

WebGestalt by examples

<http://genereg.ornl.gov/webgestalt>

Examples

1. Candidate gene identification
2. Chromosome distribution
3. Gene Ontology analysis
4. Pathway analysis
5. Boolean operations
6. Connecting from WebQTL

Examples

1. Candidate gene identification

- Starting from a genomic region (whole chromosome 15 in the example), try to narrow down the number of candidate genes that may contribute to a cerebellum-related phenotype

2. Chromosome distribution

3. Gene Ontology analysis

4. Pathway analysis

5. Boolean operations

6. Connecting from WebQTL

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD By location RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLE From file Select gene set A Select geen set B Select Operator

By location

By Gene Ontology

Select "By location" from the UPLOAD options

Done Internet

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

Gene Set Upload Form

Gene set name:

Gene set description:

Organism:

Chromosome number:

Begin(bp):

End(bp):

UPLOAD

Gene set uploading form
(By location)

Done Internet

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

Gene Set Upload Form

Gene set name:

Gene set description:

Chromosome number: 15
Begin(bp):1
End(bp):300000000

Organism:

Chromosome number:

Begin(bp):

End(bp):

UPLOAD

Click the "UPLOAD" button after complete the form

Done Internet

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

>>Active gene set description

- >Gene set name:Mm_chr_15
- >Gene set description:All genes on mouse chromosome 15
- >There are 581 input IDs in this gene set. These IDs represent 581 known genes

>>Gene set information retrieval tool

- >Default IDs Input ID LocusID
- >Nomenclature Gene symbol Symbol alias Gene name Name alias
- >Other IDs Refseq_NM Refseq_NP UNIGENE ENSEMBL SwissProt
- >Map Info Cytogenetic Physical
- >Function Info Domain OMIM PubMed GRIF
- Gene Ontology KEGG Biocarta Phenotype

Information Retrieval

>>Gene set organization tool

GO Tree Tissue Expression Bar Chart Chromosome Distribution Chart Protein Domain Table

GRIF Table PubMed Table BioCarta Table and Maps

>>Gene list(Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol	Name
104445	104445	Cdc42ep1	CDC42 effector protein (Rho GTPase binding) 1
105722	105722	Tmem16f	transmembrane protein 16F
105727	105727	Slc38a1	solute carrier family 38, member 1
105732	105732	AA409316	expressed sequence AA409316
105734	105734	AA409802	expressed sequence AA409802
105782	105782	Scrib	scribbled homolog (Drosophila)

Done Internet

User interface with various gene set analysis tools

Click the "Tissue Expression Bar Chart" button



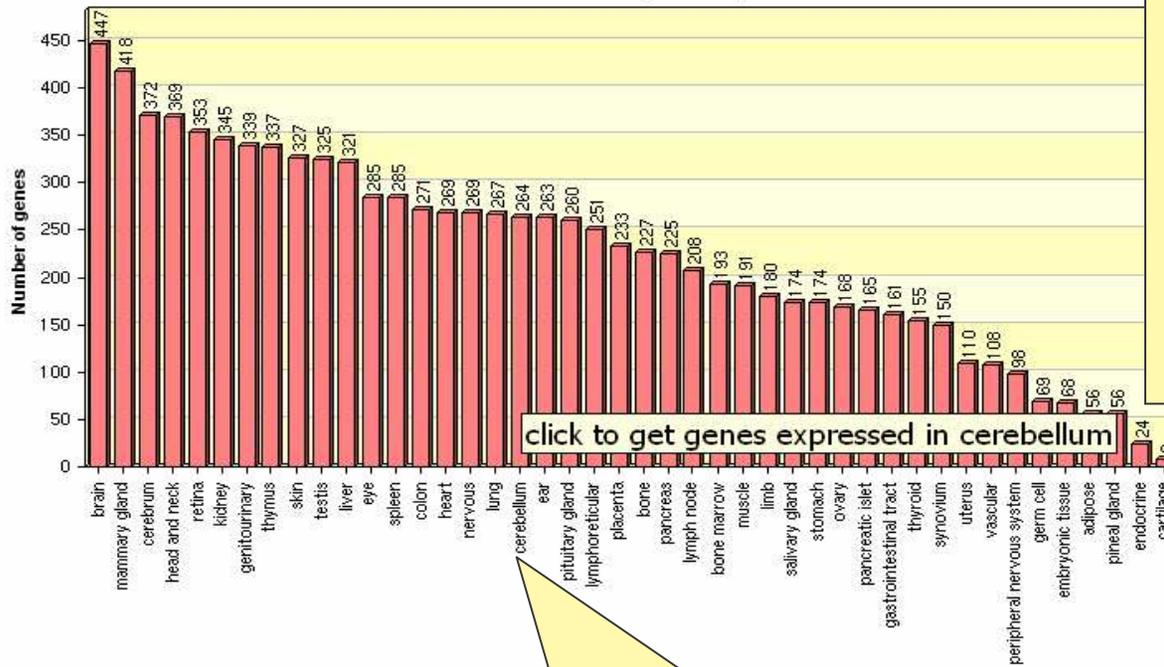
WebGestalt

Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>

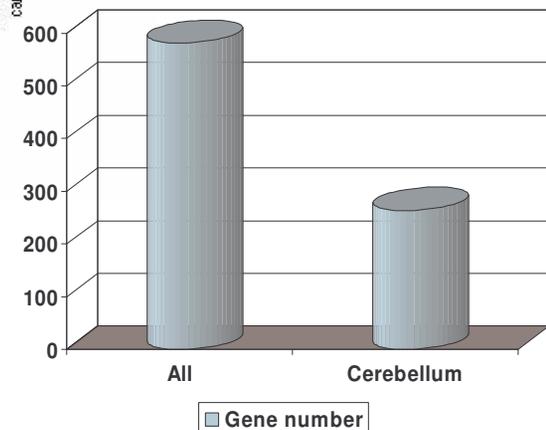
University of Tennessee and Oak Ridge National Laboratory

Bar chart of the tissue expression pattern



If you are looking for candidate genes from chromosome 15, and may contribute to a cerebellum-related phenotype, this simple tool can reduce the candidate gene numbers by 55%.

