

# **Gene Set Analysis –Methods and Tools**

## **Exercise 2.1**

**Antonio Mora**

**Lv Xuanyi**

**(Xie Chengshu)**

**December, 2019**



# Exercise 1. Using DAVID

DAVID is the acronym for “The Database for Annotation, Visualization and Integrated Discovery”. You can find it at: <https://david.ncifcrf.gov/home.jsp>

The picture below is its main page, which contains some general information about this platform. DAVID provides four main tools (details on the website):

1. Functional Annotation
2. Gene Functional Classification
3. Gene ID Conversion
4. Gene Name Batch Viewer

DAVID Bioinformatics Resources 6.8  
Laboratory of Human Retrovirology and Immunoinformatics (LHRI)

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service Why DAVID? About Us

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

Recommending: A [paper](#) published in *Nature Protocols* describes step-by-step procedure to use DAVID!

**Welcome to DAVID 6.8**

2003 - 2018

**What's Important in DAVID?**

- [Cite DAVID](#)
- [IDs of Affy Exon and Gene arrays supported](#)
- [Novel Classification Algorithms](#)
- [Pre-built Affymetrix and Illumina backgrounds](#)
- [User's customized gene background](#)
- [Enhanced calculating speed](#)

**Statistics of DAVID**

DAVID Citations (2003-2017)

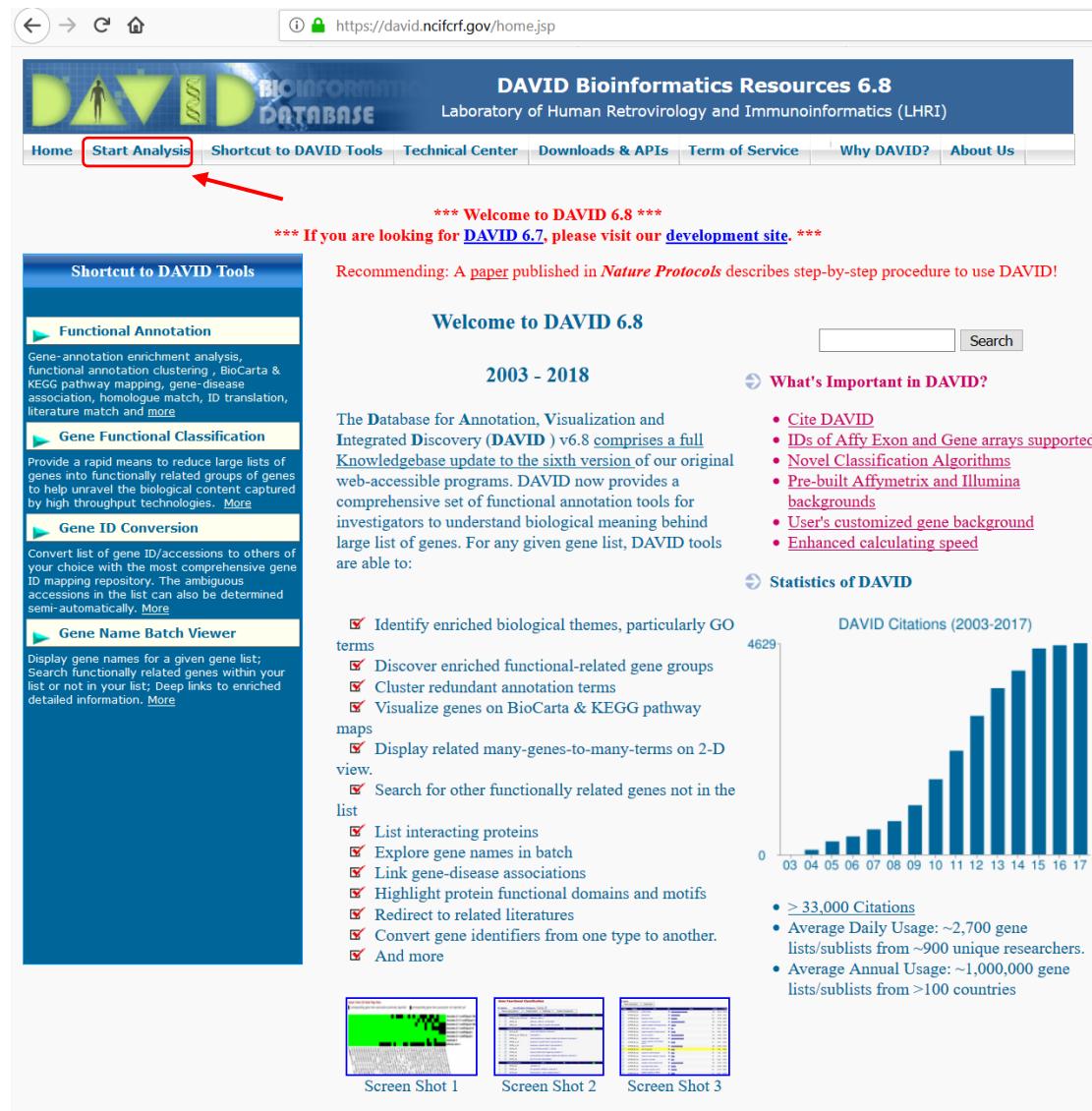
Year	Citations
2003	~10
2004	~20
2005	~30
2006	~40
2007	~50
2008	~60
2009	~70
2010	~100
2011	~200
2012	~300
2013	~400
2014	~450
2015	~4600
2016	~4600
2017	~4600

• ≥ 33,000 Citations  
• Average Daily Usage: ~2,700 gene lists/sublists from ~900 unique researchers.  
• Average Annual Usage: ~1,000,000 gene lists/sublists from >100 countries

Screen Shot 1 Screen Shot 2 Screen Shot 3

## 1. Upload datasets

Click on the “Start Analysis” button.



The screenshot shows the DAVID Bioinformatics Resources 6.8 homepage. At the top, there is a navigation bar with links for Home, Start Analysis, Shortcut to DAVID Tools, Technical Center, Downloads & APIs, Term of Service, Why DAVID?, and About Us. The 'Start Analysis' button is highlighted with a red box and a red arrow pointing to it. Below the navigation bar, there is a banner with the text "DAVID Bioinformatics Resources 6.8" and "Laboratory of Human Retrovirology and Immunoinformatics (LHRI)". A red arrow points from the 'Start Analysis' button towards this banner. The main content area includes a section titled "Welcome to DAVID 6.8" with a search bar, a timeline from 2003 to 2018, and sections for "What's Important in DAVID?" and "Statistics of DAVID". On the left side, there is a sidebar with links for Functional Annotation, Gene Functional Classification, Gene ID Conversion, and Gene Name Batch Viewer. At the bottom, there are three screenshots labeled Screen Shot 1, Screen Shot 2, and Screen Shot 3.

On the left panel of the page, there will be 3 steps:

1. Paste the gene list or choose a gene list file to upload. There are two ways to upload your gene list. One is to load a gene list from a file, another is to paste a gene list to the text box. Here we can upload the “affy\_id.txt” file. Regarding the limitations of gene lists, please see DAVID FAQs. (<https://david.ncifcrf.gov/content.jsp?file=FAQs.html>).
2. Select the ID format, according to the format of the gene list. Here we use “affymetrix ID”.
3. The list type may be a gene list or using a list as background. We choose the “gene list”.

At last, click on the “Submit List” button.

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

**Analysis Wizard**

DAVID Bioinformatics Resources 6.8, NIAID/NIH

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service Why DAVID? About Us

**Upload** List Background

**Upload Gene List**

Demolist 1 Demolist 2  
Upload Help

**Step 1: Enter Gene List**

A: Paste a list

```
55705_at
55872_at
56197_at
56256_at
56748_at
```

Or

B: Choose From a File

affy\_id.txt  
 Multi-List File

**Step 2: Select Identifier**

AFFYMETRIX\_3PRIME\_INV\_ID

**Step 3: List Type**

Gene List   
Background

**Step 4: Submit List**

Tell us how you like the tool  
Contact us for questions

## 2. Use DAVID tools

After task submission, the left panel shows the summary of the submitted gene list. The different available tools can be found under “Step 2”.

The screenshot shows the DAVID Bioinformatics Resources 6.8 interface. On the left, the "Gene List Manager" panel is visible, featuring a dropdown menu for species selection (with "Homo sapiens(2499)" highlighted) and a list manager for gene lists. A red oval highlights this section. On the right, the "Analysis Wizard" panel displays a welcome message and navigation links. Below, "Step 1" shows a successfully submitted gene list ("List\_1") with a background set to "Homo sapiens". "Step 2" lists various analysis tools: Functional Annotation Tool (with sub-options for Clustering, Chart, and Table), Gene Functional Classification Tool, Gene ID Conversion Tool, and Gene Name Batch Viewer. A blue arrow points from the "Functional Annotation Tool" section to a red speech bubble containing the text "Analysis tool types".

