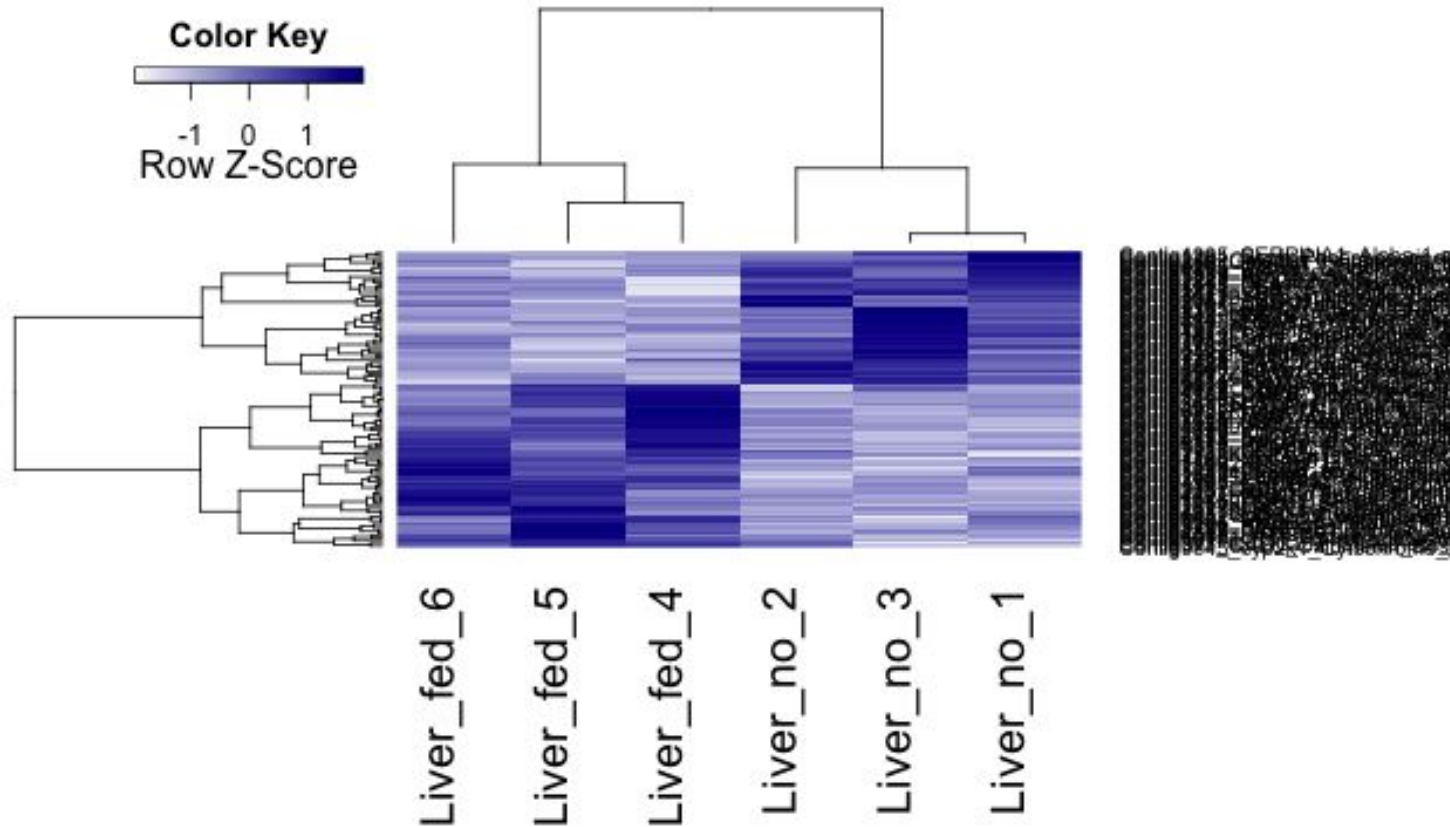


Liver: Group 3

Caroline Naso, Morgan Shannon, Sarah Paddon

What we've done so far

Generated a list of 110 differentially expressed genes between fed and unfed ($p=0.05$)