Among the 581 genes on chromosome 15, only 264 are expressed in cerebellum



Examples

1. Candidate gene identification
2. Chromosome distribution
 - Starting from a functional category, do neurotransmitter regulators form clusters on the chromosome?
3. Gene Ontology analysis
4. Pathway analysis
5. Boolean operations
6. Connecting from WebQTL

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

Select method
 Select gene set
 Select gene set

Select method
 From file
 By location
 By Gene Ontology

>>Active gene set description

>Gene set name:Mm_chr_15
 >Gene set description:All genes on mouse chromosome 15
 >There are 581 input IDs in this gene set. These IDs represent 581 known genes

>>Gene set information retrieval

>Default IDs Input ID
 >Nomenclature Gene
 >Other IDs Refseq
 >Map Info Cyto
 >Function Info Dom
 Gene

Information Retrieval

>>Gene set organization tool

>>Gene list(Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol	Name
104445	104445	Cdc42ep1	CDC42 effector protein (Rho GTPase binding) 1
105722	105722	Tmem16f	transmembrane protein 16F
105727	105727	Slc38a1	solute carrier family 38, member 1
105732	105732	AA409316	expressed sequence AA409316
105734	105734	AA409802	expressed sequence AA409802
105782	105782	Scrib	scribbled homolog (Drosophila)

Done Internet

Select "By Gene Ontology" from the UPLOAD options

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

Gene Set Upload Form

Gene set name:

Gene set description:

Organism:

GO ID: *Don't know the GO ID? Query here:

Done

Gene set uploading form
(By GO category)

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

Gene Set Upload Form

Gene set name:

Gene set description:

Organism:

GO ID: *Don't know the GO ID? Query here:

Query the GO IDs for "neuro"

Done Internet

GO_term search - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

neurotransmitter uptake: [GO:0001504](#)
[regulation of neurotransmitter levels: GO:0001505](#)
 neuromedin U receptor activity: [GO:0001607](#)
 neuronal migration: [GO:0001764](#)
 ciliary neurotrophic factor receptor activity: [GO:0004897](#)
 neuropeptide Y receptor activity: [GO:0004983](#)
 neurotrophin TRK receptor activity: [GO:0005013](#)
 neurotrophin receptor activity: [GO:0005030](#)
 ciliary neurotrophic factor receptor binding: [GO:0005127](#)
 neurotrophin receptor binding: [GO:0005165](#)
 neurotrophin p75 receptor binding: [GO:0005166](#)
 neurotrophin TRKB receptor binding: [GO:0005169](#)
 neuropeptide hormone activity: [GO:0005184](#)
 neurohypophyseal hormone activity: [GO:0005185](#)
 neurotransmitter transporter activity: [GO:0005326](#)
 neurotransmitter:sodium symporter activity: [GO:0005328](#)
 neurofilament: [GO:0005883](#)
 neurotransmitter transport: [GO:0006836](#)
 neuronal cell adhesion: [GO:0007158](#)
 neuropeptide signaling pathway: [GO:0007218](#)
 neurotransmitter secretion: [GO:0007269](#)
 neuromuscular synaptic transmission: [GO:0007274](#)
 neurogenesis: [GO:0007399](#)
 neuroblast cell fate determination: [GO:0007400](#)
 neuroblast proliferation: [GO:0007405](#)
 negative regulation of neuroblast proliferation: [GO:0007406](#)
 fasciculation of neuron: [GO:0007413](#)
 defasciculation of neuron: [GO:0007414](#)
 defasciculation of motor neuron: [GO:0007415](#)
 neuromuscular junction development: [GO:0007528](#)
 establishment of synaptic specificity at neuromuscular junction: [GO:0007529](#)
 neuronal pentraxin receptor activity: [GO:0008030](#)
 neuronal cell recognition: [GO:0008038](#)
 neuropeptide receptor activity: [GO:0008188](#)
 regulation of synaptic growth at neuromuscular junction: [GO:0008582](#)
 glial cell line-derived neurotrophic factor receptor activity: [GO:0016167](#)
 neuronal remodeling: [GO:0016322](#)
 neurotensin receptor activity, G-protein coupled: [GO:0016492](#)
 glial cell line-derived neurotrophic factor receptor binding: [GO:0030116](#)
 neuron differentiation: [GO:0030182](#)
 neurotransmitter receptor activity: [GO:0030594](#)
 neuronal lineage restriction: [GO:0042055](#)
 long-term strengthening of neuromuscular junction: [GO:0042062](#)

GO categories and the GO IDs matching the query “neuro”, Let’s assume that we are interested in genes related to “regulation of neurotransmitter levels” (GO:0001505)