Analysis Wizard  
DAVID Bioinformatics Resources 6.8, NIAID/NIH

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service Why DAVID? About Us

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

Upload List Background

**Gene List Manager**

Select to limit annotations by one or more species [Help](#)

- Use All Species -  
Homo sapiens(2499)  
Unknown(1)

Select Species

List Manager Help

List\_1

Select List to:  
Use Rename  
Remove Combine  
Show Gene List  
View Unmapped Ids

**Analysis Wizard**

Tell us how you like the tool  
[Contact us for questions](#)

Step 1. Successfully submitted gene list  
Current Gene List: List\_1  
Current Background: Homo sapiens

**Step 2. Analyze above gene list with one of DAVID tools**

[Which DAVID tools to use?](#)

Functional Annotation Tool

- Functional Annotation Clustering
- Functional Annotation Chart
- Functional Annotation Table

Gene Functional Classification Tool

Gene ID Conversion Tool

Gene Name Batch Viewer

Analysis tool types

## Tool 1. ID conversion

We can click on the “Gene ID Conversion Tool”, go to the new page, and select a new ID format (Entrez\_Gene\_ID). In the left panel we find that there are 2499 genes from our uploaded gene list that can be found in the DAVID database, and 1 that cannot be found. Click on the “Submit to conversion tool” button.

The screenshot shows the DAVID Bioinformatics Resources 6.8 interface. The top navigation bar includes links for Home, Start Analysis, Shortcut to DAVID Tools, Technical Center, Downloads & APIs, Term of Service, Why DAVID?, and About Us. The main title is "Gene ID Conversion Tool" with the subtitle "DAVID Bioinformatics Resources 6.8, NIAID/NIH". A welcome message states "\*\*\* Welcome to DAVID 6.8 \*\*\*" and "\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*". On the left, the "Gene List Manager" panel shows a dropdown menu for species selection, currently set to "Homo sapiens(2499)". Below it is a list manager for "List\_1" and options to "Select List to:" (Use, Rename, Remove, Combine) and "Show Gene List" or "View Unmapped Ids". The right panel is titled "Gene ID Conversion Tool" and contains two options: "Option 1: Convert the gene list being selected in left panel to" (with a dropdown menu showing "ENTREZ\_GENE\_ID (Default)" and a "Submit to Conversion Tool" button) and "Option 2: Go Back to Submission Form".

On the left side of the ID conversion result page, there is a summary table of the gene list conversion. In the table, there are 2499 affymetrix IDs converted into Entrez Gene IDs in DAVID database. On the right top corner, a “download file” option allows to download the whole conversion file.

DAVID Bioinformatics Resources 6.8  
Laboratory of Human Retrovirology and Immunoinformatics (LHRI)

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

**Gene Accession Conversion Tool**

Gene Accession Conversion Statistics						
Conversion Summary	Submit Converted List to DAVID as a Gene List		Submit Converted List to DAVID as a Background			
ID Count	In DAVID DB	Conversion	From	To	Species	David Gene Name
2499	Yes	Successful	201903_at	7384	Homo sapiens	ubiquinol-cytochrome c reductase core protein I(UQCRC1)
0	Yes	None	203765_at	25801	Homo sapiens	grancalcin(GCA)
0	No	None	201193_at	341	Homo sapiens	isocitrate dehydrogenase (NADP(+)) 1, cytosolic(IDH1)
0	Ambiguous	Pending	203254_s_at	7004	Homo sapiens	talin 1(TLN1)
<b>Total Unique User IDs: 2499</b>			202614_at	10463	Homo sapiens	solute carrier family 30 member 9(SLC30A9)
<b>Summary of Ambiguous Gene IDs</b>			202602_s_at	27336	Homo sapiens	HIV-1 Tat specific factor 1(HTATSF1)
<b>All Possible Sources For Ambiguous IDs</b>			201196_s_at	262	Homo sapiens	adenosylmethionine decarboxylase 1(AMD1)
			201746_at	7157	Homo sapiens	tumor protein p53(TP53)
			201141_at	10457	Homo sapiens	glycoprotein nmb(GPNMB)
			202215_s_at	4802	Homo sapiens	nuclear transcription factor Y subunit gamma(NFYC)

**Save results** [Help](#) [Download File](#)

## Tool 2. Gene Name Batch Viewer

This tool converts gene list IDs into gene names directly. Click the “Gene Name Batch Viewer” under the list of “Shortcut to DAVID tools”.

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

**Gene List Report**

Current Gene List: List\_1  
Current Background: Homo sapiens  
2490 DAVID IDs

**Gene name**

**Save results**  
[Help and Manual](#)

**Input genes**

AFFYMETRIX_3PRIME_IVT_ID	Gene Name	Related Genes	Species
1007_s_at	microRNA 4640 (MIR4640)	RG	Homo sapiens
1053_at	replication factor C subunit 2 (RFC2)	RG	Homo sapiens
117_at	heat shock protein family A (Hsp70) member 6 (HSPA6)	RG	Homo sapiens
121_at	paired box 8 (PAX8)	RG	Homo sapiens
1294_at	microRNA 5193 (MIR5193)	RG	Homo sapiens
1316_at	thyroid hormone receptor, alpha (THRA)	RG	Homo sapiens
1431_at	cytochrome P450 family 2 subfamily E member 1 (CYP2E1)	RG	Homo sapiens
1487_at	estrogen related receptor alpha (ESRRA)	RG	Homo sapiens
1494_f_at	cytochrome P450 family 2 subfamily A member 6 (CYP2A6)	RG	Homo sapiens
1598_g_at	growth arrest specific 6 (GAS6)	RG	Homo sapiens
160020_at	matrix metallopeptidase 14 (MMP14)	RG	Homo sapiens
177_at	phospholipase D1 (PLD1)	RG	Homo sapiens
179_at	DTX2P1-UPK3BP1-PMS2P11 readthrough, transcribed pseudogene (DTX2P1-UPK3BP1-PMS2P11)	RG	Homo sapiens
1861_at	BCL2 associated agonist of cell death (BAD)	RG	Homo sapiens
200000_s_at	pre-mRNA processing factor 8 (PRPF8)	RG	Homo sapiens
200001_at	calpain small subunit 1 (CAPNS1)	RG	Homo sapiens
200002_at	ribosomal protein L35 (RPL35)	RG	Homo sapiens
200003_s_at	microRNA 6805 (MIR6805)	RG	Homo sapiens
200004_at	eukaryotic translation initiation factor 4 gamma 2 (EIF4G2)	RG	Homo sapiens
200005_at	eukaryotic translation initiation factor 3 subunit D (EIF3D)	RG	Homo sapiens
200006_at	Parkinsonism associated deglycase (PARK7)	RG	Homo sapiens
200007_at	signal recognition particle 14 (SRP14)	RG	Homo sapiens
200008_c_at	GDP dissociation inhibitor 2 (GDI2)	RG	Homo sapiens

**Q1: What are the gene names of the genes with Affy\_id : “1053\_at” and “200010\_at”?**

### Tool 3. Functional Annotation Tool

Go back to the previous page or choose “shortcut to DAVID Tools”—“Functional Annotation Tool”.

DAVID BIOINFORMATICS DATABASE

Analysis Wizard  
DAVID Bioinformatics Resources 6.8, NIAID/NIH

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service Why DAVID? About Us

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

Upload List Background

Gene List Manager

Select to limit annotations by one or more species [Help](#)

- Use All Species -  
Homo sapiens(2499)  
Unknown(1)

Select Species

List Manager Help

List\_1

Select List to:  
Use Rename  
Remove Combine  
Show Gene List  
View Unmapped Ids

Step 1. Successfully submitted gene list

Step 1. Successfully submitted gene list  
Current Gene List: List\_1  
Current Background: Homo sapiens

Step 2. Analyze above gene list with one of DAVID tools

Which DAVID tools to use?

Functional Annotation Tool

- Functional Annotation Clustering
- Functional Annotation Chart
- Functional Annotation Table

Gene Functional Classification Tool  
Gene ID Conversion Tool  
Gene Name Batch Viewer

Tell us how you like the tool  
Contact us for questions

The functional annotation tool includes three options: Functional Annotation Clustering, chart and table. Click the “Functional Annotation Tool”, go to the new page, and choose the annotation we want (Gene Ontology and KEGG pathway for this exercise).


