

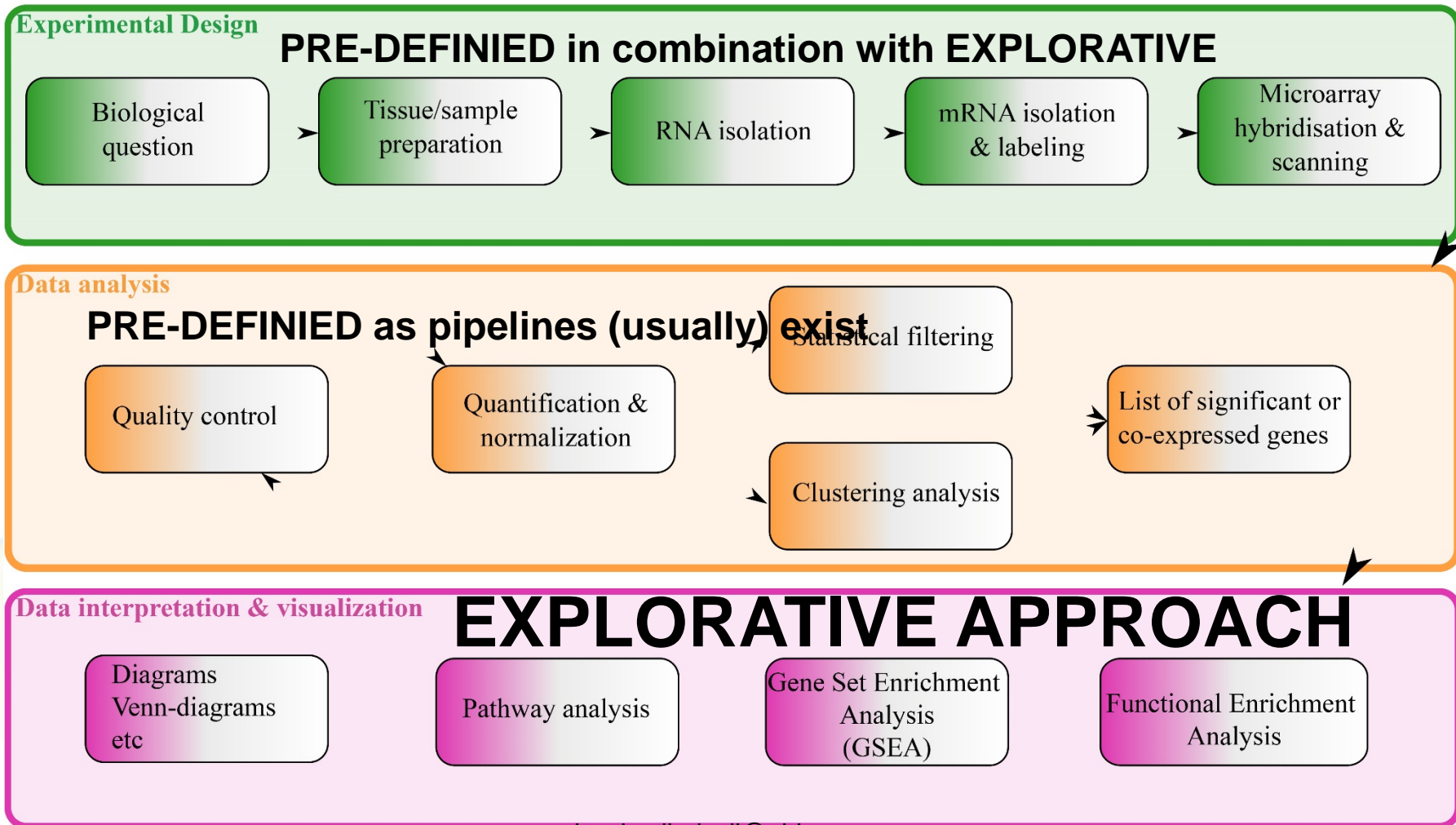
Downstream analysis – How to see the biology in your lists

