Differential expression in Babelomics 5

Matías Marín Falco

September 29th,2016



GDA

International Course on

Genomic Data Analysis

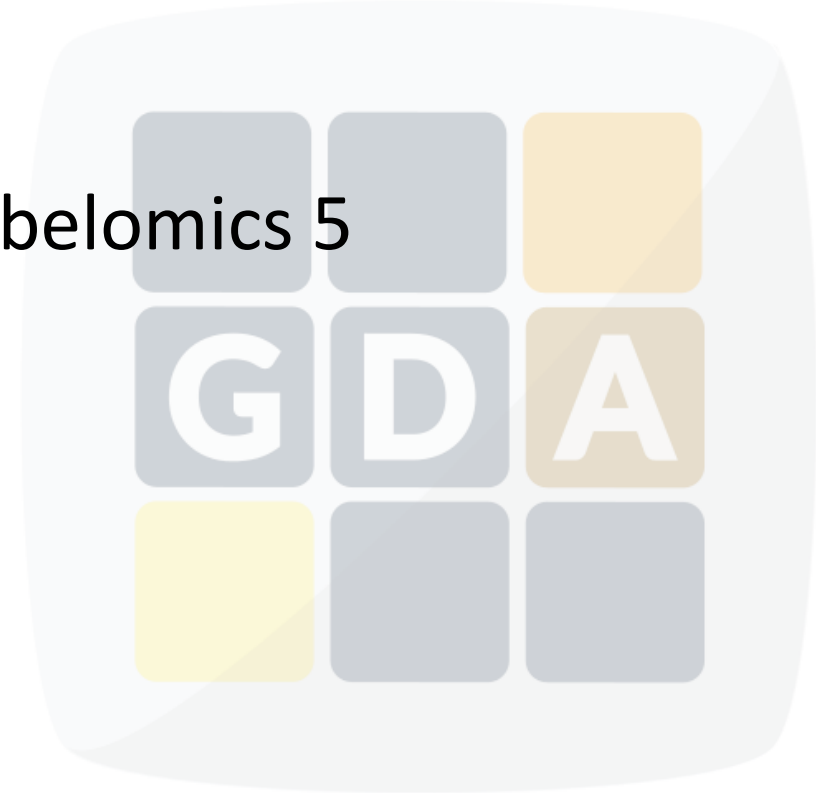


PRINCIPE FELIPE

CENTRO DE INVESTIGACION

Outline

- Differential expression
- Differential expression in Babelomics 5
- Exercises



Differential expression

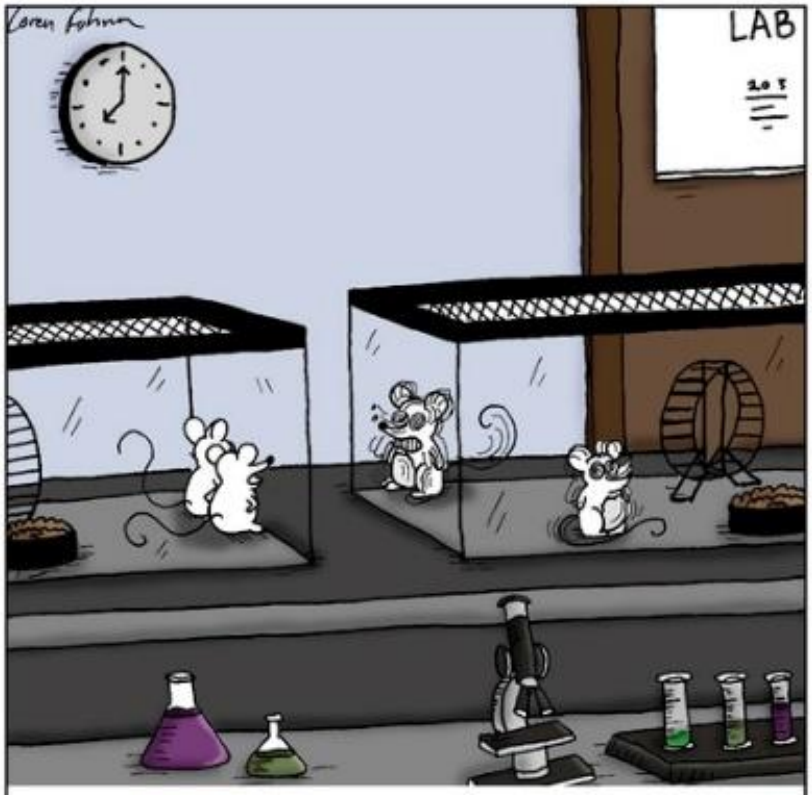
Test an hypothesis $\Rightarrow$ comparative experiment