**Functional Annotation Tool**  
 DAVID Bioinformatics Resources 6.8, NIAID/NIH

[Home](#) [Start Analysis](#) [Shortcut to DAVID Tools](#) [Technical Center](#) [Downloads & APIs](#) [Term of Service](#) | [Why DAVID?](#) [About Us](#)

\*\*\* Welcome to DAVID 6.8 \*\*\*  
 \*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

[Upload](#) [List](#) [Background](#)

### Annotation Summary Results

[Help and Tool Manual](#)

**Current Gene List:** List\_1      **2490 DAVID IDs**  
**Current Background:** Homo sapiens      [Check Defaults](#)  [Clear All](#)

Disease (0 selected)  
 Functional\_Categories (0 selected)  
 Gene\_Ontology (0 selected)  
 General\_Annotations (0 selected)  
 Literature (0 selected)  
 Main\_Accessions (0 selected)  
 Pathways (1 selected)

<input type="checkbox"/> BBID	3.9%	98	Chart
<input type="checkbox"/> BIOCARA	18.0%	449	Chart
<input type="checkbox"/> EC_NUMBER	29.2%	727	Chart
<input checked="" type="checkbox"/> KEGG_PATHWAY	55.1%	1372	Chart
<input type="checkbox"/> REACTOME_PATHWAY	65.9%	1642	Chart

Protein\_Domains (0 selected)  
 Protein\_Interactions (0 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](#)

Choose the KEGG pathway analysis only, and open the KEGG pathway chart.

Q2: What are the 3 most significant KEGG pathways? What are their p-values? Open them in the KEGG website.

Now choose “Functional Annotation Clustering”. The results show that pathways can be combined into 9 clusters.

\*\*\* Welcome to DAVID 6.8 \*\*\*  
 \*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

### Functional Annotation Clustering

[Help and Manual](#)

Current Gene List: List\_1  
 Current Background: Homo sapiens  
 1889 DAVID IDs

Options    Classification Stringency: Medium ▾

**Overall enrichment score**  
 The higher, the more enriched

The smaller, the more enriched.

Download File

Annotation Cluster 1		Enrichment Score: 7.86		G			Count	P_Value	Benjamini
<input type="checkbox"/>	KEGG_PATHWAY	Parkinson's disease	RT	■	■		53	2.6E-10	2.5E-8
<input type="checkbox"/>	KEGG_PATHWAY	Huntington's disease	RT	■	■		63	1.8E-9	1.3E-7
<input type="checkbox"/>	KEGG_PATHWAY	Alzheimer's disease	RT	■	■		56	8.9E-9	4.3E-7
<input type="checkbox"/>	KEGG_PATHWAY	Non-alcoholic fatty liver disease (NAFLD)	RT	■	■		49	2.2E-7	8.0E-6
<input type="checkbox"/>	KEGG_PATHWAY	Oxidative phosphorylation	RT	■	■		44	5.6E-7	1.6E-5
Annotation Cluster 2		Enrichment Score: 1.4		G			Count	P_Value	Benjamini
<input type="checkbox"/>	KEGG_PATHWAY	Mismatch repair	RT	■	■		9	1.7E-2	1.1E-1
<input type="checkbox"/>	KEGG_PATHWAY	DNA replication	RT	■	■		11	3.9E-2	1.9E-1
<input type="checkbox"/>	KEGG_PATHWAY	Nucleotide excision repair	RT	■	■		12	9.4E-2	3.2E-1
Annotation Cluster 3		Enrichment Score: 0.88		G			Count	P_Value	Benjamini
<input type="checkbox"/>	KEGG_PATHWAY	Long-term potentiation	RT	■	■		19	8.6E-3	7.3E-2
<input type="checkbox"/>	KEGG_PATHWAY	Oxytocin signaling pathway	RT	■	■		30	1.7E-1	4.5E-1
<input type="checkbox"/>	KEGG_PATHWAY	Vascular smooth muscle contraction	RT	■	■		23	2.0E-1	4.9E-1
<input type="checkbox"/>	KEGG_PATHWAY	cGMP-PKG signaling pathway	RT	■	■		30	2.5E-1	5.5E-1
<input type="checkbox"/>	KEGG_PATHWAY	cAMP signaling pathway	RT	■	■		31	5.6E-1	8.1E-1
Annotation Cluster 4		Enrichment Score: 0.72		G			Count	P_Value	Benjamini
<input type="checkbox"/>	KEGG_PATHWAY	Chronic myeloid leukemia	RT	■	■		22	2.0E-3	2.3E-2
<input type="checkbox"/>	KEGG_PATHWAY	Central carbon metabolism in cancer	RT	■	■		19	6.1E-3	5.4E-2
<input type="checkbox"/>	KEGG_PATHWAY	Hepatitis B	RT	■	■		31	4.8E-2	2.1E-1
<input type="checkbox"/>	KEGG_PATHWAY	Pancreatic cancer	RT	■	■		16	6.3E-2	2.5E-1

Q3: What do pathways have in common for annotation cluster1? What about annotation cluster 2?

Now go back and select “Functional Annotation Chart”.

The “Functional Annotation Chart” provides the clustering of genes’ annotations (KEGG pathway or others). It shows 77 chart records, which means that all the 1889 genes are included in 77 KEGG pathways.

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

### Functional Annotation Chart

[Help and Manual](#)

Current Gene List: List\_1  
Current Background: Homo sapiens  
1889 DAVID IDs

Options  
Thresholds: Count 2 EASE 0.1  
Display:  Fold Enrichment  Bonferroni  Benjamini  FDR  Fisher Exact  LT,PH,PT # of Records 1000

Rerun Using Options Create Sublist

77 chart records

Sublist	Category	Term	WT	Genes	Count	%	P-Value	Benjamini
	KEGG_PATHWAY	Proteasome	RT	30	1.6	3.7E-14	1.1E-11	
	KEGG_PATHWAY	Ribosome	RT	53	2.8	4.0E-11	5.8E-9	
	KEGG_PATHWAY	Parkinson's disease	RT	53	2.8	2.6E-10	2.5E-8	
	KEGG_PATHWAY	Huntington's disease	RT	63	3.3	1.8E-9	1.3E-7	
	KEGG_PATHWAY	Biosynthesis of antibiotics	RT	67	3.5	3.0E-9	1.7E-7	
	KEGG_PATHWAY	Alzheimer's disease	RT	56	3.0	8.9E-9	4.3E-7	
	KEGG_PATHWAY	Spliceosome	RT	46	2.4	6.8E-8	2.8E-6	
	KEGG_PATHWAY	Non-alcoholic fatty liver disease (NAFLD)	RT	49	2.6	2.2E-7	8.0E-6	
	KEGG_PATHWAY	Protein processing in endoplasmic reticulum	RT	53	2.8	2.2E-7	7.2E-6	
	KEGG_PATHWAY	Oxidative phosphorylation	RT	44	2.3	5.6E-7	1.6E-5	
	KEGG_PATHWAY	Epstein-Barr virus infection	RT	56	3.0	9.1E-7	2.4E-5	
	KEGG_PATHWAY	Focal adhesion	RT	57	3.0	6.3E-6	1.5E-4	

Related genes [Download File](#) Save files

Count Threshold (Minimum Count): The threshold of minimum gene counts belonging to an annotation term. Default value is 2. In short, you do not trust the term only having one gene involved.

Pathways are ordered by ascending p-value but can be ordered by any other column by clicking on the header of the column.

Q4: What is the pathway with a higher gene count?

Now choose the “Functional Annotation Table”.

The “Functional Annotation Table” shows that 1059 genes are annotated with one or more annotations (here, KEGG pathways).