Jessica M. Lindvall, PhD

NBIS Bioinformatics expert, Training coordinator
jessica.lindvall@nbis.se



Microarray Data Analysis Workflow



List of significantly differentially expressed genes

| PROBEID | ENTREZID | SYMBOL | GENENAME | logFC | AveExpr | t | PValue | adj.P.Val | B | E1 | E2 | E3 | C1 | C2 | C3 |
|-----------------|-----------|------------|---|------------|-----------|------------|--------------|--------------|--------------|-----------|-----------|-----------|-----------|-----------|-----------|
| TC0100017.hg.1 | 400726 | FAM87B | family with sequence similarity 87 member B | -0.5108676 | 5.202430 | -5.803297 | 7.381490e-04 | 0.0334267866 | -0.057279721 | 5.082846 | 5.092984 | 4.905159 | 5.502273 | 5.673162 | 5.438157 |
| TC01000113.hg.1 | 9563 | H6PD | hexose-6-phosphate dehydrogenase (glucose 1-deh... | -0.4735258 | 7.655717 | -5.053993 | 1.648450e-03 | 0.0467099016 | -0.895082448 | 7.502889 | 7.309566 | 7.444405 | 7.936681 | 8.012458 | 7.728300 |
| TC01000151.hg.1 | 5351 | PLOD1 | procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1 | -0.5105864 | 6.892207 | -5.476685 | 1.054191e-03 | 0.0386410376 | -0.429817745 | 6.790599 | 6.673173 | 6.446968 | 7.162915 | 7.158692 | 7.120893 |
| TC01000184.hg.1 | 23254 | KAZN | kazn, perleppin interacting protein | 0.3506924 | 6.551819 | 5.353727 | 1.197669e-03 | 0.0409859801 | -0.563584091 | 6.746468 | 6.774337 | 6.660690 | 6.430708 | 6.293886 | 6.404824 |
| TC01000193.hg.1 | 6248 | RSC1A1 | regulatory solute carrier protein, family 1, member 1 | -0.2682336 | 6.626843 | -5.023315 | 1.704401e-03 | 0.0477554758 | -0.934185613 | 6.500803 | 6.488452 | 6.488923 | 6.745079 | 6.733196 | 6.804605 |
| TC01000371.hg.1 | 677774 | SCARNA1 | small Cajal body-specific RNA 1 | -1.2962790 | 3.983074 | -10.113995 | 2.556184e-05 | 0.0078580955 | 3.314046365 | 3.534069 | 3.035438 | 3.437998 | 4.641666 | 4.637176 | 4.617499 |
| TC01000408.hg.1 | 347738 | SEHMC2 | serine incorporator 2 | -0.3584747 | 8.278829 | -5.708592 | 8.321888e-04 | 0.0353099204 | -0.182467479 | 6.115187 | 6.130025 | 6.057484 | 6.519004 | 6.481094 | 6.576961 |
| TC01000424.hg.1 | 8685 | EIF3I | eukaryotic translation initiation factor 3 subunit I | -0.3588492 | 8.095274 | -5.177178 | 1.443555e-03 | 0.0458687814 | -0.759565140 | 7.950683 | 7.993479 | 7.820388 | 8.522752 | 8.207841 | 8.284705 |
| TC01000497.hg.1 | 64054 | OXCT2 | 3-oxoacid CoA-transferase 2 | -0.3896637 | 4.672600 | -5.810016 | 9.199577e-04 | 0.0366517120 | -0.287254935 | 4.544900 | 4.477226 | 4.411360 | 4.949566 | 4.794054 | 4.958857 |
| TC01000535.hg.1 | 339539 | LOC339539 | uncharacterized LOC339539 | 0.6883077 | 5.218392 | 10.184796 | 2.445011e-05 | 0.0077400436 | 3.355894713 | 5.618360 | 5.462320 | 5.606956 | 4.822785 | 4.874720 | 4.925208 |
| TC01000550.hg.1 | 6487 | ST3GAL3 | ST3 beta-galactoside alpha-2,3-sialyltransferase 3 | 0.3884467 | 5.964524 | 5.974156 | 6.413111e-04 | 0.0316322927 | 0.099253120 | 6.145311 | 6.078243 | 6.252688 | 5.746718 | 5.764427 | 5.799757 |
| TC01000564.hg.1 | 6202 | RPS6 | ribosomal protein S6 | -0.3657236 | 6.695883 | -6.843647 | 7.879934e-04 | 0.0217561857 | 0.316921727 | 6.520853 | 6.478339 | 6.539872 | 6.870361 | 6.860791 | 6.905082 |
| TC01000565.hg.1 | 94101 | SNORD46 | small nuclear RNA, C/D box 46 | -0.4004855 | 11.325241 | -7.370803 | 1.840225e-04 | 0.0177412513 | 1.373998443 | 11.117220 | 11.003706 | 11.134068 | 11.510699 | 11.545139 | 11.632612 |
| TC01000668.hg.1 | 26027 | ACOT11 | acyl-CoA thioesterase 11 | 0.4607009 | 5.485165 | 8.152343 | 9.907859e-05 | 0.0139638874 | 1.906975995 | 5.601272 | 5.727505 | 5.727680 | 5.266671 | 5.300785 | 5.106988 |
| TC01000730.hg.1 | 3953 | LEPR | leptin receptor | 0.2800242 | 5.489537 | 5.210109 | 1.398693e-03 | 0.0439250668 | -0.722656530 | 5.326151 | 5.383075 | 5.339350 | 5.631077 | 5.602786 | 5.600285 |
| TC01000745.hg.1 | 1847 | GADD45A | growth arrest and DNA damage inducible alpha | 0.3319442 | 8.718726 | 5.826295 | 7.425322e-04 | 0.0342678866 | -0.083655075 | 6.905872 | 6.926172 | 6.824051 | 6.572825 | 6.524428 | 6.561011 |
| TC01000781.hg.1 | 256435 | ST5GALNAC3 | ST6 (alpha-N-acetylneuraminyl-2,3-beta-galactosyl-... | -0.7128701 | 3.674617 | -6.840652 | 3.448840e-04 | 0.0231394560 | 0.732236807 | 3.521368 | 3.242929 | 3.192050 | 4.139107 | 3.906683 | 4.047368 |
| TC01000874.hg.1 | 729970 | LOC729970 | hCG2020352-like | -0.5080706 | 5.747263 | -5.652394 | 8.800092e-04 | 0.0358800206 | -0.240834250 | 5.326493 | 5.645737 | 5.507452 | 5.951125 | 6.046520 | 6.006250 |
| TC01000883.hg.1 | 58155 | PTBP2 | polypyrimidine tract binding protein 2 | 0.7157137 | 5.906043 | 9.355395 | 4.191649e-05 | 0.0101177096 | 2.841605787 | 6.163183 | 6.383313 | 6.245203 | 5.617368 | 5.540461 | 5.486728 |
| TC01000892.hg.1 | 54873 | PALMD | palmdelphin | -1.5961253 | 7.475419 | -10.037111 | 2.683487e-05 | 0.0081723454 | 3.288093656 | 6.877048 | 6.704516 | 6.300505 | 6.343710 | 6.292006 | 6.304728 |
| TC01000909.hg.1 | 1901 | SIPR1 | sphingosine-1-phosphate receptor 1 | -0.7190823 | 6.329539 | -5.878584 | 7.039221e-04 | 0.0326016120 | -0.007776790 | 5.903407 | 5.941090 | 5.984688 | 6.452707 | 6.664954 | 6.949578 |
| TC01001004.hg.1 | 56944 | OLFML3 | olfactomedin like 3 | -0.3748706 | 6.466112 | -5.435179 | 1.100358e-03 | 0.0393837151 | -0.474724388 | 6.322097 | 6.200493 | 6.333499 | 6.570661 | 6.732215 | 6.657765 |
| TC01001025.hg.1 | 8458 | TTYF | transcription termination factor, RNA polymerase II | 0.3761420 | 5.344338 | 5.353874 | 1.197504e-03 | 0.0409859801 | -0.563422243 | 5.473276 | 5.511736 | 5.612214 | 5.220022 | 5.062530 | 5.185349 |
| TC01001030.hg.1 | 54855 | FAM46C | family with sequence similarity 46 member C | 0.8025345 | 5.935639 | 7.559015 | 1.578103e-04 | 0.0187272092 | 1.528641277 | 6.425488 | 6.524753 | 6.255900 | 5.884875 | 5.312871 | 5.619168 |
| TC01001053.hg.1 | 647121 | EMBP1 | embigin pseudogene 1 | -0.9099513 | 4.783384 | -5.778300 | 7.771126e-04 | 0.0343129955 | -0.110949190 | 4.425107 | 4.577769 | 3.982349 | 5.118014 | 5.375836 | 5.221229 |
| TC01001201.hg.1 | 10952 | MLLT11 | myeloid/lymphoid or mixed-lineage leukemia; transloc... | -0.3376877 | 4.805903 | -5.014541 | 6.795930e-04 | 0.0324793243 | 0.028875887 | 4.440566 | 4.440112 | 4.122501 | 4.956562 | 5.372954 | 5.494726 |
| TC01001262.hg.1 | 4881 | NPR1 | atriuretic peptide receptor 1 | 0.4207164 | 6.176434 | 5.328641 | 1.229578e-03 | 0.0414553795 | -0.591148043 | 6.452776 | 6.303448 | 6.404152 | 6.060845 | 5.996651 | 5.840730 |
| TC01001274.hg.1 | 57198 | ATP9B2 | ATPase phospholipid transporting 9B2 | 0.4062779 | 6.365881 | 5.245222 | 1.342654e-03 | 0.0433614839 | -0.683480409 | 6.688421 | 6.477253 | 6.541387 | 6.219719 | 6.191483 | 6.077025 |
| TC01001348.hg.1 | 3428 | IFI15 | interferon, gamma-inducible protein 15 | -0.7519609 | 6.766634 | -7.434972 | 1.745707e-04 | 0.0173431723 | 1.427477250 | 6.290559 | 6.514843 | 6.366559 | 6.986561 | 7.133814 | 7.307469 |
| TC01001377.hg.1 | 4817 | NT1 | nitilase 1 | 0.5312612 | 6.647760 | 6.074712 | 5.806069e-04 | 0.0302714402 | 0.192605120 | 6.700562 | 6.950099 | 7.029510 | 6.307677 | 6.298021 | 6.406090 |
| TC01001422.hg.1 | 6675 | UAP1 | UDP-N-acetylglucosamine pyrophosphorylase 1 | -0.4832461 | 6.574137 | -5.148068 | 1.489258e-03 | 0.0451432083 | -0.792327520 | 6.440064 | 6.268701 | 6.288776 | 6.908747 | 6.641255 | 6.807278 |
| TC01001423.hg.1 | 4921 | DDR2 | discoidin domain receptor tyrosine kinase 2 | 0.8498982 | 8.758723 | 7.286467 | 1.958727e-04 | 0.0181157270 | 1.310608299 | 8.573590 | 8.318368 | 8.109385 | 9.167569 | 9.153386 | 9.290063 |
| TC01001425.hg.1 | 4921 | DDR2 | discoidin domain receptor tyrosine kinase 2 | -0.8260146 | 8.148835 | -6.878513 | 3.332765e-04 | 0.0227678052 | 0.788880315 | 7.761718 | 7.793451 | 7.654735 | 8.530624 | 8.334582 | 8.816721 |
| TC01001464.hg.1 | 100313835 | MIR1255B2 | microRNA 1255b-2 | 0.7645595 | 3.698043 | 7.833916 | 1.267770e-04 | 0.0155307630 | 1.750264059 | 3.980791 | 4.022401 | 4.237778 | 3.449392 | 3.189162 | 3.314197 |
| TC01001484.hg.1 | 5395 | PRRX1 | paired related homeobox 1 | 0.3293720 | 5.456657 | 10.533450 | 1.971360e-05 | 0.0072212701 | 3.557543997 | 5.864014 | 5.815694 | 6.084322 | 4.915966 | 4.989740 | 5.076200 |

→ easy to sort and filter out significantly differentially expressed genes: $\text{adj.P.Val} < 0.05$ and $\text{logFC} > \pm 2$

jessica.lindvall@nbis.se

As it is Explorative bioinformatics....

Choice of tools and software's depends on:

- Biological question
- Number of groups/timepoints
- Number of Differentially Expressed Genes
- ...

Approaches:

- *Hypothesis-free**:
 - *As free as can be after design of experiment (biologically)
 - Explore **PATTERNS** within the data with no prior knowledge!
- *Hypothesis-driven*:
 - Prior Knowledge to your data!
 - Interesting pathways
 - Diseases

Remember the Picture!



Explorative bioinformatics

- **Combination** between **Your Biological knowledge** and **Bioinformatics** (statistics, pattern search etc)
- **DEPENDS** on the **Biological question** and prior biological knowledge
- **NOT talking about fixed pipelines!** – they do not (really) exist

How to start...

- *Hypothesis-free** (covered in this session):

