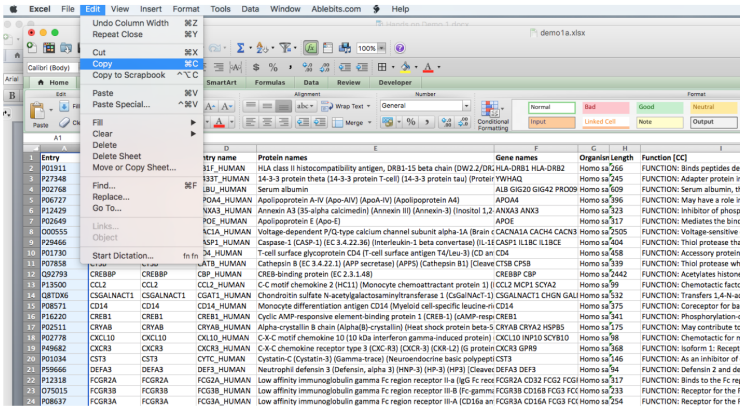


Proteomics Exercise Week 6 Part 2

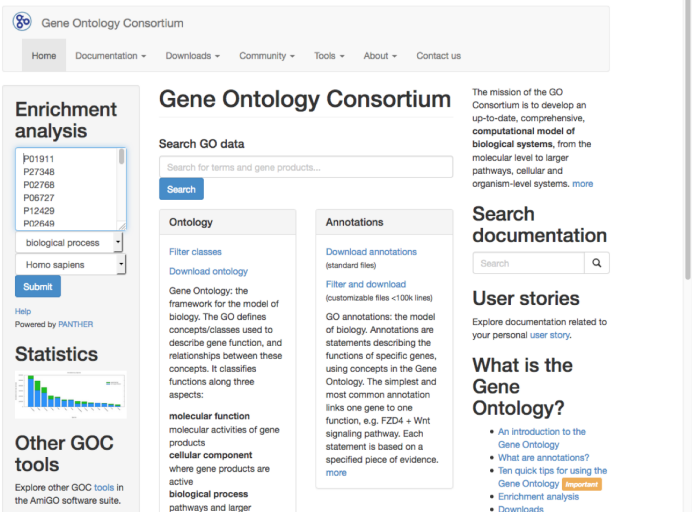
1. Take the results from our first step mapping to UniProt functional annotation and copy the 73 UniProt identifiers from your Excel file.



2. Paste them into the Enrichment Analysis box on the Gene Ontology home page at www.geneontology.org.

The tool should be set to "Biological Process" Ontology and "Homo sapiens."

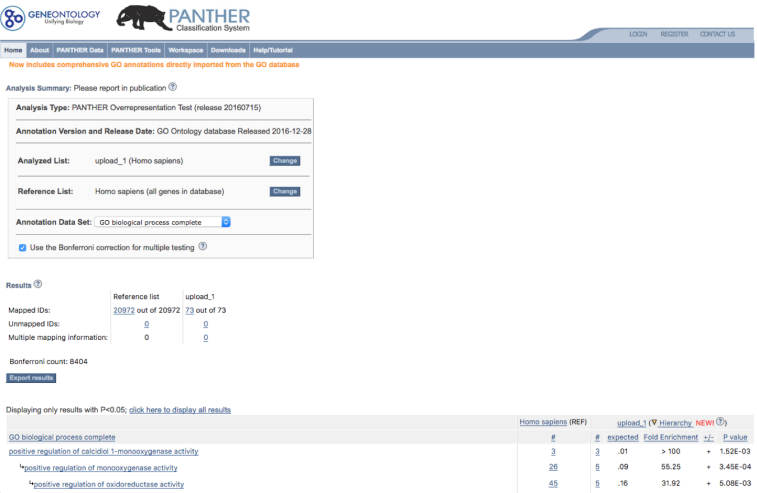
Click "Submit."