\*\*\* Welcome to DAVID 6.8 \*\*\*  
\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*

### Functional Annotation Table

Current Gene List: List\_1  
Current Background: Homo sapiens  
1889 DAVID IDs

1059 record(s)

KEGG_PATHWAY	Description	Related Genes	Homo sapiens
203282_at	<b>1,4-alpha-glucan branching enzyme 1(GBE1)</b> Starch and sucrose metabolism, Metabolic pathways,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
200862_at	<b>24-dehydrocholesterol reductase(DHCR24)</b> Steroid biosynthesis, Metabolic pathways,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
203058_s_at, 203059_s_at, 203060_s_at	<b>3'-phosphoadenosine 5'-phosphosulfate synthase 2(PAPSS2)</b> Purine metabolism, Selenocompound metabolism, Sulfur metabolism, Metabolic pathways, Biosynthesis of antibiotics,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202539_s_at	<b>3-hydroxy-3-methylglutaryl-CoA reductase(HMGCR)</b> Terpenoid backbone biosynthesis, Metabolic pathways, Biosynthesis of antibiotics, AMPK signaling pathway, Bile secretion,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202772_at	<b>3-hydroxymethyl-3-methylglutaryl-CoA lyase(HMGCL)</b> Synthesis and degradation of ketone bodies, Valine, leucine and isoleucine degradation, Butanoate metabolism, Metabolic pathways, Peroxisome,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202419_at	<b>3-ketodihydroinosine reductase(KDSR)</b> Sphingolipid metabolism, Metabolic pathways,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202780_at	<b>3-oxoacid CoA-transferase 1(OXCT1)</b> Synthesis and degradation of ketone bodies, Valine, leucine and isoleucine degradation, Butanoate metabolism,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202464_s_at	<b>6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3(PFKB3)</b> Fructose and mannose metabolism, HIF-1 signaling pathway, AMPK signaling pathway,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202123_s_at	<b>ABL proto-oncogene 1, non-receptor tyrosine kinase(ABL1)</b> ErbB signaling pathway, Ras signaling pathway, Cell cycle, Axon guidance, Neurotrophin signaling pathway, Pathogenic Escherichia coli infection, Shigellosis, Pathways in cancer, MicroRNAs in cancer, Chronic myeloid leukemia, Viral myocarditis,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202604_x_at, 202603_at	<b>ADAM metallopeptidase domain 10(ADAM10)</b> Alzheimer's disease, Epithelial cell signaling in Helicobacter pylori infection,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
200734_s_at, 200011_s_at	<b>ADP ribosylation factor 3(ARF3)</b> Endocytosis,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
201526_at	<b>ADP ribosylation factor 5(ARF5)</b> Endocytosis,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
203312_x_at	<b>ADP ribosylation factor 6(ARF6)</b> Ras signaling pathway, Endocytosis, Fc gamma R-mediated phagocytosis,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202211_at	<b>ADP ribosylation factor GTPase activating protein 3(ARFGAP3)</b> Endocytosis,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
202956_at, 202955_s_at	<b>ADP ribosylation factor guanine nucleotide exchange factor 1(ARFGEF1)</b> Endocytosis,	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>
201924_at	<b>AF4/FMR2 family member 1(AFF1)</b> AF4/FMR2 family member 1(AFF1)	<a href="#">Related Genes</a>	<a href="#">Homo sapiens</a>

[Help and Manual](#)

**Save results**

**Download File**

#### Tool 4. Gene Functional Classification

Click the “Gene Functional Classification tool” under the list of “Shortcut to DAVID tools”. The results show 106 clusters of annotations. This tool is used to cluster the functionally related genes as a group and give a score to this cluster.

The screenshot shows the DAVID Bioinformatics Resources 6.8 interface for gene functional classification. The top navigation bar includes links for Home, Start Analysis, Shortcut to DAVID Tools, Technical Center, Downloads & APIs, Term of Service, Why DAVID?, and About Us. A banner at the top reads "Gene Functional Classification Tool" and "DAVID Bioinformatics Resources 6.8, NIAID/NIH". Below the banner, a message says "\*\*\* Welcome to DAVID 6.8 \*\*\*" and "\*\*\* If you are looking for DAVID 6.7, please visit our [development site](#). \*\*\*". The main content area is titled "Gene Functional Classification Result". On the left, there is a "Gene List Manager" panel with tabs for Upload, List, and Background. It shows a list of species: "Homo sapiens(2499)" selected, "Unknown(1)", and a "Select Species" dropdown. Below this is a "List Manager" section with a list named "List\_1" and buttons for Use, Rename, Remove, Combine, and Show Gene List. There is also a link to View Unmapped Ids. The right side displays "106 Cluster(s)" in two groups. Group 1 has an enrichment score of 93.06 and contains 11 entries. Group 2 has an enrichment score of 79.17 and contains 5 entries. Each entry lists a gene ID, name, and description. A "Download File" button is located at the top right of the cluster table. A "Help and Tool Manual" link is also present.

Gene Group 1	Enrichment Score: 93.06	RG	T	NC
1 203262_s_at	family with sequence similarity 50 member A(FAM50A)			
2 203023_at	NOP16 nucleolar protein(NOP16)			
3 201922_at	NSA2, ribosome biogenesis homolog(NSA2)			
4 202579_x_at	high mobility group nucleosomal binding domain 4(HMGN4)			
5 201414_s_at	nucleosome assembly protein 1 like 4(NAP1L4)			
6 200053_at	sperm associated antigen 7(SPAG7)			
7 203119_at	coiled-coil domain containing 86(CCDC86)			
8 204528_s_at	nucleosome assembly protein 1 like 1(NAP1L1)			
9 203831_at	R3H domain containing 2(R3HDM2)			
10 202882_x_at	nucleolar protein 7(NOL7)			
11 204805_s_at	H1 histone family member X(H1FX)			

Gene Group 2	Enrichment Score: 79.17	RG	T	NC
1 202791_s_at	protein phosphatase 6 regulatory subunit 2(PPP6R2)			
2 201309_x_at	neuronal regeneration related protein(NREP)			
3 201462_at	secremin 1(SCRN1)			
4 204837_at	myotubularin related protein 9(MTMR9)			
5 204793_at	G_protein-coupled receptor associated sorting protein 1(GPRASP1)			

Q6: What differences can you see between gene groups?

# Exercise 2. Using Enrichr

Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>) accepts either BED format or a list of genes with gene symbols.

## 1. Upload your gene list

Enrichr uses a list of gene symbols as input data. You can upload the list by either selecting the text file that contains the list or just simply pasting the list into the text box. It is better to enter a description for the gene list so that multiple lists can be differentiated from each other.

We will use the same genes from the previous exercise.

The screenshot shows the 'Input data' section of the Enrichr web interface. At the top, there is a navigation bar with links: Analyze, What's New?, Libraries, Find a Gene (circled in red), About (circled in red), and Help. Below the navigation bar, the title 'Input data' is displayed. To the left of the main input area, there is a text box with the placeholder 'Choose an input file to upload. Either in BED format or a list of genes...'. Below this text box is a 'Browse...' button and a message 'No file selected.'. A red oval labeled 'Text file including genes' points to this area. To the right of the input area, there is a text box with the placeholder 'Or paste in a list of gene symbols optionally followed by a comma and levels of membership. Try two examples: crisp set example, fuzzy set example'. Below this text box is a scrollable list of gene symbols: 204820\_s\_at, 204824\_at, 204831\_at, 204832\_s\_at, 204834\_at, 204835\_at, 204837\_at, 204838\_s\_at, 204839\_at, and 204841\_s\_at. A red oval labeled 'Input gene symbols' points to this list. At the bottom of the input area, there is a text box for 'Enter a brief description for the list in case you want to share it. (Optional)' containing the text 'GSE3585'. A red oval labeled 'Description of this dataset' points to this text box. On the far right, there is a red 'Submit' button.

## 2. Results page

On the results page, the analysis is divided into different categories of enrichment (Transcription, Pathways, ontologies and so on). The first category is shown. Within each category, the enrichment analyses of various gene-set libraries are listed. We open the pathway analysis as an example, presenting a multitude of visualizations. If you want to change the category, just tap the other category name.

