

Supplementary Figure 1

Figure S1A: MOET provides flexible options for beginning an analysis. Select from all of RGD's species (rat, human, mouse, dog, pig, chinchilla, squirrel, naked mole-rat, green monkey and bonobo) and/or ontologies (disease, phenotype, pathway, GO and ChEBI) (S1A1) and enter a list of identifiers (Affymetrix Array ID, GenBank Nucleotide, Ontology Term ID, Ensembl Gene ID, GenBank Protein ID, RGD ID, Ensembl Protein ID, Gene Symbol, dbSNP ID, EntrezGene ID, KEGG Pathway ID) (S1A2). The default is to enter gene symbols but any of the listed ID types are acceptable (S1A3). Alternatively, choose a sequence assembly for your selected species and enter the position of a genomic region (S1A4). Clicking 'Continue' (S1A5) initiates the enrichment analysis.

A MOET - Multi Ontology Enrichment Tool

The New MOET Algorithm (v.2 released in May 2021)

Select a Species to view enrichment for all RGD ontologies

Select an Ontology to view enrichment in all RGD species

Enter Symbols

Please select an identifier type

Affymetrix Array ID GenBank Nucleotide Ontology Term ID

Ensembl Gene GenBank Protein RGD ID

Ensembl Protein Gene Symbol dbSNP ID

EntrezGene ID Kegg Pathway

Enter a Genomic Region

Chr: 1 Start: Stop: Assembly: mRatBN7

Continue

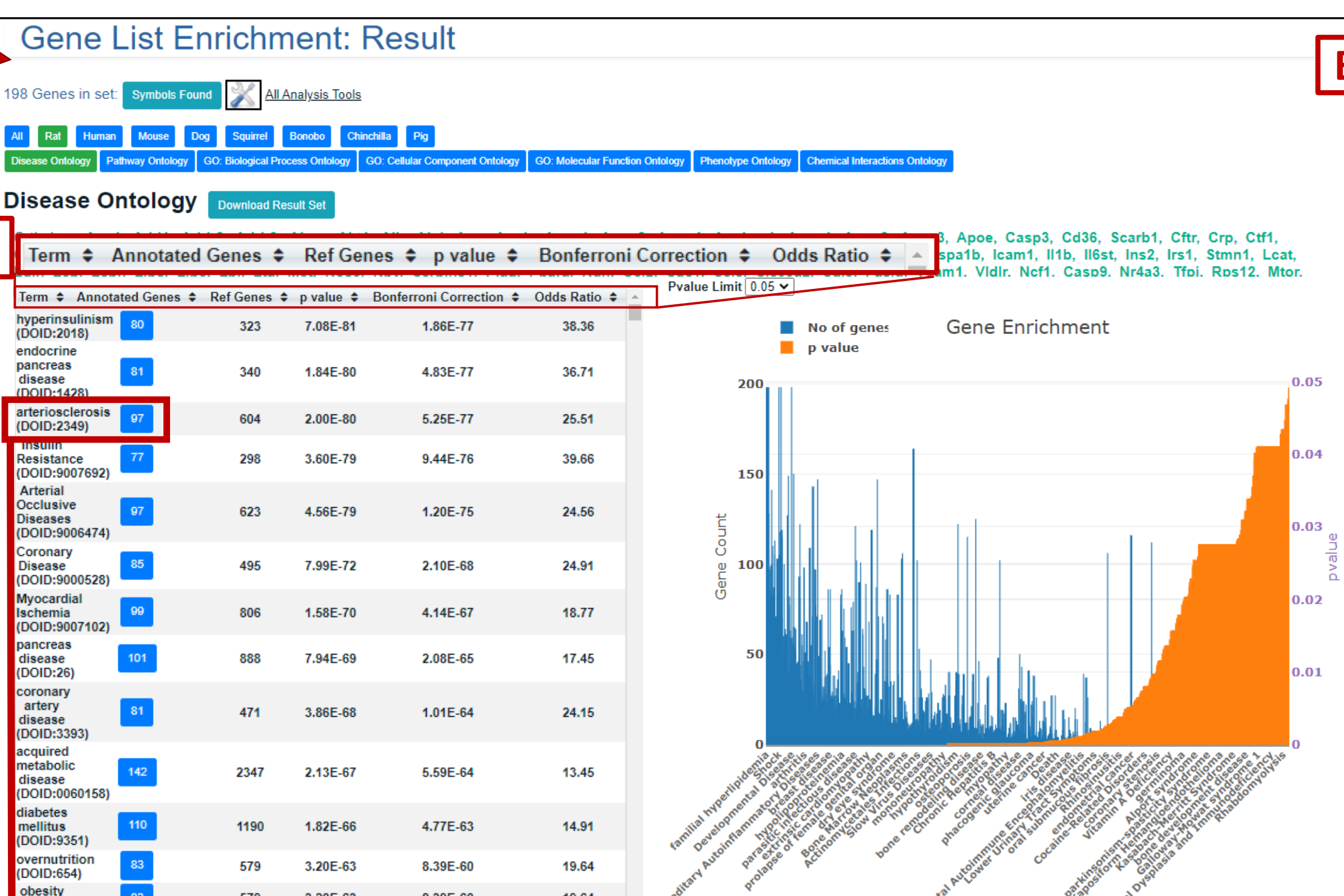


Figure S1B: The MOET result page. MOET displays the results as a list and a graph of overrepresented terms from all of the ontologies used by RGD for gene curation that are over-represented in the annotations for those genes, or orthologs in other species. (S1B1) The results in the list are sortable and are displayed with the number of Annotated genes, reference genes, p-value, Bonferroni Correction and Odds Ratio for each term.

B1 Explore this Gene Set

Abca1 ["coronary artery disease", "atherosclerosis"]

Abcb1a ["ST Elevation Myocardial Infarction"]

Abcc6 ["coronary artery disease"]

Abcg5 ["atherosclerosis", "coronary artery disease"]

Abcg8 ["coronary artery disease"]

