WEEK 8 ASSIGNMENT

Introduction to Computational Biology – BIOL 509000 | Fall 1 2020

Christina Morgenstern

Question #1: How many rows of protein identifiers are included in this data set?

There are 199 rows of protein identifiers included in the data set.

Question #2: How many IDs are reported as not present in the DAVID DB (database)? How many are listed as "Ambiguous" and are awaiting approval for conversion?

72 IDs are reported as being not present in DAVID DB.
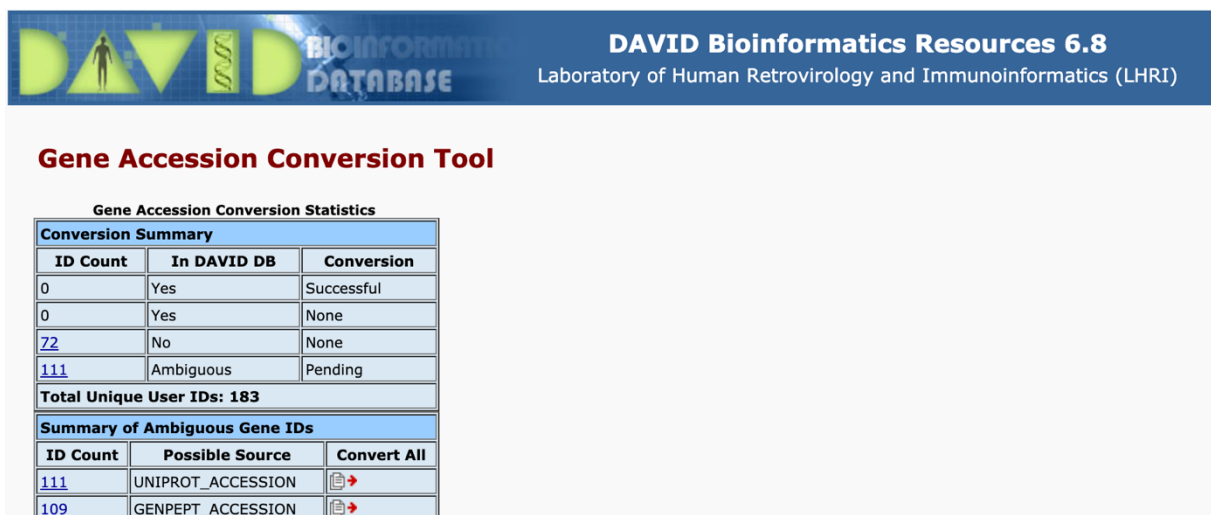111 IDs are reported as being ambiguous.

Figure 1. Conversion summary of DAVID database query.

Question #3: What species is listed in the box above the List Manager?

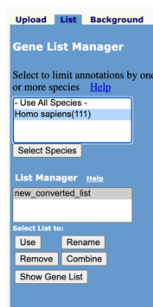The species *Homo sapiens* is listed in the Gene List Manager (see Fig. 2)

Figure 2. Gene List Manager of DAVID DB.

Question #4: Find "lectin, mannose-binding 2" and click on the gene name. What chromosome is this gene located on in humans?

The gene encoding lectin, mannose-binding 2 protein is located on chromosome 5 (see Fig. 3).

Gene Report

List Id: 10960	lectin, mannose binding 2(LMAN2)	Related Genes	Homo sapiens
CHROMOSOME	5,		
CYTOBAND	5q35.3,		
ENSEMBL_GENE_ID	ENSG00000169223,		
ENTREZ_GENE_ID	10960,		
GENERIF_SUMMARY	Live imaging of the canine homologous protein., VIP36 interacts with molecular chaperone BiP and has a role in the quality control of secretory proteins, selective interaction of VIPL and VIP36 with the deglucosylated trimannose in the D1 branch of high-mannose-type oligosaccharides but with different pH dependence., Results identify the glycoprotein alpha1-antitrypsin as a cargo of VIP36., Interaction with VIP36 is dependent on glycosylation at the same sites that allow GC-C to fold and bind ligand.,		
OFFICIAL_GENE_SYMBOL	LMAN2,		
SP_COMMENT	cofactor:Binds 2 calcium ions per subunit., function:Plays a role as an intracellular lectin in the early secretory pathway. Interacts with N-acetyl-D-galactosamine and high-mannose type glycans and may also bind to O-linked glycans. Involved in the transport and sorting of glycoproteins carrying high mannose-type glycans., similarity:Contains 1 L-type lectin-like domain., tissue specificity:Ubiquitous.,		

Figure 3. Result for lectin, mannose-binding 2 gene.

Question #5: What is the official gene symbol for this gene?

The official gene symbol for lectin, mannose-binding 2 gene is *LMAN2* (see Fig. 3).

Question #6: For "lectin, mannose-binding 2," click on "RG" to find related genes. How many related genes are reported, and what are they?

There are two functional related genes listed: transmembrane p24 trafficking protein 3 (*TMED3*) and translocation associated membrane protein 1 (*TRAM1*) (see Fig. 4).

Functional Related Genes

Search Scope: User's List

⊞ Options

Rerun Using Options

111 gene(s) were searched. 3 genes passed the filter.

Similarity Score: ■ Very High (0.75-1) ▨ High (0.5-0.75) □ Moderate (0.25-0.5) □ Low (<0.25)

#	Functionally Related Gene	Species	Kappa
1	lectin, mannose binding 2(LMAN2)	Homo sapiens	1.00
2	transmembrane p24 trafficking protein 3(TMED3)	Homo sapiens	0.52
3	translocation associated membrane protein 1(TRAM1)	Homo sapiens	0.37

Figure 4. Related genes for lectin, mannose-binding 2 gene.

Question #7: For this dataset and the criteria stated above, how many functional clusters were determined by the software?

The Functional Annotation Clustering resulted in 30 clusters (see Fig. 5).

Figure 5. Result of Functional Annotation Clustering.

Question #8: What is the enrichment score for this cluster?

Annotation cluster 4 has an enrichment score of 2.71 (see Fig. 6).

Figure 6. Enrichment Score for Annotation Cluster 4.

Question #9: How many genes from your list are not in the output?

63 genes from my list are not in the output when classification stringency is set to low (see Fig. 7).

Figure 7. Result of Gene Functional Classification Tool.

Question #10: How many term records are listed on this term report?

The term report lists 123 term records (see Fig. 8).

Term Report

Gene Cluster 2
Current Gene List: new_converted_list
Current Background: Homo sapiens
5 DAVID IDs
123 term records

Category	Term	RT	Genes in Group	Count	LT	PH	PT	%	P-Value	Fold Enrichment	Genes not in Group
UP_KEYWORDS	Spliceosome	RT		5	5	127	20581	100,0	1,4E-9	1.3E2	(1)
UP_KEYWORDS	mRNA splicing	RT		5	5	260	20581	100,0	2,5E-8	6.3E1	(3)
GOTERM_BP_DIRECT	mRNA splicing, via spliceosome	RT		5	5	222	16792	100,0	3,0E-8	6.1E1	(6)
UP_KEYWORDS	mRNA processing	RT		5	5	332	20581	100,0	6,7E-8	5.0E1	(3)
KEGG_PATHWAY	Spliceosome	RT		4	5	133	6879	80,0	7,1E-6	3.9E1	(2)
GOTERM_BP_DIRECT	RNA splicing, via transesterification reactions	RT		3	5	25	16792	60,0	1,3E-5	2.7E2	

Figure 8. Term record for gene group 2.