

Functional analysis of RNA-seq data : answers to questions

Céline Keime
keime@igbmc.fr

Question 1

- Use “Functional Annotation Chart” tool

Annotation Summary Results

[Help and Tool Manual](#)

Current Gene List:
siMitfvssiLuc_upgenes_lfc1_padj005

780 DAVID IDs

Current Background: Homo sapiens

Check Defaults



Clear All

- Disease** (2 selected)
- Functional_Annotations** (5 selected)
- Gene_Ontology** (3 selected)
- General_Annotations** (0 selected)
- Interactions** (1 selected)
- Literature** (0 selected)
- Pathways** (3 selected)
- Protein_Domains** (4 selected)
- Tissue_Expression** (0 selected)

Red annotation categories denote DAVID defined defaults

Combined View for Selected Annotation

Functional Annotation Clustering

Functional Annotation Chart

Functional Annotation Table

Question 1

- 10 most significantly enriched functional annotation terms among the over-expressed genes annotations
- Number of genes annotated with each of these terms
- Most enriched biological process GO term

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamin
<input type="checkbox"/>	UP_SEQ_FEATURE	CARBOHYD:N-linked (GlcNAc...) asparagine	RT		256	32,8	4,1E-20	1,2E-16
<input type="checkbox"/>	UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		55	7,1	3,9E-16	4,1E-14
<input type="checkbox"/>	UP_KW_PTM	Glycoprotein	RT		269	34,5	6,9E-15	2,2E-13
<input type="checkbox"/>	GOTERM_CC_DIRECT	extracellular matrix	RT		36	4,6	2,2E-12	1,1E-9
<input type="checkbox"/>	GOTERM_CC_DIRECT	extracellular region	RT		131	16,8	4,4E-12	1,1E-9
<input type="checkbox"/>	GOTERM_BP_DIRECT	cell adhesion	RT		55	7,1	4,4E-12	1,5E-8
<input type="checkbox"/>	UP_KW_DOMAIN	Signal	RT		233	29,9	4,2E-11	1,2E-9
<input type="checkbox"/>	GOTERM_MF_DIRECT	extracellular matrix structural constituent	RT		25	3,2	9,1E-11	7,6E-8
<input type="checkbox"/>	GOTERM_CC_DIRECT	plasma membrane	RT		254	32,6	1,3E-10	2,1E-8
<input type="checkbox"/>	UP_KW_CELLULAR_COMPONENT	Secreted	RT		133	17,1	4,1E-10	9,5E-9

- To view these genes click on :

ENSEMBL_GENE_ID	GENE NAME	Related Genes	Species
ENSG00000114948	ADAM metallopeptidase domain 23(ADAM23)	RG	Homo sapiens
ENSG00000108691	C-C motif chemokine ligand 2(CCL2)	RG	Homo sapiens
ENSG00000272398	CD24 molecule(CD24)	RG	Homo sapiens
ENSG00000135218	CD36 molecule(CD36)	RG	Homo sapiens
ENSG00000087589	Cas scaffold protein family member 4(CASS4)	RG	Homo sapiens
ENSG00000164176	EGF like repeats and discoidin domains 3(EDIL3)	RG	Homo sapiens
ENSG00000044524	EPH receptor A3(EPHA3)	RG	Homo sapiens
ENSG00000150893	FRAS1 related extracellular matrix 2(FREM2)	RG	Homo sapiens
ENSG00000049130	KIT ligand(KITLG)	RG	Homo sapiens

Question 2

- Use “Functional Annotation Table” tool

Annotation Summary Results [Help and Tool Manual](#)

Current Gene List:
siMitfvssiLuc_upgenes_lfc1_padj005

Current Background: Homo sapiens

780 DAVID IDs

Check Defaults [Clear All](#)

- Disease** (2 selected)
- Functional_Annotations** (5 selected)
- Gene_Ontology** (3 selected)
- General_Annotations** (0 selected)
- Interactions** (1 selected)
- Literature** (0 selected)
- Pathways** (3 selected)
- Protein_Domains** (4 selected)
- Tissue_Expression** (0 selected)