Take two randomized groups

- Disease/Treatment group
- Control group

$\Downarrow$

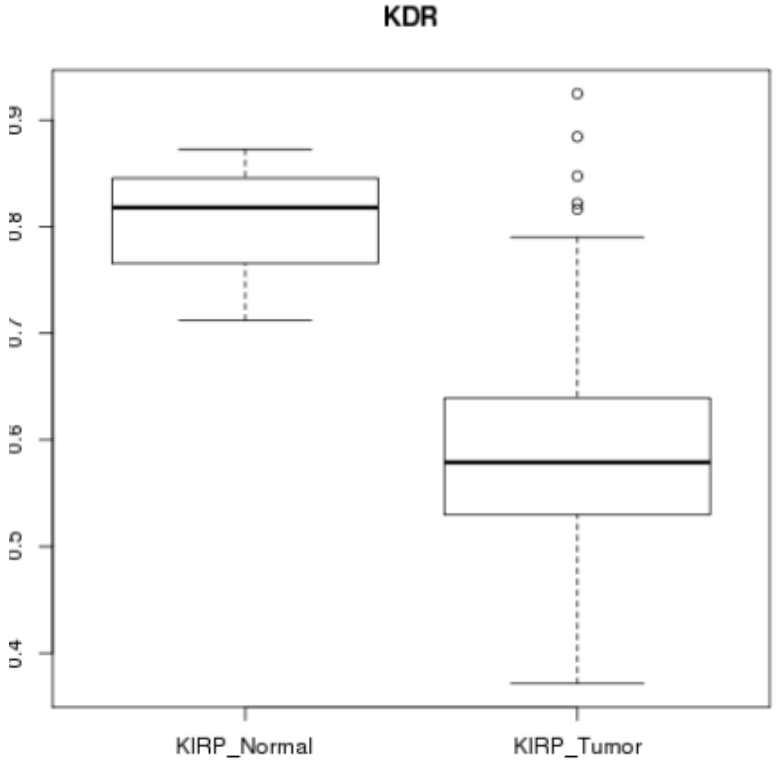
Find significant differences

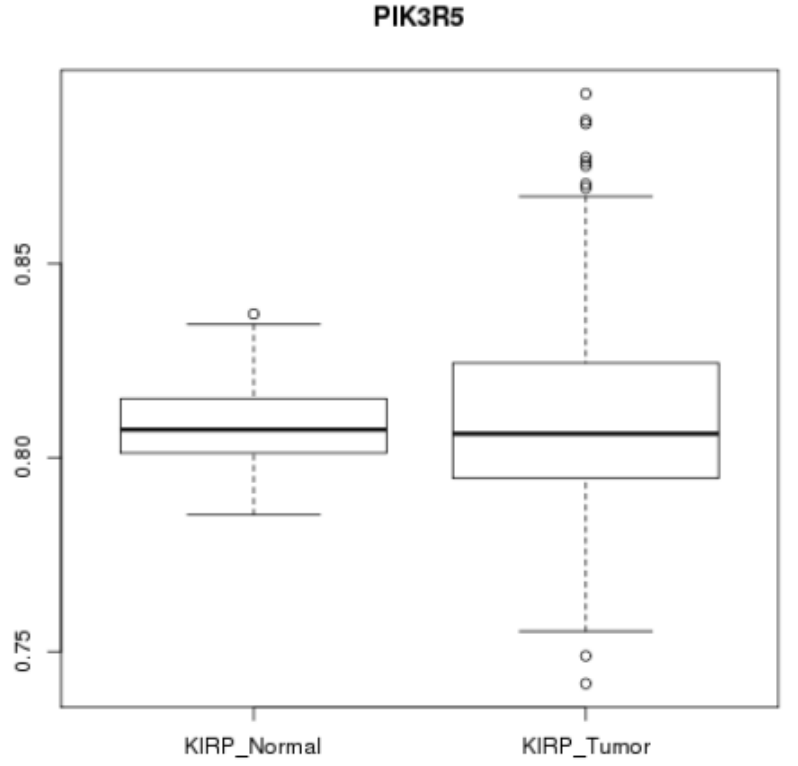


"Well, I guess we're the control group."

Differential expression

Genes that show statistically significant differences in expression level between conditions





Differential expression

- We obtain some data related to expression levels under two experimental conditions

$$X \sim N(\mu_X, \sigma^2)$$

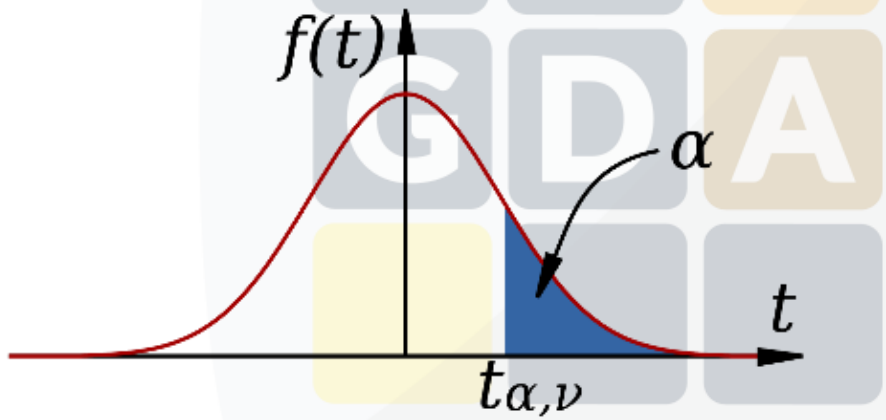
$$Y \sim N(\mu_Y, \sigma^2)$$

→ Our hypothesis: $H_0 : \mu_X = \mu_Y$

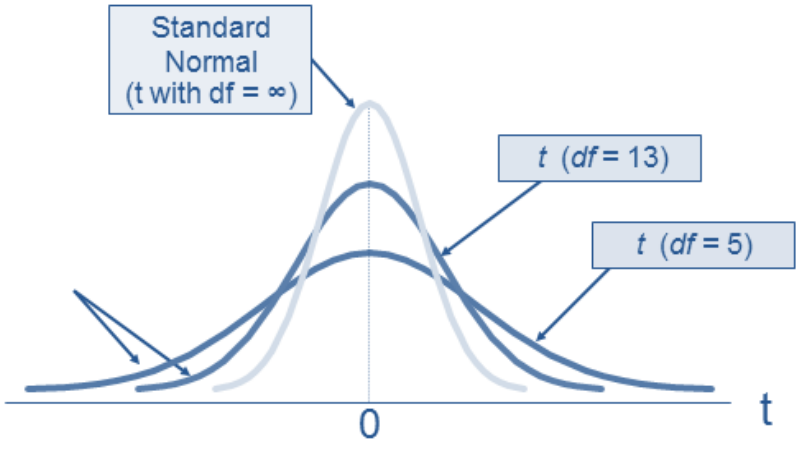
$H_1 : \mu_X \neq \mu_Y$

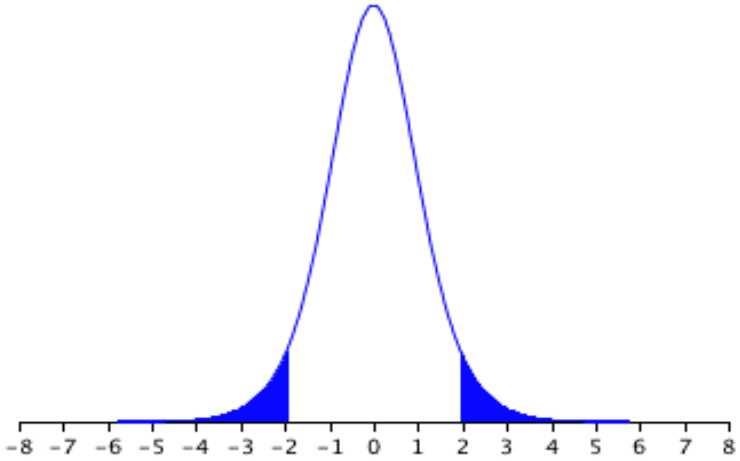
- We calculate:

$$T = \frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{n} + \frac{1}{m}}}$$