Internet

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

Gene Set Upload Form

Gene set name:

Gene set description:

Organism:

GO ID: *Don't know the GO ID? Query here:

Done Internet

Complete the Form with the GO ID GO:0001505 (neurotransmitter , the Click the UPLOAD button

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

>>Active gene set description

>Gene set name:Mm_RL_NT_levels
>Gene set description:All genes in the GO category "regulation of neurotransmitter levels" in mou
>There are 35 input IDs in this gene set. These IDs represent 35 known genes

>>Gene set information retrieval tool

>Default IDs Input ID LocusID

>Nomenclature Gene symbol Symbol alias Gene name Name alias

>Other IDs Refseq_NM Refseq_NP UNIGENE ENSEMBL SwissProt

>Map Info Cytogenetic Physical

>Function Info Domain OMIM PubMed GRIF
 Gene Ontology KEGG Biocarta Phenotype

Information Retrieval

>>Gene set organization tool

GO Tree Tissue Expression Bar Chart **Chromosome Distribution Chart** Protein Domain Table

GRIF Table PubMed Table BioCarta Table and Maps

>>Gene list(Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol Name	
104245	104245	Slc6a5	solute carrier family 6 (neurotransmitter transporter, glycine), me
109647	109647	For5	formamidase 5
12287	12287	Cacna1b	calcium channel, voltage-dependent, N type, alpha 1B subunit
12322	12322	Camk2a	calcium/calmodulin-dependent protein kinase II alpha
12647	12647	Chat	choline acetyltransferase
12846	12846	Comt	catechol-O-methyltransferase

Done Internet

User interface with various gene set analysis tools

Click the "Chromosome Distribution Chart" button

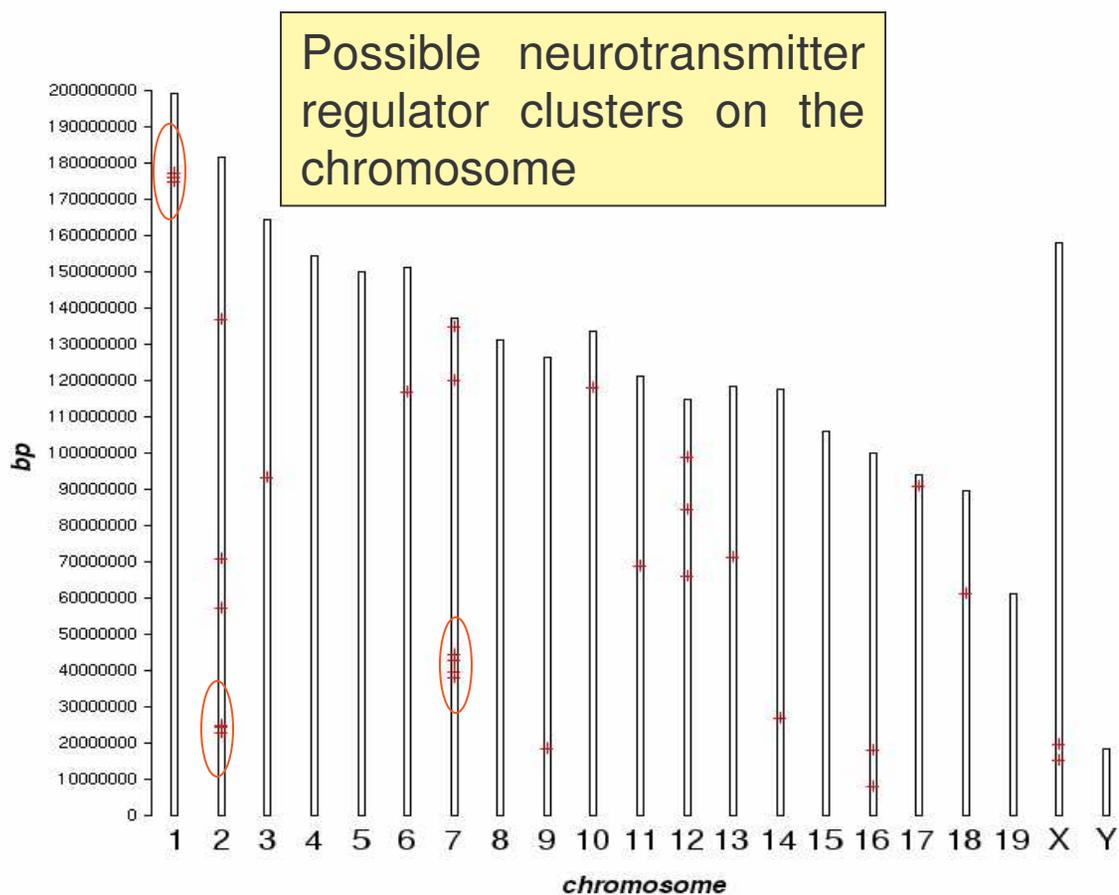


WebGestalt

Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>

University of Tennessee and Oak Ridge National Laboratory



Examples

1. Candidate gene identification
2. Chromosome distribution
3. Gene Ontology analysis
 - Starting from a gene list generated by high-throughput techniques, what are the major functional categories of the ethanol responsive genes identified from my microarray experiment?
4. Pathway analysis
5. Boolean operations
6. Connecting from WebQTL

