THE OHIO STATE UNIVERSITY
UNIVERSITY MEDICAL CENTER

Biomedical Informatics

3184 Graves Hall 333 W. 10th Ave. Columbus, OH 43210 614-292-4778 (phone) 614-292-7659 (fax)

Statistical Microarry Data Analysis

Review of Microarray

Elements of Gene Expression Data Analysis

- Comparative study
- Clustering

Introduction to Pathway and Gene Ontology Enrichment Analysis

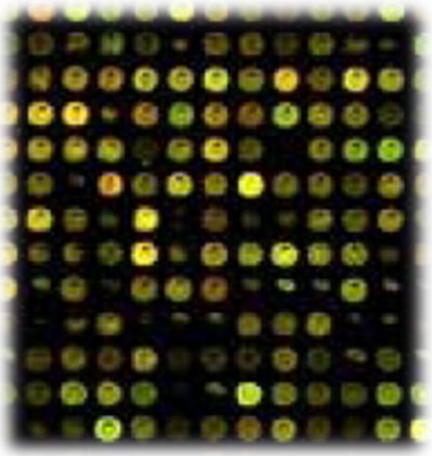
Review of Microarray

Elements of Gene Expression Data Analysis

- Comparative study
- Clustering

Introduction to Pathway and Gene Ontology Enrichment Analysis

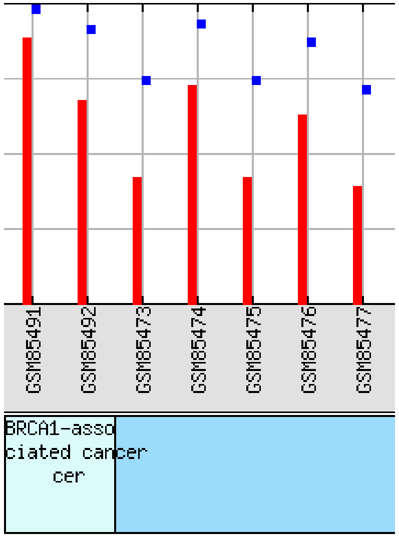
Representation



10	500	30	100
20	28	7	42
53	11	10	40
1	1000	200	51

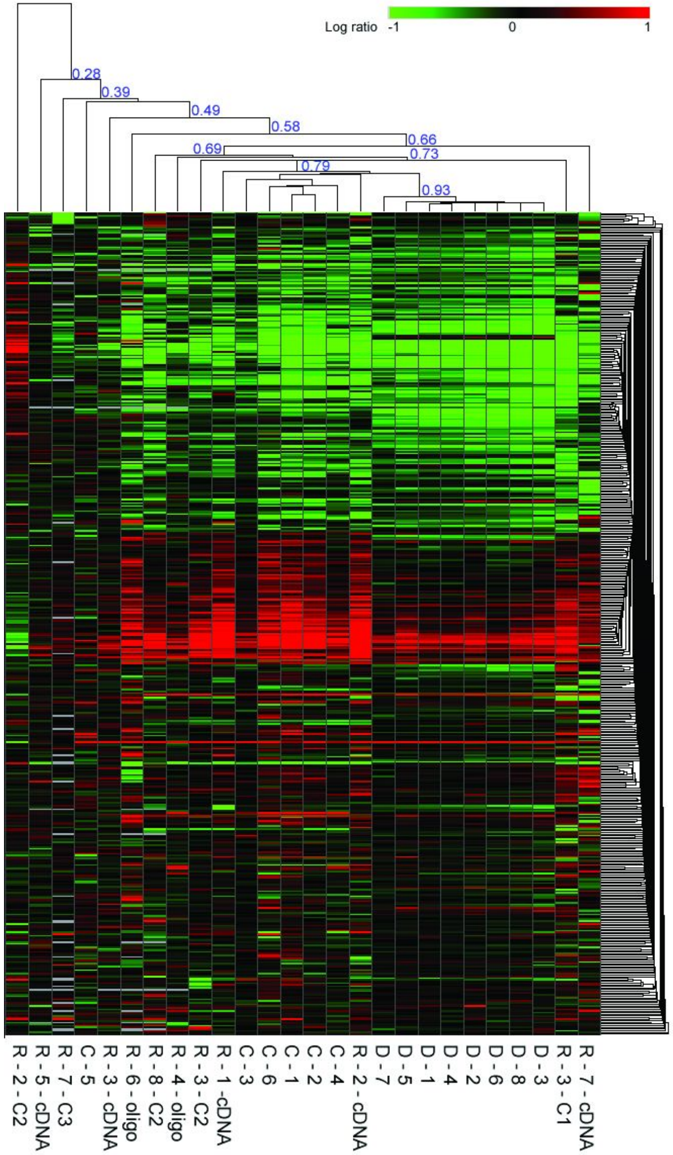
Gene 1	10
Gene 2	500
Gene 3	30
Gene 4	100
Gene 5	20
Gene 6	28
Gene 7	7
…	…

Gene 1	10	50	60	…
Gene 2	500	400	800	…
Gene 3	30	38	35	…
Gene 4	100	107	120	…
Gene 5	20	50	70	…
Gene 6	28	42	33	…
Gene 7	7	15	8	…
…	…	…	…	…



How do we use microarray?

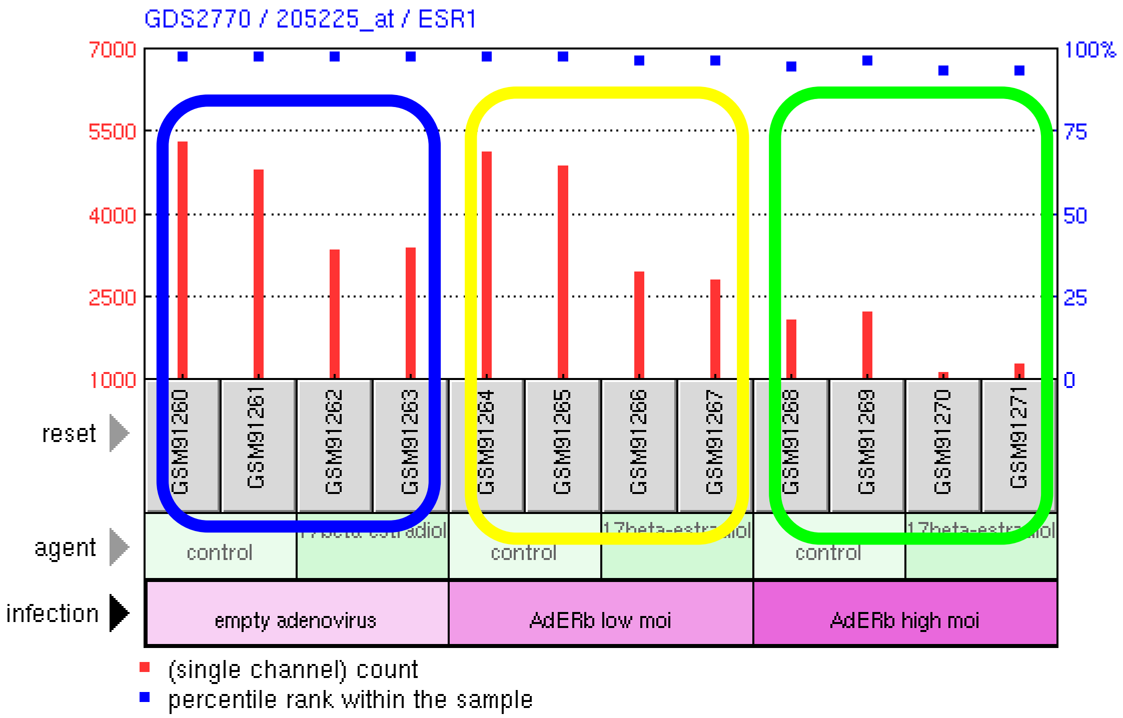
- Profiling
- Comparative study
- Clustering
- Inference



Supplementary Figure 1: Clustering of laboratory/platform combinations using log ratio values of commone genes

Hypothesis Testing

- Two set of samples sampled from two distributions (N=2)