Differential expression

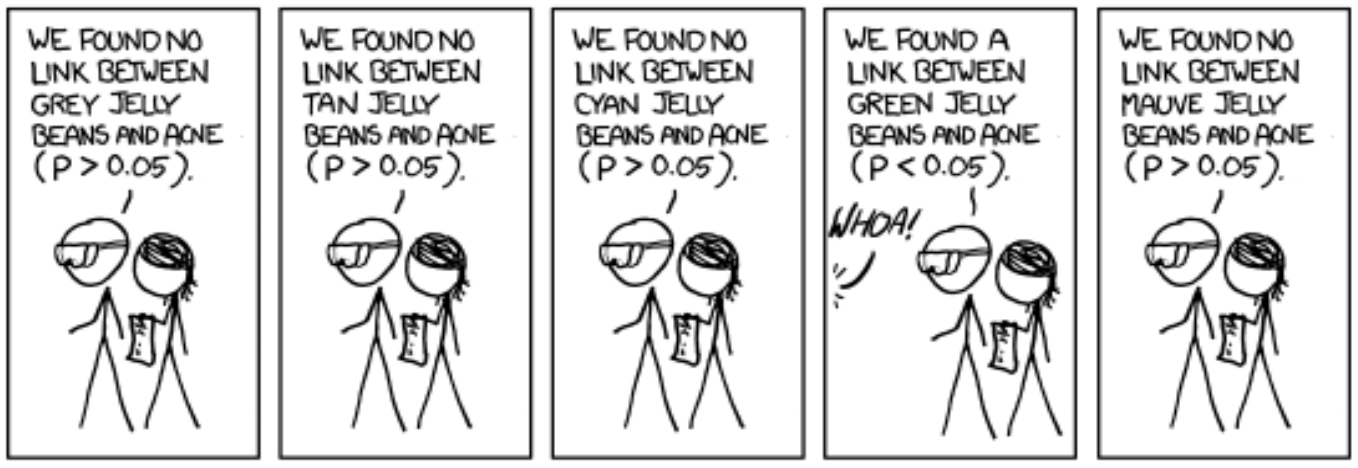




one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965

Differential expression

In order to avoid find randomly significant results, we **must correct p-value**



Normalization methods for DE

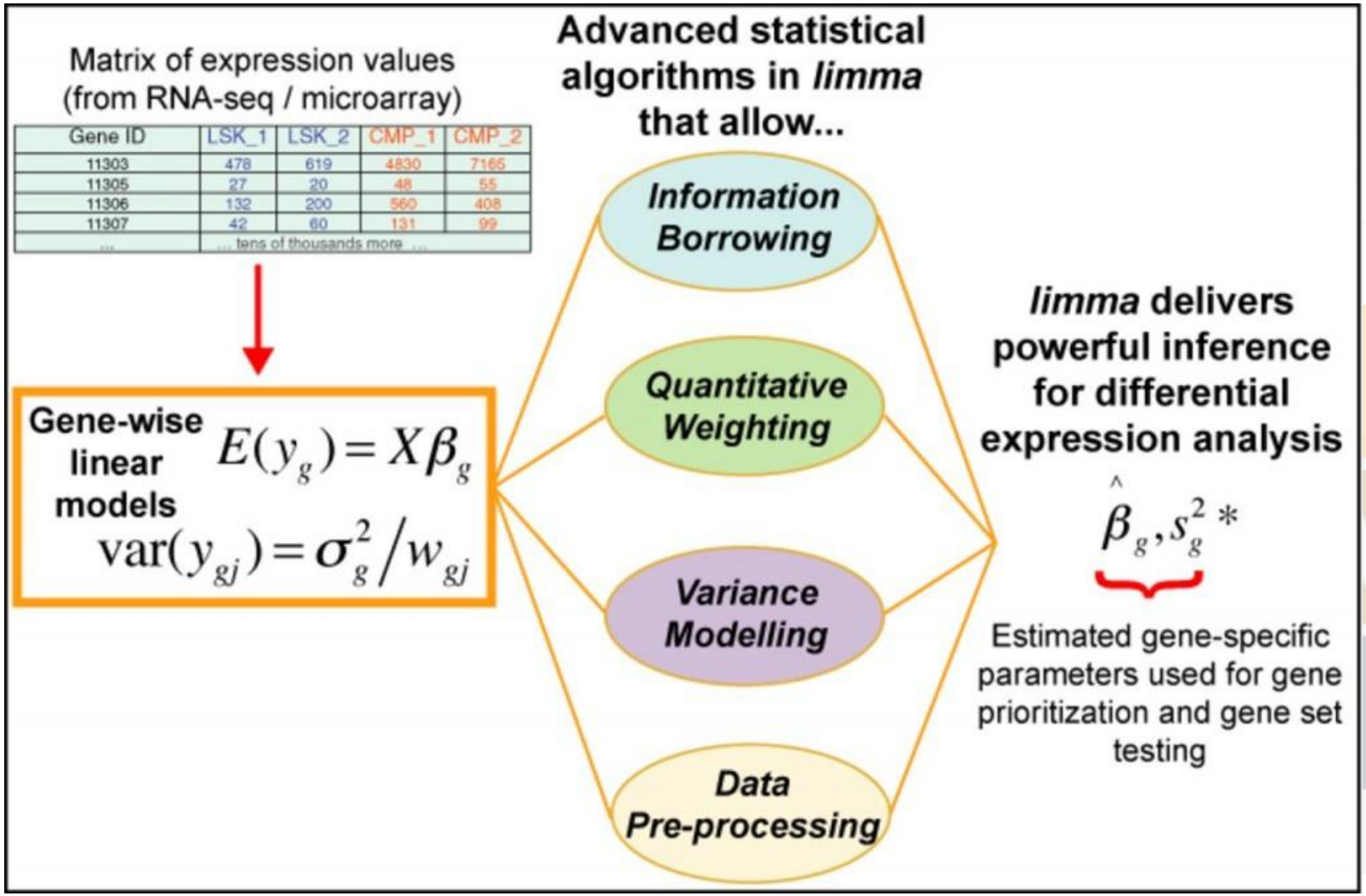
	Gene length	Library depth	RNA composition
Gene length	✓	✗	✗
Library depth	✗	✓	✗
RPKM	✓	✓	✗
TMM	✗	✓	✓
Quantiles	✗	✓	✓