Exploration

- Began exploring gene list

Gene Name	KO Number	Check if TF	pval	log2 fold change	Notes
Contig255_Gm6484_Hepatocellular_carcinoma-associated_protein_TD26_homolog_Mus_musculus					
Contig2459_FASN_Fatty_acid_synthase_Gallus_gallus			0.027594606	-1.18538324	Protein coding gene related to formation of long-chain saturated fat
Contig2371_probable_hydrolase_PNKD_isoform_X3					
Contig2337_Protease_inhibitor_Austrelaps_labialis					
Contig230_mki67ipl_MKI67_FHA_domain-interacting_nucleolar_phosphoprotein-like_Xenopus_tropicalis	K14838		1.85E-02		nucleolar protein interacting with the FHA domain of MKI6
Contig227_ORAI1_Calcium_release-activated_calcium_channel_protein_1_Homo_sapiens			0.040325344	-1.236404738	calcium influx into t cells (immune system)
Contig223_ST7L_Suppressor_of_tumorigenicity_7_protein-like_Gallus_gallus					
Contig2223_RPL18A_60S_ribosomal_protein_L18a_Bos_taurus			4.42E-02	-5.70E-01	gene encoding ribosomal protein
Contig2157_tmem177_Transmembrane_protein_177_Xenopus_laevis					look at cited articles
Contig2157_MRAS_Ras-related_protein_M-Ras_Homo_sapiens					cell division
Contig2125_myof_Myoferlin_Xenopus_tropicalis			0.019122432	1.11669952	repair mechanical stress
Contig2084_CCDC106_Coiled-coil_domain-containing_protein_106_Bos_taurus			0.017003472	-1.250872671	
Contig205_SLC35B1_Solute_carrier_family_35_member_B1_Gallus_gallus					
Contig2-snap-gene-11.29-mRNA-1_transcript_Name:Aatk_Serine/threonine-protein_kinase_LMTK1_Mus_musculus					
Contig197_Protein_of_unknown_function_3288					
Contig1965_CHMP4C_Charged_multivesicular_body_protein_4c_Homo_sapiens			0.003695138	1.714572122	Multivesicular bodies (MVBs) are a specialised subset of en
Contig1939_NUS1_Nogo-B_receptor_Homo_sapiens			0.033131713	0.872580269	inhibiting neurite outgrowth
Contig193_RAP2C_Ras-related_protein_Rap-2c_Bos_taurus					
Contig1859_mboat7_Lysophospholipid_acyltransferase_7_Xenopus_laevis_4					
Contig1849_Gcdh_Glutaryl-CoA_dehydrogenase_mitochondrial_Mus_musculus					
Contig1779_APC2_Adenomatous_polyposis_coli_protein_2_Homo_sapiens_5					
Contig1738_Protein_of_unknown_function_2671					
Contig1722_Stxbp3_Syntaxin-binding_protein_3_Mus_musculus					SNARE, syntaxin, vesicle release?
Contig167_SLC26A4_Pendrin_Homo_sapiens					anion exchanger protein
Contig15632_TSMF_Elongation_factor_Ts_mitochondrial_Homo_sapiens					mitochondrial translational elongation factor
Contig1555_OLA1_Obg-like_ATPase_1_Gallus_gallus					GTP bindin, reg cell matrix adhesion, atpase, regulator of oxidative st
Contig1533_Vamp8_Vesicle-associated_membrane_protein_8_Rattus_norvegicus			0.013153212	0.922190446	SNARE, vesicle associated membrane protein 8

Kegg Pathways



[Home](#) [Help](#)

Query info

ID :	1457900362
query name :	Liver Fed vs. Unfed 2
program :	BLAST
method :	BBH
GENES data set :	hsa, mmu, gga, cfa, bta, ptr, xla, xtr, rno, ssc, tgu, ocu, cge, dme, mcf, pbi, acs, cmy, pss, amj, asn

[Feedback](#) [KEGG2](#) [KEGG](#) [GenomeNet](#) [Kanehisa Lab.](#)

Bi-directional is more accurate than single-directional best hit (SBH)

KO list created using organisms represented in 110 significant genes and all reptiles