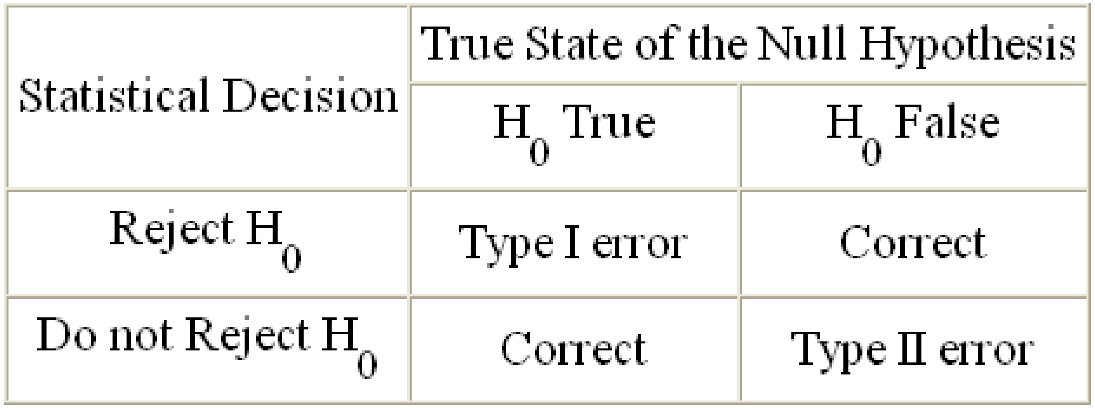
Hypothesis Testing

- Two set of samples sampled from two distributions (N=2)
- Hypothesis

$H_0 : \mu_1 - \mu_2 = 0$ Null hypothesis

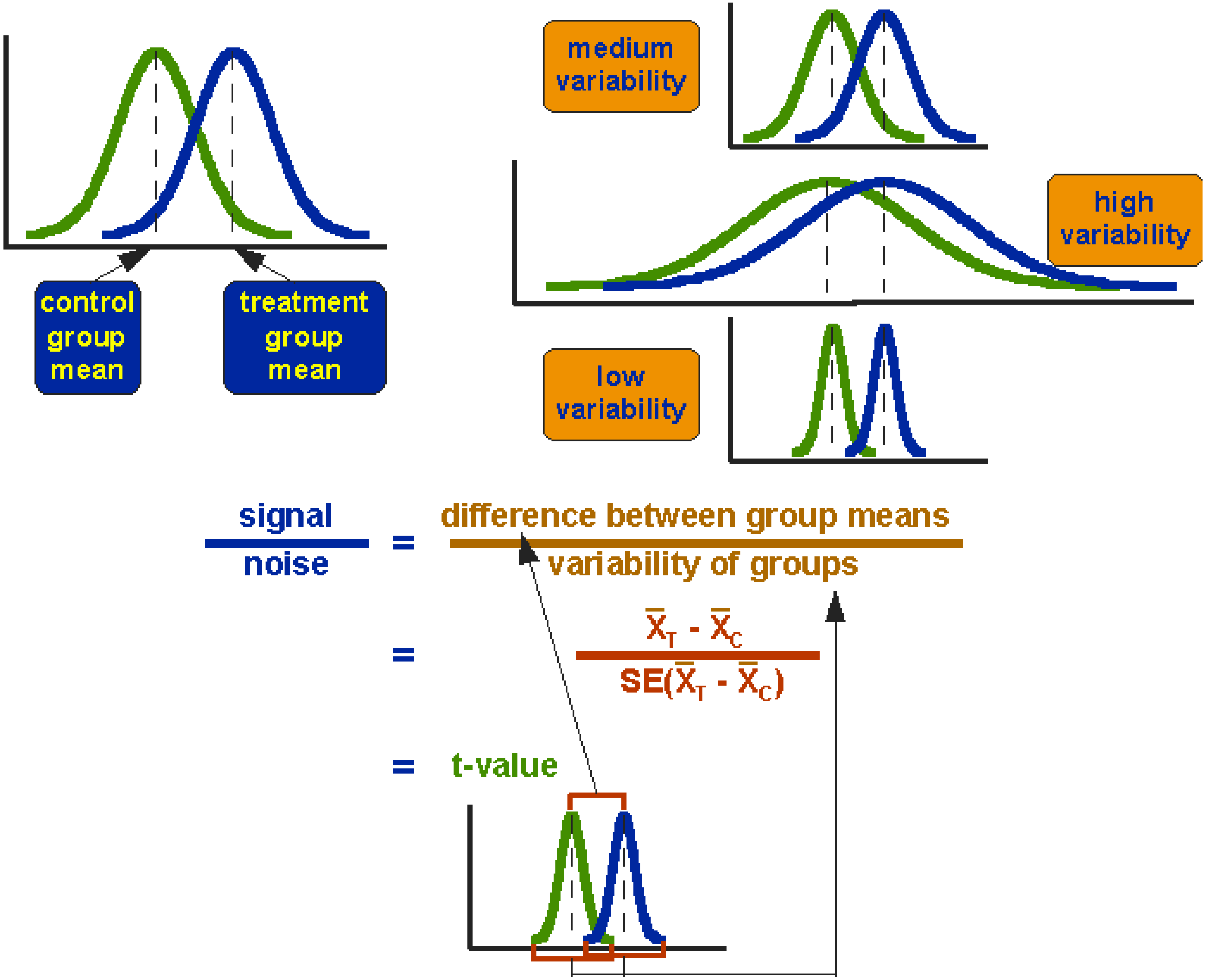
$H_1 : \mu_1 - \mu_2 \neq 0$ Alternative hypothesis

μ_1 and μ_2 are the means of the two distributions.



Statistical Decision	True State of the Null Hypothesis	
	H_0 True	H_0 False
Reject H_0	Type I error	Correct
Do not Reject H_0	Correct	Type II error

Student’s t-test

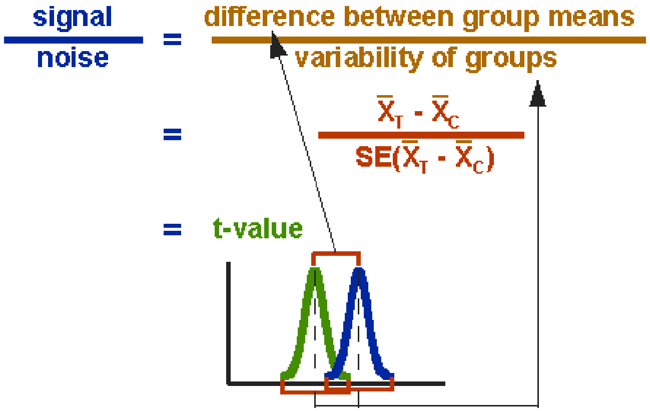


$$\frac{\text{signal}}{\text{noise}} = \frac{\text{difference between group means}}{\text{variability of groups}}$$

$$= \frac{\bar{X}_T - \bar{X}_C}{SE(\bar{X}_T - \bar{X}_C)}$$

$$= \text{t-value}$$

Student’s t-test



$$\frac{\text{signal}}{\text{noise}} = \frac{\text{difference between group means}}{\text{variability of groups}} = \frac{\bar{X}_T - \bar{X}_C}{SE(\bar{X}_T - \bar{X}_C)} = \text{t-value}$$

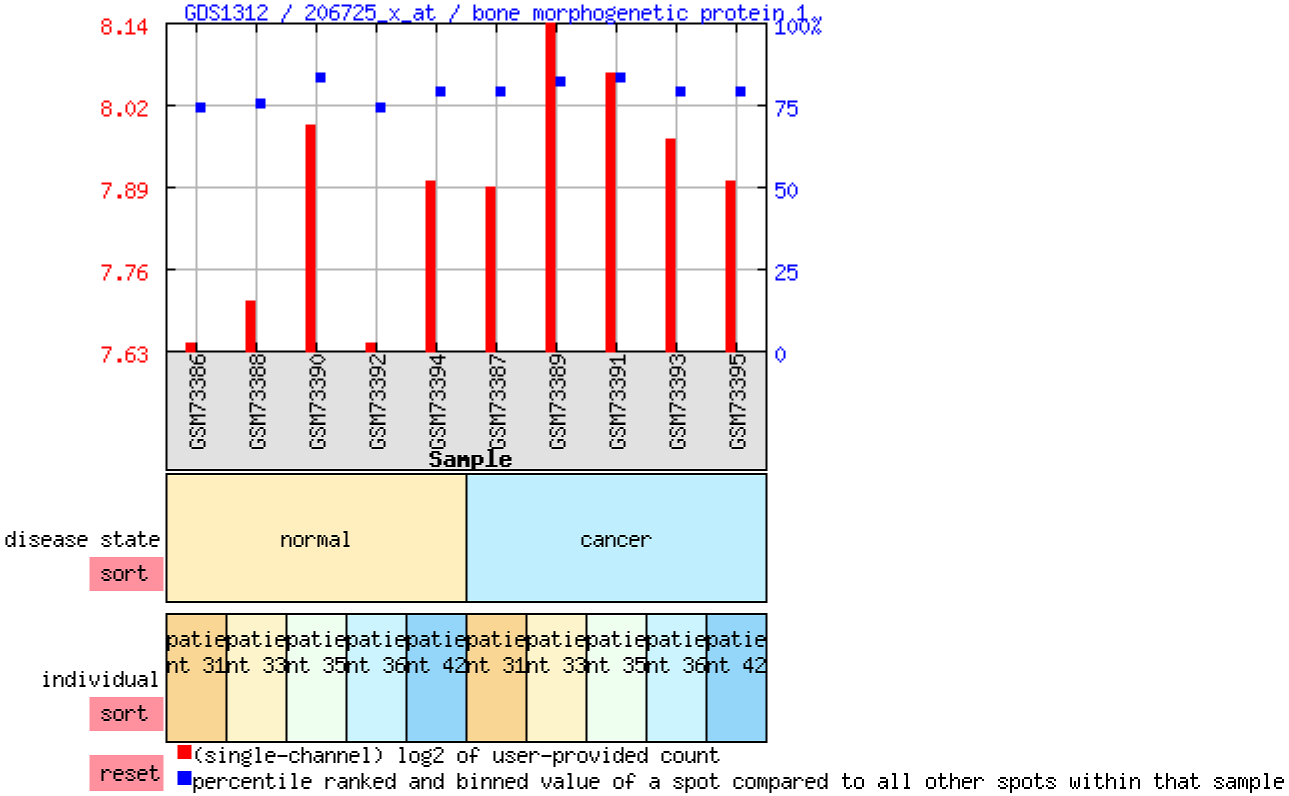
$$t = \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{var_T}{n_T} + \frac{var_C}{n_C}}}$$

p-value can be computed from t-value and number of freedom (related to number of samples) to give a bound on the probability for type-I error (claiming insignificant difference to be significant) assuming normal distributions.

Student’s t-test

- Dependent (paired) t-test

$$t = \frac{\sum D_i}{\sqrt{\frac{n\sum D_i^2 - (\sum D_i)^2}{n-1}}} \qquad D_i = X_2^i - X_1^i$$



Permutation (t-)test

T-test relies on the parametric distribution assumption (normal distribution). Permutation tests do not depend on such an assumption. Examples include the permutation t-test and Wilcoxon rank-sum test.

Perform regular t-test to obtain t-value t_0. The randomly permute the N_1+N_2 samples and designate the first N_1 as group 1 with the rest being group 2. Perform t-test again and record the t-value t. For all possible $K = \binom{N_1+N_2}{N_1}$ permutations, count how many t-values are larger than t_0 and write down the number K_0.

$$p = \frac{1+K_0}{1+K}$$

Multiple Classes (N>2)

F-test

- The null hypothesis is that the distribution of gene expression is the same for all classes.
- The alternative hypothesis is that at least one of the classes has a distribution that is different from the other classes.
- Which class is different cannot be determined in F-test (ANOVA). It can only be identified post hoc.

Example

- GEO Dataset Subgroup Effect

Gene Discovery and Multiple T-tests

Controlling False Positives

- p-value cutoff = 0.05 (probability for false positive - type-I error)
- 22,000 probesets
- False discovery 22,000X0.05=1,100
- Focus on the 1,100 genes in the second speciman. False discovery 1,100X0.05 = 55

Gene Discovery and Multiple T-tests

Controlling False Positives

- State the set of genes explicitly before the experiments
    - Problem: not always feasible, defeat the purpose of large scale screening, could miss important discovery
- Statistical tests to control the false positives

Gene Discovery and Multiple T-tests

Controlling False Positives

- Statistical tests to control the false positives
  - Controlling for no false positives (very stringent, e.g. Bonferroni methods)
  - Controlling the number of false positives (
  - Controlling the proportion of false positives
  - Note that in the screening stage, false positive is better than false negative as the later means missing of possibly important discovery.

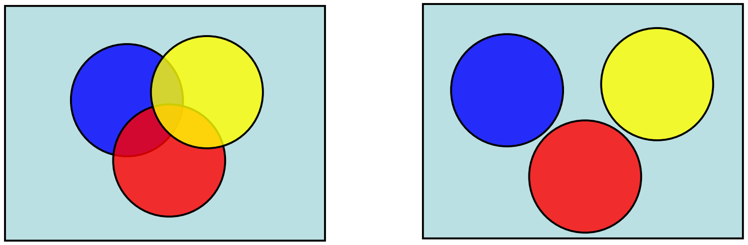
Gene Discovery and Multiple T-tests

Controlling False Positives

- Statistical tests to control the false positives
  - Controlling for no false positives (very stringent)
  - Bonferroni methods and multivariate permutation methods

Bonferroni inequality

$$Prob(E_1 \cup E_2 \cup \cdots \cup E_K) \leq \sum_{i=1}^{K} Prob(E_i)$$



Area of union < Sum of areas

$$Prob(E_i) = 0.05, K = 20$$

$$Prob(E_1 \cup E_2 \cup \cdots \cup E_K) \leq \sum_{i=1}^{K} Prob(E_i) = 1$$

Gene Discovery and Multiple T-tests

Bonferroni methods

- Bonferroni adjustment

$$Prob(E_i) = 0.05, K = 20$$

$$Prob(E_1 \cup E_2 \cup \cdots \cup E_K) \leq \sum_{i=1}^{K} Prob(E_i) = 1$$

- If E_i is the event for false positive discovery of gene I, conservative speaking, it is almost guaranteed to have false positive for K > 19.
- So change the p-value cutoff line from $\mathbf{p_0}$ to $\mathbf{p_0/K}$. This is called ***Bonferroni adjustment.***
- If K=20, p_0=0.05, we call a gene i is significantly differentially expressed if pi<0.0025.