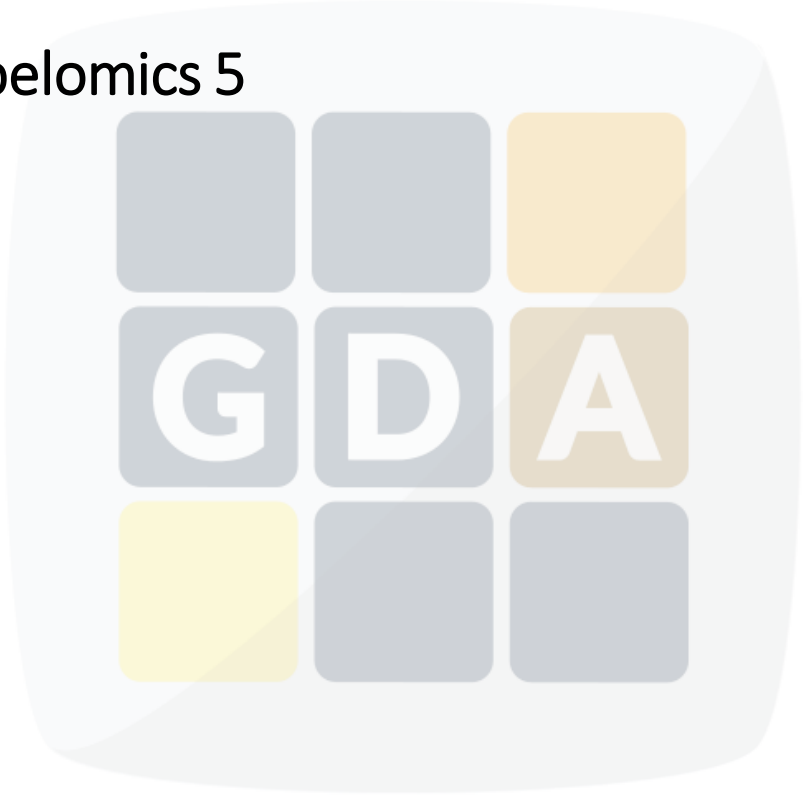
Differential expression methods

- **edgeR** (Robinson et al., 2010): Exact test based on Negative Binomial distribution
- **DESeq** (Anders & Huber, 2010): Exact test based on Negative Binomial distribution
- **NOISeq** (Tarazona et al., 2015): Non-parametric test, replicates not needed
- **voom+limma** (Law et al., 2014): Bayesian method for microarrays, adapted with VOOM