3. Here is the top part of the results page that you should see.

Databases change over time, so it is possible there might be small differences in the results on this page from what was seen in the demo video, these slides, and your results. These are probably not significant changes, but improvements in the databases and tools over time.

Scroll down the page.



4. Initial results are presented in GO hierarchy with more specific nodes on the left and parental nodes indented.

On the right side of the page, we see

- The number of all human proteins annotated with that term
- The number of proteins in our list with that term
- The expected number if random
- A fold increase and P-Value

By default, all displayed nodes are at the $p < 0.05$ threshold. Formally the p-value tests if a null-hypothesis can be rejected or not at a particular significance level (commonly set at $p < 0.05$). In this case and similar cases

with pathway tools and genetic tests, the null-hypothesis is that the result was a random event. A result less than 0.05 means that either the null-hypothesis is false or we happened to see a very improbable but not impossible result.

See the demo video for additional discussion.

Displaying only results with P<0.05; click here to display all results

	Homo sapiens (REF)	#	expected	Fold Enrichment	+/-	P value
GO biological process complete		3	3	.01	> 100	+ 1.52E-03
positive regulation of catalytic 1-monooxygenase activity		26	5	.09	55.25	+ 3.45E-04
positive regulation of monooxygenase activity		45	5	.16	31.92	+ 5.08E-03
positive regulation of oxidoreductase activity		6214	64	21.63	2.03	+ 4.66E-04
regulation of metabolic process		11037	65	38.42	1.72	+ 3.29E-08
regulation of biological process		11553	67	40.60	1.65	+ 1.04E-07
biological regulation		2397	27	8.34	3.24	+ 1.10E-04
regulation of catalytic activity		2918	30	10.16	2.96	+ 9.64E-05
regulation of molecular function		3064	37	10.67	3.47	+ 3.67E-09
positive regulation of metabolic process		5301	55	18.45	2.88	+ 3.86E-15
positive regulation of biological process		1552	19	5.51	3.45	+ 1.17E-02
positive regulation of catalytic activity		1674	23	6.52	3.53	+ 4.17E-04
positive regulation of molecular function		57	5	.20	25.20	+ 1.61E-02
regulation of monooxygenase activity		1784	28	6.21	4.51	+ 2.37E-08
positive regulation of biosynthetic process		4356	35	15.16	2.31	+ 1.77E-03
regulation of biosynthetic process		7	3	.02	> 100	+ 1.91E-02
regulation of catalytic 1-monooxygenase activity		9	3	.03	95.76	+ 4.04E-02
regulation of vitamin D biosynthetic process		4305	34	14.98	2.27	+ 4.64E-03
regulation of cellular biosynthetic process		5895	63	20.52	2.10	+ 3.18E-04
regulation of cellular metabolic process		10453	65	36.39	1.79	+ 1.13E-08
regulation of lipid biosynthetic process		137	7	.48	14.68	+ 4.77E-03
regulation of lipid metabolic process		300	9	1.04	8.82	+ 8.96E-03
regulation of primary metabolic process		5845	62	20.35	2.06	+ 8.88E-04
positive regulation of vitamin D biosynthetic process		3	3	.01	> 100	+ 1.52E-03
positive regulation of steroid metabolic process		25	5	.09	57.46	+ 2.84E-04
positive regulation of lipid metabolic process		136	8	.47	16.90	+ 2.43E-04

5. Now, look at the sorted p-value table.

The hierarchy is destroyed, but all terms are here and we see a lot of very significant parental terms near the top with their still significant children down the table.

We see lots of related terms: Response to bacterium; Immune system processes; Defense responses.

As we look through the list, it seems that this set of proteins is definitely related to an immune response, probably to some external stimulus like a bacterial or other infectious agent stimulating cytokines and other defense responses.

This is still very general, but you have quickly narrowed the options while knowing nothing about the origin of the data.