Gene Discovery and Multiple T-tests

Bonferroni methods

- Bonferroni adjustment
- Too conservative. Excessive stringency leads to increased false negative (type II error).
- Has problem with metaanalysis.
- Variations: sequential Bonferroni test (Holm-Bonferroni test)
  - Sort the K p-values from small to large to get $p_1 \leq p_2 \leq \ldots \leq p_{K.}$
  - So change the p-value cutoff line for the *i*th p-value to be $\mathbf{p_0/(K-i+1)}$ (ie, $p_1 \leq p_0/K$, $p_2 \leq p_0/(K-1)$, …, $p_K \leq p_0$.
  - If $p_j \leq \mathbf{p_0/(K-j+1)}$ for all $j \leq i$ but $p_{i+1} > \mathbf{p_0/(K-i+1+1)}$**,** reject all the alternative hypothesis from i+1 to K, but keep the hypothesis from 1 to i.

Gene Discovery and Multiple T-tests

Controlling False Positives

- Statistical tests to control the false positives
  - Controlling the number of false positives
    - Simple approach – choose a cutoff for p-values that are lower than the usual 0.05 but higher than that from Bonferroni adjustment
    - More sophisticated way: a version of multivariate permutation.

Gene Discovery and Multiple T-tests

Controlling False Positives

- Statistical tests to control the false positives
  - Controlling the proportion of false positives

Let γ be the portion (percentage) of false positive in the total discovered genes.

$$p_1 \leq p_2 \leq \cdots \leq p_D \leq \cdots \leq p_K$$

$$D = \arg\max(\underbrace{p_i \cdot K}_{\text{False positive}} / \underset{\substack{|\\ \text{Total positive}}}{i} < \gamma)$$

p_D is the choice. There are other ways for estimating false positives. Details can be found in Tusher et. al. PNAS 98:5116-5121.

Review of Microarray

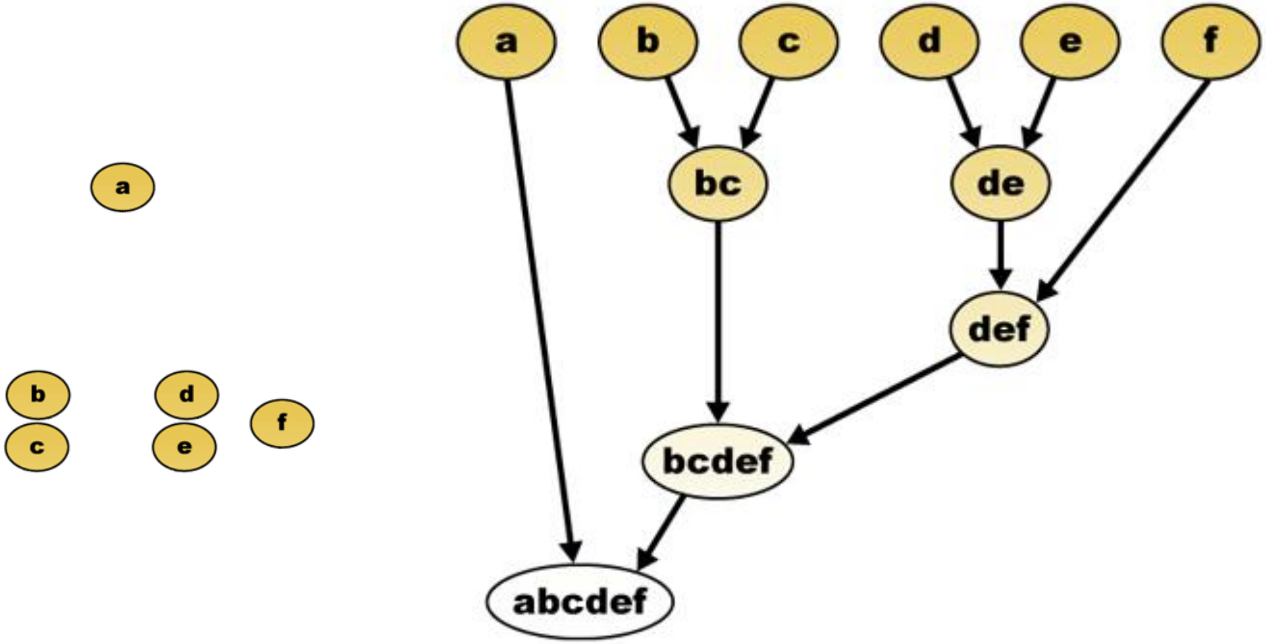
Elements of Gene Expression Data Analysis

- Comparative study
- Clustering

Introduction to Pathway and Gene Ontology Enrichment Analysis

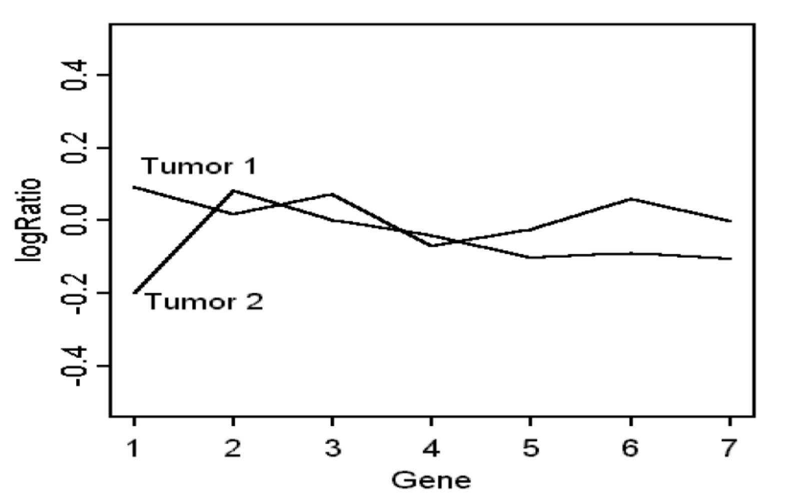
How do we process microarray data (clustering)?

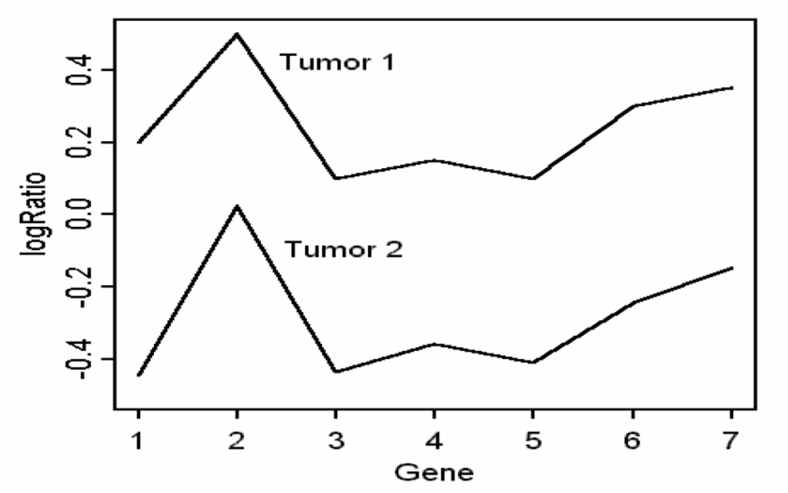
–Unsupervised Learning – Hierarchical Clustering



Distance Measure (Metric?)

- What do you mean by “similar”?
- Euclidean
- Uncentered correlation
- Pearson correlation





Distance Metric

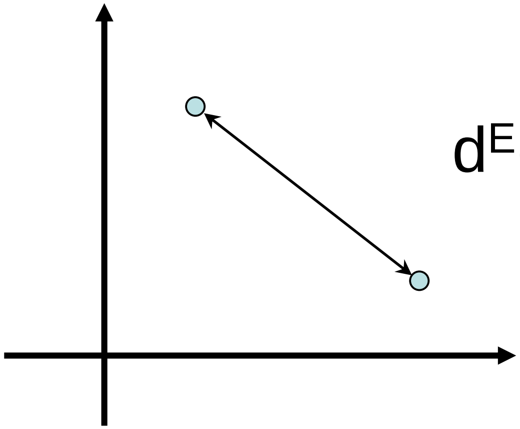
- Euclidean

$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$$

$$\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^T$$

$$d^E(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots (x_n - y_n)^2}$$

102123_at	Lip1	1596.000	2040.900	1277.000	4090.500	1357.600	1039.200	1387.300	3189.000	1321.300	2164.400	868.600	185.300	266.400	2527.800
160552_at	Ap1s1	4144.400	3986.900	3083.100	6105.900	3245.800	4468.400	7295.000	5410.900	3162.100	4100.900	4603.200	6066.200	5505.800	5702.700



d^E(Lip1, Ap1s1) = 12883

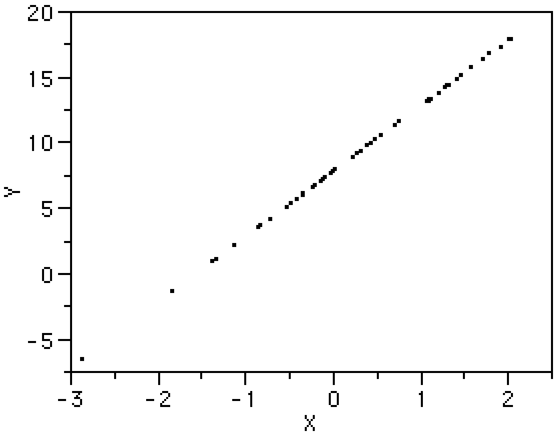
Distance Metric

- Pearson Correlation

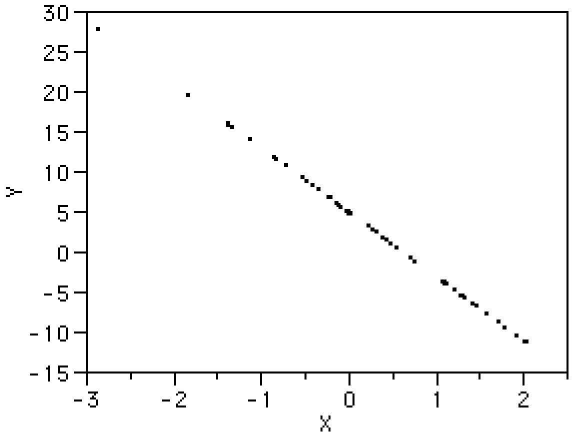
$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$$
$$\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^T$$

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Ranges from 1 to -1.



r = 1

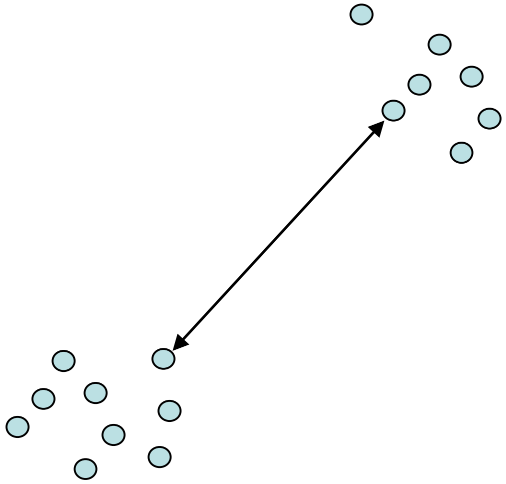


r = -1

How do we process microarray data (clustering)?