Differential expression methods



Differential expression in Babelomics 5



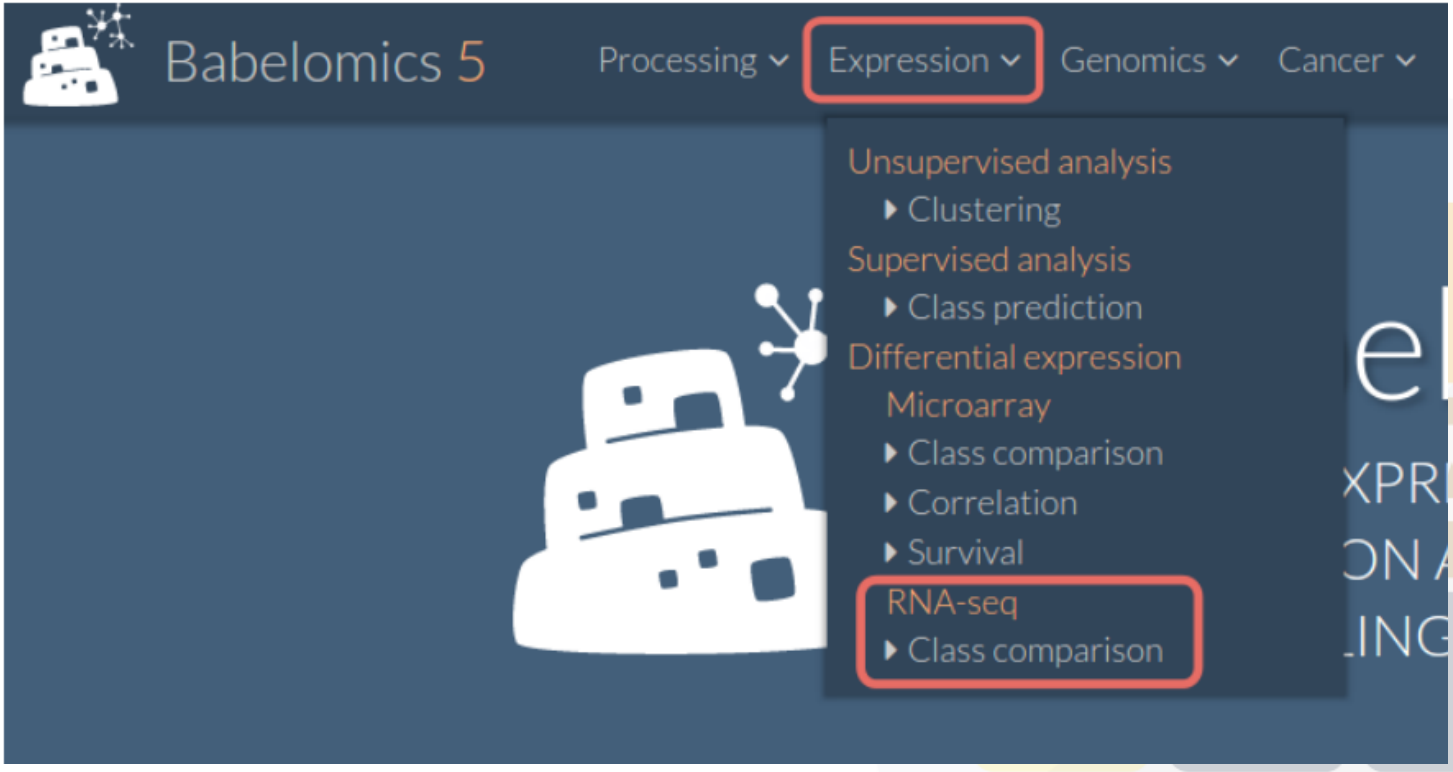
Differential expression in Babelomics 5

RNA-Seq Differential Expression

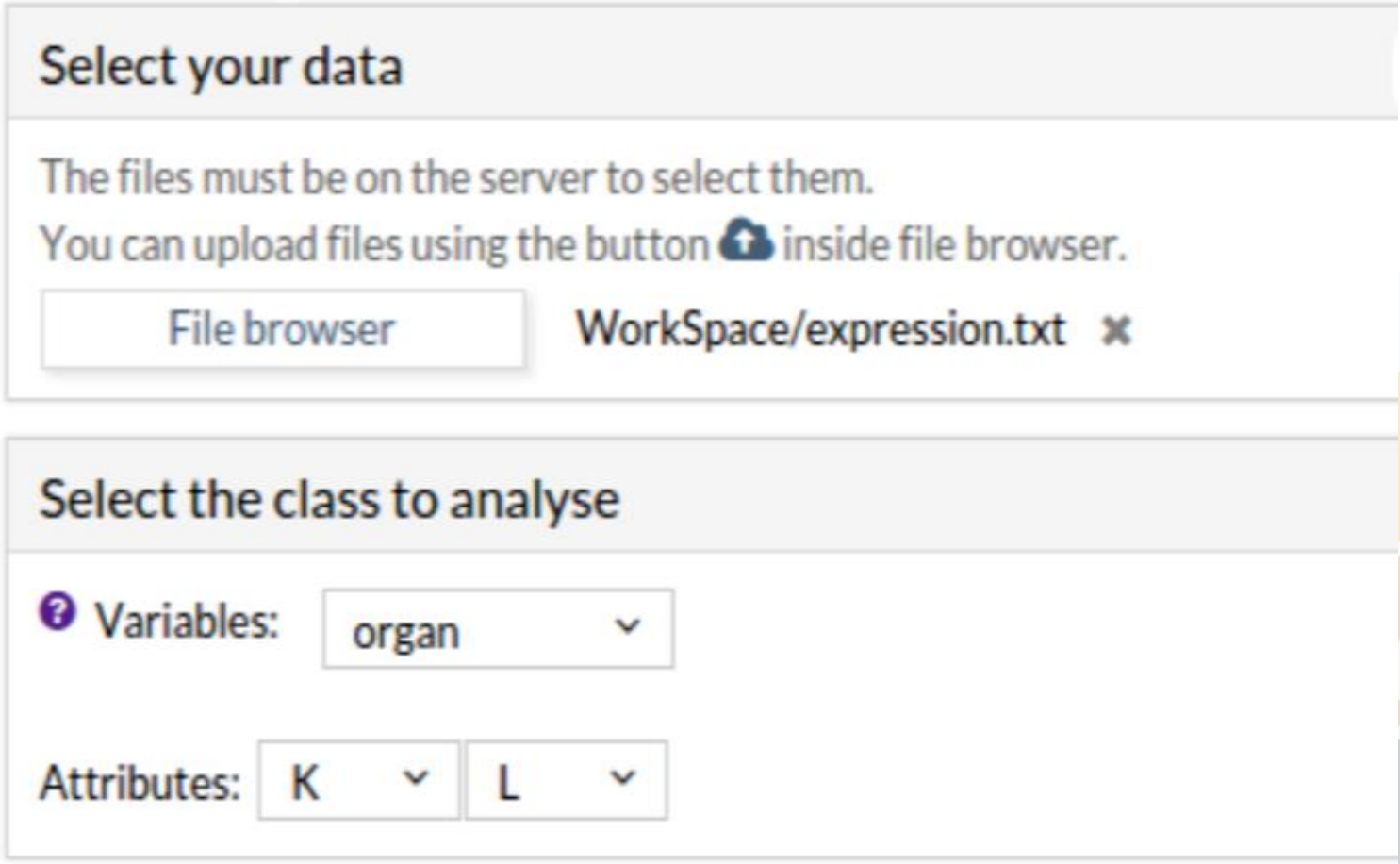
- Two class comparison (data must be edited)
- Differential expression method: **voom+limma**
  - performed well under many conditions
  - relatively unaffected by outliers
  - computationally fast

Differential expression in Babelomics 5

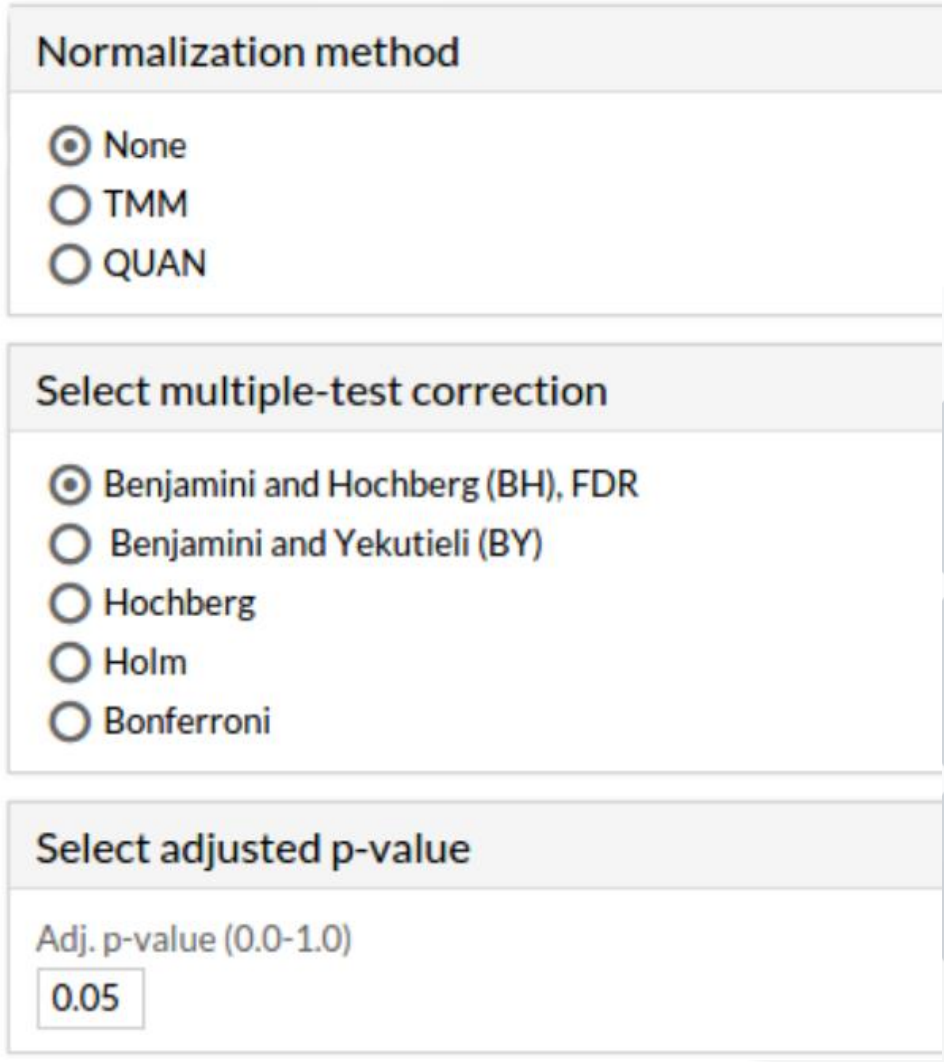
Where can we find RNA-Seq Differential Expression in Babelomics 5?