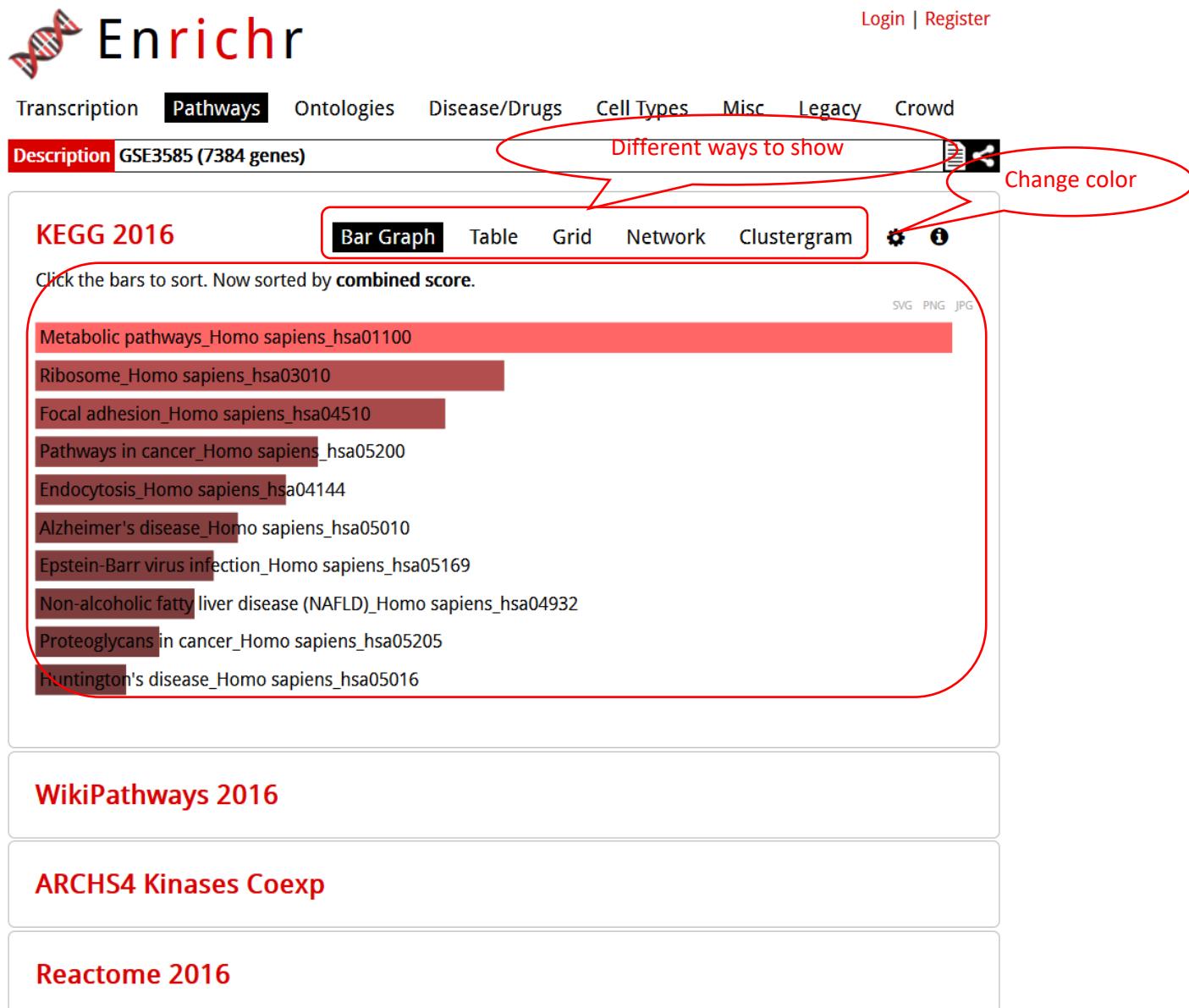
The screenshot shows the Enrichr results page with a red speech bubble highlighting the "Categories" section at the top right. Below the categories, a navigation bar includes links for Transcription, Pathways, Ontologies, Disease/Drugs, Cell Types, Misc, Legacy, and Crowd. A sub-navigation bar below shows "Description GSE3585 (7384 genes)" and icons for download and share.

The main content area displays 12 categories of gene-set libraries:

- KEGG 2016**: Metabolic pathways\_Homo sapiens\_hsa0110, Pathways in cancer\_Homo sapiens\_hsa0520, Focal adhesion\_Homo sapiens\_hsa04510, Endocytosis\_Homo sapiens\_hsa04144, Epstein-Barr virus infection\_Homo sapiens\_hsa04010, XPodNet - protein-protein interactions in the mRNA processing\_Mus musculus\_WP310, PodNet: protein-protein interactions in the Cytoplasmic Ribosomal Proteins\_Homo sapiens\_WP411, YES1\_human\_kinase\_ARCHS4\_coexpression, UHMK1\_human\_kinase\_ARCHS4\_coexpression, TGFBR2\_human\_kinase\_ARCHS4\_coexpression, RYK\_human\_kinase\_ARCHS4\_coexpression, MAPK6\_human\_kinase\_ARCHS4\_coexpression.
- WikiPathways 2016**: mCalpain and friends in Cell motility\_Homo sapiens\_WP00000, Role of ERBB2 in Signal Transduction and Or, Skeletal muscle hypertrophy is regulated via Mechanism of Gene Regulation by Peroxisor, Transcription factor CREB and its extracellular.
- ARCHS4 Kinases Coexp**: superpathway of conversion of glucose to ac, protein ubiquitylation\_Homo sapiens\_WPY-7, 3-phosphoinositide biosynthesis\_Homo sapiens\_WPY-10, TCA cycle\_Homo sapiens\_WPY66-398, superpathway of inositol phosphate compo.
- Reactome 2016**: Metabolism\_Homo sapiens\_R-HSA-1430728, Gene Expression\_Homo sapiens\_R-HSA-7416, Infectious disease\_Homo sapiens\_R-HSA-5616, Disease\_Homo sapiens\_R-HSA-1643685, Metabolism of proteins\_Homo sapiens\_R-HS.
- BioCarta 2016**: Integrin signalling pathway\_Homo sapiens\_P00000, EGF receptor signalling pathway\_Homo sapiens\_P00001, Ubiquitin proteasome pathway\_Homo sapiens\_P00002, CCKR signalling map ST\_Homo sapiens\_P0069, Angiogenesis\_Homo sapiens\_P00005.
- HumanCyc 2016**: RRS1, SNRNP27, FGB, RPL18A, PSMB9.
- NCI-Nature 2016**: PDGFR-beta signaling pathway\_Homo sapiens\_P00003, ErbB1 downstream signaling\_Homo sapiens\_P00004, Signaling events mediated by VEGFR1 and VIL, mTOR signaling pathway\_Homo sapiens\_P00005, TGF-beta receptor signaling\_Homo sapiens\_P00006.
- Panther 2016**: SLC2A4, ESR1, GABARAPL1, GABARAPL2, CSNK2A1.
- BioPlex 2017**: CDK2, MAPK14, GSK3B, MAPK1, CDK1.
- huMAP**: RPL19, RPS2, RPS18, RPL5, RPS16.
- PPI Hub Proteins**: SLC2A4, ESR1, GABARAPL1, GABARAPL2, CSNK2A1.
- KEA 2015**: CDK2, MAPK14, GSK3B, MAPK1, CDK1.
- LINCS L1000 Kinase Perturbations down**.
- LINCS L1000 Kinase Perturbations up**.
- Kinase Perturbations from GEO down**.

Click on “KEGG 2016” to view the detailed results. They include: “Bar Graph”, “Table”, “Grid”, “Network”, and “Clustergram”. When you click on the bars, you get different ranks by other score methods. Notice that it takes longer time to open “Clustergram”.

Bar Graph:



The length of the bar represents the significance of that specific gene-set or term. In addition, the brighter the color, the more significant that term is.

Table:



[Login](#) | [Register](#)

Transcription   **Pathways**   Ontologies   Disease/Drugs   Cell Types   Misc   Legacy   Crowd

Description GSE3585 (7384 genes)



KEGG 2016

Bar Graph

**Table**

Grid

Network

Clustergram



Hover each row to see the overlapping genes.

10 entries per page

Search:

Index	Name	P-value	Adjusted p-value	Z-score	Combined score
1	Metabolic pathways_Homo sapiens_hsa01100	3.221e-32	9.439e-30	-2.01	145.87
2	Ribosome_Homo sapiens_hsa03010	5.393e-28	7.901e-26	-1.72	108.11
3	Focal adhesion_Homo sapiens_hsa04510	1.167e-24	1.140e-22	-1.87	103.15
4	Alzheimer's disease_Homo sapiens_hsa05010	8.657e-22	6.341e-20	-1.77	85.69
5	Endocytosis_Homo sapiens_hsa04144	1.212e-21	7.105e-20	-1.86	89.74
6	Pathways in cancer_Homo sapiens_hsa05200	4.883e-21	2.385e-19	-1.98	92.43
7	Epstein-Barr virus infection_Homo sapiens_hsa05169	6.539e-21	2.737e-19	-1.80	83.65
8	Non-alcoholic fatty liver disease_Homo sapiens_hsa05169	CBLB, FGF2, ACTB, ACTG1, IGF1R, PPP1CB, PPP1CC, CCND1, PLAU, AKT3, KDR, AKT1, PLCε1, PRKACA, PRKACB, MAP2K1, MAP2K2, PRKCB, HGF, WNT5A, RPS6, PRKCA, ANK2, ANK3, ANK1, HSPG2, TIAM1, EZR, RAF1, TP53, DDX5, SDC4, SDC2, PXN, ITPR1, TWIST1, ITPR2, PIK3R3, PIK3R2, PIK3R1, IQGAP1, HIF1A, PIK3R5, (NAFLD)_Homo sapiens_hsa05169	82.04		
9	Huntington's disease_Homo sapiens_hsa05169	VTN, WNT6, RRAS, PLCG2, FZD1, SMAD2, FZD3, TGFβ2, TGFβ1, FZD5, FZD4, CAV2, FZD7, PTCH1, CAV1, FZD6, RDX, IGF2, FN1, MSN, BRAF, IGF1, ESR1, PTK2, GRB2, FGFR1, ITGB1, CDKN1A, ITGB5, ITGB3, HSPB2, PIK3CD, PIK3CB, CTSL, CASP3, TIMP3, ITGAV, RAC1, HRAS, ARHGEF12, PPP1R12A, PDPK1, MMP2, GAB1, PLAUR, RRAS2, MMP9, RHOA, DCN, MRAS, CTTN, PIK3CA, HCLS1, COL21A1, ITGA5, PPP1R12B, SOS1, TLR4, SOS2, MET, CD44, HBEGF, CAMK2B, CD63, ROCK1, ROCK2, THBS1, EGFR, CDC42, NRAS, ERBB3, ERBB4, GPC1, ERBB2, GPC3, DROSHA, FLNA, MAPK1, FLNB, FLNC, CAMK2G, EIF4B, MAPK3, LUM, STAT3, PTPN11, MAPK14, MAPK13, VEGFA, PPP1CA, RPS6KB1, PDCD4, CTNNB1, FAS, PTPN6, KRAS	76.28		
10	Proteoglycans in cancer_Homo sapiens_hsa05205	5.132e-20	1.504e-18	-1.78	79.08