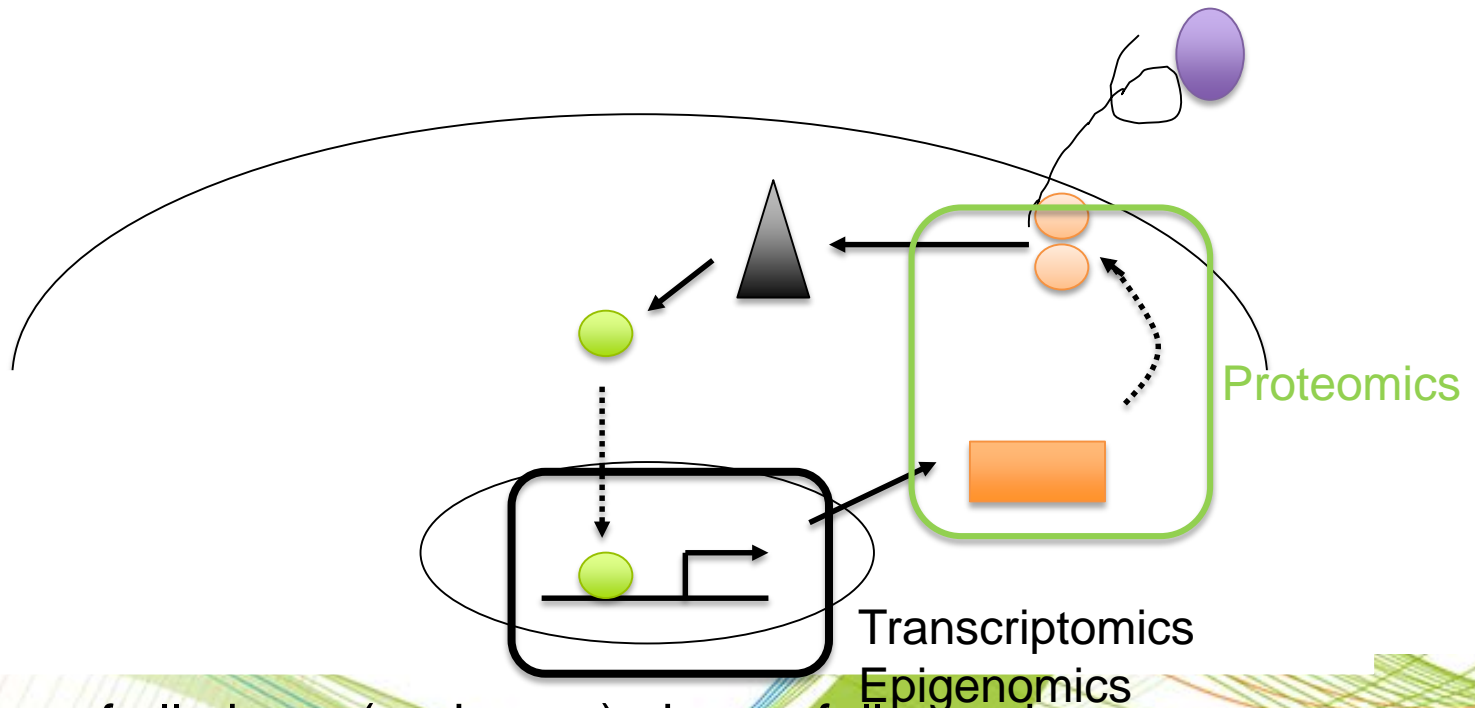
- *As free as can be (biologically)
- Explore PATT (without prior knowledge!

- *Hypothesis-driven* (with prior knowledge):

- Prior Knowledge
 - Interesting
 - Diseases



Tools, Softwares, Databases...



- A mixture of all above (and more) gives a fuller and more comprehensive view of the biological picture

Example 1 – Large list of DEGs ~1000-2000

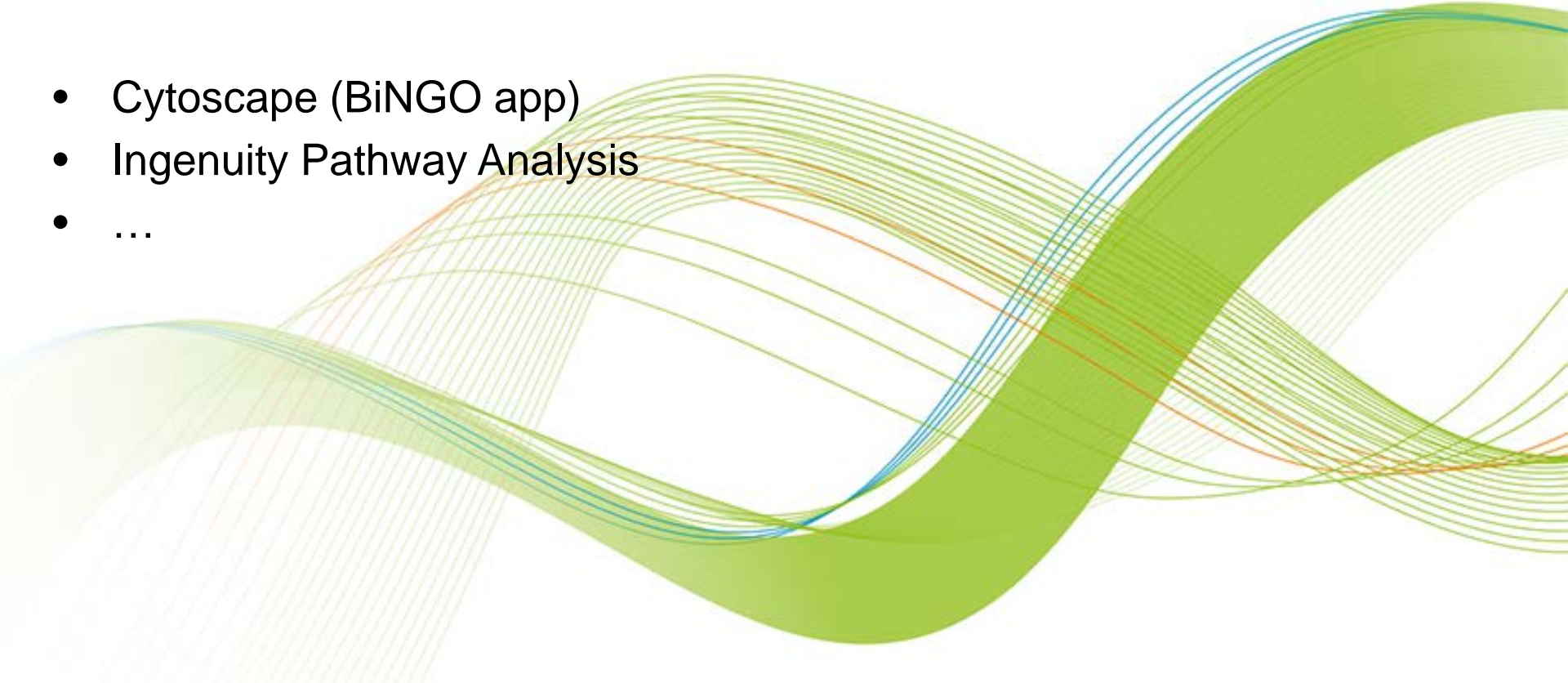
Workflow QC Normalisation Clustering analysis Statistical filtering DE