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

>>Gene set management tool

From file
 Select gene set
 Select gene set

Select method
 From file
 By location
 By Gene Ontology

>>Active gene set description

>Gene set name: Mm_RL_NT_Le
 >Gene set description: All genes in GO category "regulation of neurotransmitter levels" in mouse
 >There are 35 input IDs in this gene set. These IDs represent 35 known genes

>>Gene set information retrieval

>Default IDs Input ID
 >Nomenclature Gene
 >Other IDs Refseq Uniprot
 >Map Info Cytoband
 >Function Info Domains Gene Ontology

>>Gene set organization tool

>>Gene list (Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol Name	Description
104245	104245	Slc6a5	solute carrier family 6 (neurotransmitter transporter, glycine), member 5
109647	109647	For5	formamidase 5
12287	12287	Cacna1b	calcium channel, voltage-dependent, N type, alpha 1B subunit
12322	12322	Camk2a	calcium/calmodulin-dependent protein kinase II alpha
12647	12647	Chat	choline acetyltransferase
12846	12846	Comt	catechol-O-methyltransferase

Internet

Select "From file" from the UPLOAD options

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

Gene Set Upload Form

Gene set name:

Gene set description:

Organism:

ID type:

Upload your file of gene set:

Internet

Upload a gene list with 330 Affymetrix probe set ids that are responsive to ethanol treatment, as identified from a microarray experiment

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select geen set B Select Operator

>>Active gene set description

- >Gene set name: Ethanol_response
- >Gene set description: Genes responds to ethanol treatment in wild type mouse striatum
- >There are 519 input IDs in this gene set. These IDs represent 494 known genes

>>Gene set information retrieval tool

- >Default IDs Input ID LocusID
- >Nomenclature Gene symbol Symbol alias Gene name Name alias
- >Other IDs Refseq_NM
- >Map Info Cytogenetic
- >Function Info Domain Gene Ontol

Information Retrieval

>>Gene set organization tool

GO Tree Tissue Expression Bar Chart Chromosome Distribution Chart Protein Domain Table

GRIF Table PubMed Table BioCarta Table and Maps

>>Gene list (Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol	Name
100033_at	17685	Msh2	mutS homolog 2 (E. coli)
100043_f_at	76943	2310020A21Rik	RIKEN cDNA 2310020A21 gene
100056_at	30050	Fbxw2	F-box and WD-40 domain protein 2
100073_at	68043	2510005D08Rik	RIKEN cDNA 2510005D08 gene
100086_at	16976	Lrpap1	low density lipoprotein receptor-related protein associated protein 1
100150_f_at	16334	Ins2	insulin II

Done Internet

User interface with various gene set analysis tools

Click the "GO Tree" button

Add a reference? - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

If you just want to organize the active gene set by GO, you can press the "Make GO Tree" button directly. However, if you want to do statistical analysis to identify GO categories with enriched gene numbers, you need to select a reference gene set from the list. The list includes some prestored gene sets in WebGestalt, including human genome, mouse genome and some affy arrays (prefixed with "WEBGESTALT_"), as well as all of your uploaded gene sets. WebGestalt provides two statistical methods, hypergeometric test and Fisher's exact test. If you know which method is better for your analysis, you can choose it from the dropdown menu. If you use MOE430 as a reference, as it is a special combination of MOE430A and MOE430B, hypergeometric test will be used. If you are not sure which method is better for you, WebGestalt will try to choose an appropriate one for you based on the reference gene set you select. However, this will take some time.

Select a reference set:

WEBGESTALT_MG_U74AV2

Select a statistical method:

Hypergeometric test

Make GO Tree

Select a reference gene set and a statistical method, then click the "Make GO Tree" button

Done Internet

Go Tree Structure - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

There are totally 4590 categories in your GOTree.
WebGestalt will use **Hypergeometric test** to calculate the statistic for each category.
It may take a while if the number of categories is large. Please be patient.
If you didn't pick the statistical method by yourself, WebGestalt has picked the Hypergeometric test for you because:
All genes in the active gene set are included in the reference gene set.

[.....] progress scale
[.....]

It may take a while if the gene list is large, waiting for the processing...

Opening page <http://genereg.ornl.gov/webgestalt/gotree.php...> Internet

Go Tree Structure - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

There are totally 4590 categories in your GOTree.
WebGestalt will use **Hypergeometric test** to calculate the statistic for each category.
It may take a while if the number of categories is large. Please be patient.
If you didn't pick the statistical method by yourself, WebGestalt has picked the Hypergeometric test for you because:
All genes in the active gene set are included in the reference gene set.

[.....] progress scale
[.....]

Check Go Tree

When it is done, click the "Check Go Tree" button

Done Internet

Go Tree Structure - Microsoft Internet Explorer

File Edit

WebGestalt genereg.ornl.gov/webgestalt
University of Tennessee and Oak Ridge National Laboratory

GO term search

Bar chart at selected GO annotation level

Export GOTree

GO Term Search Key Word Search **biological_process** Level 4

Bar Chart Export GOTree Enriched DAG

Close

GOTree for the gene set **tom_ECvsSC**
All genes from **MG_U74AV2** were used as the reference for the statistical analysis
Use **GO Term Search** to search for one or more GO categories, **separated by comma only**
Use **Key Word Search** when you don't know the exact GO term, **one word or phase each time**
Use **Enriched DAG** to get the graphical network view of enriched GO categories ($p < 0.01$) in a DAG (Directed Acyclic Graph)
Use **Bar Chart** to get the bar chart for the GO categories at the selected annotation level
Use **Export GOTree** to export GOTree in html format
Click on + to browse the GOTree, click on each node to check gene list