Download results

When you put the cursor on the "Name", there will be a list of related genes

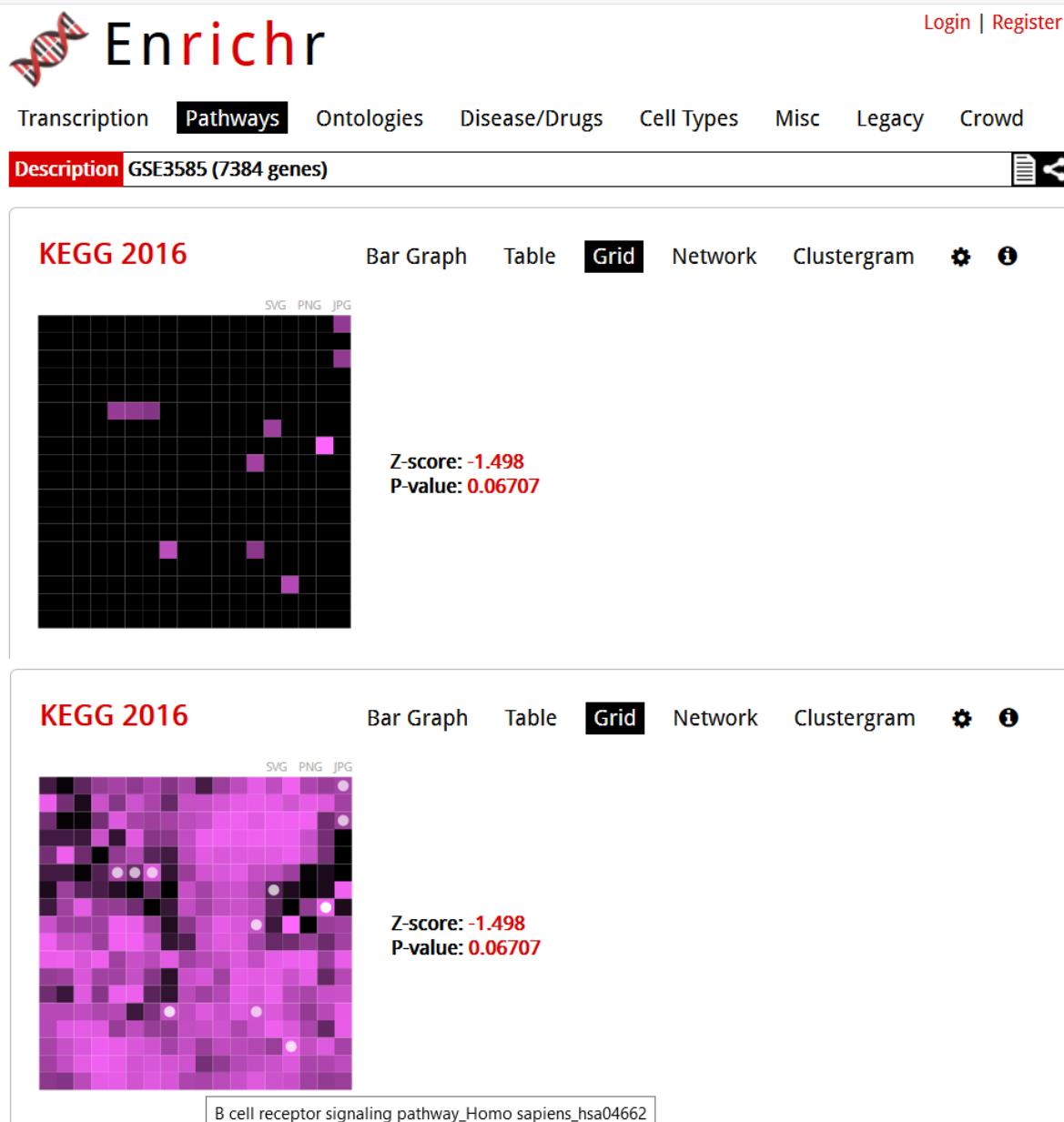
Showing 1 to 10 of 293 entries | [Export entries to table](#)

Terms marked with an \* have an overlap of less than 5

Previous Next

By clicking on the column header, you can sort the table by the term, p-value, z-score, or combined score. You can also download the table information by clicking on the "Export entries to table" button.

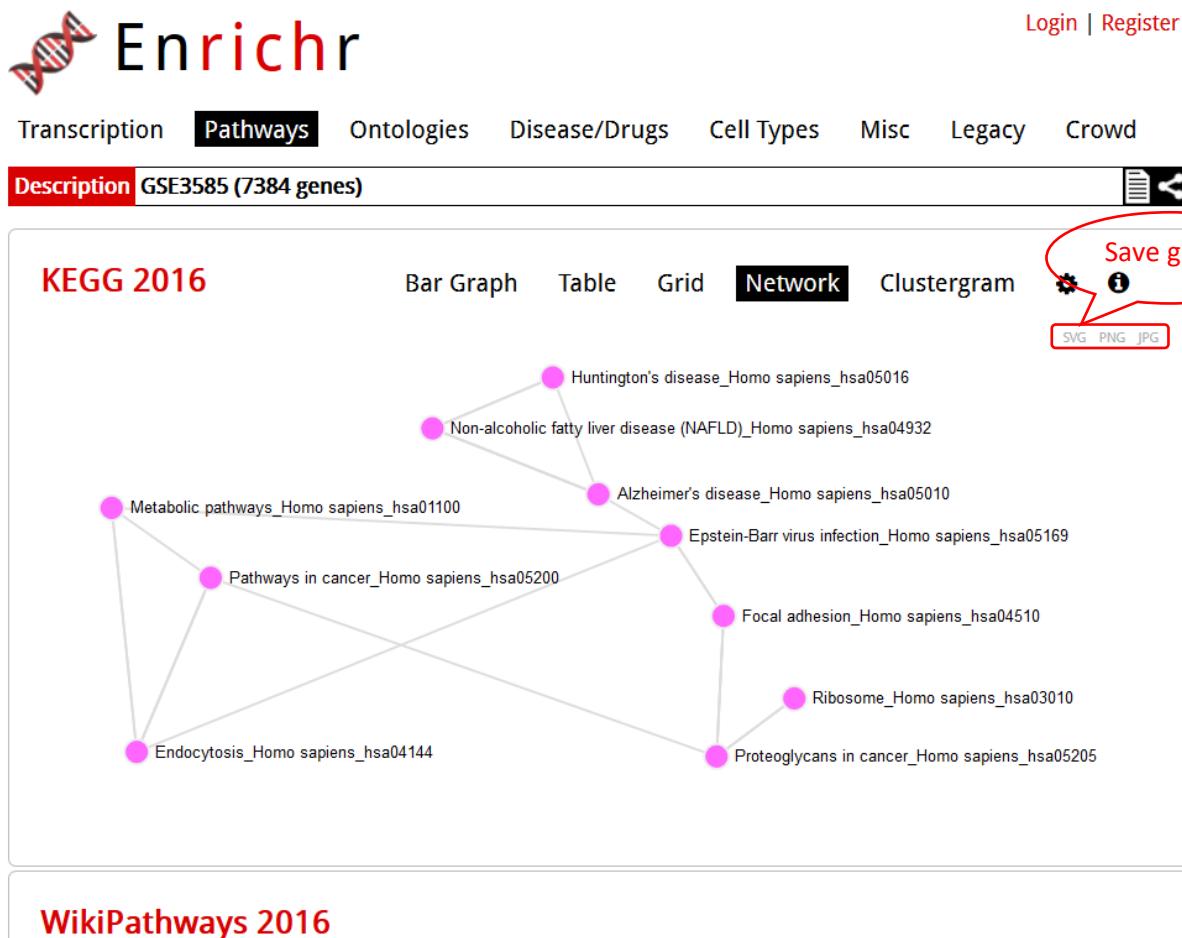
Grid:



Each grid square represents a term and is arranged based on its gene-set similarity with other terms. It shows only the top 10 terms sorted by combined score. The brighter the square, the more significant that term is. Clicking on the grid allows you to another view that colors the grid based on its correlation score with neighbors with white dots representing the significant terms. The z-score and p-value is a measure of how clustered the top 10 terms are on the grid.

