HTML

Static html where you can browse the genes, modules and statistics. It is useful to share the analyses with collaborators.



Export heat-maps as images

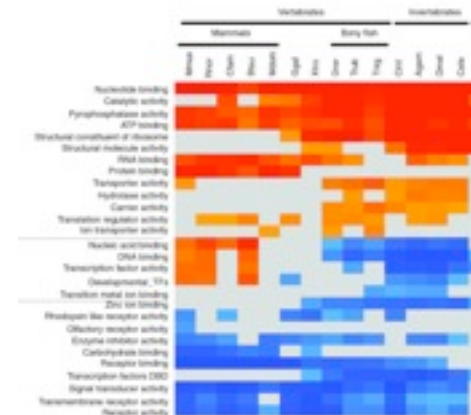


Fig2 in Lopez-Bigas et al Genome Biology 2008

After selecting rows and columns with the browser and choosing the labels, colors and other properties of your results heat-map you can export it as an image for presentations or manuscripts.