List of significantly differentially expressed genes

| PROBEID | ENTREZID | SYMBOL | GENENAME | logFC | AveExpr | t | P.Value | adj.P.Val | B | E1 | E2 | E3 | C1 | C2 | C3 |
|----------------|-----------|------------|--|------------|-----------|------------|--------------|---------------|--------------|-----------|-----------|-----------|-----------|-----------|-----------|
| TC1000017.hg.1 | 409728 | FAM87B | family with sequence similarity 87 member B | -0.5108676 | 5.204200 | -5.830297 | 7.381450e-04 | 0.034267066 | -0.957279721 | 5.020246 | 5.092904 | 4.905159 | 5.502270 | 5.673162 | 5.436157 |
| TC1000113.hg.1 | 9563 | HGF | heparanase-6-phosphate dehydrogenase (glucose 1-deh... | -0.4735258 | 7.655717 | -5.053993 | 1.648450e-03 | 0.046709901 | -0.895062448 | 7.502889 | 7.309566 | 7.444405 | 7.936681 | 8.012458 | 7.728300 |
| TC1000151.hg.1 | 5351 | PLD1 | phospholipase, class C, 1 | -0.5105864 | 6.892207 | -5.476685 | 1.054191e-03 | 0.0386410376 | -0.429817745 | 6.790599 | 6.673173 | 6.464968 | 7.162915 | 7.158692 | 7.126893 |
| TC1000184.hg.1 | 23254 | KA2N | kazh1, periplasmic interacting protein | 0.3566924 | 6.551819 | 5.353727 | 1.197606e-03 | 0.040859801 | -0.583564091 | 6.746466 | 6.774337 | 6.660690 | 6.430708 | 6.259306 | 6.404874 |
| TC1000193.hg.1 | 6248 | RSC1A1 | regulatory solute carrier protein, family 1, member 1 | -1.2062236 | 6.620643 | -5.023315 | 1.704401e-03 | 0.0477554756 | -0.934185613 | 6.500003 | 6.480452 | 6.480923 | 6.745079 | 6.733196 | 6.004605 |
| TC1000193.hg.1 | 67774 | SCARNA1 | small Cajal body-specific RNA 1 | -1.2062290 | 3.983074 | -10.113995 | 3.556184e-05 | 0.0078580055 | 3.314046365 | 3.534069 | 3.035438 | 3.437998 | 4.641664 | 4.637176 | 4.617499 |
| TC1000406.hg.1 | 347748 | SNRNP2 | snRNA associated protein 2 | -0.8854747 | 8.279829 | -8.708982 | 8.131888e-03 | 0.0458090216 | -0.182607479 | 8.115187 | 8.100328 | 8.187488 | 8.511064 | 8.481098 | 8.534061 |
| TC1000424.hg.1 | 8668 | EFS1 | eukaryotic translation initiation factor 3 subunit 1 | -0.3568492 | 8.093274 | -5.177178 | 1.443555e-03 | 0.0458687614 | -0.759565140 | 7.930683 | 7.995479 | 7.820358 | 8.322752 | 8.207641 | 8.284705 |
| TC1000497.hg.1 | 64954 | DKC2 | 3'-capicbin Cap-transporter 2 | -0.3896637 | 4.672660 | -5.016016 | 3.199577e-04 | 0.0366517126 | -0.297254955 | 4.544900 | 4.477226 | 4.411360 | 4.949566 | 4.794054 | 4.958057 |
| TC1000536.hg.1 | 339530 | LOC339539 | uncharacterized LOC339539 | 0.6883677 | 5.218992 | 10.184795 | 2.445011e-05 | 0.0077400436 | 3.355594713 | 5.618360 | 5.462320 | 5.666956 | 4.822785 | 4.874720 | 4.925209 |
| TC1000550.hg.1 | 6487 | ST3GAL3 | ST3 beta-galactosidase alpha-2,3-sialyltransferase 3 | 0.3884467 | 5.964524 | 5.974156 | 6.413111e-04 | 0.0336329297 | 0.089253120 | 6.145311 | 6.078243 | 6.252688 | 5.746718 | 5.764427 | 5.759757 |
| TC1000564.hg.1 | 6292 | RP50 | ribosomal protein S0 | -0.3057236 | 6.959603 | -6.047642 | 2.879954e-04 | 0.0212501857 | 0.916921737 | 6.920853 | 6.478339 | 6.539877 | 6.870361 | 6.660791 | 6.905082 |
| TC1000565.hg.1 | 94101 | SNORD46 | small nuclear RNA, C/D box 46 | -0.4604855 | 11.325241 | -7.376033 | 1.840225e-04 | 0.0177412913 | 1.373980443 | 11.172220 | 11.003706 | 11.114069 | 11.510699 | 11.545139 | 11.032612 |
| TC1000648.hg.1 | 26027 | ACOT11 | acyl-CoA thioesterase 11 | 0.4607609 | 5.485165 | 8.153243 | 9.007850e-05 | 0.0139638874 | 1.906675995 | 5.601272 | 5.727505 | 5.727680 | 5.266671 | 5.300785 | 5.106988 |
| TC1000730.hg.1 | 3858 | LEPR | leptin receptor | 0.2801242 | 5.409637 | 5.216109 | 1.399692e-03 | 0.0429250668 | 0.732656593 | 5.326151 | 5.383075 | 5.339959 | 5.631077 | 5.657286 | 5.606286 |
| TC1000745.hg.1 | 1647 | GADD45A | growth arrest: and DNA damage inducible alpha | 0.5519442 | 6.718726 | 5.824295 | 7.425522e-04 | 0.0354267866 | -0.084655076 | 6.903872 | 6.926172 | 6.824061 | 6.572626 | 6.524426 | 6.561011 |
| TC1000781.hg.1 | 256435 | ST6GALNAC3 | ST6 feline-N-acetylglucosaminyl-2,3-beta-galactosyl... | -0.7128701 | 3.674617 | -6.046052 | 3.449840e-04 | 0.0231394566 | 0.7322266097 | 3.523369 | 3.242929 | 3.150259 | 4.139107 | 3.906603 | 4.041366 |
| TC1000874.hg.1 | 729970 | LOC729970 | LOC729970-like | -0.5007066 | 5.747263 | -5.653954 | 8.800902e-04 | 0.0358000205 | -0.240824250 | 5.326493 | 5.645737 | 5.567452 | 5.951325 | 6.046520 | 6.006250 |
| TC1000883.hg.1 | 58155 | PIBF2 | poly(pyrimidine tract binding protein 2 | 0.7157137 | 5.906403 | 9.355395 | 4.191649e-05 | 0.0101177096 | 2.841605787 | 6.163183 | 6.383313 | 6.245203 | 5.617368 | 5.540461 | 5.486728 |
| TC1000892.hg.1 | 54873 | PALMD | palmitolein | -1.5861253 | 7.425419 | -10.037111 | 2.683487e-05 | 0.0081728454 | 3.268093656 | 6.877048 | 6.704516 | 6.300505 | 6.343710 | 6.292006 | 6.034728 |
| TC1000909.hg.1 | 1901 | SIPIR1 | sphingosine-1-phosphate receptor 1 | -0.7180823 | 6.339539 | -5.078584 | 0.039221e-04 | 0.0326010126 | -0.007776795 | 5.905407 | 5.941050 | 5.964605 | 4.527207 | 6.664954 | 6.949578 |
| TC1001004.hg.1 | 56944 | OLFML3 | olfactomedin like 3 | -0.3748706 | 4.603112 | -4.425179 | 1.003560e-03 | 0.0363837151 | -0.474742388 | 6.322097 | 6.200433 | 6.313499 | 5.570661 | 6.732215 | 6.057765 |
| TC1001026.hg.1 | 8458 | TFE2 | transcription termination factor, RNA polymerase II | 0.7671420 | 5.349238 | 8.583874 | 1.197504e-03 | 0.0409859803 | 0.563422248 | 5.473276 | 5.511736 | 5.612211 | 5.220022 | 5.662530 | 5.188349 |
| TC1001030.hg.1 | 54855 | FAM48C | family with sequence similarity 48 member C | 0.8025145 | 5.945639 | 7.555014 | 1.578103e-04 | 0.0187222092 | 1.250861277 | 6.426485 | 6.524753 | 6.263900 | 5.686476 | 5.312671 | 5.619168 |
| TC1001055.hg.1 | 647123 | EMR1P1 | embigin pseudogene 1 | -0.3099513 | 4.783584 | -5.778500 | 7.771126e-04 | 0.0345129955 | -0.110949190 | 4.425107 | 4.577769 | 3.982349 | 5.118034 | 5.375896 | 5.221229 |
| TC1001201.hg.1 | 10992 | MLL1T1 | myeloblastin/lymphoid or mixed-lineage leukemia trans... | -0.3076877 | 4.805503 | -5.014541 | 6.795030e-04 | 0.0324783243 | 0.020875897 | 4.440566 | 4.440112 | 4.122501 | 4.950562 | 5.372954 | 5.454726 |
| TC1001262.hg.1 | 4881 | NPR1 | natriuretic peptide receptor 1 | 0.4207164 | 6.176434 | 5.328641 | 1.229578e-03 | 0.0434553795 | -0.591148043 | 6.452776 | 6.303448 | 6.464152 | 6.060845 | 5.996651 | 5.846730 |
| TC1001274.hg.1 | 57198 | AT9B2 | AT9ase phospholipid transporting BB2 | 0.4062779 | 6.369881 | 5.245222 | 1.342654e-03 | 0.04313614839 | -0.683480409 | 6.688421 | 6.477253 | 6.541387 | 6.219719 | 6.191483 | 6.077025 |
| TC1001348.hg.1 | 3420 | IFIT5 | interferon, gamma-inducible protein 15 | 0.7515869 | 6.760674 | 7.434972 | 1.745707e-04 | 0.0173431723 | 1.427472750 | 6.930559 | 6.514045 | 6.360559 | 6.806651 | 7.133814 | 7.307469 |
| TC1001377.hg.1 | 4817 | NTL1 | nitric oxide synthase 1 | 0.2310612 | 6.477004 | 6.077412 | 8.800606e-04 | 0.0302714402 | 0.182605120 | 6.700562 | 6.950099 | 7.029510 | 6.030767 | 6.298023 | 6.460690 |
| TC1001422.hg.1 | 6675 | UMP1 | UDP-N-acetylglucosamine pyrophosphorylase 1 | 0.4823461 | 6.574137 | 8.181068 | 1.489258e-03 | 0.0451432083 | 0.732227520 | 6.440064 | 6.248701 | 6.288776 | 6.908747 | 6.641258 | 6.807278 |
| TC1001428.hg.1 | 4921 | DDR2 | discoidin domain receptor tyrosine kinase 2 | 0.8908982 | 6.758223 | 7.205497 | 1.958727e-01 | 0.0181157270 | 1.810608290 | 8.573590 | 8.318948 | 8.109385 | 9.157669 | 9.153386 | 9.226043 |
| TC1001435.hg.1 | 4921 | DDR2 | discoidin domain receptor tyrosine kinase 2 | -0.8240146 | 6.148635 | -8.678513 | 3.332786e-04 | 0.02217678052 | 0.788880515 | 7.761718 | 7.794851 | 7.854745 | 8.530624 | 8.334582 | 8.816721 |
| TC1001464.hg.1 | 100313035 | MIR1255B2 | microRNA 1255b-2 | 0.7645595 | 3.690843 | 7.833916 | 1.267770e-04 | 0.0155307630 | 1.750264059 | 3.980791 | 4.022401 | 4.237778 | 3.443932 | 3.189102 | 3.314197 |
| TC1001484.hg.1 | 5395 | PBX1 | paired related homeobox 1 | 0.3020370 | 5.456657 | 10.533458 | 1.971360e-05 | 0.0072212701 | 3.557543997 | 5.064041 | 5.015604 | 6.064322 | 4.915066 | 4.989740 | 5.076200 |

→ easy to sort and filter out significantly differentially expressed genes: $adj.P.Val < 0.05$ and $logFC > \pm 2$