Displaying only results with P<0.05; click here to display all results

	Homo sapiens (REF)	#	expected	Fold Enrichment	+/-	P value
GO biological process complete		550	32	1.91	16.71	+ 3.72E-27
response to bacterium		939	36	3.27	11.01	+ 3.87E-25
positive regulation of immune system process		875	35	3.05	11.49	+ 7.45E-25
response to other organism		876	35	3.05	11.48	+ 7.74E-25
response to external biotic stimulus		1282	39	4.39	8.88	+ 2.00E-24
defense response		1169	38	4.07	9.34	+ 2.23E-24
immune response		1475	41	5.13	7.99	+ 2.37E-24
regulation of immune system process		908	35	3.16	11.07	+ 2.56E-24
response to biotic stimulus		2105	45	7.23	6.14	+ 6.89E-23
immune system process		723	31	2.70	11.49	+ 2.79E-21
regulation of defense response		1911	42	6.65	6.31	+ 3.65E-21
response to external stimulus		3386	51	11.79	4.33	+ 1.99E-20
response to stress		2720	47	9.47	4.96	+ 2.15E-20
regulation of multicellular organismal process		921	32	3.21	9.98	+ 2.63E-20
regulation of immune response		345	23	1.20	19.10	+ 2.18E-19
response to molecule of bacterial origin		588	27	2.05	13.19	+ 2.18E-19
regulation of cytokine production		332	22	1.16	19.04	+ 2.42E-18
response to lipopolysaccharide		2030	40	7.07	5.66	+ 5.48E-18
positive regulation of response to stimulus		924	30	3.22	9.33	+ 7.55E-18
response to lipid		463	24	1.68	14.28	+ 1.65E-17
inflammatory response		622	26	2.17	12.01	+ 1.74E-17
positive regulation of immune response		379	22	1.32	16.68	+ 3.99E-17
positive regulation of cytokine production		507	24	1.76	13.60	+ 5.01E-17
defense response to other organism						

6. Go back to the home page and try the Molecular Function Ontology.

Gene Ontology Consortium

Home Documentation Downloads Community Tools About Contact us

Enrichment analysis

P01911
P27348
P02768
P06727
P12429
P02649

molecular function

Homo sapiens

Submit

Help
Powered by PANTHER

Statistics

Other GOC tools

Gene Ontology Consortium

Search GO data

Search for terms and gene products...

Search

Ontology

Filter classes

Download ontology

Gene Ontology: the framework for the model of biology. The GO defines concepts/classes used to describe gene function, and relationships between these concepts. It classifies functions along three aspects:

- molecular function
- cellular component
- biological process

where gene products are active pathways and larger processes made up of the activities of multiple gene products.

Annotations

Download annotations (standard files)

Filter and download (customizable files <100k lines)

GO annotations: the model of biology. Annotations are statements describing the functions of specific genes, using concepts in the Gene Ontology. The simplest and most common annotation links one gene to one function, e.g. FZD4 + Wnt signaling pathway. Each statement is based on a specified piece of evidence. [more](#)

The mission of the GO Consortium is to develop an up-to-date, comprehensive, **computational model of biological systems**, from the molecular level to larger pathways, cellular and organism-level systems. [more](#)

Search documentation

Search

User stories

Explore documentation related to your personal [user story](#).

What is the Gene Ontology?

- An introduction to the Gene Ontology
- What are annotations?
- Ten quick tips for using the Gene Ontology [\[download\]](#)
- Enrichment analysis
- Downloads

7. A much smaller list of results here.

You see more immunoglobulin binding, cytokine-related functions, and "lipopolysaccharide receptor activity," which is commonly seen in response to bacterial infections.

Now go back and try the cellular component.