Acat2 ["coronary artery disease"]

Acp1 ["coronary artery disease"]

Add1 ["myocardial infarction", "atherosclerosis"]

Adipoq ["atherosclerosis", "coronary artery disease"]

Adrb2 ["myocardial infarction"]

Adrb3 ["coronary artery disease"]

Ahsg ["myocardial infarction"]

Akt1 ["intermediate coronary artery disease"]

97 Genes in set: Symbols Found All Analysis Tools

Disease Ontology

Term	Annotated Genes	Ref Genes	p value	Bonferroni Correction	Odds Ratio
carbohydrate metabolism disease (DOID:000013)	88	1328	6.15E-81	1.47E-77	102.28
glucose metabolism disease (DOID:4194)	88	1328	6.15E-81	1.47E-77	102.28
hypercholesterolemia (DOID:900808)	48	115	3.00E-78	7.19E-75	206.80
inherited metabolic disorder (DOID:655)	97	2497	6.02E-75	1.44E-71	Infinity
vascular disease (DOID:178)	97	2661	3.24E-72	7.77E-69	Infinity
diabetes mellitus (DOID:9351)	80	1190	3.15E-70	7.55E-67	55.54
endocrine pancreas disease (DOID:1428)	88	340	3.80E-70	9.11E-67	73.46
Insulin Resistance (DOID:9007692)	95	298	6.30E-70	1.51E-66	78.84
hyperinsulinism (DOID:2318)	57	323	1.16E-69	2.78E-66	74.71
acquired metabolic disease	93	2347	3.63E-68	8.70E-65	123.34

Gene Enrichment graph showing Gene Count vs p value.

(S1B3) Use the 'All Analysis Tools' link and the tool box icon to analyze your gene list from MOET with the other RGD tools

(S1B2) Click on the button that shows the number of annotated genes to see the list of genes annotated to that term. The gene list can be sent again to MOET for analysis using the 'Explore this Gene Set' option. The 97 genes annotated to 'arteriosclerosis' are selected here.

Figure S1C: MOET provides the option to compare across species and ontologies, and to view and modify the results graph to view the results in different ways (S1C1). Click on the required species and ontology to view the list and graph of results for the selected species and ontology (S1C2). Below are the results with Human species and Pathway ontology overrepresented in the gene list entered in S1A2.