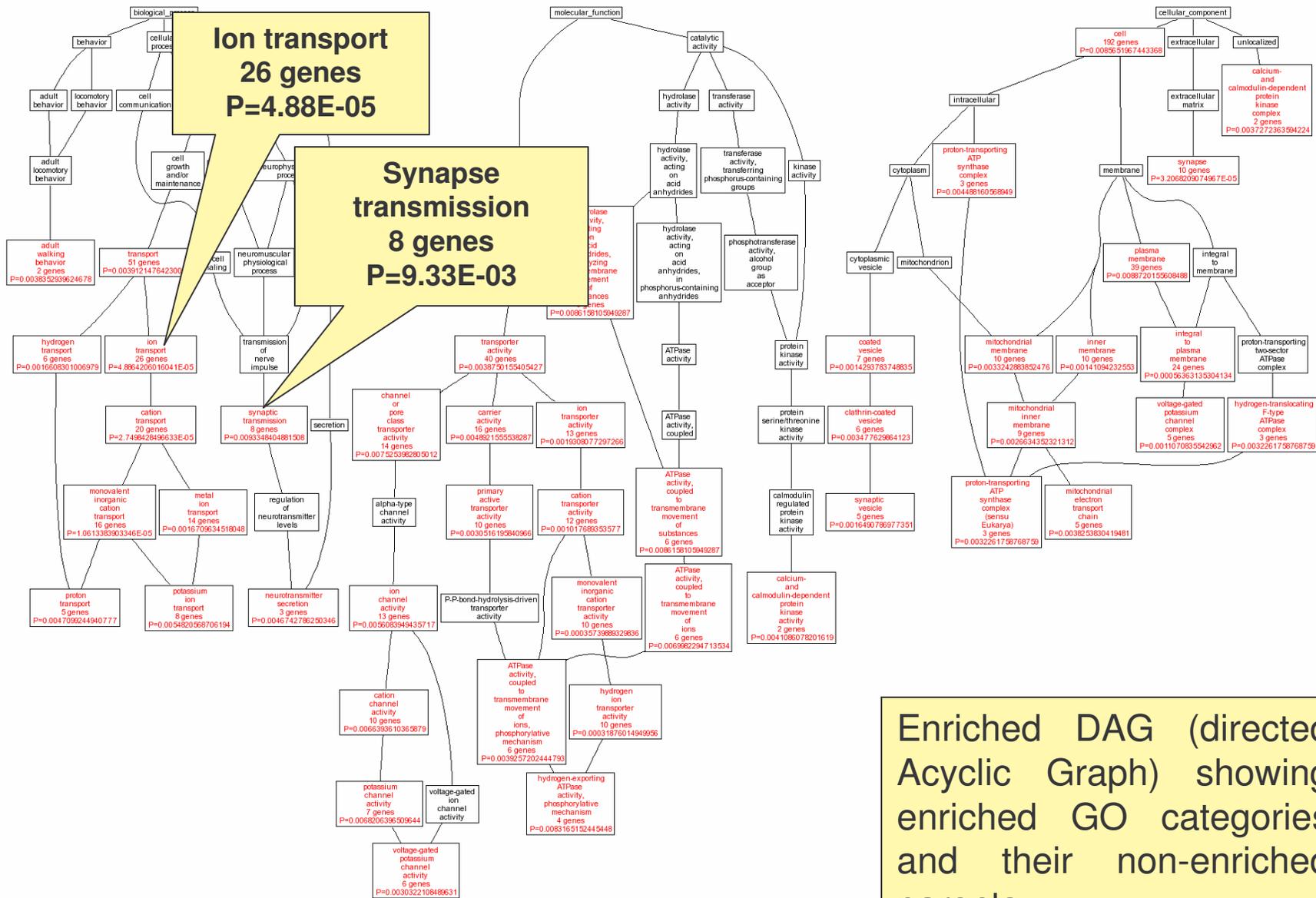
Enriched DAG

Important categories are colored red

Browse the GO Tree

- biological_process(O=214;E=214;R=1)
 - behavior(O=3;E=2.36;R=1.27;P=0.42138232209814)
 - adult behavior(O=2;E=0.4;R=5;P=0.058059951578858)
 - adult locomotory behavior(O=2;E=0.29;R=6.9;P=0.031738944712094)
 - adult walking behavior*(O=2;E=0.11;R=18.18;P=0.0038352939624678)**
 - locomotory behavior(O=2;E=0.76;R=2.63;P=0.17569968038461)
 - rhythmic behavior(O=1;E=0.58;R=1.72;P=0.44672514821446)
 - biological_process unknown(O=10;E=11.28;R=0.89)
 - cellular process(O=125;E=115.85;R=1.08;P=0.11317722333308)
 - development(O=32;E=37.03;R=0.86)
 - physiological process(O=182;E=178.96;R=1.02;P=0.32198167785308)
 - regulation of biological process(O=38;E=46.93;R=0.81)
 - molecular_function(O=239;E=239;R=1)
 - binding(O=139;E=146.48;R=0.95)
 - catalytic activity(O=104;E=89.93;R=1.16;P=0.033171730611191)
 - enzyme regulator activity(O=8;E=9.46;R=0.85)
 - molecular_function unknown(O=8;E=10.25;R=0.78)
 - motor activity(O=4;E=2.22;R=1.8;P=0.17989232120059)
 - obsolete molecular function(O=1;E=1.43;R=0.7)
 - signal transducer activity(O=39;E=43.37;R=0.9)
 - structural molecule activity(O=19;E=14.49;R=1.31;P=0.13537721446544)
 - transcription regulator activity(O=17;E=23.02;R=0.74)
 - translation regulator activity(O=2;E=3.04;R=0.66)
 - transporter activity*(O=40;E=26.21;R=1.53;P=0.0038750155405427)**
 - cellular_component(O=215;E=215;R=1)
 - cell*(O=192;E=179.25;R=1.07;P=0.0085651967443368)**
 - cell fraction(O=6;E=6.26;R=0.96)