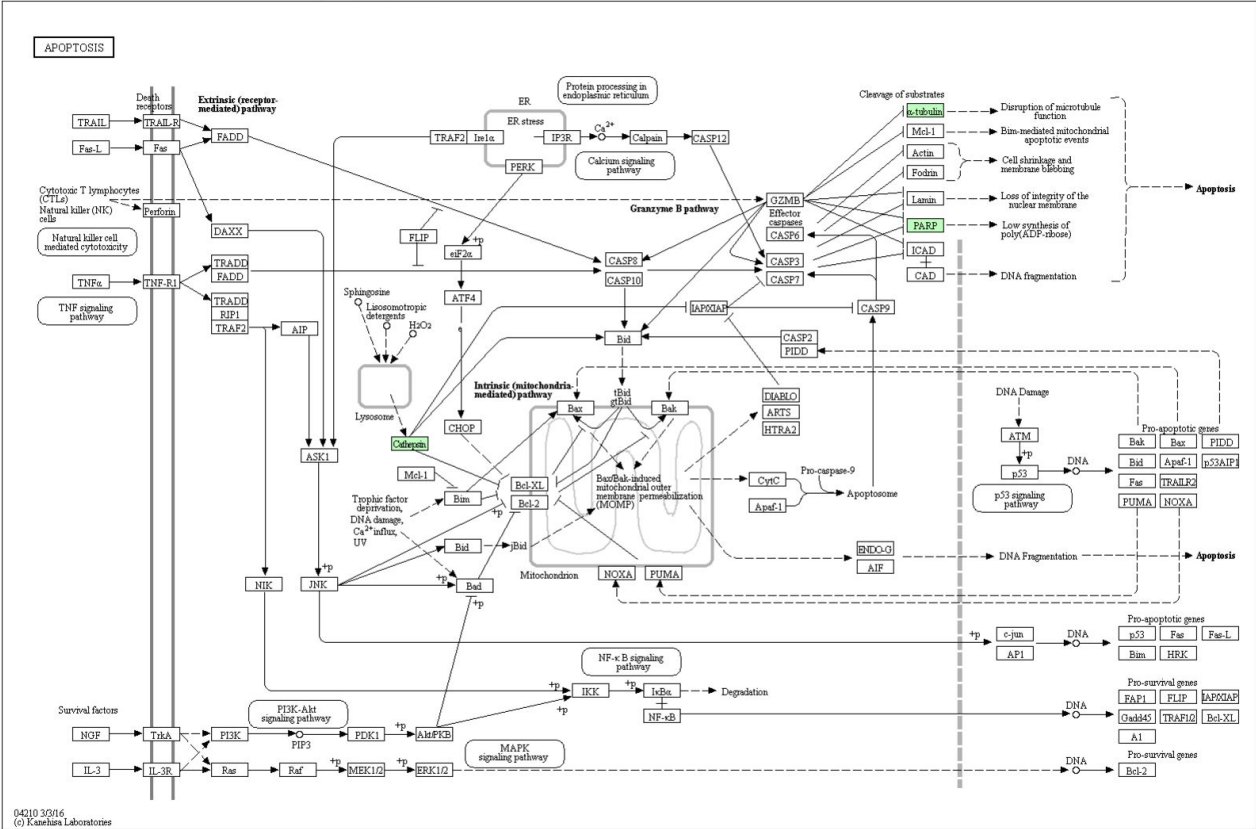
Pathways

[01200](#) Carbon metabolism (1)
[01212](#) Fatty acid metabolism (1)
[01230](#) Biosynthesis of amino acids (2)
[00680](#) Methane metabolism (1)
[00061](#) Fatty acid biosynthesis (1)
[00071](#) Fatty acid degradation (1)
[00100](#) Steroid biosynthesis (1)
[00561](#) Glycerolipid metabolism (1)
[00564](#) Glycerophospholipid metabolism (1)
[00230](#) Purine metabolism (1)
[00240](#) Pyrimidine metabolism (1)
[00260](#) Glycine, serine and threonine metabolism (2)
[00270](#) Cysteine and methionine metabolism (2)
[00310](#) Lysine degradation (1)
[00380](#) Tryptophan metabolism (1)
[00430](#) Taurine and hypotaurine metabolism (2)
[00450](#) Selenocompound metabolism (1)
[00460](#) Cyanoamino acid metabolism (1)
[00480](#) Glutathione metabolism (1)
[00533](#) Glycosaminoglycan biosynthesis - keratan sulfate (1)
[00601](#) Glycosphingolipid biosynthesis - lacto and neolacto series (1)
[00790](#) Folate biosynthesis (2)
[00670](#) One carbon pool by folate (1)
[00830](#) Retinol metabolism (2)
[00900](#) Terpenoid backbone biosynthesis (1)
[00909](#) Sesquiterpenoid and triterpenoid biosynthesis (1)
[00362](#) Benzoate degradation (1)
[00627](#) Aminobenzoate degradation (1)
[03010](#) Ribosome (3)
[04120](#) Ubiquitin mediated proteolysis (3)
[03018](#) RNA degradation (1)
[03410](#) Base excision repair (1)
[02010](#) ABC transporters (1)
[04014](#) Ras signaling pathway (1)
[04015](#) Rap1 signaling pathway (1)
[04010](#) MAPK signaling pathway (1)
[04630](#) Jak-STAT signaling pathway (1)
[04668](#) TNF signaling pathway (1)
[04020](#) Calcium signaling pathway (2)

