

Resources for Multiscale Analysis of Cellular Networks and ECM

Yves Lussier

Andrea Califano, Olivier Bodenreider, Eric Newmann





MIND THE GAP

- Accurate Cell Network Computing requires executable models of cell interactions to
 - Cells
 - ECM
 - Environmental factors (biomolecular, xeno/nanotechnologies)



Grand Challenges in Science

What don't we know?"

July 1st, 2005, *Science Journal*

- Question 1.

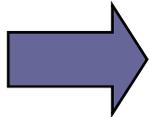
What Is the universe made of?

(...)

- Question 4.

To what extent are genetic variation and personal health linked?

(...)



- Question 17.

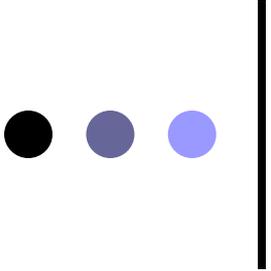
How will big pictures emerge from the sea of biological data?

(...)



More than one Cell Ontology?

- Cell ontogenic classification
- Functional groups
 - E.g. angiogenesis cells
- Disease affected groups (Pathophysiology, morphology)
 - E.g. neovascularisation cells
- Topological Location
- Etc.



Cell Function Classification may require a Directed Acyclic Graph Structure

- Normative

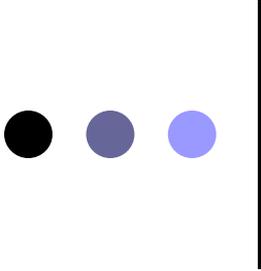
- Description of tissues

- Phylogenetic

- Muscle cells “function” derived from distinct tissues in vertebrates and invertebrates

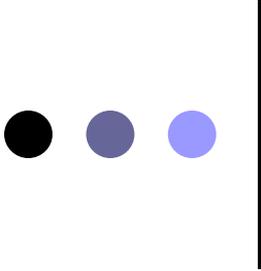
- Aberrant

- E.g. same neoplasms achieved from distinct initial cell types...



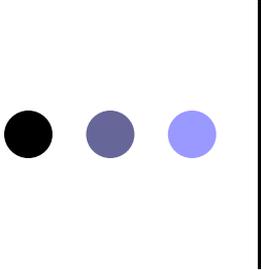
Consider Compositional Models of Cells

- Cell assembly instances
- Different cellular states, e.g., activities, cell communications, gene/prot expression, infectious attacks
- Clonal mutations compared to neighboring cells
- Synaptic basis of memory
- Physical integrity and physical relations (pressures and gradients)
- Various to different functions of same cytoskeletal components (tubulin, actin, myosin) in different contexts: roles of molecules depend on type and role of surrounding cell



Healthcare and Cellular Networks

- Growth / tissue pattern formation
 - e.g. tumorigenesis
 - e.g., gastrulation: normal and aberrant
- correspondence between histological/morphological classifications and possible cellular type
- cell to tissue to system effects
 - liver processing of substances arriving from the portal vein to rest of vascular system
- inflammatory states and responses
 - E.g. infection
- Imaging
 - neuroscience
 - cancer
- Nanotechnology in medicine



National Center for Biomedical Computing: Initiatives in MAGNet

(Multiscale Analysis of Cellular and Genomic Networks)

- geWorkbench, caBIG-caGRID-accessible
 - open-source software platform for biologists. genomic data integration, bringing together analysis and visualization tools for gene expression, sequences, pathways, and other biomedical data
 - <https://cabig.nci.nih.gov/tools/geWorkbench>
- Communication of cellular-molecular knowledge and data (previous work)
- PhenoGO (<http://www.phenogo.org>)
 - A Resource for the Multiscale Integration of Clinical and Biological Data