Red annotation categories denote DAVID defined defaults

Combined View for Selected Annotation

Functional Annotation Clustering

Functional Annotation Chart

Functional Annotation Table

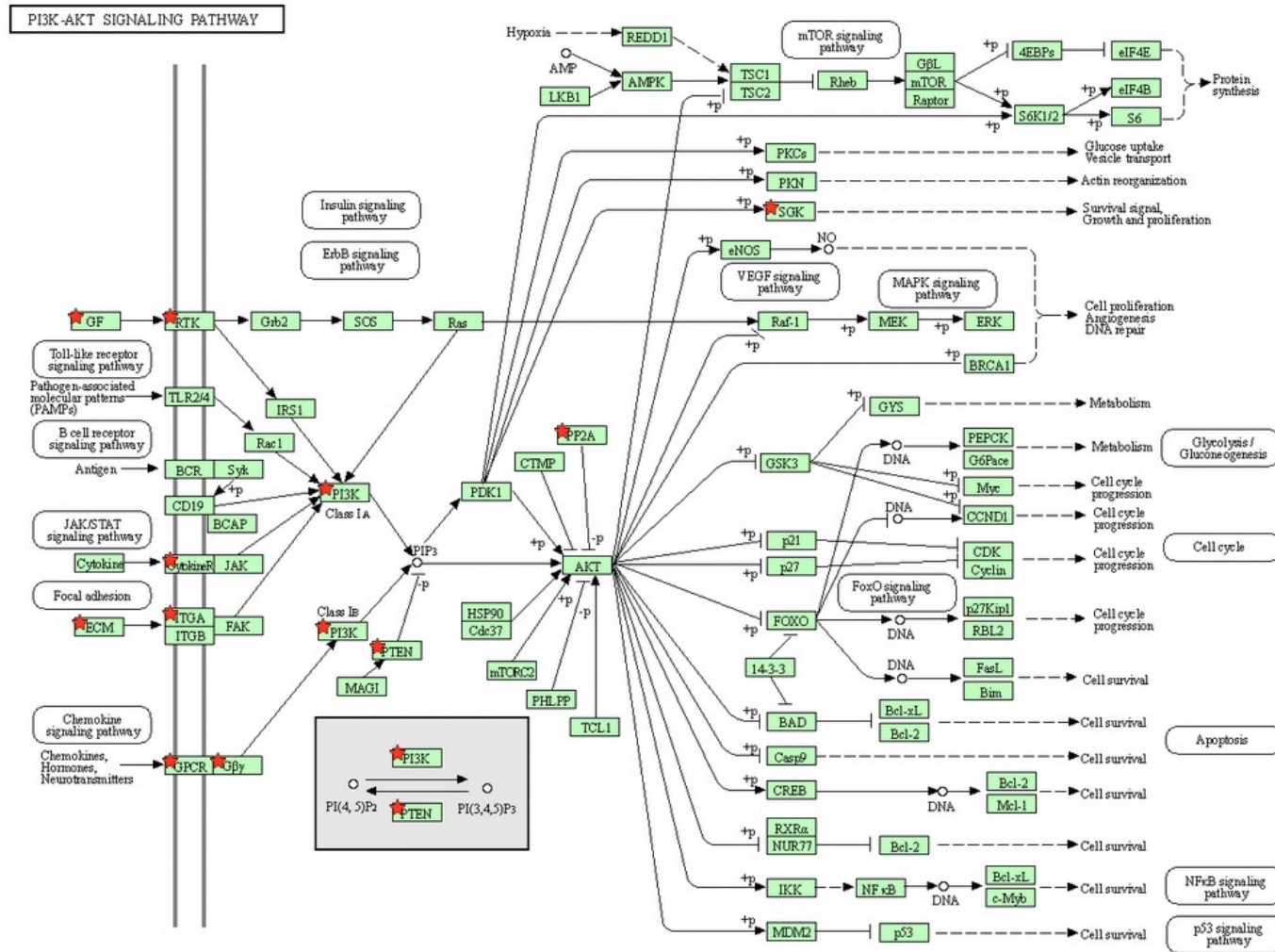
Question 2

- Search for *KITLG* in the page to see all associated annotations

ENSG00000049130	KIT ligand(KITLG)	Related Genes	Homo sapiens
BBID	124.GH-IGF1_thymopoiesis,		
BIOCARTA	Regulation of BAD phosphorylation , Erythrocyte Differentiation Pathway , IL 17 Signaling Pathway , CDK Regulation of DNA Replication , Melanocyte Development and Pigmentation Pathway ,		
GOTERM_BP_DIRECT	ovarian follicle development , neural crest cell migration , hematopoietic progenitor cell differentiation , myeloid leukocyte differentiation , positive regulation of leukocyte migration , positive regulation of myeloid leukocyte differentiation , cell adhesion , Ras protein signal transduction , positive regulation of cell proliferation , male gonad development , response to organic cyclic compound , peptidyl-tyrosine phosphorylation , mast cell apoptotic process , negative regulation of mast cell apoptotic process , embryonic hemopoiesis , ectopic germ cell programmed cell death , T cell proliferation , positive regulation of T cell proliferation , positive regulation of MAP kinase activity , positive regulation of melanocyte differentiation , positive regulation of Ras protein signal transduction , positive regulation of peptidyl-tyrosine phosphorylation , mast cell proliferation , positive regulation of mast cell proliferation , extrinsic apoptotic signaling pathway in absence of ligand , melanocyte migration , mast cell migration , positive regulation of hematopoietic progenitor cell differentiation , positive regulation of hematopoietic stem cell proliferation ,		
GOTERM_CC_DIRECT	extracellular region , extracellular space , cytoplasm , cytoskeleton , plasma membrane , membrane , integral component of membrane , lamellipodium , filopodium ,		
GOTERM_MF_DIRECT	cytokine activity , stem cell factor receptor binding , protein binding , growth factor activity ,		
INTERPRO	Stem cell factor , Four-helical cytokine-like_core ,		
KEGG_PATHWAY	MAPK signaling pathway , Ras signaling pathway , Rap1 signaling pathway , Phospholipase D signaling pathway , PI3K-Akt signaling pathway , Hematopoietic cell lineage , Melanogenesis , Pathways in cancer ,		
OMIM_DISEASE	Hyperpigmentation with or without hypopigmentation , Skin/hair/eye pigmentation 7, blond/brown hair , Deafness, autosomal dominant 69, unilateral or asymmetric , Waardenburg syndrome, type 2F ,		
PIR_SUPERFAMILY	Kit ligand precursor/stem cell factor ,		
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion ,		
UP_KW_CELLULAR_COMPONENT	Cytoskeleton , Membrane , Cytoplasm , Secreted , Cell projection , Cell membrane ,		