Hypothesis-free

- Cytoscape (BiNGO app)
 - Ingenuity Pathway Analysis
 - ...
- 

Workflow

QC

Normalisation

Clustering analysis

Statistical filtering

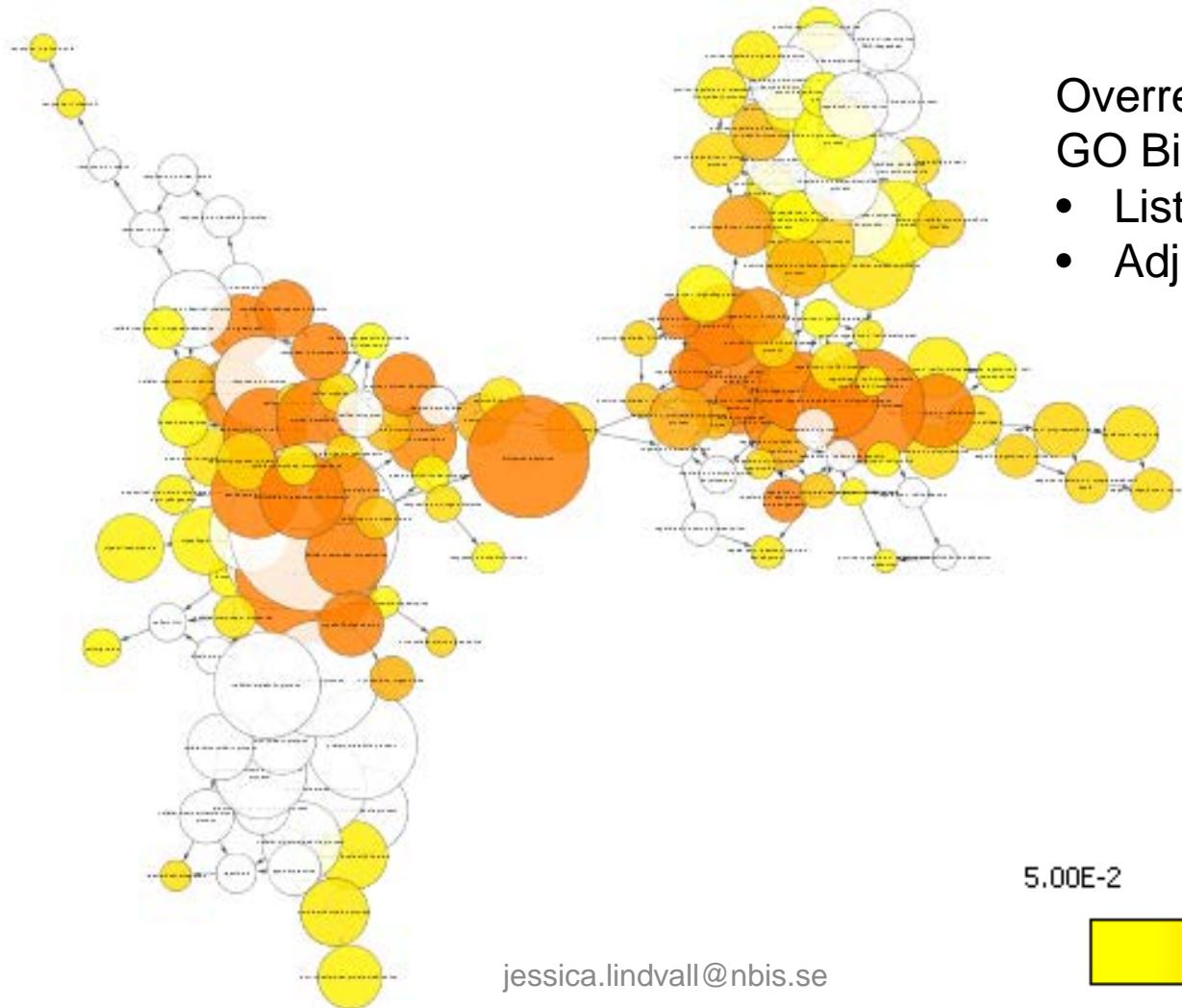
DE

List of significantly differentially expressed genes

~2000 Differentially Expressed Genes

| PROBEID | ENTREZID | SYMBOL | GENENAME | logFC | AveExpr | t | PValue | adj_P.Val | B | E1 | E2 | E3 | C1 | C2 | C3 |
|-----------------|----------|----------|--|------------|----------|------------|--------------|---------------|--------------|----------|----------|----------|----------|----------|----------|
| TC0100017.hg.1 | 400726 | FAM87B | family with sequence similarity 87 member B | -0.5108676 | 5.282430 | -5.830297 | 7.381490e-04 | 0.0324267066 | -0.057279721 | 5.002846 | 5.092904 | 4.905159 | 5.502273 | 5.673162 | 5.436157 |
| TC01000113.hg.1 | 9563 | HPFD | hexose 6-phosphate dehydrogenase (glucose 1-deh... | -0.4735258 | 7.655717 | -5.053993 | 1.648450e-03 | 0.0467099019 | -0.899082448 | 7.502889 | 7.309566 | 7.444406 | 7.936661 | 8.012458 | 7.728300 |
| TC01000151.hg.1 | 5351 | PCOD1 | procollagen-lysine, 2-oxoglutarate 5-lyoxigenase 1 | -0.5105864 | 6.892207 | -5.476685 | 1.054191e-03 | 0.03864410376 | -0.429817745 | 6.790599 | 6.673173 | 6.444958 | 7.162915 | 7.158692 | 7.120893 |
| TC01000184.hg.1 | 73254 | KAZN | kazrin, perlecan interacting protein | 0.3506924 | 6.551819 | 5.353727 | 1.197689e-03 | 0.0409859601 | -0.563564091 | 6.746468 | 6.774357 | 6.660690 | 6.430708 | 6.293806 | 6.404824 |
| TC01000193.hg.1 | 6248 | RSC1A1 | regulatory sulfate carrier protein, family 1, member 1 | -0.2662336 | 6.626843 | -5.023315 | 1.704401e-03 | 0.0477554758 | -0.934185613 | 6.500803 | 6.488452 | 6.468923 | 6.745079 | 6.733196 | 6.904605 |
| TC01000371.hg.1 | 677774 | SCARNA1 | small Cajal body-specific RNA 1 | -1.2062790 | 3.983074 | -10.113995 | 2.555184e-05 | 0.0078580055 | 3.314046365 | 3.534069 | 3.035438 | 3.437908 | 4.641666 | 4.637176 | 4.617490 |
| TC01000405.hg.1 | 347755 | SPRINC2 | serine incorporator 2 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000406.hg.1 | 347756 | SPRINC1 | serine incorporator 1 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000407.hg.1 | 347757 | SPRINC3 | serine incorporator 3 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000408.hg.1 | 347758 | SPRINC4 | serine incorporator 4 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000409.hg.1 | 347759 | SPRINC5 | serine incorporator 5 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000410.hg.1 | 347760 | SPRINC6 | serine incorporator 6 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000411.hg.1 | 347761 | SPRINC7 | serine incorporator 7 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000412.hg.1 | 347762 | SPRINC8 | serine incorporator 8 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000413.hg.1 | 347763 | SPRINC9 | serine incorporator 9 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000414.hg.1 | 347764 | SPRINC10 | serine incorporator 10 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000415.hg.1 | 347765 | SPRINC11 | serine incorporator 11 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000416.hg.1 | 347766 | SPRINC12 | serine incorporator 12 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000417.hg.1 | 347767 | SPRINC13 | serine incorporator 13 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000418.hg.1 | 347768 | SPRINC14 | serine incorporator 14 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000419.hg.1 | 347769 | SPRINC15 | serine incorporator 15 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000420.hg.1 | 347770 | SPRINC16 | serine incorporator 16 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000421.hg.1 | 347771 | SPRINC17 | serine incorporator 17 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000422.hg.1 | 347772 | SPRINC18 | serine incorporator 18 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000423.hg.1 | 347773 | SPRINC19 | serine incorporator 19 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000424.hg.1 | 347774 | SPRINC20 | serine incorporator 20 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000425.hg.1 | 347775 | SPRINC21 | serine incorporator 21 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000426.hg.1 | 347776 | SPRINC22 | serine incorporator 22 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000427.hg.1 | 347777 | SPRINC23 | serine incorporator 23 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000428.hg.1 | 347778 | SPRINC24 | serine incorporator 24 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000429.hg.1 | 347779 | SPRINC25 | serine incorporator 25 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000430.hg.1 | 347780 | SPRINC26 | serine incorporator 26 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000431.hg.1 | 347781 | SPRINC27 | serine incorporator 27 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000432.hg.1 | 347782 | SPRINC28 | serine incorporator 28 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000433.hg.1 | 347783 | SPRINC29 | serine incorporator 29 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000434.hg.1 | 347784 | SPRINC30 | serine incorporator 30 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000435.hg.1 | 347785 | SPRINC31 | serine incorporator 31 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000436.hg.1 | 347786 | SPRINC32 | serine incorporator 32 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000437.hg.1 | 347787 | SPRINC33 | serine incorporator 33 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000438.hg.1 | 347788 | SPRINC34 | serine incorporator 34 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000439.hg.1 | 347789 | SPRINC35 | serine incorporator 35 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000440.hg.1 | 347790 | SPRINC36 | serine incorporator 36 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000441.hg.1 | 347791 | SPRINC37 | serine incorporator 37 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000442.hg.1 | 347792 | SPRINC38 | serine incorporator 38 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000443.hg.1 | 347793 | SPRINC39 | serine incorporator 39 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000444.hg.1 | 347794 | SPRINC40 | serine incorporator 40 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000445.hg.1 | 347795 | SPRINC41 | serine incorporator 41 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000446.hg.1 | 347796 | SPRINC42 | serine incorporator 42 | -0.5554747 | 6.779629 | -5.705992 | 8.371898e-04 | 0.0583090264 | -0.184664749 | 6.115187 | 6.130025 | 6.074464 | 6.519064 | 6.481098 | 6.529961 |
| TC01000447.hg.1 | 347797 | SPRINC43 | serine incorporator 43 | | | | | | | | | | | | |