Results ?	Reference list	upload_1
Mapped IDs:	20972 out of 20972	73 out of 73
Unmapped IDs:	0	0
Multiple mapping information:	0	0
Bonferroni count: 2788		
Export results		

Displaying only results with P<0.05; [click here to display all results](#)

	Homo sapiens (REF)		upload_1 (▼ Hierarchy NEW! ?)		
GO molecular function complete	#	#	expected	Fold Enrichment	+/- P value
lipopolysaccharide receptor activity	5	3	.02	> 100	+ 2.32E-03
↳ signaling pattern recognition receptor activity	17	4	.06	67.60	+ 1.25E-03
↳ pattern recognition receptor activity	17	4	.06	67.60	+ 1.25E-03
IgG binding	12	3	.04	71.82	+ 3.15E-02
↳ immunoglobulin binding	24	4	.08	47.88	+ 4.89E-03
↳ protein complex binding	769	12	2.68	4.48	+ 3.76E-02
↳ protein binding	11027	66	38.38	1.72	+ 1.04E-08
↳ binding	14535	68	50.59	1.34	+ 2.02E-03
↳ macromolecular complex binding	1361	16	4.74	3.38	+ 4.11E-02
cytokine activity	221	13	.77	16.90	+ 2.64E-09
↳ receptor binding	1550	29	5.40	5.38	+ 3.18E-11
cytokine receptor binding	280	15	.97	15.39	+ 1.49E-10
protein homodimerization activity	752	12	2.62	4.58	+ 3.01E-02
↳ protein dimerization activity	1174	17	4.09	4.16	+ 1.25E-03
↳ identical protein binding	1396	20	4.86	4.12	+ 1.10E-04
Unclassified	3745	1	13.04	< 0.2	- 0.00E00

8. Here is a partial list of the Cellular Component results sorted by p-value. Please look at the entire list in the hierarchy and p-value sorted displays. Most of the proteins seem to be membrane bound, external side or extracellular. Also, many located in vesicles related to the immune response or to transport to the extracellular region.

Please click on the terms to educate yourselves to some of the meanings.

Also note the "lipopolysaccharide receptor complex," which functions as a lipopolysaccharide (LPS) receptor that primes the innate immune response against bacterial pathogens.

Also, see "myelin sheath," an electrically insulating fatty layer that surrounds the axons of many neurons.

We did not notice a nerve system connection before though. If you go back to biological process you will see a neuron projection regeneration term listed closer to the bottom of the sorted list.

So is this all helpful in any way...?

Displaying only results with P<0.05; [click here to display all results](#)

	Homo sapiens (REF)		upload_1 (Hierarchy) NEW! ?		
GO cellular component complete	#	#	expected	Fold Enrichment	+/- Δ P value
extracellular space	1495	32	5.20	6.15	+ 6.99E-15
extracellular region	4578	49	15.94	3.07	+ 1.98E-13
extracellular region part	3852	43	13.41	3.21	+ 2.20E-11
cell surface	786	21	2.74	7.68	+ 2.40E-10
external side of plasma membrane	253	14	.88	15.90	+ 3.43E-10
side of membrane	454	15	1.58	9.49	+ 6.24E-08
myelin sheath	181	10	.63	15.87	+ 1.13E-06
vesicle	4071	37	14.17	2.61	+ 2.83E-06
cytoplasmic vesicle	1799	22	6.26	3.51	+ 1.49E-04
intracellular vesicle	1808	22	6.29	3.50	+ 1.62E-04
extracellular exosome	2750	27	9.57	2.82	+ 3.02E-04
extracellular vesicle	2765	27	9.62	2.81	+ 3.37E-04
extracellular organelle	2766	27	9.63	2.80	+ 3.40E-04
cell body	550	12	1.91	6.27	+ 5.75E-04
cytoplasmic vesicle part	939	15	3.27	4.59	+ 9.33E-04
lipopolysaccharide receptor complex	6	3	.02	> 100	+ 1.88E-03
endocytic vesicle	266	8	.93	8.64	+ 5.65E-03
secretory granule	375	9	1.31	6.89	+ 8.44E-03
neuronal cell body	485	10	1.69	5.92	+ 9.39E-03
plasma membrane part	2703	24	9.41	2.55	+ 1.10E-02
membrane raft	302	8	1.05	7.61	+ 1.41E-02
membrane microdomain	303	8	1.05	7.59	+ 1.45E-02
neuron projection	1042	14	3.63	3.86	+ 1.67E-02