UP_KW_DISEASE	Deafness , Disease variant , Waardenburg syndrome , Non-syndromic deafness ,		
UP_KW_DOMAIN	Signal , Transmembrane , Transmembrane helix ,		
UP_KW_MOLECULAR_FUNCTION	Growth factor ,		
UP_KW_PTM	Glycoprotein , Disulfide bond ,		
UP_SEQ_FEATURE	CARBOHYD:N-linked (GlcNAc...) asparagine, CARBOHYD:N-linked (GlcNAc...) asparagine; partial, CARBOHYD:O-linked (GalNAc...) serine, CARBOHYD:O-linked (GalNAc...) threonine, REGION:Disordered, SITE:Not glycosylated, TOPO_DOM:Cytoplasmic, TOPO_DOM:Extracellular, TRANSMEM:Helical,		

- Click on “PI3K-Akt signalling pathway” to see all genes from this pathway

Question 2



Genes from your list are highlighted with a red star






Question 3.2



2. Import `pi3k_akt_signalling_genes.txt` file on Galaxy




Download from web or upload from disk

Regular Composite Collection Rule-based

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.




Name	Size	Type	Genome	Settings	Status
 pi3k_akt_signalling.	2 KB	tabular 	unspecified (?) 		0% 

Type (set all): Auto-detect  Genome (set all): unspecified (?) 







 Choose local files  Choose remote files  Paste/Fetch data **Start** Pause Reset Close

Question 3


3. Join `siMitfvssiLuc.up.annot.txt` with `pi3k_akt_signalling_genes.txt` on their common column (Ensembl gene ID)

 Join two Datasets side by side on a specified field (Galaxy Version 2.1.3)  







Join

   28: siMitfvssiLuc.up.annot.txt   


using column

Column: 1 


with

   32: pi3k_akt_signalling_genes.txt   


and column

Column: 1 


Keep lines of first input that do not join with second input

No 


Keep lines of first input that are incomplete

No 

Fill empty columns

No 

Keep the header lines

Yes 

Question 3

4. Use the **Cut** tool to retain only columns of interest (Gene name, norm.siLuc2, norm.siLuc3, norm.siMitf3, norm.siMitf4)

Cut columns from a table (Galaxy Version 1.0.2)

Cut columns
c29,c6-c9

Delimited by
Tab

From
33: Join two Datasets on data 32 and data 28

Email notification
 No
Send an email notification when the job completes.



Gene name	norm.siLuc2	norm.siLuc3	norm.siMitf3	norm.siMitf4
LAMC1	18515	19104	50030	53404
LAMA4	474	507	2261	2286
PTEN	1651	1603	4900	4571
LAMB1	2137	2136	4690	4760
LAMB2	3919	3750	7696	8207
KITLG	391	428	1121	1093
PIK3R3	507	438	1091	1114
FGF12	69	72	339	354
COL6A1	341	336	796	919
FGFR3	52	37	252	229
ITGA3	105	131	333	381
ITGA1	62	61	223	257
EFNA1	165	131	351	331
VTN	8	4	65	55
GNG2	101	127	243	235
COL6A3	43	27	98	129
VEGFD	57	50	137	138
CHRM2	2	1	36	29
FGF1	1	0	39	36
LPAR2	40	38	102	97
SGK2	7	9	46	36
FGFR2	30	24	79	74
IL7R	14	11	54	38
FGF7	0	1	21	13
COL5A2	16	26	48	63
PPP2R2B	7	5	24	26
PDGFRA	9	6	28	23
PIK3R5	4	5	12	27

5. Download this file, change file extension to txt and the first column name to NAME

→ [siMitfvssiLuc_up_pi3k_akt.txt](#)

Question 3

6. Heatmap using <http://www.heatmapper.ca/expression/>