Internet



Enriched DAG (directed Acyclic Graph) showing enriched GO categories and their non-enriched parents

Examples

1. Candidate gene identification
2. Chromosome distribution
3. Gene Ontology analysis
4. **Pathway analysis**
 - Starting from the same gene list as in example 3, what are the pathways in which the ethanol responsive genes involve and where are the genes located in the pathways?
5. Boolean operations
6. Connecting from WebQTL

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

WebGestalt

>>Gene set management tool

UPLOAD Select method RETRIEVE ST_ECvsSC_homolog DELETE Select gene set LOG OFF

BOOLEAN OPERATION Select gene set A Select gene Select Operator

>>Active gene set description

- >Gene set name:ST_ECvsSC_homolog
- >Gene set description:homolog of ST_ECvsSC
- >There are 257 input IDs in this gene set. These IDs represent 257 known genes

>>Gene set information retrieval tool

- >Default
- >Nomenclature
- >Other
- >Map Information
- >Function Information
 - OMIM
 - PubMed
 - GRIF
 - KEGG
 - Biocarta
 - Phenotype

Information Retrieval

>>Gene set organization tool

GO Tree Tissue Expression Bar Chart Chromosome Distribution Chart

GRIF Table PubMed Table **KEGG Table and Maps** BioCarta Table and Maps

>>Gene list(Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol	Name
103172	400916	C22orf16	chromosome 22 open reading frame 16
103742	84299	C17orf37	chromosome 17 open reading frame 37
105083	53918	PELO	pelota homolog (Drosophila)
107094	23223	KIAA0690	KIAA0690
107373	63901	FLJ22794	FLJ22794 protein
108911	55920	TD-60	RCC1-like

Internet

Click the "KEGG Table and Maps" button

It's a pity that WebGestalt doesn't have the mouse KEGG pathway information in the database yet, however, when it uploads the gene list, a corresponding human homologous gene list is automatically uploaded, which can be retrieved for the KEGG pathway analysis

KEGG Table - Microsoft Internet Explorer

File Edit View Favorites Tools Help


GeneReg
<http://genereg.ornl.gov>
 University of Tennessee

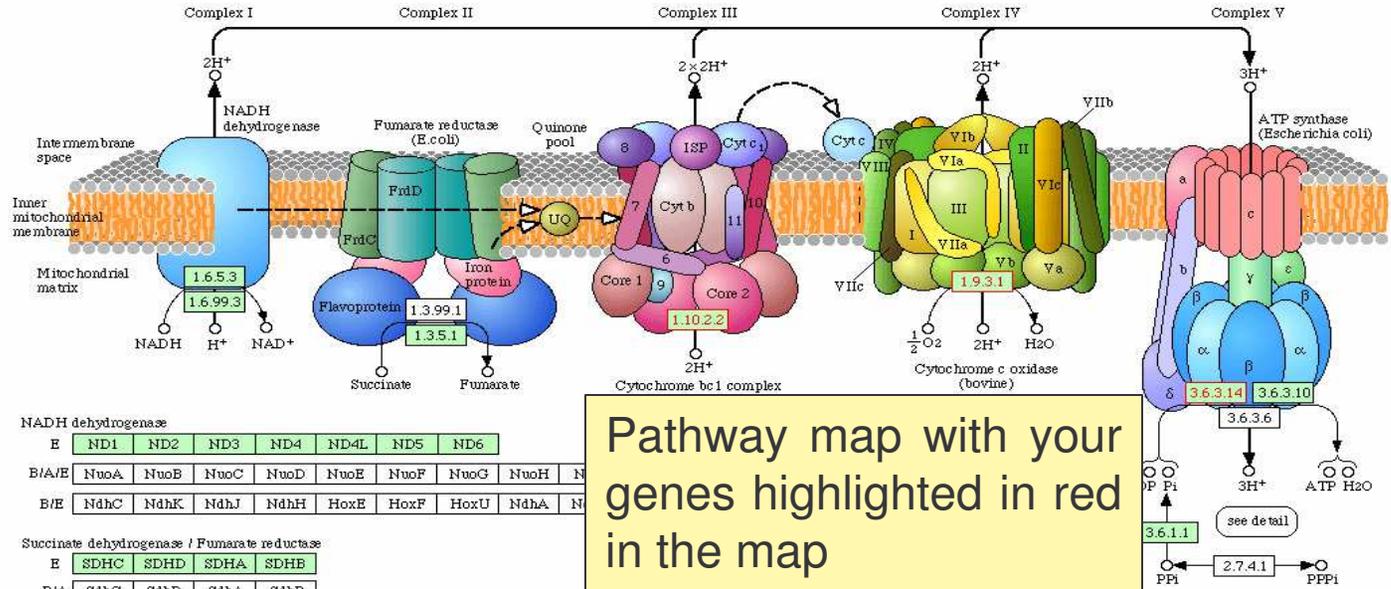
WebGestalt

All pathways related to the gene list. Click on the "Oxidative Phosphorylation" if this is the map you are interested in