**Q7: Where can we find the significant gene terms?**

Network:



Each node represents a term and a link between two nodes means that the two terms have some gene content similarity.

**Q8: How to find the pathway with your genes of interest?**

# Exercise 3. Using WebGestalt

WebGestalt(<http://www.webgestalt.org/option.php>) is a functional enrichment analysis web tool that supports three well-established and complementary methods for enrichment analysis: Over-Representation Analysis (ORA), Gene Set Enrichment Analysis (GSEA), and Network Topology-based Analysis (NTA).

The screenshot shows the 'WEB-based GEne SeT AnaLysis Toolkit' (WebGestalt) interface. At the top, there is a logo and the text 'Translating gene lists into biological insights...'. Below the logo, a navigation bar includes links for ORA Sample Run, GSEA Sample Run, NTA Sample Run, External Examples, Manual, Citation, User Forum, GOView, WebGestaltR, and WebGestalt 2013. The main area is titled 'Basic Parameters' and contains several input fields:

- Select Organism of Interest:** dropdown menu for Organisms.
- Select Method of Interest:** dropdown menu for Methods.
- Select Functional Database:** dropdown menu for Functional Database Class and Functional Database Name.
- Gene List:** section with 'Select Gene ID Type' dropdown, 'Choose File' button, and 'Reset' button. It also includes a text input field for 'Please enter gene ids...' and a 'Clear' button.
- Reference Gene List:** section with 'Select Reference Set for Enrichment Analysis' dropdown, 'Reference Gene Set' dropdown, and 'Reset' button.
- Advanced parameters:** section with a 'Submit' button.

At the bottom of the page, there is a note about browser support: 'Browser support: PC: Google Chrome 56.0 or later, Mac: Google Chrome 56.0, Safari 10.0 or later. We strongly recommend upgrading to the latest version of the supported browsers. For Safari users, please enable Flash for network visualization. Detailed information on how to enable Flash can be found here.'

## 1. Setting parameters

Set the parameters and upload the gene list, as in the following picture, and click the “Submit” button.

We are using ORA. If we change the method to “GSEA”, then we need a ranked gene list.

 **WEB-based GEne SeT AnaLysis Toolkit**

WebGestalt *Translating gene lists into biological insights...*

[ORA Sample Run](#) | [GSEA Sample Run](#) | [NTA Sample Run](#) | [External Examples](#) | [Manual](#) | [Citation](#) | [User Forum](#)  
[GOView](#) | [WebGestaltR](#) | [WebGestalt 2013](#)

---

**» Basic Parameters**

Select Organism of Interest [?](#)

Select Method of Interest [?](#)

Select Functional Database [?](#)

**Gene List**

Select Gene ID Type [?](#)

Choose File No file chosen Reset

Upload Gene List (max size: 5 MB) [?](#)  
OR

**Reference Gene List**

Select Reference Set for Enrichment Analysis [?](#)  Reset

OR

Upload User Reference Set File (max size: 5 MB)  
and Select ID Type [?](#)

Choose File No file chosen Reset

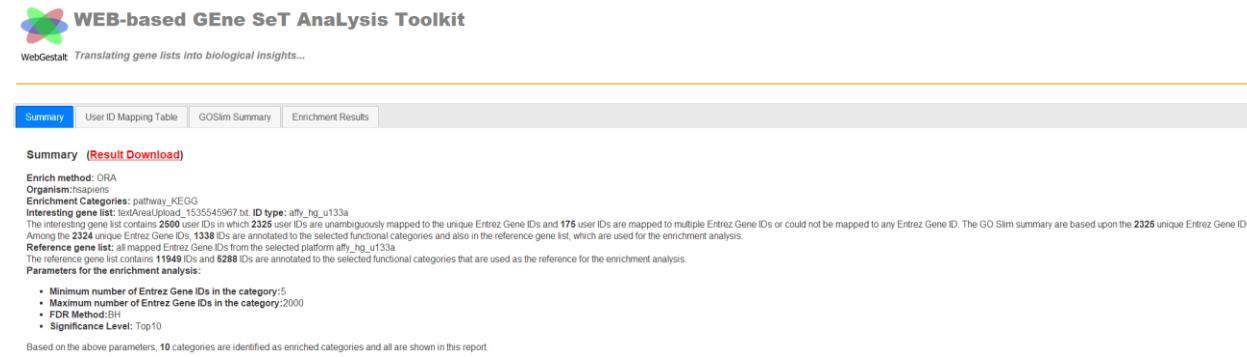
---

**» Advanced parameters**

**Submit**

## 2. Results

After we submit the task, the summary comes into being at first. It contains enrichment method, organism, enrichment category, gene list with ID type, reference gene list, and parameters for enrichment analysis. We also get: “User ID Mapping Table”, “GOSlim Summary” and “Enrichment Results”.



WEB-based GENE SeT Analysis Toolkit  
WebGestalt Translating gene lists into biological insights...

Summary User ID Mapping Table GOSlim Summary Enrichment Results

Summary (Result Download)

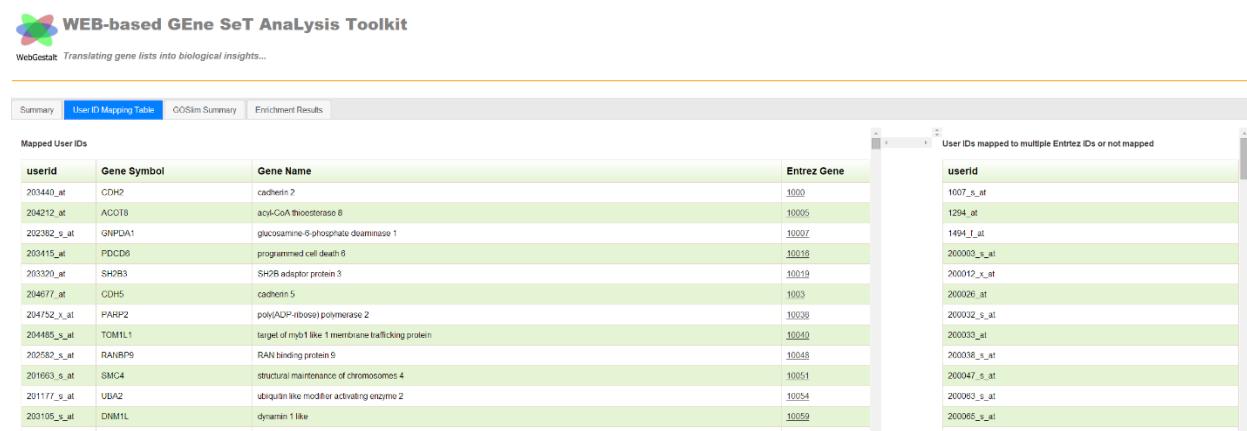
Enrich method: ORA  
Organism:hsapens  
Enrichment Categories: pathway\_KEGG  
Interesting gene list: testAreaUpload\_1535545967.txt. ID type: affy\_hg\_u133a  
The interesting gene list contains 2500 user IDs in which 2325 user IDs are unambiguously mapped to the unique Entrez Gene IDs and 175 user IDs are mapped to multiple Entrez Gene IDs or could not be mapped to any Entrez Gene ID. The GO Slim summary are based upon the 2325 unique Entrez Gene IDs. Among the 2325 unique Entrez Gene IDs, 1338 IDs are annotated to the selected functional categories and also in the reference gene list, which are used for the enrichment analysis.  
Reference gene list: all mapped Entrez Gene IDs from the selected platform affy\_hg\_u133a  
The reference gene list contains 11949 IDs and 5298 IDs are annotated to the selected functional categories that are used as the reference for the enrichment analysis.  
Parameters for the enrichment analysis:

- Minimum number of Entrez Gene IDs in the category: 5
- Maximum number of Entrez Gene IDs in the category: 2000
- FDR Method: BH
- Significance Level: Top 10

Based on the above parameters, 10 categories are identified as enriched categories and all are shown in this report.