Cytoscape (BiNGO app) – Overrepresented GO BP



Overrepresented
GO Biological Process

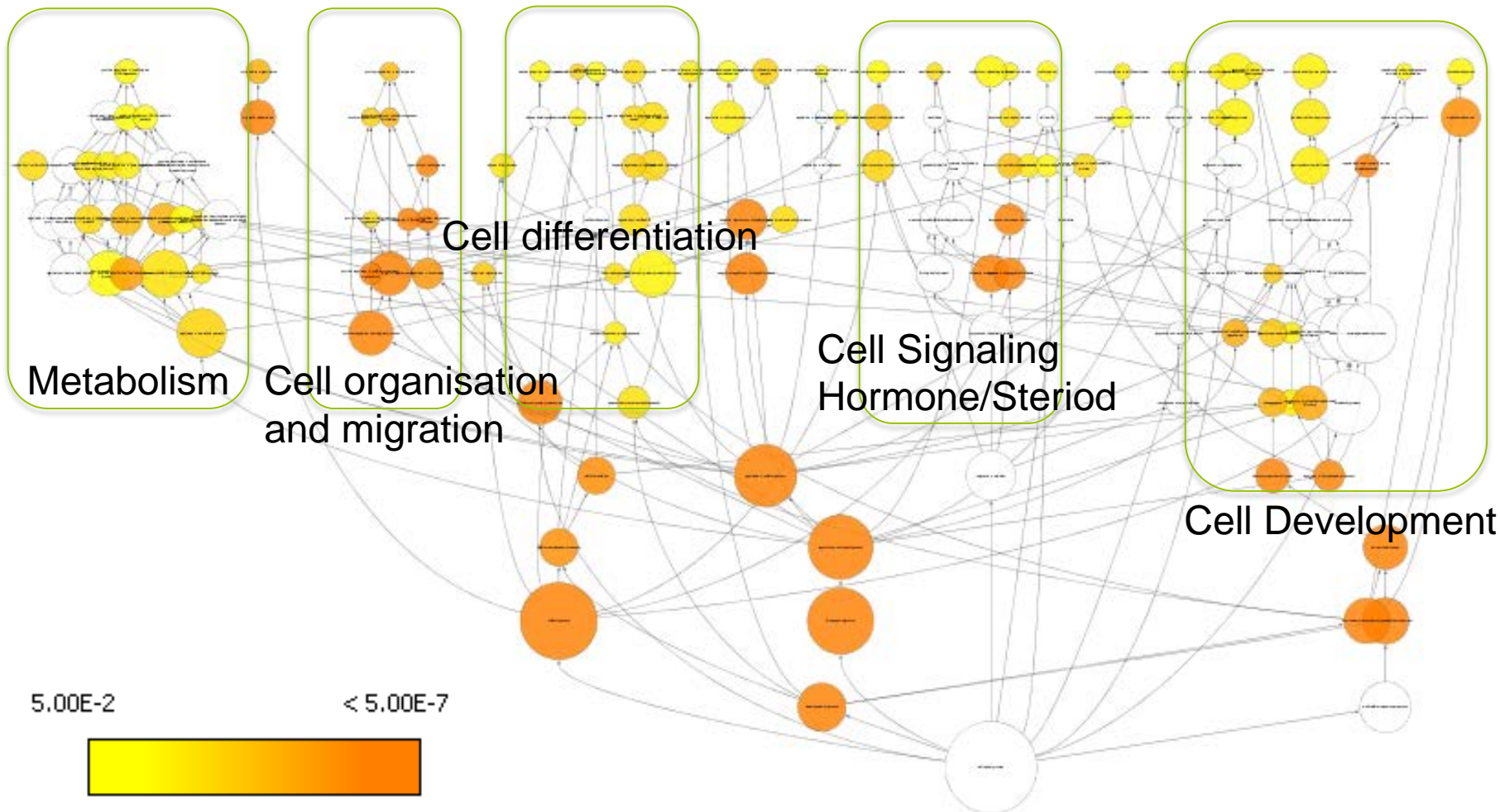
- List of ~2000 genes
- Adj. $p < 0.0001$