KEGG map	Gene number	Locus link IDs
Oxidative Phosphorylation	9	10632 10975 1337 23545 27089 514 521 535 7384
Phosphatidylinositol Signaling System	6	11266 140885 3632 5286 5795 8760
ATP Synthesis	5	10632 23545 514 521 535
Type III Secretion System	5	10632 23545 514 521 535
Purine Metabolism	4	2987 5437 9061 9583
Glycerolipid Metabolism	4	10554 4023 5049 8760
Pyrimidine Metabolism	3	1841 5437 9583
Glycolysis_Gluconeogenesis	2	1737 97
Ubiquitin-mediated Proteolysis	2	7328 997
Pyruvate Metabolism	2	1737 97
Fructose and Mannose Metabolism	2	5208 5372
Arginine and Proline Metabolism	2	1159 4842
Inositol Phosphate Metabolism	2	3632 5286
Citrate Cycle (TCA Cycle)	1	2271
Aminosugars Metabolism	1	55907
RNA Polymerase	1	5437
Aminoacyl-tRNA Biosynthesis	1	7453
Sulfur Metabolism	1	9061
Nitrogen Metabolism	1	762
Porphyrin and Chlorophyll Metabolism	1	3163
Sphingoglycolipid Metabolism	1	7357
Phospholipid Degradation	1	11313
Keratan Sulfate Biosynthesis	1	8704
Chondroitin / Heparan Sulfate Biosynthesis	1	2137
Proteasome	1	5691
Glycoprotein Biosynthesis	1	8704
Galactose Metabolism	1	8704
Tryptophan Metabolism	1	7453
Selenoamino Acid Metabolism	1	9061
Glutathione Metabolism	1	2944
Urea Cycle and Metabolism of Amino Groups	1	1159

Internet

OXIDATIVE PHOSPHORYLATION



Pathway map with your genes highlighted in red in the map

NADH dehydrogenase

E	ND1	ND2	ND3	ND4	ND4L	ND5	ND6		
B/A/E	NuoA	NuoB	NuoC	NuoD	NuoE	NuoF	NuoG	NuoH	NuoI
B/E	NdhC	NdhK	NdhJ	NdhH	HoxE	HoxF	HoxU	NdhA	NdhB

Succinate dehydrogenase / Fumarate reductase

E	SDHC	SDHD	SDHA	SDHE
B/A	SdhC	SdhD	SdhA	SdhB
	FrdA	FrdE	FrdC	FrdD

Cytochrome c reductase

E/B/A	ISP	Cyt b	Cyt c1				
E	COR1	QCR2	QCR6	QCR7	QCR8	QCR9	QCR10

Cytochrome c oxidase

E	COX10	COX3	COX1	COX2	COX4	COX5A	COX5B	COX6A	COX6B	COX6C	COX7A	COX7B	COX7C	COX8	E/B/A	COX11	COX15	COX17
B/A	CyoE	CyoD	CyoC	CyoB	CyoA	CoxD	CoxC	CoxA	CoxB	QoxD	QoxC	QoxB	QoxA					

Cytochrome c oxidase, cbb3-type

B	I	II	IV	III
---	---	----	----	-----

Cytochrome bd complex

B/A	CydA	CydB
-----	------	------

Examples

1. Candidate gene identification
2. Chromosome distribution
3. Gene Ontology analysis
4. Pathway analysis
5. Boolean operations
 - More than 500 genes are responsive to ethanol treatment in the wild-type mouse, however, only 47 genes are responsive in the mutant. Are there any common genes in the two responsive gene sets?
6. Connecting from WebQTL

WebGestalt - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Click the "Boolean operation" button after complete the selections

Select gene set II: responsive genes in the mutant

Select gene set I: responsive genes in the wild type

Select the operation of "A and B, Intersection" as we are looking for commonalities

>>Gene set management tool

UPLOAD Select method RETRIEVE Select gene set DELETE Select gene set LOG OFF

BOOLEAN OPERATION ST_ECvsNC ST_EMvsNM A and B (Intersection)

>>Active gene set description

>Gene set name:ST_EMvsNM

>Gene set description:Gene list generated from microarray data. Genes that were differentially expressed in the ethanol treated sample and non- ethanol treated sample.