### 2.1 User ID Mapping Table

In the table, the left contains the mapped ID, gene symbol, gene names, and Entrez gene ID. The right contains the “User IDs mapped to multiple Entrez IDs or not mapped”.



WEB-based GENE SeT Analysis Toolkit  
WebGestalt Translating gene lists into biological insights...

Summary User ID Mapping Table GOSlim Summary Enrichment Results

Mapped User IDs

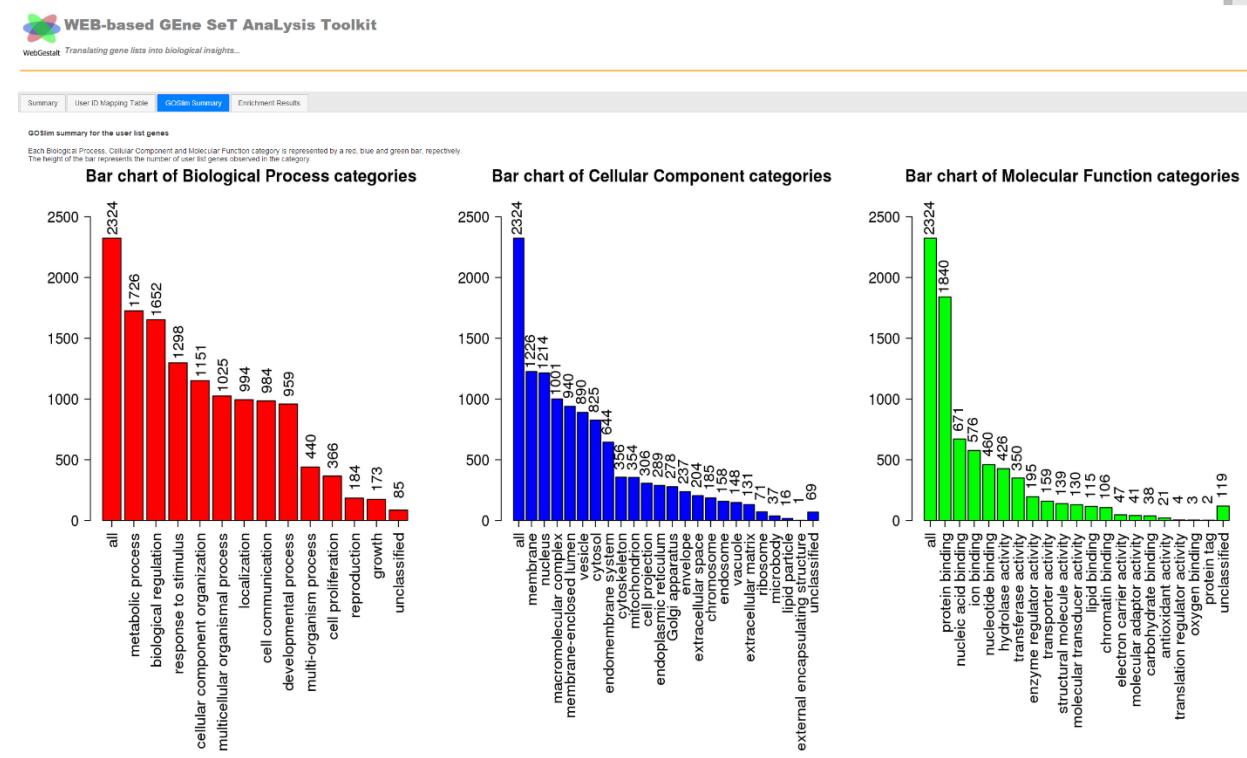
userid	Gene Symbol	Gene Name	Entrez Gene
20340_at	CINCH	cashen 2	1000
20421_at	ACOTB	acyl-CoA thioesterase 8	10005
202382_s_at	GNPDA1	glucosamine-6-phosphate deaminase 1	10007
203415_at	PDCD6	programmed cell death 6	10018
203320_at	SH2B3	SH2B adaptor protein 3	10019
204677_at	CDH5	cadherin 5	1003
20452_x_at	PAPR2	poly(ADP-ribose) polymerase 2	10038
204465_s_at	TOMM1	target of myb 1 like 1 membrane trafficking protein	10040
202582_s_at	RAB18P9	RAN binding protein 9	10048
201663_s_at	SMC4	structural maintenance of chromosomes 4	10051
201177_s_at	UBA2	ubiquitin like modifier activating enzyme 2	10054
203050_s_at	DHM1L	dynamin 1 like	10059

User IDs mapped to multiple Entrez IDs or not mapped

userid
1007_s_at
1294_at
1494_f_at
200003_s_at
200012_x_at
200026_at
200032_s_at
200035_at
200038_s_at
200047_s_at
200063_s_at
200068_s_at

## 2.2 GOSlim Summary

The three charts represent Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories, in red, blue and green bars, respectively. The height of the bar represents the number of user list genes observed in the category.



Q9: Based on the pictures, how would you describe the genes in your dataset in your own words?

## 2.3 Enrichment Results

The left table is the summary table, and the right one is the detailed information table.

In the right table:

"C": the number of reference genes in the category

"O": the number of genes in the uploaded gene list and also in the category

"E": the expected number in the category

"R": ratio of enrichment

"P-Value": p-value from hypergeometric test

"FDR": FDR from BH

The screenshot shows the WebGestalt interface with the following elements highlighted:

- KEGG id**: A red oval highlights the "KEGG id" column in the main table.
- pathway name**: A red oval highlights the "Name" column in the main table.
- Related genes**: A red oval highlights the "Gene" column in the detailed information table.
- FDR**: A red oval highlights the "FDR" column in the detailed information table.
- Save results**: A red oval highlights the "Save results" button at the top right of the detailed information table.

**Main Table (Summary of enriched categories):**

ID	Name	Gene	FDR
hsa03050	Proteasome - Homo sapiens (human)	30	6.22e-09
hsa05106	Huntington's disease - Homo sapiens (human)	73	5.12e-08
hsa03010	Ribosome - Homo sapiens (human)	63	7.64e-08
hsa05012	Parkinson's disease - Homo sapiens (human)	57	7.64e-08
hsa03040	Spliceosome - Homo sapiens (human)	55	1.1e-07
hsa04190	Oxidative phosphorylation - Homo sapiens (human)	51	1.17e-05
hsa05010	Alzheimer's disease - Homo sapiens (human)	66	2.99e-05
hsa04032	Non-alcoholic fatty liver disease (NAFLD) - Homo sapiens (human)	58	3.75e-05
hsa04142	Lysosome - Homo sapiens (human)	48	2.39e-04
hsa04141	Protein processing in endoplasmic reticulum - Homo sapiens (human)	58	6.49e-04

**Detailed Information Table:**

ID	Name	user_id	Gene Symbol	Gene Name	Entrez Gene
hsa03050	Proteasome - Homo sapiens (human)	200907_x_at	PSMB3	proteasome activator subunit 3	51917
		201679_x_at	PSMA1	proteasome subunit alpha 1	5682
		201510_x_at	PSMA2	proteasome subunit alpha 2	5653
		201532_x_at	PSMA3	proteasome subunit alpha 3	5604
		200398_x_at	PSMA4	proteasome subunit alpha 4	5635
		201271_x_at	PSMA5	proteasome subunit alpha 5	5605
		201114_x_at	PSMA7	proteasome subunit alpha 7	5688
		200975_x_at	PSMB1	proteasome subunit beta 1	5622
		200939_x_at	PSMB2	proteasome subunit beta 2	5620
		201400_x_at	PSMB3	proteasome subunit beta 3	5601
		202243_x_at	PSMB4	proteasome subunit beta 4	5622
		200780_x_at	PSMB7	proteasome subunit beta 7	5695
		204719_x_at	PSMB9	proteasome subunit beta 9	5688
		202059_x_at	PSMB10	proteasome subunit beta 10	5699

Q10: What are the top 10 significant pathways?

## **Exercise 4. Compare the three websites in terms of KEGG pathways enrichment**

What are the most significant pathways in each of the GSA websites?

How well do they agree?

Which website uses more databases? Which website uses more GSA methods?

Which website gives you better summary tables and figures?

What was your favorite GSA website?