5.00E-2

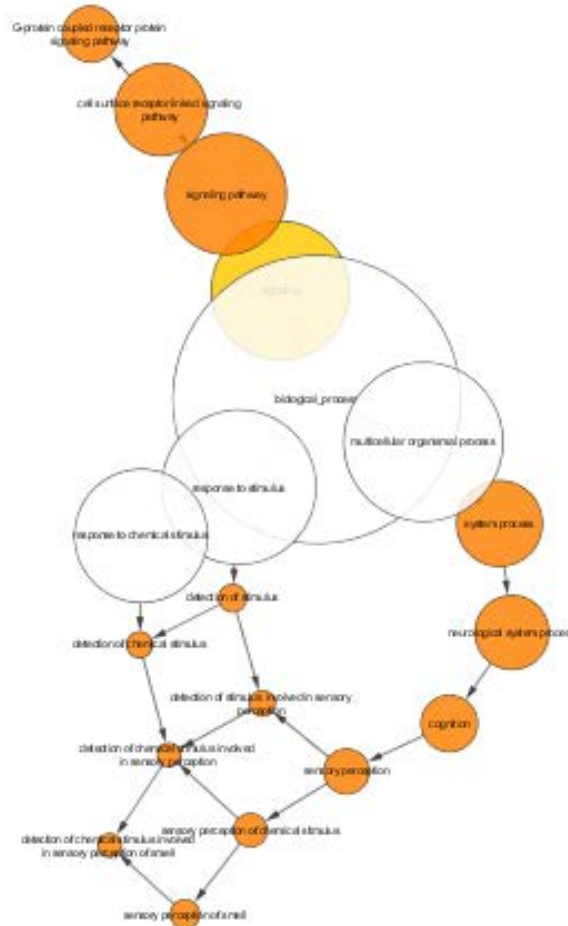
< 5.00E-7



STRUCTURING the list -> explorative patterns



Cytoscape (BiNGO app) – Underrepresented GO BP



Underrepresented
GO Biological Process

- List of ~2000 genes
- Adj. $p < 0.05$

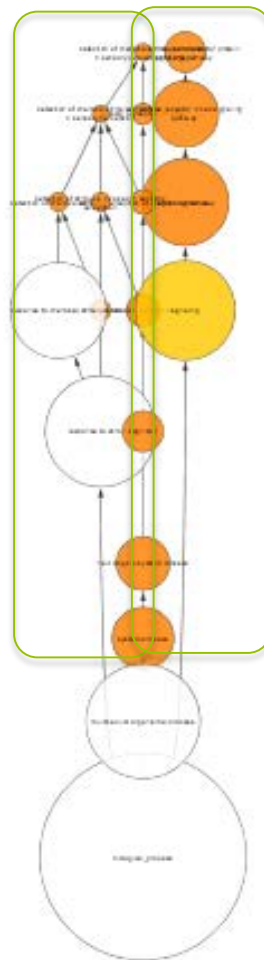
5.00E-2

< 5.00E-7



STRUCTURING the list -> explorative patterns

Cell signalling
Neurological



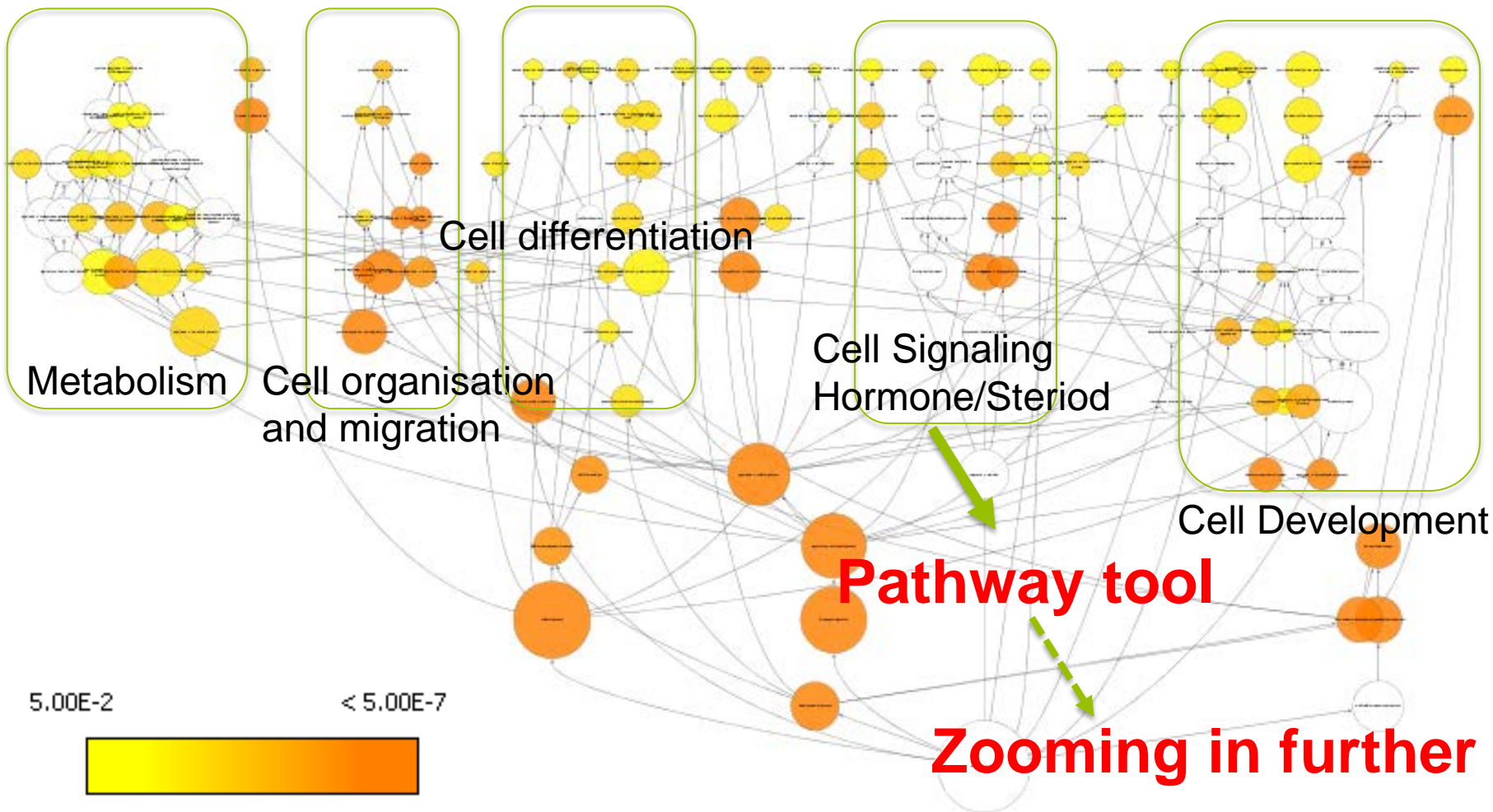
Cell signalling
G-protein coupled receptor

5.00E-2

< 5.00E-7



Further explorative...



Ingenuity Pathway Analysis (commercial)

- Canonical pathway – overrepresented
- Upstream regulators – common upstream Transcriptional regulator having impact on overrepresented genes in DEGs list
- ... And much more
- www.ingenuity.com

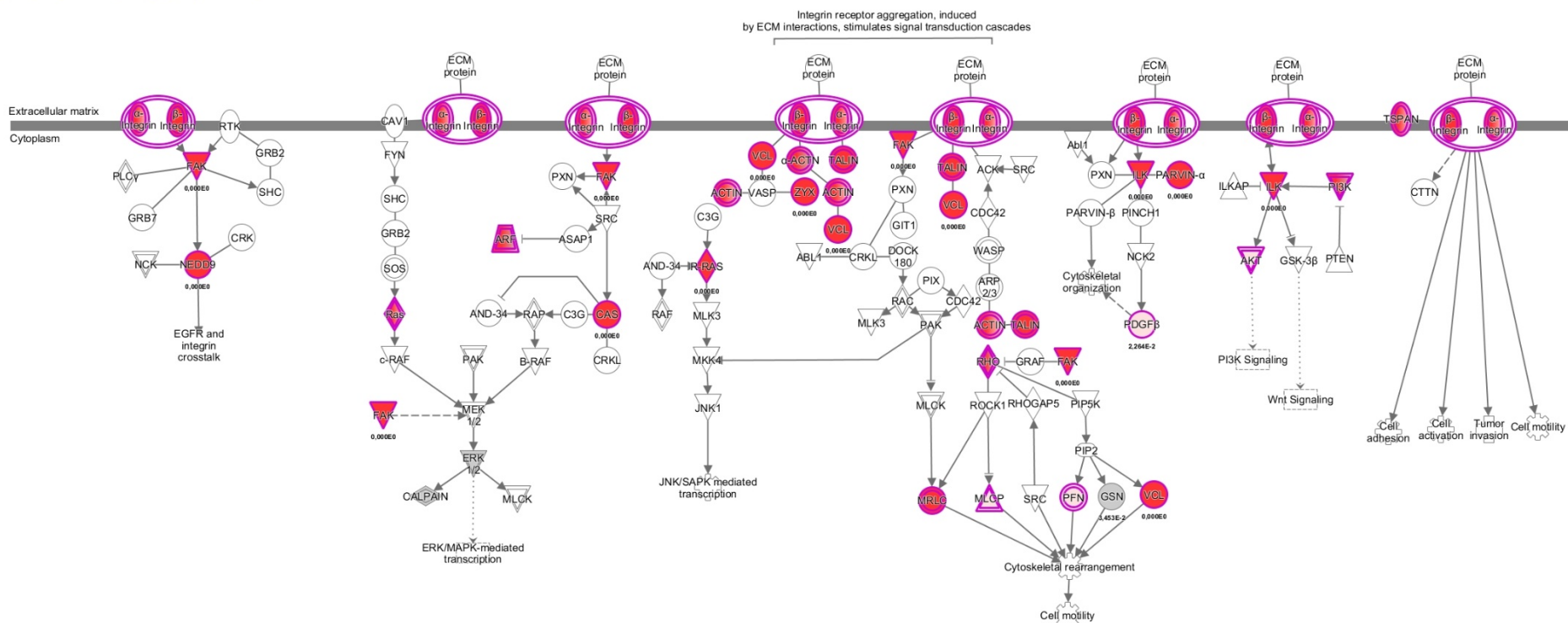
Canonical pathways – two examples

ILK Signaling : Webinar_DEGs_long : Expr Fold Change

Extracellular space



Integrin Signaling : Webinar_DEGs_long : Expr Fold Change



Upstream Transcriptional regulators

IPA

File Edit View Window Help Provide Feedback | Support Jessica Lindvall Close IPA

Genes and Chemicals Diseases and Functions Pathways and Tox Lists

NEW SEARCH Advanced Search

Webinar_DEGs_long - Core analysis

Summary Canonical Pathways Upstream Analysis Diseases & Functions Regulator Effects Networks Lists Molecules

ADD TO MY PATHWAY ADD TO MY LIST DISPLAY AS NETWORK CUSTOMIZE TABLE MECHANISTIC NETWORKS p-value of over... 1,62E-26 - 2,70E-04 (p1 of 4) More Info

| Upstream Regulator | Expr Fold Change | Molecule Type | Predicted Activation ... | Activation z-score | Flags | p-value of overlap | Target molecules ... | Mechanistic Netw... |
|--------------------|------------------|-------------------------|--------------------------|--------------------|-------|--------------------|---------------------------|---------------------|
| TP53 | | transcription regulator | Activated | 3,199 | | 1,62E-26 | ↑ACTN1, ↑AC... all 178 | 389 (19) |
| MYC | | transcription regulator | Activated | 2,867 | bias | 8,05E-21 | ↑ABCE1, ↑ACT... all 131 | 468 (22) |
| HIF1A | ↑1,182 | transcription regulator | Activated | 5,167 | bias | 1,17E-12 | ↑ADAMTS1, ↑A... all 57 | 377 (18) |
| FOS | | transcription regulator | | 1,131 | bias | 2,30E-10 | ↑ACTG1, ↑ADA... all 68 | 361 (16) |
| NFKBIA | | transcription regulator | Activated | 3,266 | bias | 1,34E-09 | ↑ANKRD1, ↑AU... all 56 | 402 (16) |
| EPAS1 | | transcription regulator | Activated | 4,239 | bias | 1,68E-09 | ↑AXL, ↑BHLHE... all 33 | 354 (18) |
| ATF4 | | transcription regulator | Activated | 4,317 | bias | 4,85E-09 | ↑AARS, ↑ANKR... all 29 | 200 (7) |
| MTPN | ↑1,322 | transcription regulator | Activated | 4,411 | bias | 1,97E-08 | ↑ANXA1, ↑Cal... all 20 | 343 (11) |
| HTT | | transcription regulator | Activated | 2,302 | | 7,63E-08 | ↑ACTA1, ↑ADA... all 72 | 358 (21) |
| JUN | | transcription regulator | Activated | 2,930 | bias | 1,03E-07 | ↑AKT1, ↑AKT3, ↑... all 50 | 369 (17) |
| TBX2 | | transcription regulator | Activated | 3,638 | bias | 1,45E-07 | ↑AURKB, ↑BHL... all 18 | 213 (7) |
| CTNNB1 | | transcription regulator | Activated | 5,472 | bias | 2,35E-07 | ↑ACTC1, ↑ACT... all 70 | 364 (19) |
| XBP1 | | transcription regulator | Activated | 4,413 | bias | 2,89E-07 | ↑BFAR, ↑Ccl2, ↑... all 31 | 125 (4) |
| FOXO3 | | transcription regulator | | 0,714 | bias | 6,08E-07 | ↑ACVR1, ↑AGF... all 35 | 391 (17) |
| VHL | | transcription regulator | | -1,016 | bias | 6,92E-07 | ↑AKT1, ↑BHLH... all 22 | 369 (17) |
| FOXM1 | ↑1,893 | transcription regulator | Activated | 4,263 | bias | 1,52E-06 | ↑ANXA1, ↑AUR... all 20 | 327 (17) |
| PAX3 | | transcription regulator | | | | 1,90E-06 | ↑ANXA1, ↑CD1... all 28 | 300 (7) |
| HIF3A | | transcription regulator | | -1,756 | bias | 2,16E-06 | ↑APLN, ↑EDN1, ↑... all 7 | 198 (8) |
| SMAD7 | | transcription regulator | Inhibited | -3,989 | bias | 3,70E-06 | ↑ACVR1, ↑BGN... all 21 | 322 (17) |
| CREB1 | | transcription regulator | Activated | 4,404 | bias | 4,09E-06 | ↑AK4, ↑AKT3, ↑... all 57 | 344 (17) |
| SRF | | transcription regulator | Activated | 4,414 | bias | 4,98E-06 | ↑ACTA1, ↑ACT... all 38 | 366 (13) |
| CDKN2A | | transcription regulator | | -1,206 | | 6,97E-06 | ↑ANAPC5, ↑AU... all 35 | 368 (19) |
| SMARCA4 | | transcription regulator | Activated | 4,048 | bias | 8,86E-06 | ↑ACTA1, ↑ACT... all 55 | 370 (22) |