>There are 47 input IDs in this gene set

>>Gene set information retrieval

>Default IDs Input ID

>Nomenclature Gene symbol

>Other IDs Refseq_NM Refseq_NP UNIGENE ENSEMBL SwissProt

>Map Info Cytogenetic Physical

>Function Info Domain OMIM PubMed GRIF Gene Ontology KEGG Biocarta Phenotype

Information Retrieval

>>Gene set organization tool

GO Tree Tissue Expression Bar Chart Chromosome Distribution Chart Protein Domain Table

GRIF Table PubMed Table BioCarta Table and Maps

>>Gene list(Click on each LLID to get gene information record for individual genes)

Input ID	LLID	Symbol	Name
100029	56273	Dax1	peroxisomal biogenesis factor 14

Done Internet

Boolean operation - Microsoft Internet Explorer

File Edit View Favorites Tools Help



Gene Set Analysis Toolkit

<http://genereg.ornl.gov/webgestalt>
University of Tennessee and Oak Ridge National Laboratory

ST_ECvsNC 'INTERSECT' ST_EMvsNM generates:

LLID	Symbol Name
11980	Atp8a1 ATPase, aminophospholipid transporter (APLT), class I, type 8A, member 1
16572	Kif5a kinesin family member 5A
16976	Lrpap1 low density lipoprotein receptor-related protein associated protein 1
258803	Olf136 olfactory receptor 136

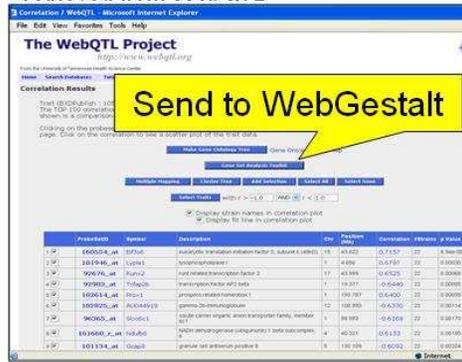
There are 4 genes that are common in the two gene sets

Done Internet

Examples

1. Candidate gene identification
2. Chromosome distribution
3. Gene Ontology analysis
4. Pathway analysis
5. Boolean operations
6. Connecting from WebQTL
 - Analyses you can do for a gene set with 100 genes whose expression patterns are highly correlated with the “ethanol acceptance-male raw mean consumption” phenotype, as identified from WebQTL (<http://www.webqtl.org>)

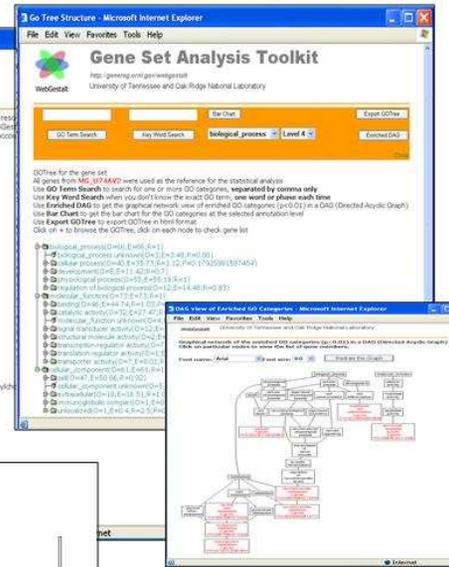
A. The top 100 genes whose expression patterns are highly correlated with the "ethanol acceptance-male raw mean consumption" phenotype were retrieved from WebQTL



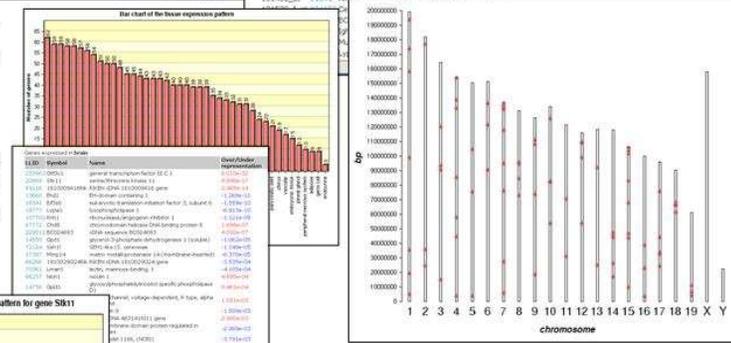
B. WebGestalt receives the gene set from WebQTL



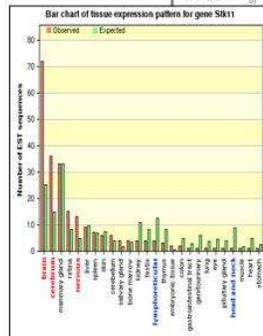
C. GO Tree generated for the set of 100 genes received from WebQTL



E. Tissue expression bar chart for the set of 100 genes

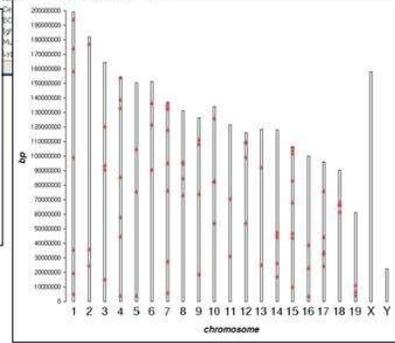


F. Genes expressing in brain. One brain enriched gene is serine/threonine kinase 11 (Stk11)



G. Tissue expression bar chart for Stk11. This gene is over-represented in brain, cerebrum and nervous while under-represented in lymphoreticular and head/neck

H. Chromosome distribution of the set of 100 genes



D. Enriched DAG. Interestingly, two major enriched areas are "neurotransmitter metabolism" and "neutral lipid metabolism"

I. BioCarta table and map

BioCarta map

MAPKinase Signaling Pathway

Cell Cycle: G2/M Checkpoint

Growth Hormone Signaling Pathway

Regulation of BAD phosphorylation

Erk and PI-3 Kinase Are Necessary for Collagen Binding in Corneal

Gene number	Locus link IDs
2	13163 20111
1	20111
1	20111
1	20111
1	18643