C Disease Ontology

C1 Pathway Ontology

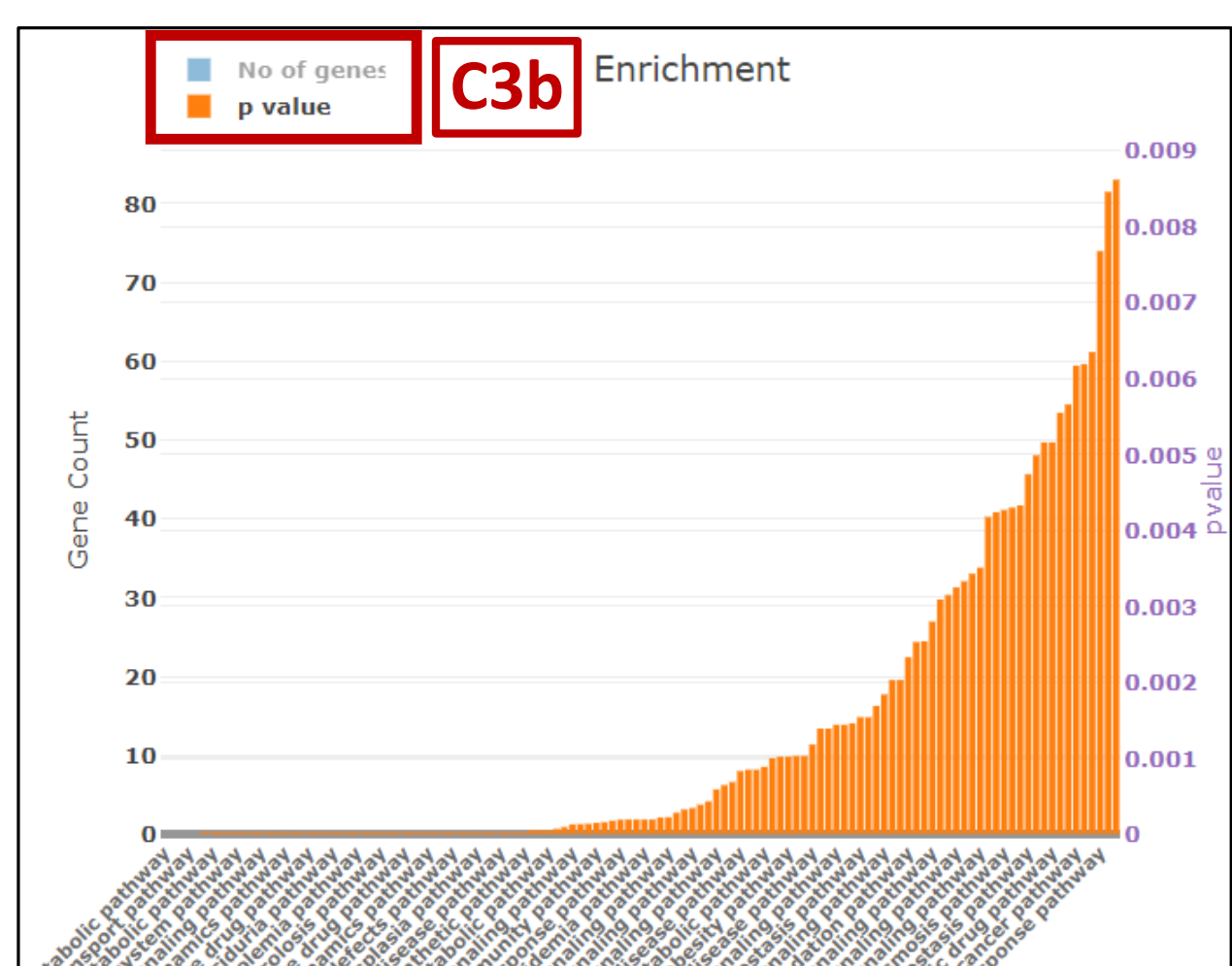
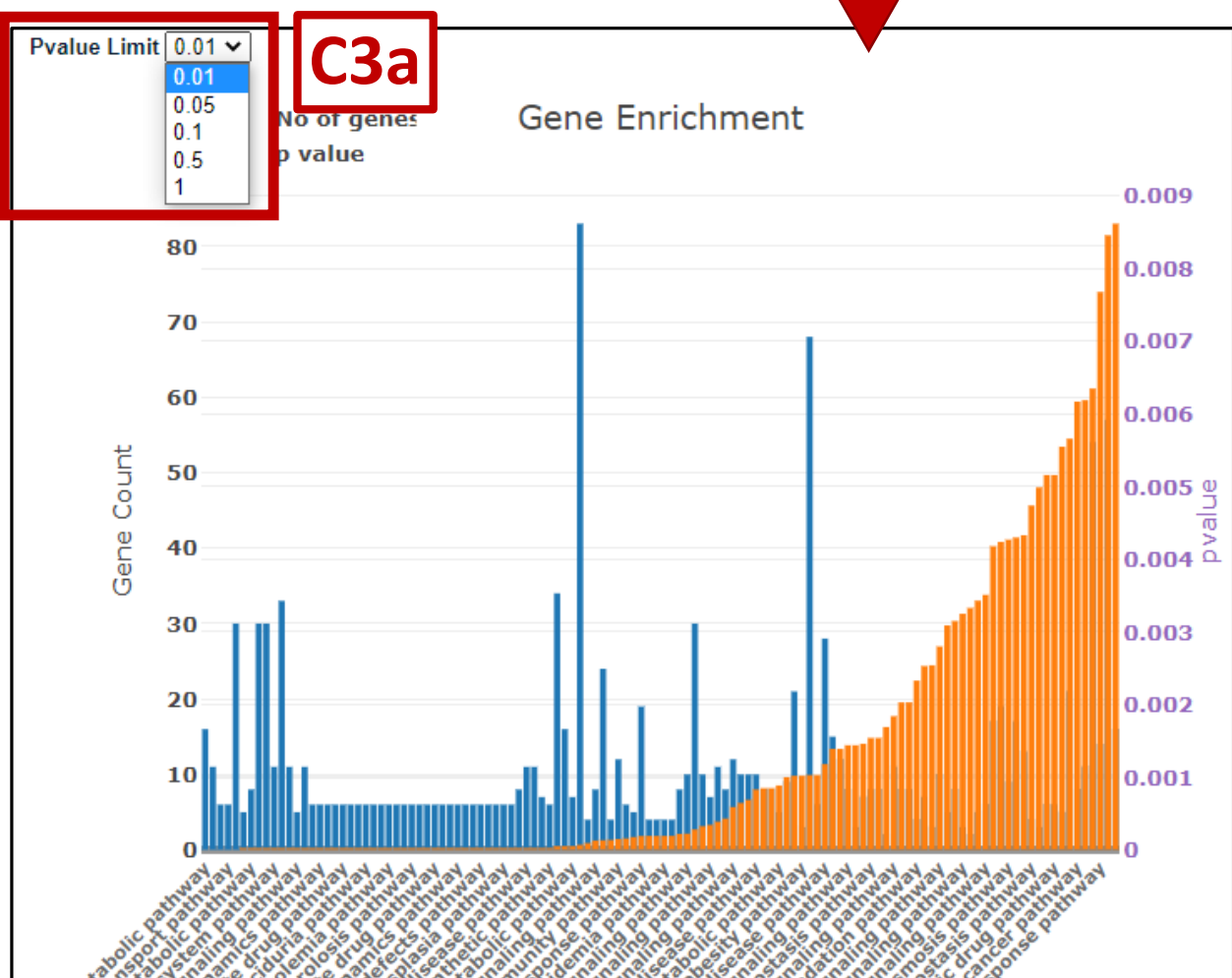
Orthologs: ADP1, MTKR, STMN1, GCG, SERPINE1, SREBF1, LEP, GSKR, CASP9, GHR, CYP7A1, SREBF1, ICAM1, FGF2, MET, LPL, CD36, CFTR, NPY, SOCS3, EDN1, SDC1, PLIN2, ABCB11, SOAT2, DNMT2, SREBF2, IL1B, VCAM1, CASP3, PCSK9, SOAT1, CCL2, DGAT1, EPHX2, CSF1, VEGFA, IL6ST, HMGCR, CPB2, RGN, BCHE, ABCG5, GSR, NOX1, PDE5A, NR4A1, ABCG8, KL, PHKG2, AKT1, HMBS, GPD1, PHKA1, IRS1, ADRB2, SLC30A2, HSPD1, USE1, ADD1, VLDLR, MIF, ABCG6, RPS12, ALB, INSIG2, APOC1, LDLR.