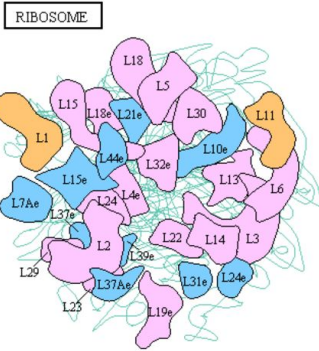
[04024](#) cAMP signaling pathway (2)
[04022](#) cGMP - PKG signaling pathway (1)
[04151](#) PI3K-Akt signaling pathway (1)
[04152](#) AMPK signaling pathway (1)
[04080](#) Neuroactive ligand-receptor interaction (1)
[04144](#) Endocytosis (2)
[04142](#) Lysosome (3)
[04110](#) Cell cycle (2)
[04210](#) Apoptosis (3)
[04510](#) Focal adhesion (1)
[04540](#) Gap junction (1)
[04610](#) Complement and coagulation cascades (2)
[04611](#) Platelet activation (2)
[04612](#) Antigen processing and presentation (1)
[04910](#) Insulin signaling pathway (2)
[04920](#) Adipocytokine signaling pathway (1)
[04914](#) Progesterone-mediated oocyte maturation (1)
[04917](#) Prolactin signaling pathway (1)
[04918](#) Thyroid hormone synthesis (1)
[04924](#) Renin secretion (1)
[04925](#) Aldosterone synthesis and secretion (1)
[04270](#) Vascular smooth muscle contraction (1)
[04975](#) Fat digestion and absorption (1)
[04728](#) Dopaminergic synapse (1)
[04380](#) Osteoclast differentiation (1)
[04212](#) Longevity regulating pathway - worm (2)
[05200](#) Pathways in cancer (1)
[05206](#) MicroRNAs in cancer (1)
[05205](#) Proteoglycans in cancer (1)
[05203](#) Viral carcinogenesis (1)
[05212](#) Pancreatic cancer (1)
[05214](#) Glioma (1)
[05220](#) Chronic myeloid leukemia (1)
[05218](#) Melanoma (1)
[05222](#) Small cell lung cancer (1)
[05223](#) Non-small cell lung cancer (1)
[05323](#) Rheumatoid arthritis (1)
[05340](#) Primary immunodeficiency (1)
[05010](#) Alzheimer's disease (1)
[05012](#) Parkinson's disease (2)

[05014](#) Amyotrophic lateral sclerosis (ALS) (1)
[05016](#) Huntington's disease (1)
[05030](#) Cocaine addiction (1)
[05031](#) Amphetamine addiction (1)
[05034](#) Alcoholism (1)
[04930](#) Type II diabetes mellitus (1)
[04932](#) Non-alcoholic fatty liver disease (NAFLD) (1)
[04931](#) Insulin resistance (1)
[05130](#) Pathogenic Escherichia coli infection (1)
[05132](#) Salmonella infection (1)
[05131](#) Shigellosis (1)
[05152](#) Tuberculosis (1)
[05100](#) Bacterial invasion of epithelial cells (1)
[05166](#) HTLV-I infection (1)
[05162](#) Measles (2)
[05164](#) Influenza A (2)
[05161](#) Hepatitis B (1)
[05160](#) Hepatitis C (1)
[05168](#) Herpes simplex infection (1)
[05169](#) Epstein-Barr virus infection (1)
[05144](#) Malaria (1)

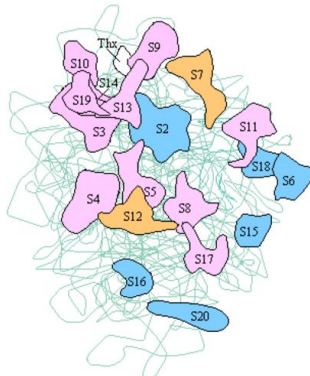
Model Pathways: Apoptosis



Model Pathways: Ribosome



Large subunit (*Haloarcula marismortui*)



Small subunit (*Thermus aquaticus*)