–Unsupervised Learning – Hierarchical Clustering

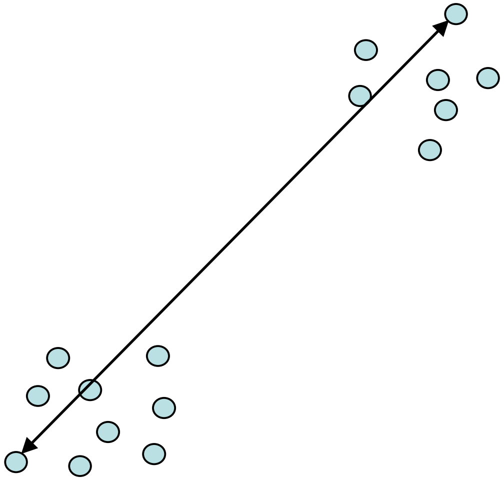
Single linkage: The linking distance is the minimum distance between two clusters.



How do we process microarray data (clustering)?

–Unsupervised Learning – Hierarchical Clustering

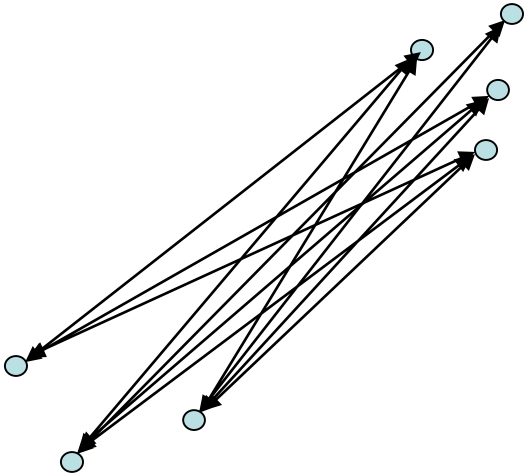
Complete linkage: The linking distance is the maximum distance between two clusters.



How do we process microarray data (clustering)?

–Unsupervised Learning – Hierarchical Clustering

Average linkage/UPGMA: The linking distance is the average of all pair-wise distances between members of the two clusters. Since all genes and samples carry equal weight, the linkage is an Unweighted Pair Group Method with Arithmetic Means (UPGMA).



Review of Microarray

Elements of Gene Expression Data Analysis

- Comparative study
- Clustering

Introduction to Pathway and Gene Ontology Enrichment Analysis

Where do I get the gene list?

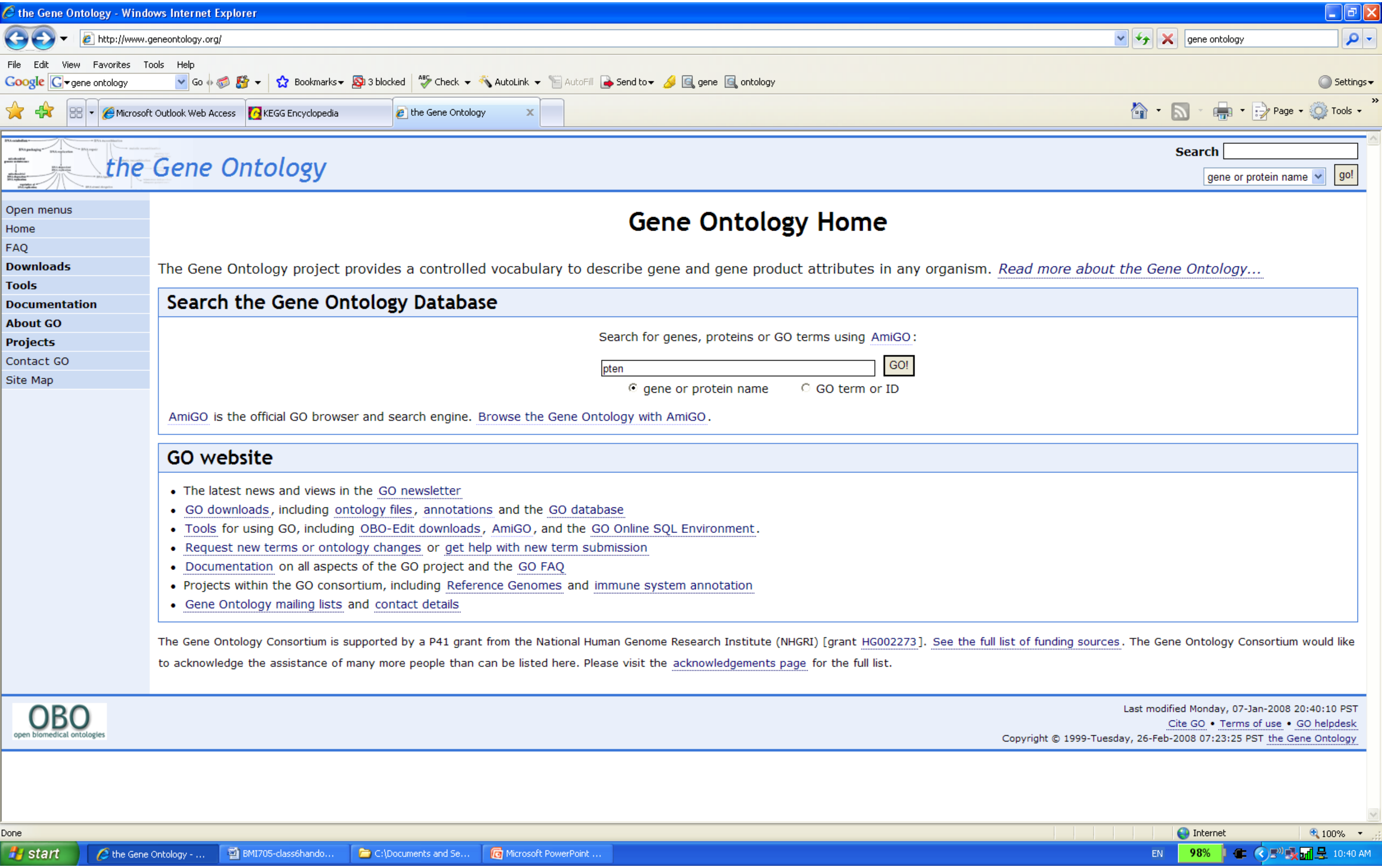
- Comparative study
  - ➢e.g., microarray experiments between two types of samples or two disease states (can also be from RT-PCA, proteomics, …)
- Clustering / classification of genes
  - ➢e.g., co-expressed genes
- Homologue analysis
  - ➢e.g., genes from BLAST
- Other sources

What do I do with the gene list – *enrichment analysis*?

- Find commonality among the gene
  - ➢Common molecular functions (GO)
  - ➢Common biological processes (GO)
  - ➢Common cellular components (GO)

  GO enrichment analysis

  - ➢Common pathways
  - ➢Interact with common genes
  - ➢Common sequences / molecular structures
  - ➢Regulated by common Transcription Factors
  - ➢Targeted by common microRNAs
  - ➢Involved in the same disease
  - ➢…
- Generate new hypothesis based on the commonality



the Gene Ontology

Search

gene or protein name

go!

- Open menus
- Home
- FAQ
- **Downloads**
- **Tools**
- **Documentation**
- **About GO**
- **Projects**
- Contact GO
- Site Map

Gene Ontology Home

The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism. *Read more about the Gene Ontology...*

Search the Gene Ontology Database

Search for genes, proteins or GO terms using AmiGO:

pten GO!

● gene or protein name ○ GO term or ID

AmiGO is the official GO browser and search engine. Browse the Gene Ontology with AmiGO.

GO website

- The latest news and views in the GO newsletter
- GO downloads, including ontology files, annotations and the GO database
- Tools for using GO, including OBO-Edit downloads, AmiGO, and the GO Online SQL Environment.
- Request new terms or ontology changes or get help with new term submission
- Documentation on all aspects of the GO project and the GO FAQ
- Projects within the GO consortium, including Reference Genomes and immune system annotation
- Gene Ontology mailing lists and contact details

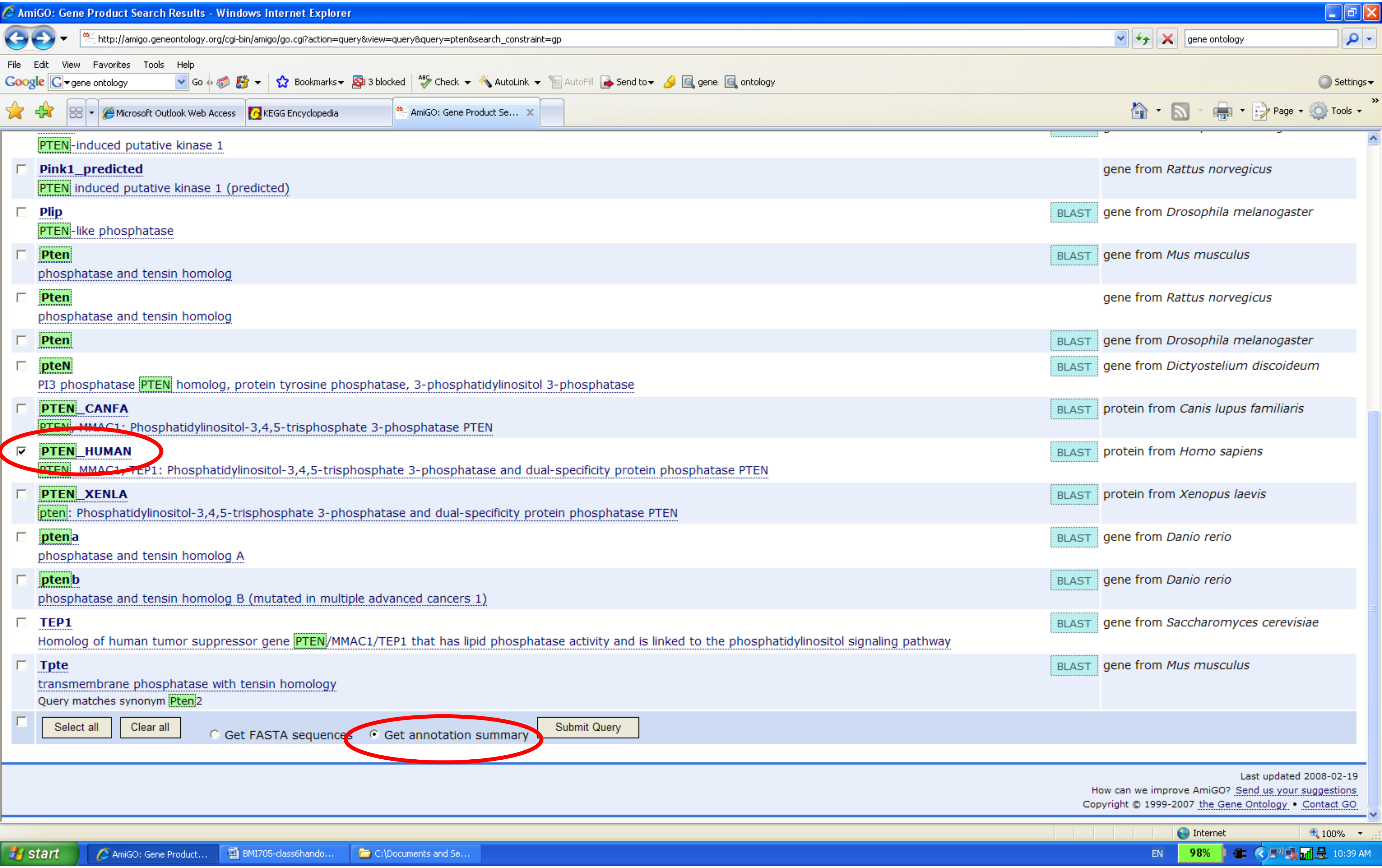
The Gene Ontology Consortium is supported by a P41 grant from the National Human Genome Research Institute (NHGRI) [grant HG002273]. See the full list of funding sources. The Gene Ontology Consortium would like to acknowledge the assistance of many more people than can be listed here. Please visit the acknowledgements page for the full list.