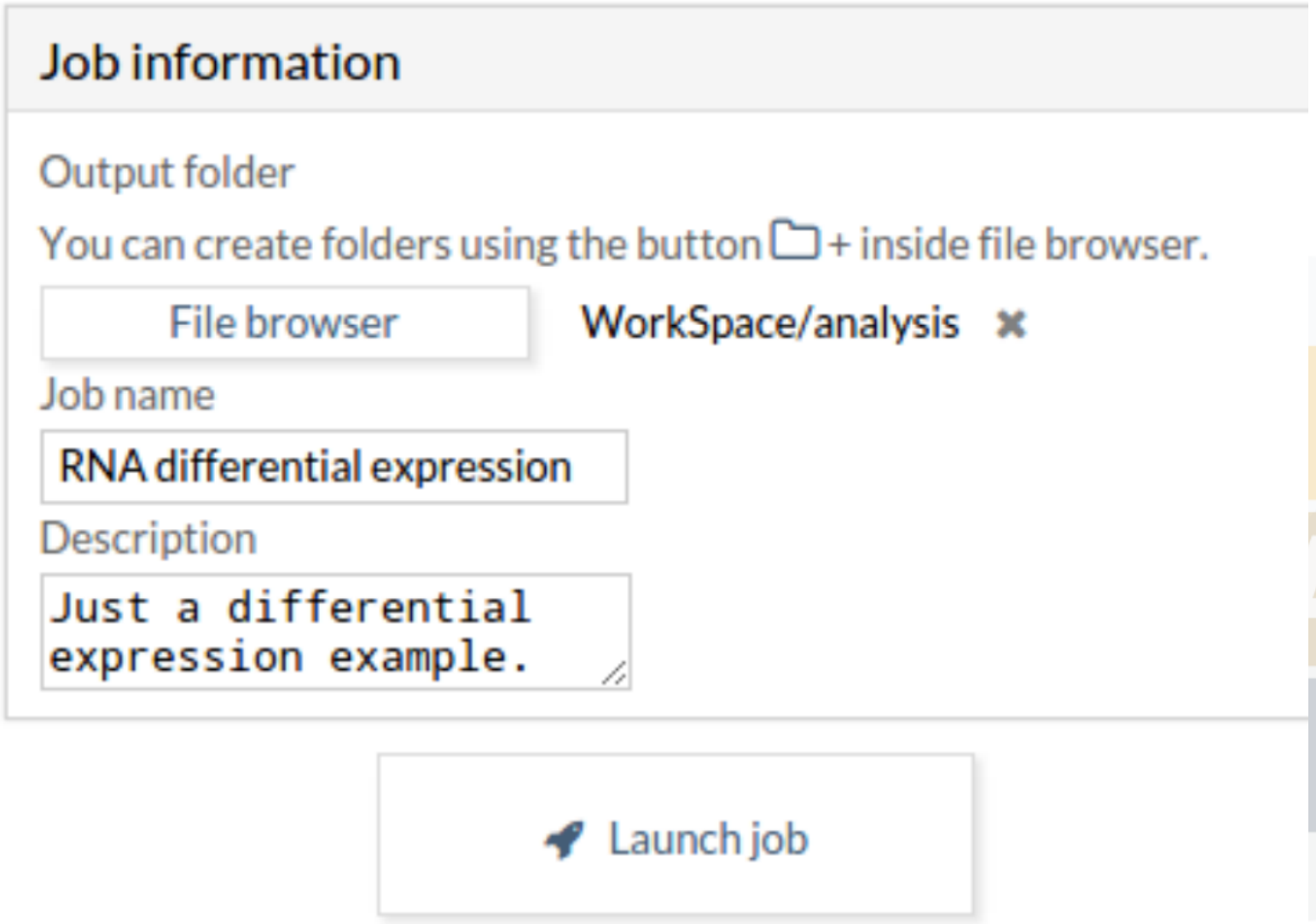
Differential expression in Babelomics 5



Differential expression in Babelomics 5

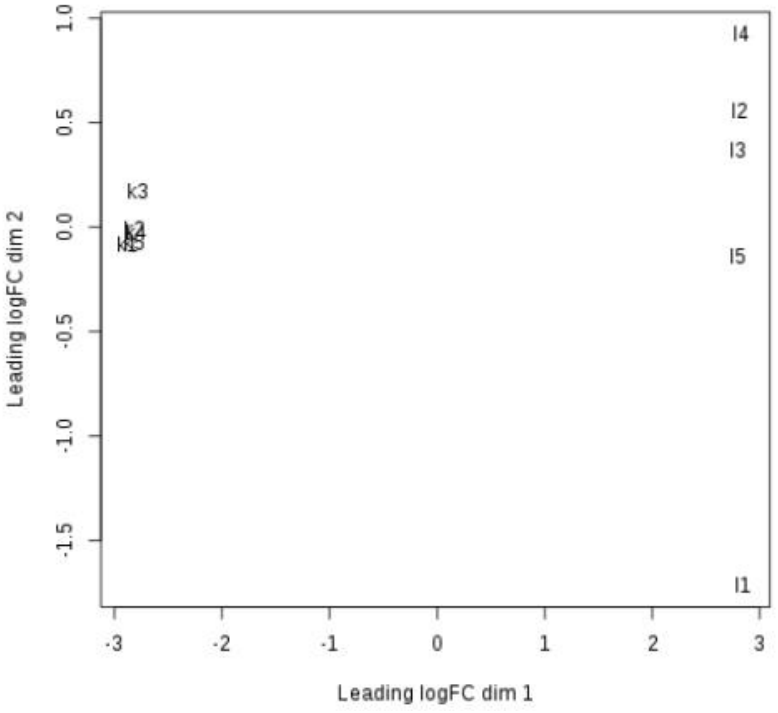


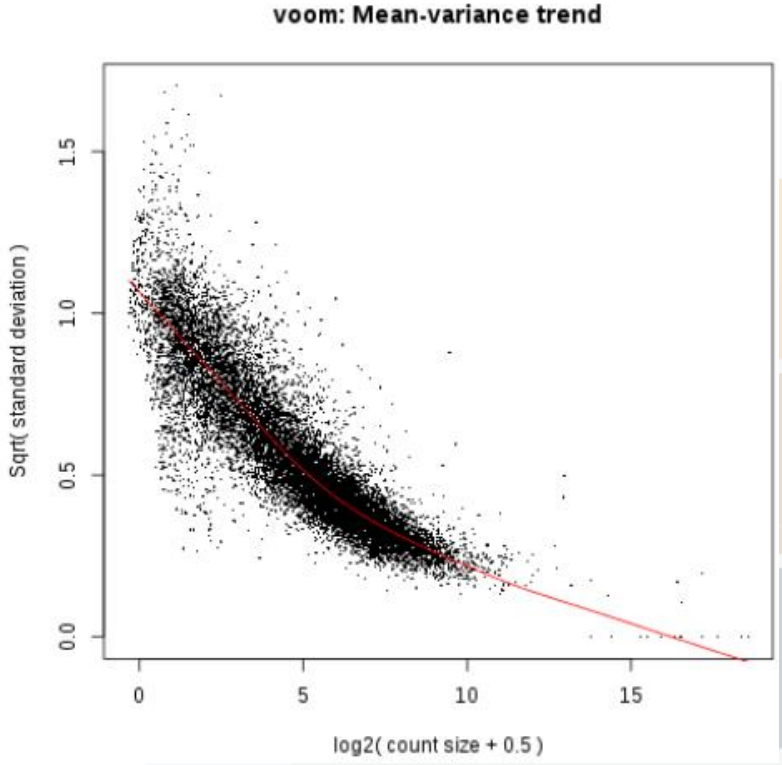
Differential expression in Babelomics 5



Results

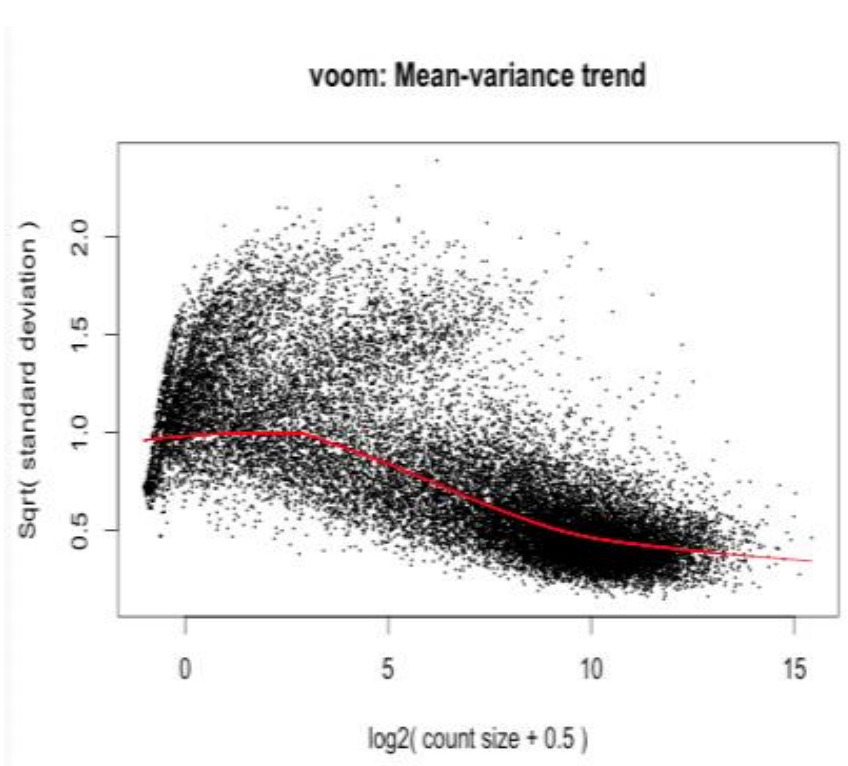
Voom graphics

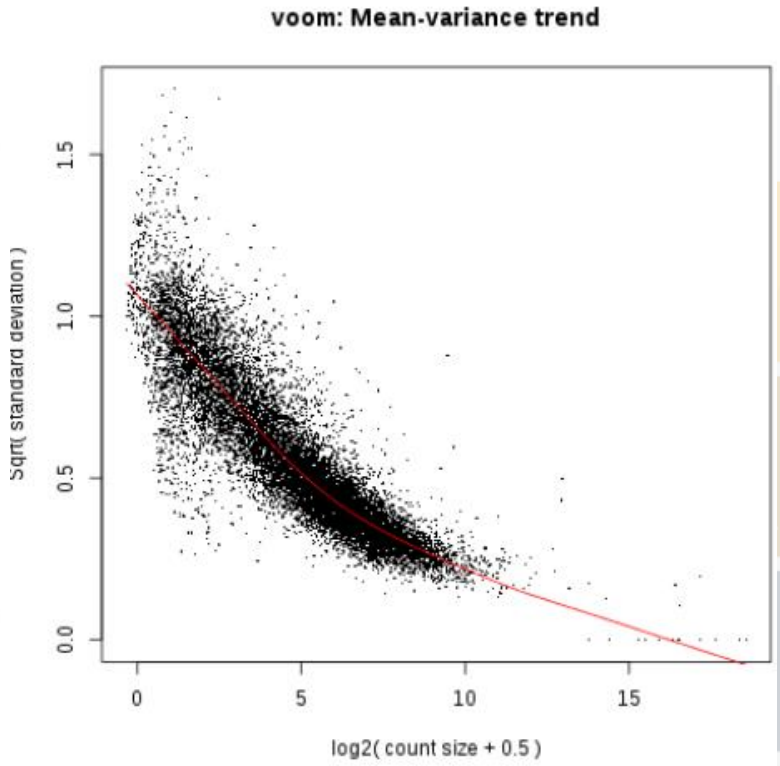




Results

Voom graphics





Results

Table of results

Significative file