Ribosomal RNAs

Bacteria / Archaea	23S	5S		16S
Eukaryotes	25S	5S	5.8S	18S

Ribosomal proteins

EF-Tu	S10	L3	L4	L23	L2	S19	L22	S3	RP-L16	L29	
	S20e	L3e	L4e	L23Ae	L8e	S15e	L17e	S3e		L35e	L7/L12 stalk
									L10e		

S17	L14	L24		L5	S14	S8	L6		L18	S5	L30	L15	SecY
S11e	L23e	L26e	S4e	L11e	S29e	S15Ae	L9e	L32e	L19e	L5e	S2e	L7e	L27Ae

IF1	L36	S13	S11	S4	RpoA	L17	L13	S9
L34e	L14e	S18e	S14e	S9e		L18e	L13Ae	S16e

EF-Tu,G	S7	S12	L7A	RpoC,B	L7/L12	L12	L10	L1	L11
	S5e	S23e	L30e	L7Ae		L12	L10Ae	L12e	
					L11	L12e	L10Ae	L12e	

S2	EF-Ts	IF2	S15	IF3	L35	L20	L34	RF1	L31	L32	L9	S18	S6
SAe			S13e										

L28	L33	L21	L27	PtsY,Ffh	S16	L19	S1	S20	S21	L25
-----	-----	-----	-----	----------	-----	-----	----	-----	-----	-----

L10e	L13e	L15e	L21e	L24e	L31e	L35Ae	L37e	L37Ae	L39e	L40e	L41e	L44e
------	------	------	------	------	------	-------	------	-------	------	------	------	------

S3Ae	S6e	S8e	S17e	S19e	S24e	S25e	S26e	S27e	S27Ae	S28e	S30e	LX
------	-----	-----	------	------	------	------	------	------	-------	------	------	----

L6e	L18Ae	L22e	L27e	L28e	L29e	L36e	L38e
-----	-------	------	------	------	------	------	------

S7e	S10e	S12e	S21e
-----	------	------	------

Isolate Transcription Factors - GO Terms

- Associated our 110 genes with GO terms
- Manually searched for genes tagged with transcription factor related GO terms
 - **GO:003677** = "DNA binding"
 - **GO:0006355** = "regulation of transcription, DNA-dependent"
 - **GO:0003700** = "Transcription Factor activity"
- Two transcription factors identified
 - **SUB1**: Activated RNA Polymerase II transcriptional coactivator p15
 - **ZNF180**: Zinc finger protein 180

```
1 #open and read in as list our file containing our genes of interest
2 genesofinterest = open('diffexpressedgenes.txt', 'r')
3 genelist = genesofinterest.readlines()
4 genesofinterest.close()
5
6 #edit each gene name to change to appropriate format
7 goodnamelist = []
8 for gene in genelist:
9     newname = gene.rstrip('\n')
10    withcarot = '>' + newname
11    goodnamelist.append(withcarot)
12
13 #read in file containing gene ID terms and GO terms and save it as a list
14 allgoterms = {}
15 allgos = open('contigs_GO.txt', 'r')
16 golist = allgos.readlines()
17
18 goodgolist = []
19 for item in golist:
20     newname2 = item.rstrip('\n')
21     goodgolist.append(newname2)
22
23 #convert list of gene ID and GO terms to dictionary
24 #gene ID is key and associated GO terms are the values
25 allgodict = {}
26 i = 0
27 while i < len(goodgolist):
28     allgodict.update({goodgolist[i]: goodgolist[i+1]})
29     i += 2
30
31 #create dictionary of gene IDs and GO terms only for our genes of interest
32 diffexpressdict = {}
33 for name in goodnamelist:
34     if name[0] == '>':
35         for term in goodgolist:
36             if term[0] == '>':
37                 if name == term:
38                     diffexpressdict.update({name: allgodict[name]})
39
40 #write our genes of interest and associated GO terms to a text file
41 output = open('genes_with_GOterms.txt', 'w')
42 for contig in diffexpressdict:
43     output.write(contig + ';' + diffexpressdict[contig] + '\n')
44 output.close()
45
```

What we plan to do

Refining Kegg Pathways

- Some genes were not assigned Kegg numbers
 - Including our 2 transcription factors
- Liter



[Home](#)

[\[KO list\]](#) [\[BRITE hierarchies\]](#) [\[Pathway map\]](#) [\[Threshold change\]](#) [\[Download KO list\]](#)