OBO open biomedical ontologies

Last modified Monday, 07-Jan-2008 20:40:10 PST

Cite GO • Terms of use • GO helpdesk

Copyright © 1999-Tuesday, 26-Feb-2008 07:23:25 PST the Gene Ontology



AmiGO: Gene Product Search Results - Windows Internet Explorer
File Edit View Favorites Tools Help
Google gene ontology Go Bookmarks 3 blocked Check AutoLink AutoFill Send to gene ontology Settings
Microsoft Outlook Web Access
KEGG Encyclopedia
AmiGO: Gene Product Se...
Page Tools
Internet 100%
start AmiGO: Gene Product... BMI705-class6hando... C:\Documents and Se...
EN 98% 10:39 AM

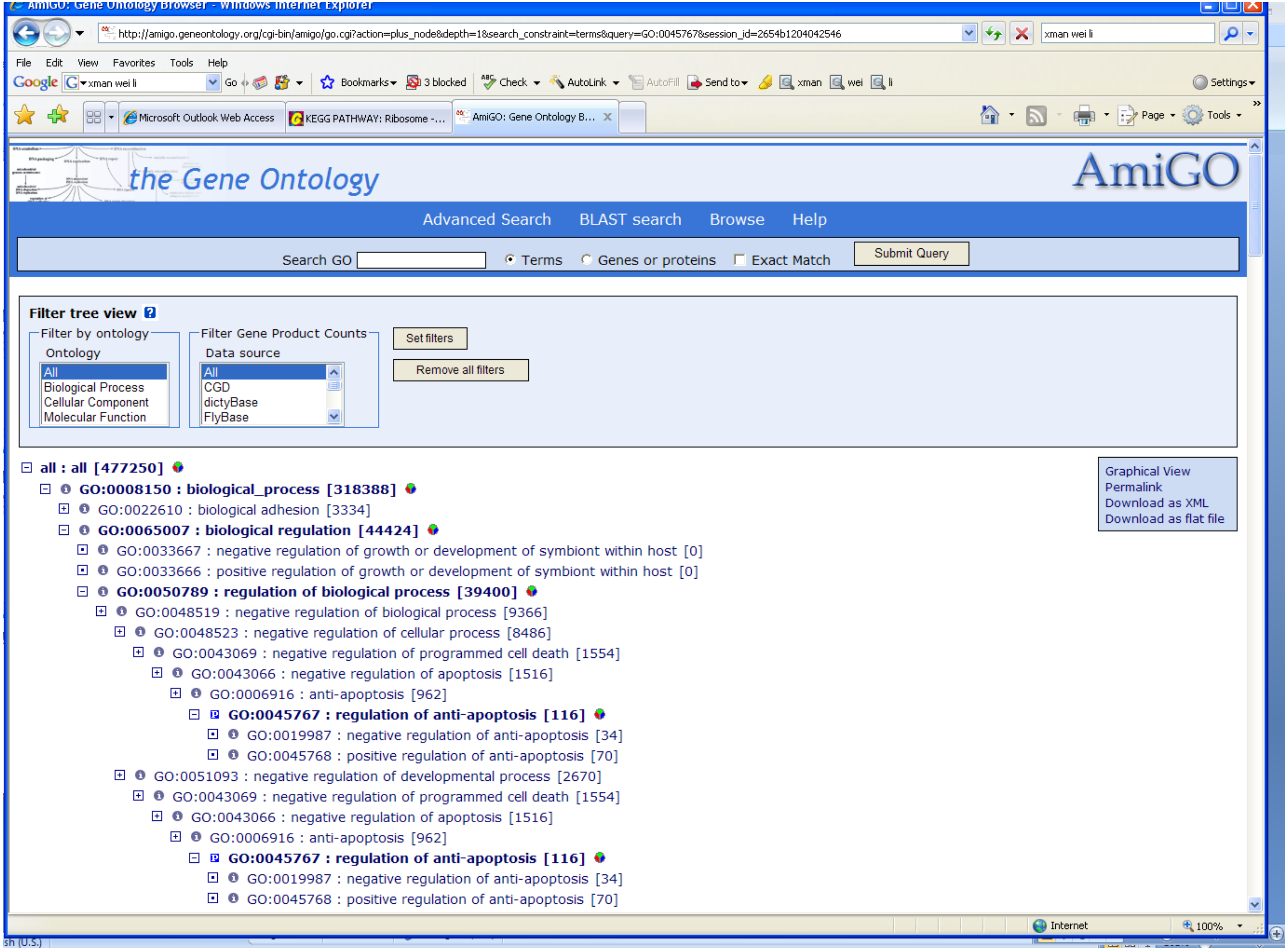
http://amigo.geneontology.org/cgi-bin/amigo/go.cgi?action=query&view=query&query=pten&search_constraint=gp

gene ontology

	PTEN-induced putative kinase 1		
□	**Pink1_predicted** PTEN induced putative kinase 1 (predicted)		gene from *Rattus norvegicus*
□	**Plip** PTEN-like phosphatase	BLAST	gene from *Drosophila melanogaster*
□	**Pten** phosphatase and tensin homolog	BLAST	gene from *Mus musculus*
□	**Pten** phosphatase and tensin homolog		gene from *Rattus norvegicus*
□	**Pten**	BLAST	gene from *Drosophila melanogaster*
□	**pteN** PI3 phosphatase PTEN homolog, protein tyrosine phosphatase, 3-phosphatidylinositol 3-phosphatase	BLAST	gene from *Dictyostelium discoideum*
□	**PTEN_CANFA** PTEN, MMAC1: Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase PTEN	BLAST	protein from *Canis lupus familiaris*
☑	**PTEN_HUMAN** PTEN, MMAC1, TEP1: Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN	BLAST	protein from *Homo sapiens*
□	**PTEN_XENLA** pten: Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN	BLAST	protein from *Xenopus laevis*
□	**ptena** phosphatase and tensin homolog A	BLAST	gene from *Danio rerio*
□	**ptenb** phosphatase and tensin homolog B (mutated in multiple advanced cancers 1)	BLAST	gene from *Danio rerio*
□	**TEP1** Homolog of human tumor suppressor gene PTEN/MMAC1/TEP1 that has lipid phosphatase activity and is linked to the phosphatidylinositol signaling pathway	BLAST	gene from *Saccharomyces cerevisiae*
□	**Tpte** transmembrane phosphatase with tensin homology Query matches synonym Pten2	BLAST	gene from *Mus musculus*

□ Select all Clear all ○ Get FASTA sequences ● Get annotation summary Submit Query

Last updated 2008-02-19
How can we improve AmiGO? Send us your suggestions
Copyright © 1999-2007 the Gene Ontology • Contact GO


AmiGO: Gene Ontology Browser - Windows Internet Explorer
http://amigo.geneontology.org/cgi-bin/amigo/go.cgi?action=plus_node&depth=1&search_constraint=terms&query=GO:0045767&session_id=2654b1204042546
xman wei li
File Edit View Favorites Tools Help
Google xman wei li Go Bookmarks 3 blocked Check AutoLink AutoFill Send to xman wei li Settings
Microsoft Outlook Web Access
KEGG PATHWAY: Ribosome -...
AmiGO: Gene Ontology B...
Page Tools
the Gene Ontology
AmiGO
Advanced Search BLAST search Browse Help
Submit Query
Set filters
Remove all filters
Graphical View
Permalink
Download as XML
Download as flat file
Internet
100%
sh (U.S.)

Search GO ● Terms ○ Genes or proteins □ Exact Match

Filter tree view

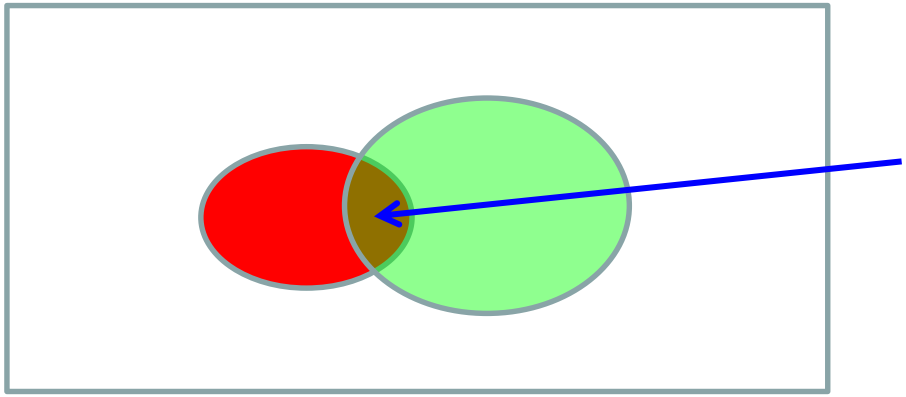
Filter by ontology
Ontology: All (selected), Biological Process, Cellular Component, Molecular Function

Filter Gene Product Counts
Data source: All (selected), CGD, dictyBase, FlyBase

- ⊟ **all : all [477250]**
  - ⊟ **GO:0008150 : biological_process [318388]**
    - ⊞ GO:0022610 : biological adhesion [3334]
    - ⊟ **GO:0065007 : biological regulation [44424]**
      - ⊞ GO:0033667 : negative regulation of growth or development of symbiont within host [0]
      - ⊞ GO:0033666 : positive regulation of growth or development of symbiont within host [0]
      - ⊟ **GO:0050789 : regulation of biological process [39400]**
        - ⊞ GO:0048519 : negative regulation of biological process [9366]
          - ⊞ GO:0048523 : negative regulation of cellular process [8486]
            - ⊞ GO:0043069 : negative regulation of programmed cell death [1554]
              - ⊞ GO:0043066 : negative regulation of apoptosis [1516]
                - ⊞ GO:0006916 : anti-apoptosis [962]
                  - ⊟ **GO:0045767 : regulation of anti-apoptosis [116]**
                    - ⊞ GO:0019987 : negative regulation of anti-apoptosis [34]
                    - ⊞ GO:0045768 : positive regulation of anti-apoptosis [70]
          - ⊞ GO:0051093 : negative regulation of developmental process [2670]
            - ⊞ GO:0043069 : negative regulation of programmed cell death [1554]
              - ⊞ GO:0043066 : negative regulation of apoptosis [1516]
                - ⊞ GO:0006916 : anti-apoptosis [962]
                  - ⊟ **GO:0045767 : regulation of anti-apoptosis [116]**
                    - ⊞ GO:0019987 : negative regulation of anti-apoptosis [34]
                    - ⊞ GO:0045768 : positive regulation of anti-apoptosis [70]

How do I find commonality from my gene list?