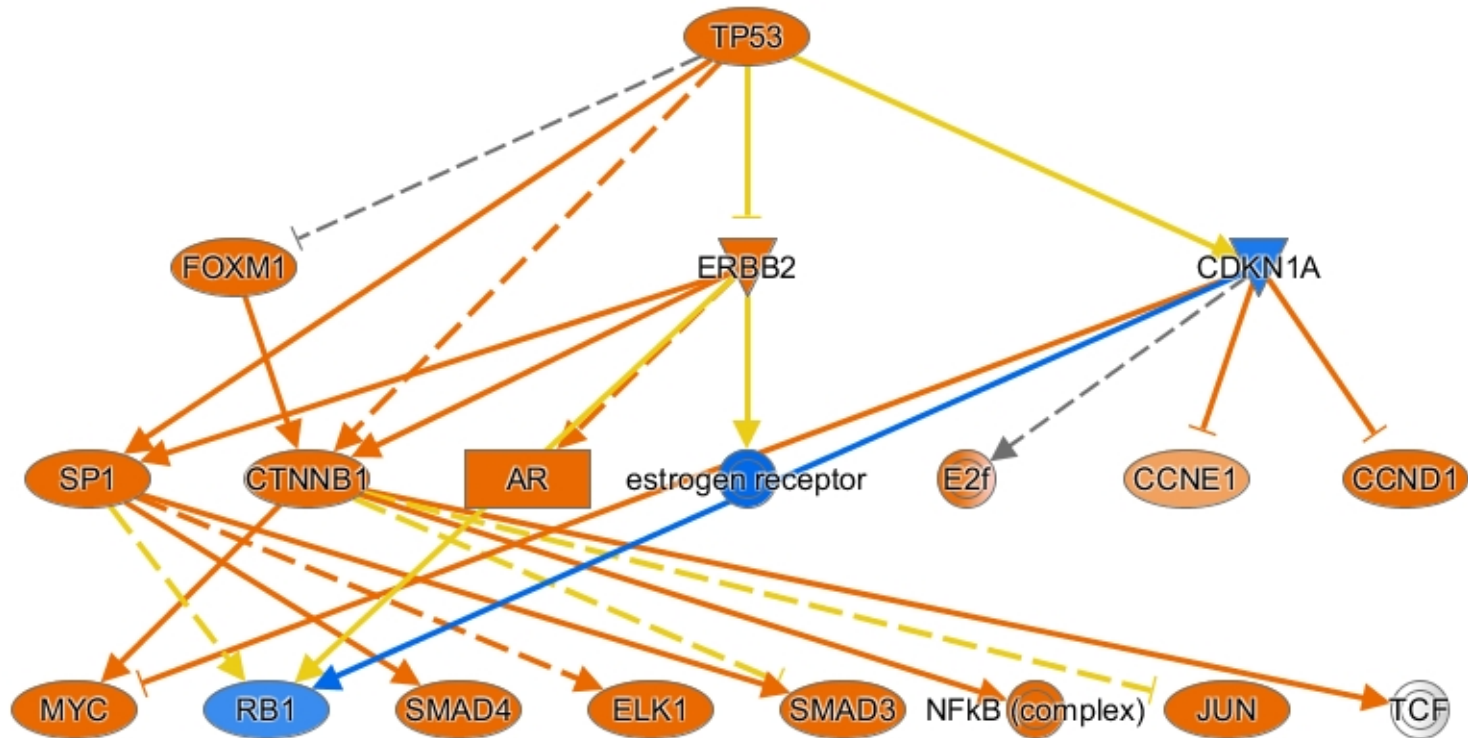
Selected/Total molecules: 0 / 174

jessica.lindvall@nbis.se

13:34 2017-02-13

Mechanistic network – from upstream regulators

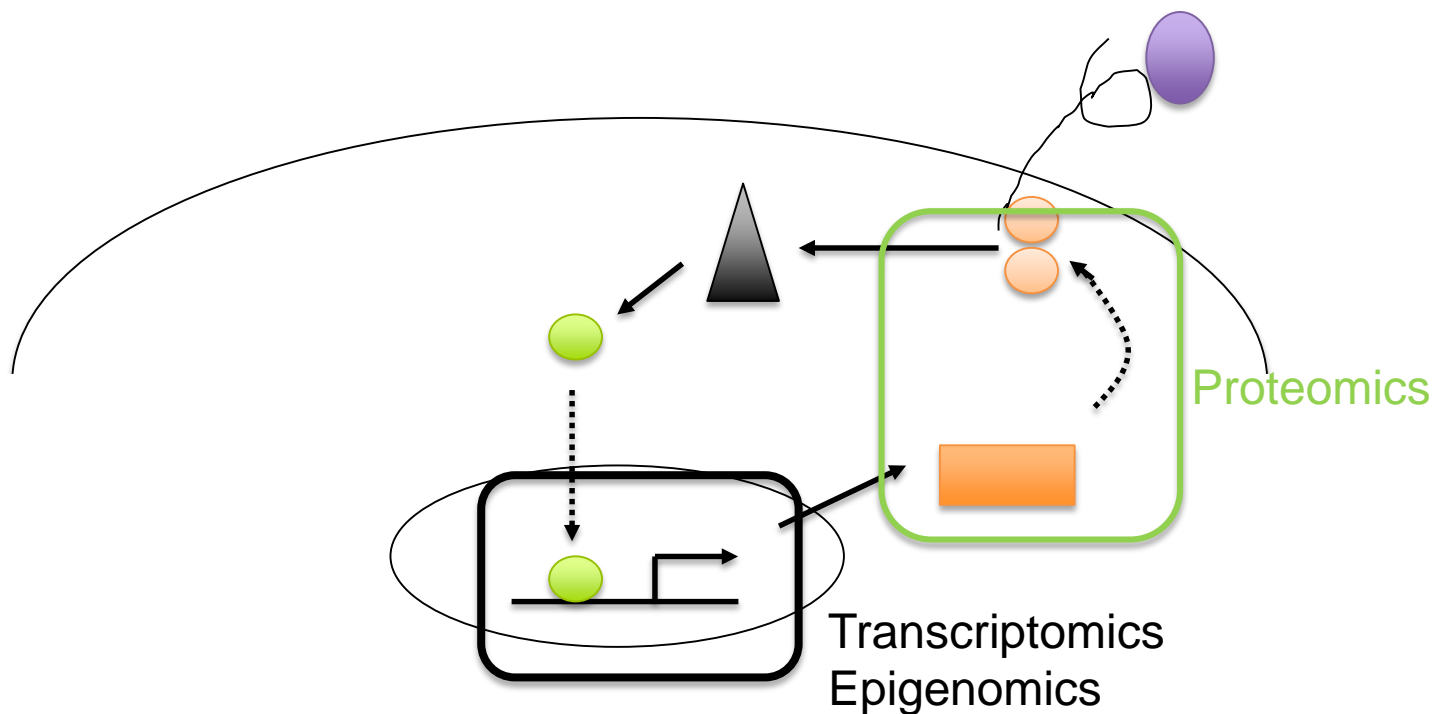
TP53 1



© 2000-2017 QIAGEN. All rights reserved.

Explorative bioinformatics leads to...

- add depth and nuances to your list of DEGs
- find biological sub-groups within your data – sorting the long lists



Questions?

List of tools, software and databases

- GenePattern:
<http://software.broadinstitute.org/cancer/software/genepattern#>
- FunCoup: <http://funcoup.sbc.su.se/search/>
- **Cytoscape**: <http://www.cytoscape.org/>
- Paradigm: <http://sbenz.github.io/Paradigm/>
- Pathifier:
<http://www.weizmann.ac.il/complex/compphys/software/yotam/pathifier/>
- **Ingenuity Pathway Analysis** (IPA-commercial software):
www.ingenuity.com
- **... there is A LOT more – Which to use depends on your biological question!**