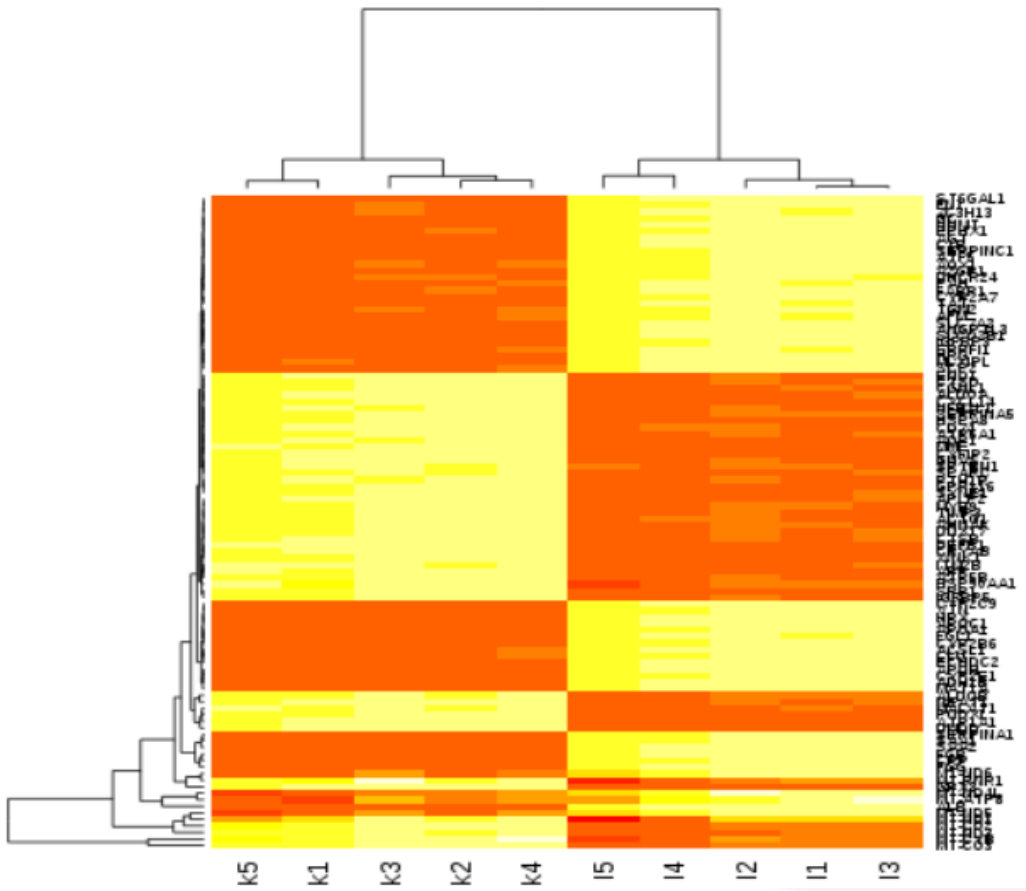
File *diffexp_Sig_results.txt*

#NAMES	statistic	p-value	adjusted p-value
MT-CO2	3595.25	2.8e-56	1.5e-52
MT-ND3	3345.54	1.1e-55	4.4e-52
MT-CO3	1100.93	1.3e-46	4.1e-43
MT-ATP8	549.62	5.8e-41	1.6e-37
MT-ND4L	492.59	4.6e-40	1.1e-36
ALDOB	318.32	1.7e-36	3e-33
MT-ND6	283.2	1.5e-35	2.4e-32
GPX3	249.23	1.6e-34	2.4e-31
MT-CYB	203.45	7.4e-33	9.2e-30
MT-RNR1	170.09	2.1e-31	2.2e-28

11285 Results

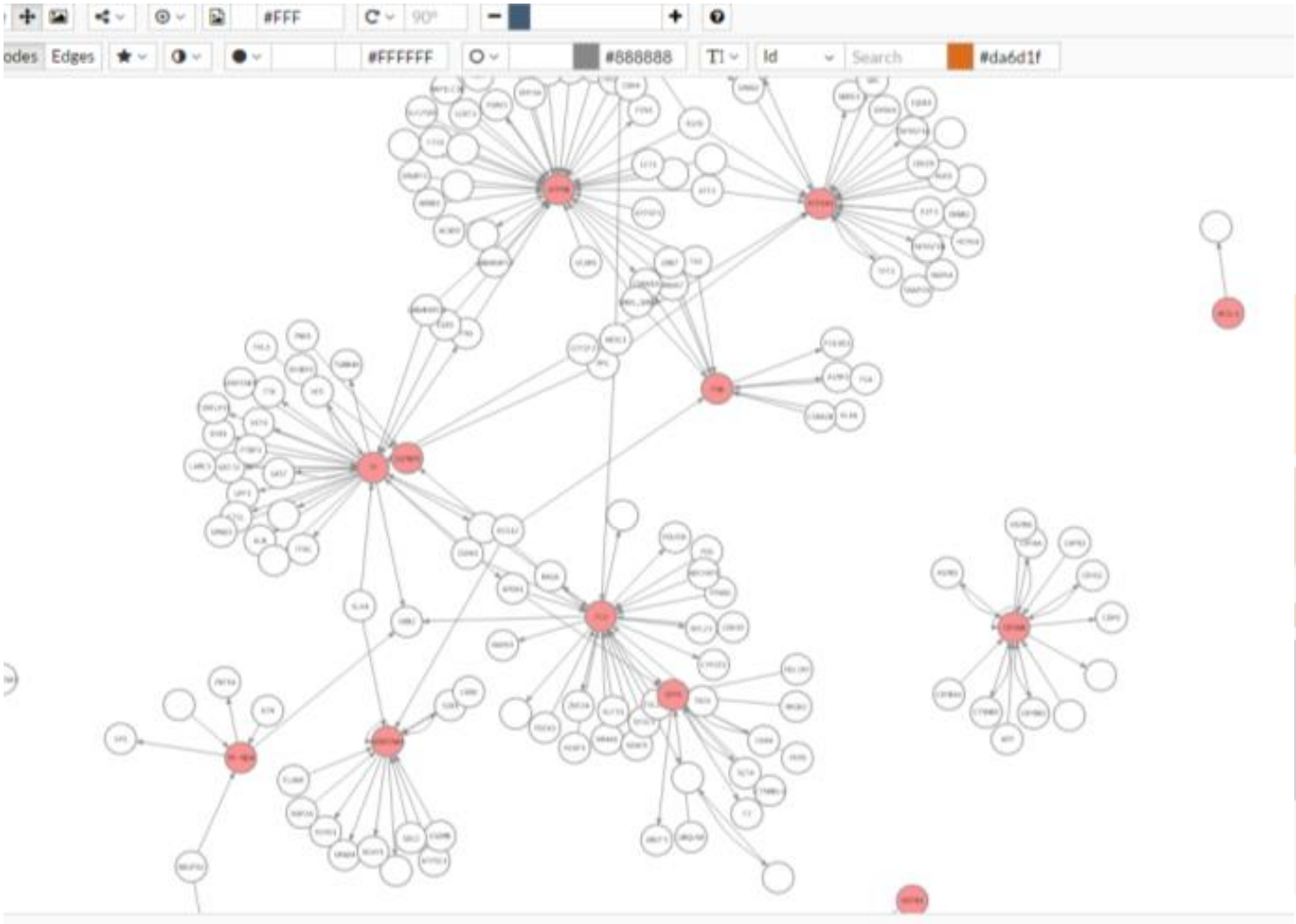
Results

Heatmap



Results

Network Viewer



Results

Continue processing

Single enrichment analysis

Top list vs Genome diffexp_toplist_results.txt

Send to Single enrichment

Bottom list vs Genome diffexp_bottomlist_results.txt

Send to Single enrichment

Top list vs Bottom list

Send to Single enrichment

Gene set enrichment analysis

T statistic results diffexp_t_results.txt

Send to Gene set enrichment

Exercises