- Using a priori knowledge (e.g., gene ontology, pathway, annotation, etc.)
- Fisher’s exact test, hypergeometric test, Bayesian-based methods, etc.

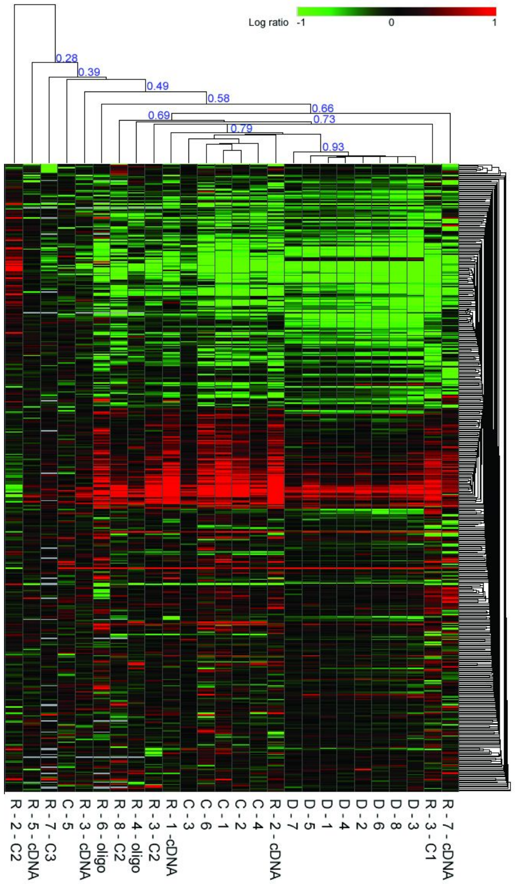


How significant is the intersection?

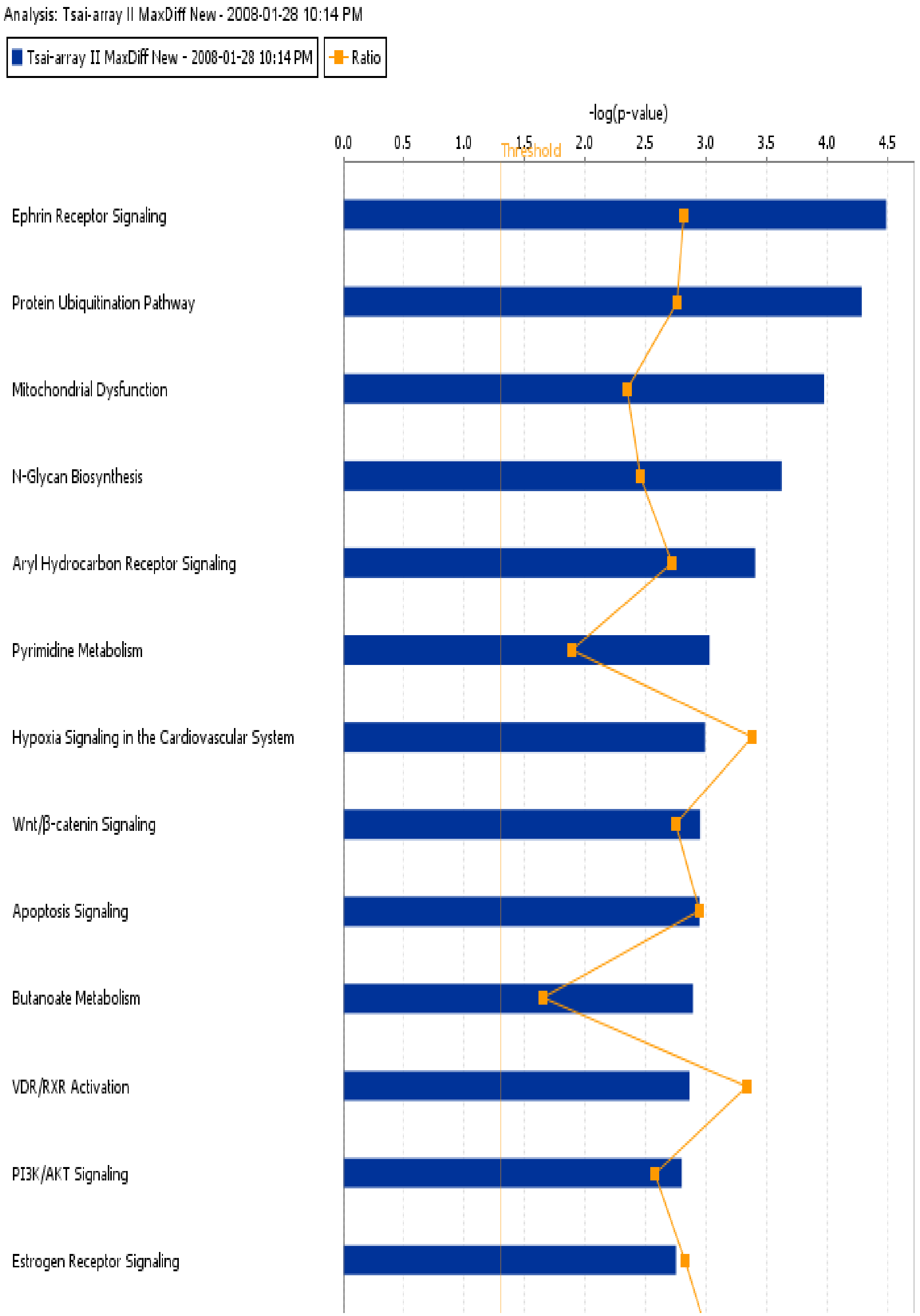
- Good news – most of the time you can use software to do it

What softwares are available?

- DAVID (http://david.abcc.ncifcrf.gov/)
- TOPPGene
- Cytoscape
  - GOTerm
  - BiNGO
- GSEA
- GenMapp (Free)
- Pathway Architect (Commercial)
- Pathway Studio (Commercial)
- Ingenuity Pathway Analysis (Commercial)
  - Manually curated
  - On-demand computation



Supplementary Figure 1: Clustering of laboratory/platform combinations using log ratio values of commone genes

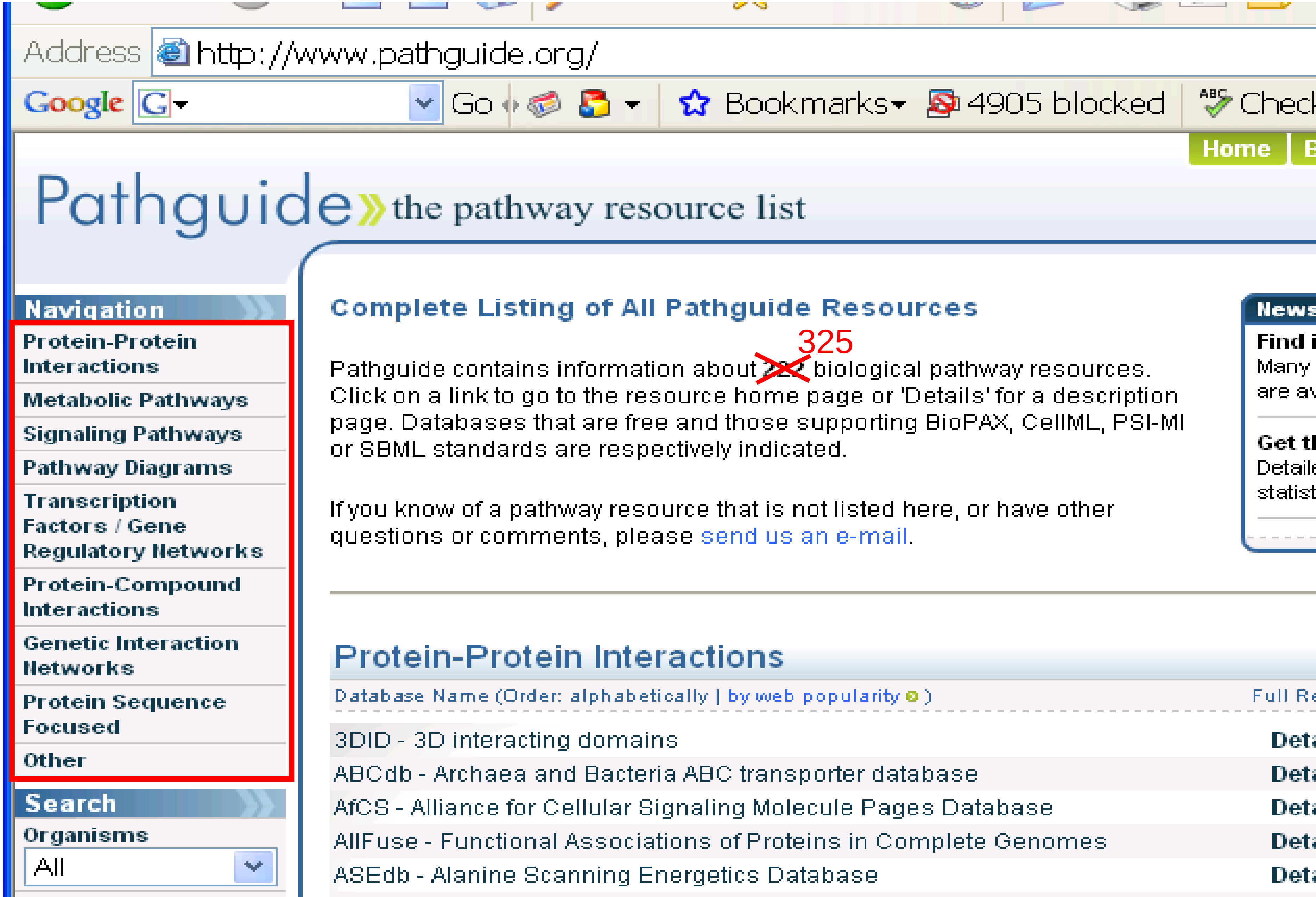


Analysis: Tsai-array II MaxDiff New - 2008-01-28 10:14 PM

Genes

Functions, pathways and networks

Pathway – What's out there?