Query gene : KO assignment

Liver Fed vs. Unfed 2

Contig1264_Psph_Phosphoserine_phosphatase_Rattus_norvegicus K01079

Contig1395_SERPINA1_Alpha-1-antitrypsin_Homo_sapiens K04525

Contig3672_A2M_Alpha-2-macroglobulin_Homo_sapiens K03910

Contig2677_ATPIF1_ATPase_inhibitor_mitochondrial_Sus_scrofa

Contig2223_RPL18A_60S_ribosomal_protein_L18a_Bos_taurus K02882

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig1264_Psph_Phosphoserine_phosphatase_catalytic_subunit_Taeniopygia_guttata K03635

Contig783_SUB1_Activated_RNA_polymerase_II_transcriptional_coactivator_p15_Gallus_gallus

Contig4401_RPLP0_60S_acidic_ribosomal_protein_P0_Gallus_gallus_1 K02941

Contig3381_RPS24_40S_ribosomal_protein_S24_Homo_sapiens_1 K02974

Contig5624_LRP1_Low-density_lipoprotein_receptor-related_protein_1_Gallus_gallus

Contig12414_Npffr1_Neuropeptide_FF_receptor_1_Rattus_norvegicus K04240

Contig3480_ZNF180_Zinc_finger_protein_180_Homo_sapiens

Improving Transcription Factor Identification

- Increase p value from 0.05 to 0.1 to widen the pool of potential transcription factors
- Gene Ontology is notoriously ineffective at annotating transcription factors
 - Transcription factor specific databases

Nature Reviews Genetics **10**, 252-263 (April 2009) | doi:10.1038/nrg2538

A census of human transcription factors: function, expression and evolution

Juan M. Vaquerizas, Sarah K. Kummerfeld, Sarah A. Teichmann & Nicholas M. Luscombe

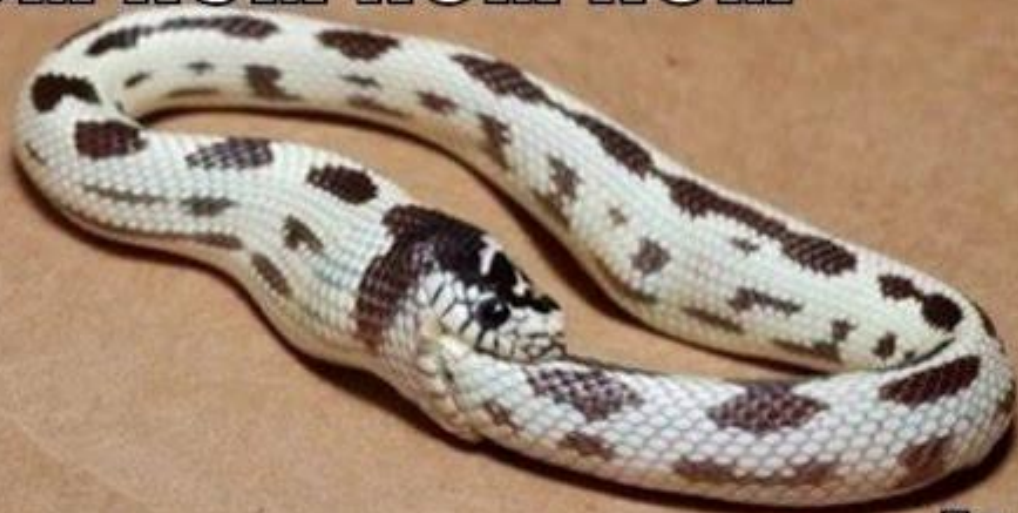
Transcription factors are key cellular components that control gene expression: their activities determine how cells function and respond to the environment. Currently, there is great interest in research into human transcriptional regulation. However, surprisingly little is known about these regulators themselves. For example, how many transcription factors does the human genome contain? How are they expressed in different tissues? Are they evolutionarily conserved? Here, we present an analysis of 1,391 manually curated sequence-specific DNA-binding transcription factors, their functions, genomic organization and evolutionary conservation. Much remains to be explored, but this study provides a solid foundation for future investigations to elucidate regulatory mechanisms underlying diverse mammalian biological processes.

Further analysis using the GO database (Fig. 1b) showed that most human TFs are unannotated, indicating that they remain uncharacterized

Supervised Clustering

- Identify genes that cluster with our genes of interest (transcription factors)
- Map those clusters to KEGG pathways

OM NOM NOM NOM



.... wait.