Term	Annotated Genes	Ref Genes	p value	Bonferroni Correction	Odds Ratio
lipoprotein metabolic pathway (PW:0000482)	16	20	2.32E-23	1.71E-20	182.99213
malaria pathway (PW:0001051)	11	51	2.17E-8	1.60E-5	12.029166
reverse cholesterol transport pathway (PW:000408)	6	10	3.35E-8	2.47E-5	63.61314
cholesterol transport pathway (PW:0001346)	6	10	3.35E-8	2.47E-5	63.61314
lipid metabolic pathway (PW:000010)	30	431	6.96E-8	5.13E-5	3.5837398
steroid regulatory element-binding protein signaling pathway (PW:0000753)	5	7	1.50E-7	1.11E-4	105.289856
altered metabolic pathway (PW:000262)	8	28	1.88E-7	1.39E-4	17.167408
transcription factor					

Gene Enrichment graph showing Gene Count vs p value.

Note: Only positive odds ratio so only positive associations, excludes high probability of a negative association.

(S1C3) The default p-value limit is 0.05. You can adjust it to modify the results in the graph accordingly. (S1C3a) The 'P-value limit' dropdown shows different options to modify the P-value limit. This is the graph when the p-value was changed to 0.01. (S1C3b) Click on the 'No of genes' or 'p value' to turn off the selected option in the result output graph.



C4 Download Result Set

Species:	Human					
Ontology:	Pathway Ontology					
No of genes in the input set:	198					
Input Genes:	Abca1 Abcb11 Abcb1a Abcc6 Abcg5 Abcg8 Acat2					
acc	refCount	oddsratio	correctedpvalue	count	term	pvalue
PW:0000482	20	182.99213	1.71E-20	16	lipoprotein metabolic pathway	2.32E-23
PW:0001051	51	12.029166	1.60E-05	11	malaria pathway	2.17E-08
PW:0000498	10	63.61314	2.47E-05	6	reverse cholesterol transport pathway	3.35E-08
PW:0001346	10	63.61314	2.47E-05	6	cholesterol transport pathway	3.35E-08
PW:0000010	431	3.5837398	5.13E-05	30	lipid metabolic pathway	6.96E-08
PW:0000753	7	105.289856	1.11E-04	5	steroid regulatory element-binding protein signaling pathway	1.50E-07
PW:000262	28	17.167408	1.39E-04	8	altered metabolic pathway	1.88E-07
PW:0000716	463	3.2992704	2.51E-04	30	transcription factor mediated signaling pathway	3.41E-07
PW:0001313	463	3.2992704	2.51E-04	30	transcription pathway via transcription factor mediated signaling	3.41E-07
PW:0000502	70	8.128531	4.92E-04	11	complement system pathway	6.68E-07
PW:0000465	563	2.9909434	5.93E-04	33	hormone signaling pathway	8.05E-07
PW:0000474	72	7.8592896	6.60E-04	11	coagulation cascade pathway	8.95E-07
PW:0001371	10	42.094204	1.25E-03	5	basic helix-loop-helix signaling pathway	1.70E-06
PW:0000475	77	7.257576	1.32E-03	11	hemostasis pathway	1.79E-06
PW:0001093	18	21.175182	1.86E-03	6	bile acid signaling pathway	2.52E-06
PW:0000728	22	15.870438	6.90E-03	6	statin pharmacodynamics pathway	9.36E-06

(S1C4) You can click on 'Download Result Set' to download the results in a .CSV format file.