Address
Google
Go
Bookmarks
4905 blocked
Check
Home
Pathguide » the pathway resource list
Navigation
Protein-Protein Interactions
Metabolic Pathways
Signaling Pathways
Pathway Diagrams
Transcription Factors / Gene Regulatory Networks
Protein-Compound Interactions
Genetic Interaction Networks
Protein Sequence Focused
Other
Search
Organisms
All
News

Address: http://www.pathguide.org/

Complete Listing of All Pathguide Resources

Pathguide contains information about ~~222~~ 325 biological pathway resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

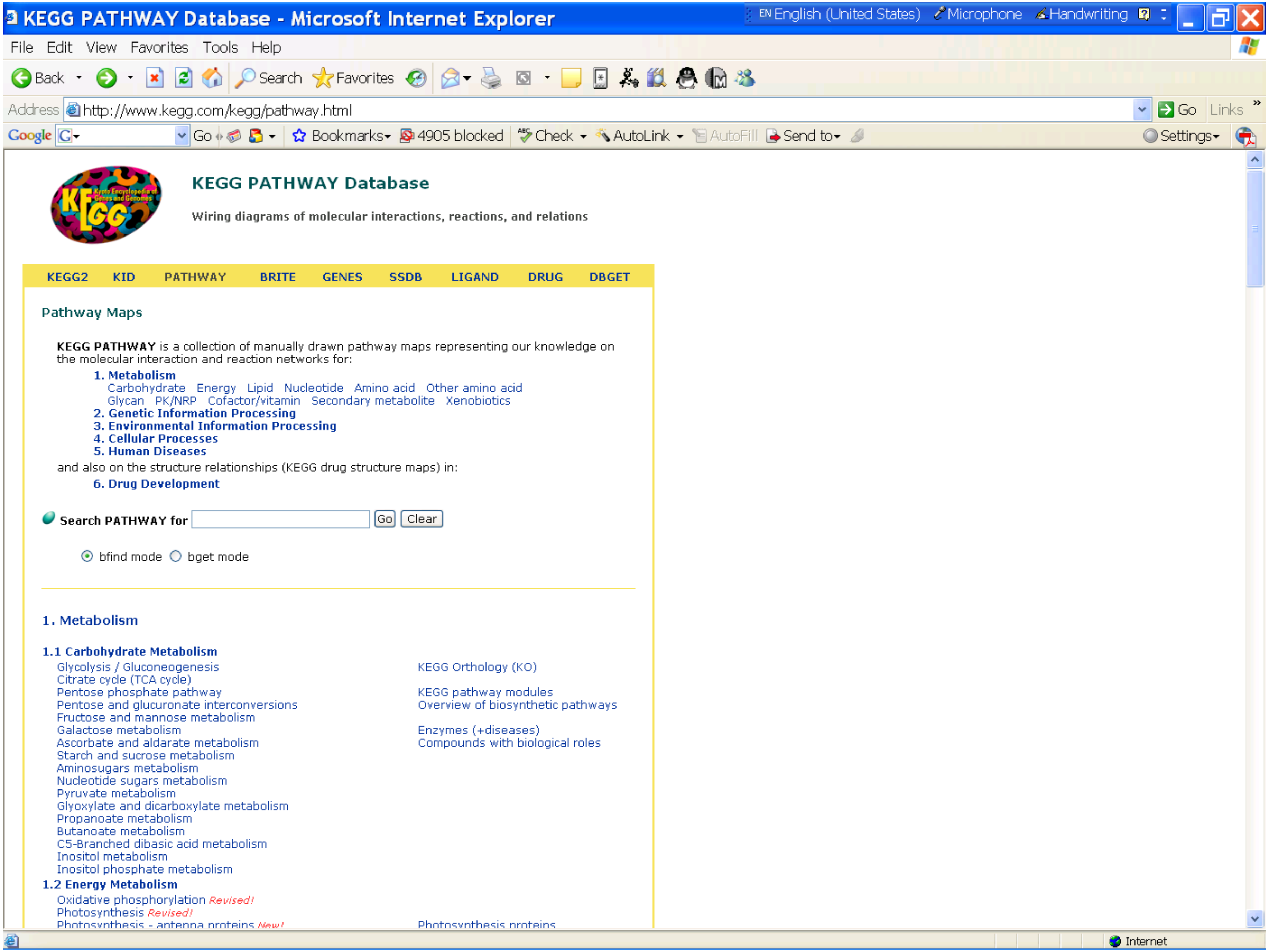
If you know of a pathway resource that is not listed here, or have other questions or comments, please send us an e-mail.

Find i
Many
are av

Get th
Detaile
statist

Protein-Protein Interactions

Database Name (Order: alphabetically | by web popularity)	Full Re
3DID - 3D interacting domains	Deta
ABCdb - Archaea and Bacteria ABC transporter database	Deta
AfCS - Alliance for Cellular Signaling Molecule Pages Database	Deta
AllFuse - Functional Associations of Proteins in Complete Genomes	Deta
ASEdb - Alanine Scanning Energetics Database	Deta


KEGG PATHWAY Database - Microsoft Internet Explorer
EN English (United States) Microphone Handwriting
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.kegg.com/kegg/pathway.html Go Links
Google Go Bookmarks 4905 blocked Check AutoLink AutoFill Send to Settings
Internet

KEGG PATHWAY Database

Wiring diagrams of molecular interactions, reactions, and relations

KEGG2 KID PATHWAY BRITE GENES SSDB LIGAND DRUG DBGET

Pathway Maps

KEGG PATHWAY is a collection of manually drawn pathway maps representing our knowledge on the molecular interaction and reaction networks for:

1. **Metabolism**
   Carbohydrate Energy Lipid Nucleotide Amino acid Other amino acid
   Glycan PK/NRP Cofactor/vitamin Secondary metabolite Xenobiotics
2. **Genetic Information Processing**
3. **Environmental Information Processing**
4. **Cellular Processes**
5. **Human Diseases**

and also on the structure relationships (KEGG drug structure maps) in:

6. **Drug Development**

Search PATHWAY for [ ] Go Clear

● bfind mode ○ bget mode

1. Metabolism

1.1 Carbohydrate Metabolism

Glycolysis / Gluconeogenesis
Citrate cycle (TCA cycle)
Pentose phosphate pathway
Pentose and glucuronate interconversions
Fructose and mannose metabolism
Galactose metabolism
Ascorbate and aldarate metabolism
Starch and sucrose metabolism
Aminosugars metabolism
Nucleotide sugars metabolism
Pyruvate metabolism
Glyoxylate and dicarboxylate metabolism
Propanoate metabolism
Butanoate metabolism
C5-Branched dibasic acid metabolism
Inositol metabolism
Inositol phosphate metabolism

KEGG Orthology (KO)

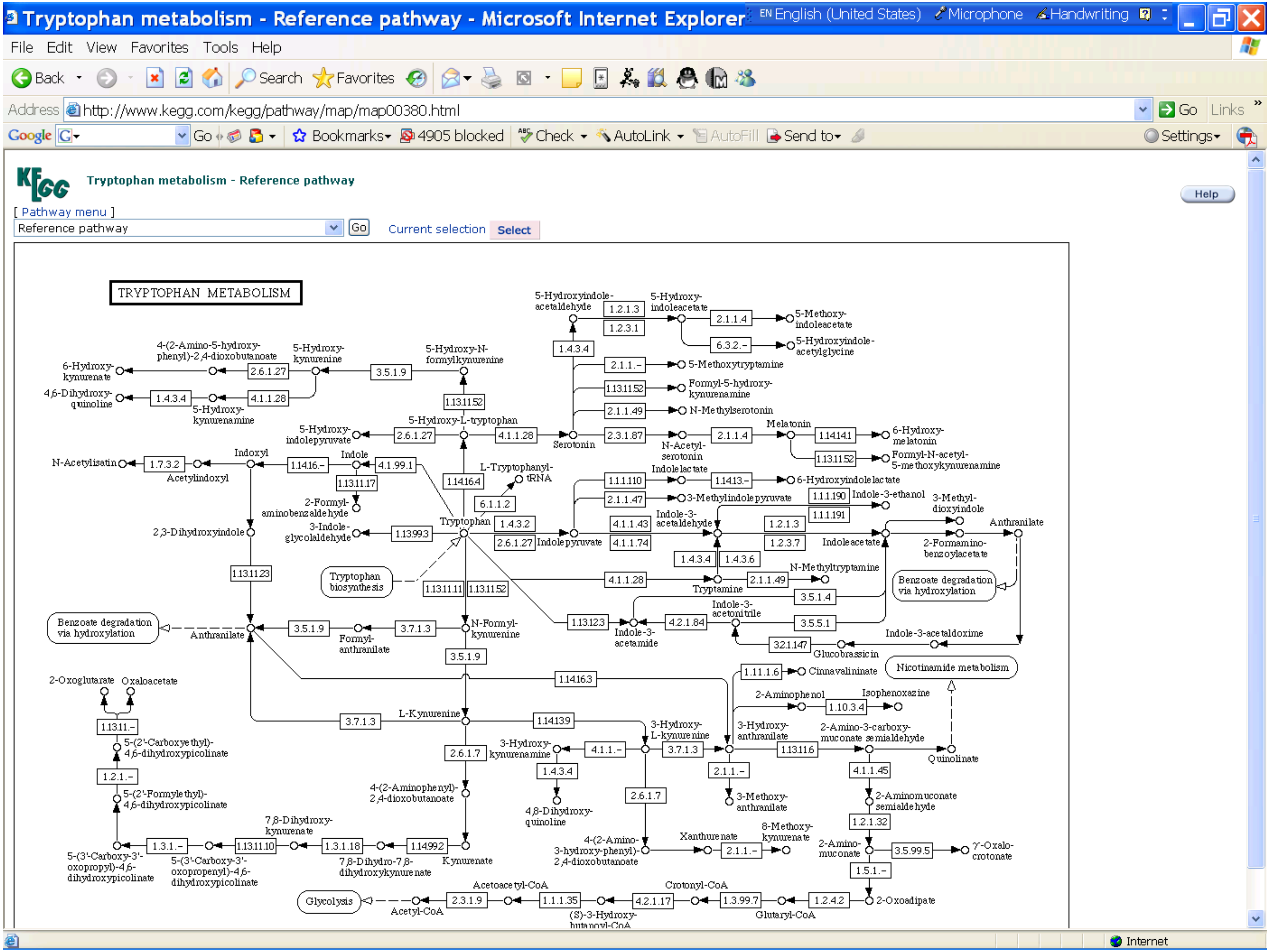
KEGG pathway modules
Overview of biosynthetic pathways

Enzymes (+diseases)
Compounds with biological roles

1.2 Energy Metabolism

Oxidative phosphorylation *Revised!*
Photosynthesis *Revised!*
Photosynthesis - antenna proteins *New!*

Photosynthesis proteins



Tryptophan metabolism - Reference pathway - Microsoft Internet Explorer

Address http://www.kegg.com/kegg/pathway/map/map00380.html

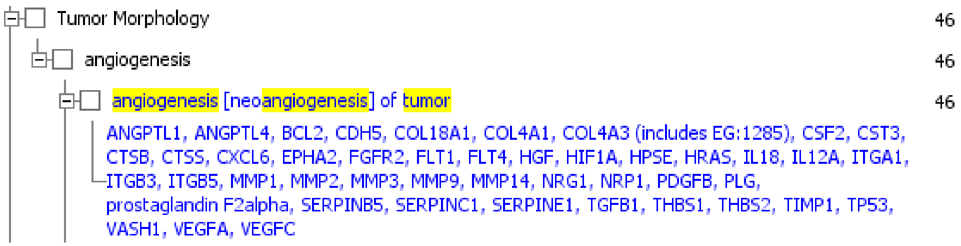
Tryptophan metabolism - Reference pathway

[Pathway menu]

Reference pathway

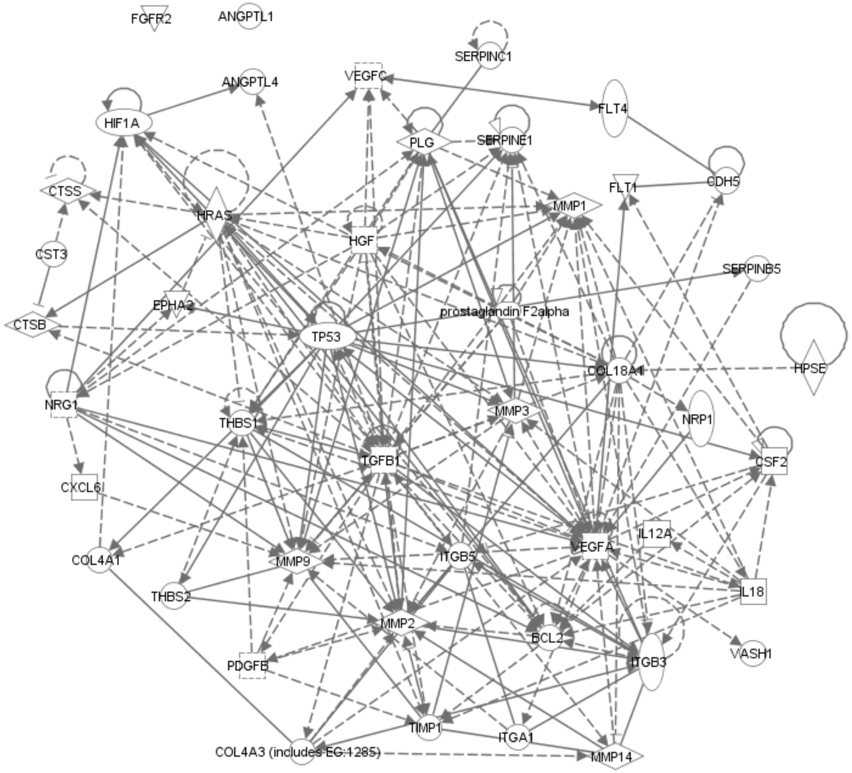
Current selection Select

Ingenuity Pathway Analysis (IPA)

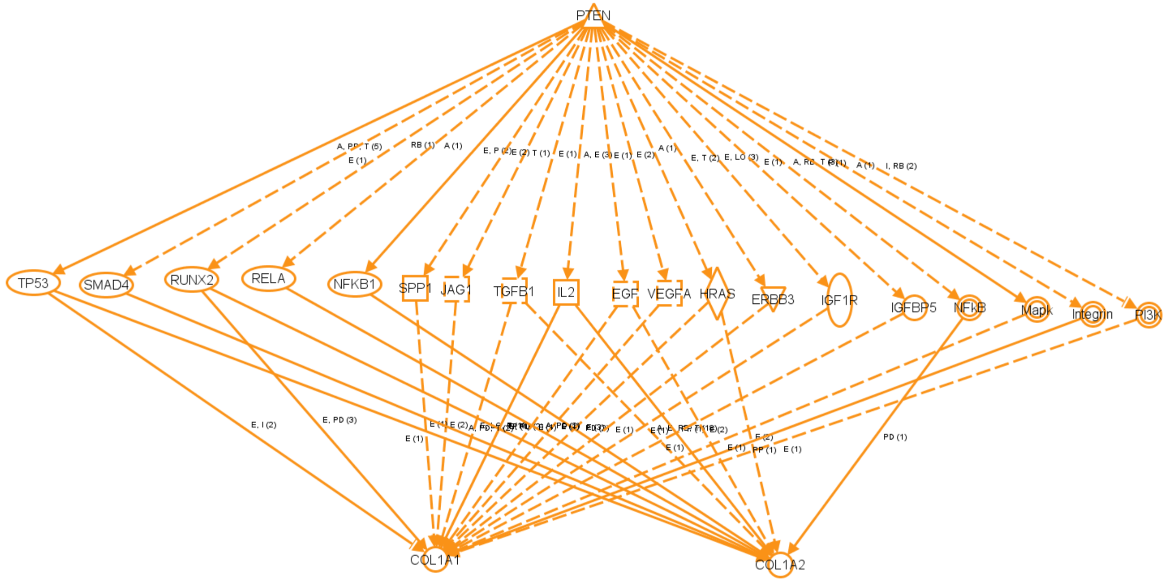


□ Tumor Morphology	46
□ angiogenesis	46
□ angiogenesis [neoangiogenesis] of tumor	46

ANGPTL1, ANGPTL4, BCL2, CDH5, COL18A1, COL4A1, COL4A3 (includes EG:1285), CSF2, CST3, CTSB, CTSS, CXCL6, EPHA2, FGFR2, FLT1, FLT4, HGF, HIF1A, HPSE, HRAS, IL18, IL12A, ITGA1, ITGB3, ITGB5, MMP1, MMP2, MMP3, MMP9, MMP14, NRG1, NRP1, PDGFB, PLG, prostaglandin F2alpha, SERPINB5, SERPINC1, SERPINE1, TGFB1, THBS1, THBS2, TIMP1, TP53, VASH1, VEGFA, VEGFC



New Pathway 6



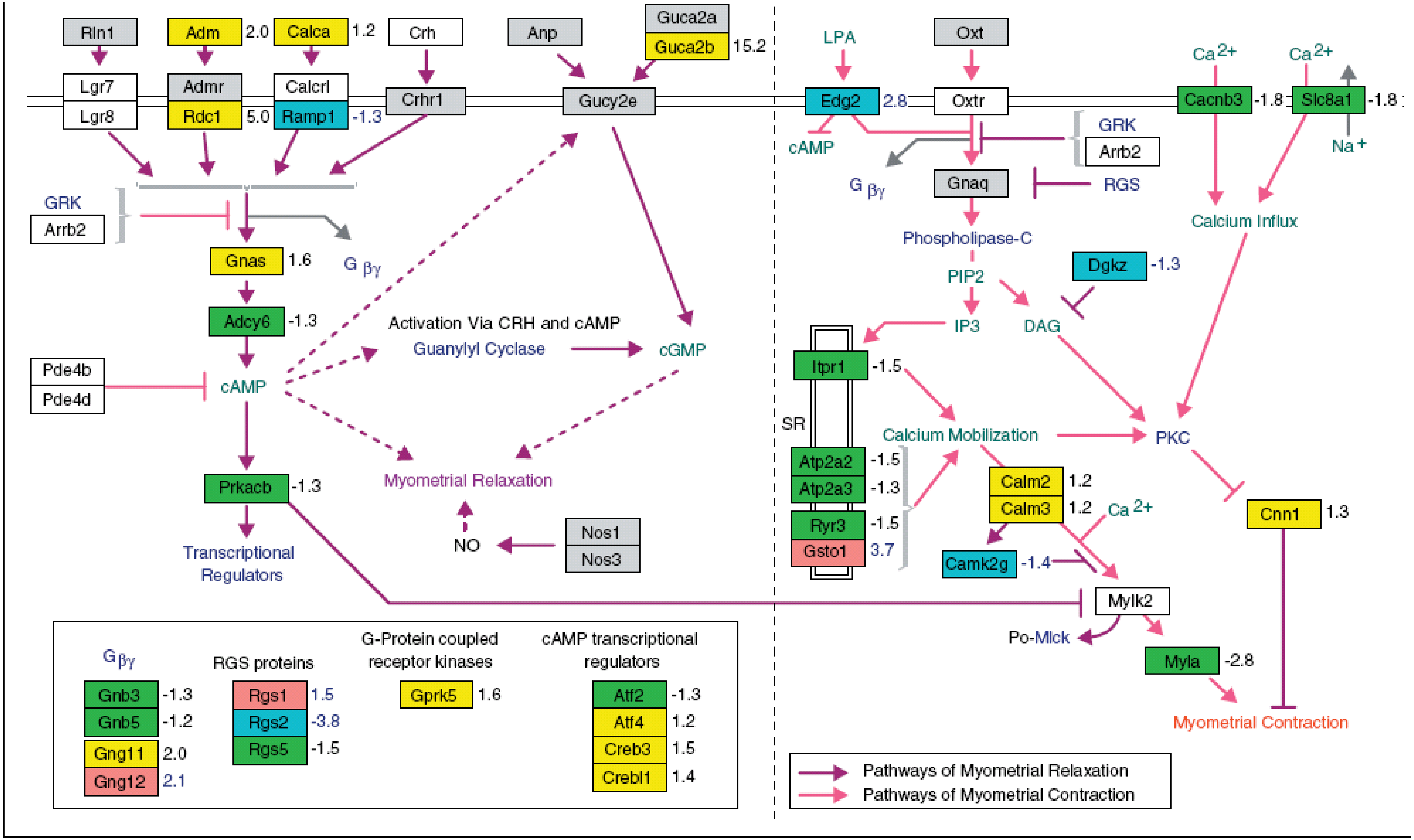


Figure 3

Analysis of pathways of uterine smooth muscle contraction. **(a)** Prostaglandin synthesis and **(b)** G-protein signaling pathways in the myometrium are overlaid with gene-expression color criterion and fold-changes from the program GenMAPP. Interactions suggested by results of this microarray analysis are included in these figures. Detailed gene-expression data, statistics and full gene annotations are available on the GenMAPP interactive version of these pathways online [40].

Demo

- DAVID (http://david.abcc.ncifcrf.gov/)
- TOPPGene
- Ingenuity Pathway Analysis

Gene List1: AURKA BIRC5 ASPM BUB1 CCNA2 CCNB2 CDC2
ACOT7 CDC20 CDC45L CDCA8 CENPE CENPF CEP55 CKS2
CHEK1 DKFZp762E1312 DLG7 DNA2L E2F8 EPR1
FANCI HMMR KIF4A LMNB1 MAD2L1 MELK NCAPG
RANBP1 RRM2 SPAG5 STIL TACC3 TPX2 TRIP13 TTK
UBE2C UBE2S

Gene List2: AI445650 CD2 CCR5 CD247 CD27 CD38 CD3D
CD3E CD3G CD79A CD8A CRTAM CST7 CTSW
CXCR6 DENND2D FAIM3 FMNL1 GZMA GZMB GZMH
GZMK HLA-DOB IL21R IL2RB IL2RG IL7R KLRK1
LAG3 LAT LAX1 MIRN650 NKG7 NM_014792
PTPN7 RASGRP1 RUNX3 SELPLG SEPT6 SERPINB9
SH2D1A SIRPG SLAMF7 SOCS1 TBX21 TRBC1 WAS XCL1
CCL4 XCL2 ZAP70