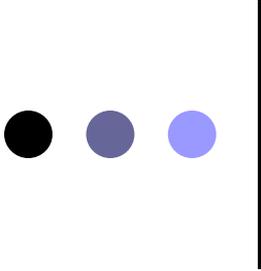


Previous work (1/3)

o Statistical co-occurrences and rule mining in GO

- (Single ontology- Review) Khatri P, Draghici S: Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* 2005, 21: 3587-3595.
- (Rule mining) Kumar A et al. Dependence relationships between Gene Ontology terms based on TIGR gene product annotations. (*CompuTerm* 2004), 31-38
- (text mining) Bada M, Hunter L. Enrichment of OBO ontologies. *J Biomed Inform.* 2007 Jun;40(3):300-15. Epub 2006 Jul 26.
- **(unadjusted GO-GO):** Bodenreider O, Aubry M, Burgun A. Non-lexical approaches to identifying associative relations in the gene ontology. *Pac Symp Biocomput.* 2005;:91-102.
- **(Multi-ontologies Adjusted Co-occurrences, Biological modules: Phenotype-GO-PFAM-KEGG)** Liu Y, Li J, Sam L, Goh CS, Gerstein M*, Lussier YA*. From phenotypes to genotypes: A meta-genomic approach to uncover the molecular mechanisms of prokaryotic traits. *PLoS Computational Biology*, 2006 2(11): 1419-1435





Previous work (2/3)

o Text-based Integration

• Computational terminology

• Lexico-semantic

- **Cimino J**, Barnett O. Automated translation between medical terminologies using semantic definitions. MD Comput. **1990** Mar-Apr;7(2):104-9.
- **(GO-UMLS)** Cantor MN, Sarkar IN, Gelman R, Hartel F, Bodenreider O, Lussier YA. An evaluation of hybrid methods for matching biomedical terminologies: mapping the gene ontology to the UMLS. Stud Health Technol Inform. 2003;95:62-7.

• Lexical: Aronson AR. Effective mapping of biomedical text to the UMLS

Metathesaurus: the MetaMap program. Proc AMIA Symp. 2001;:17-21

• Compositional – lexical - semantic

- **“Minimal Representable Units Method”**. **Zeng Q**, Cimino JJ. Mapping medical vocabularies to the Unified Medical Language System. Proc AMIA Annu Fall Symp. **1996**;:105-9.
- (GO-GO). Ogren, P.V., Cohen, K.B., Acquaaah-Mensah, G.K., Eberlein, J. & Hunter, L. The compositional structure of Gene Ontology terms. Pac Symp Biocomput, 214-25 (2004)
- (MP-SNOMED) Lussier YA, Li J. Terminological mapping for high throughput comparative biology of phenotypes. Pac Symp Biocomput. 2004;:202-13.
- (Term-SNOMED). Shah N, Rubin D, Supekar K, Musen M. AMIA Annu Symp Proc 2006; 709-13. Ontology-based Annotation and Query of Tissue Microarray Data.



• Combinatorial (GO terms): Hill DP, **Blake JA**, Richardson JE, Ringwald M.

Extension and integration of the gene ontology (GO): combining GO vocabularies with external vocabularies. Genome Res. **2002** Dec;12(12):1982-91



Previous work (3/3)

- Literature mining
 - Indexing
 - Jensen LJ, Saric J, Bork P. Literature mining for the biologist: from information retrieval to biological discovery. Nat Rev Genet. 2006 Feb;7(2):119-29. Review.
 - Hybrid: natural language processing and lexico-semantic (GO, CT, MP, MA, UMLS)
 - **PhenoGO**: assigning phenotypic context to gene ontology annotations with natural language processing. Lussier Y, Borlawsky T, Rappaport D, Liu Y, Friedman C. Pac Symp Biocomput. 2006;:64-75.



A Problem of Specificity in the GO Model: The $Foxn1^{nu}/Foxn1^{nu}$ mouse

Current GO model implies a unicellular organization

MGI: $Foxn1^{nu}/Foxn1^{nu}$



OMIM (NCBI): In 2 sisters with T-cell immunodeficiency, congenital alopecia

GO 24 annotations for FOXN1

“T-lymphocyte function”

“impaired embryonic morphogenesis”

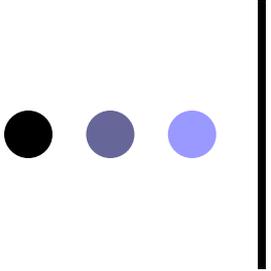
“keratinocyte differentiation”

PMID: 16232301

Nude ($Foxn1^{nu}/Foxn1^{nu}$) mice develop largely normal hair follicles and produce hair shafts. Since hair shafts fail to penetrate the epidermis, macroscopic nudity results ...

MH - Hair Follicle

MH - Mice, Nude/anatomy & histology/genetics/physiology



Overview

- Introduction/Previous Work
- Updates
- Web Interface
- Evaluation
 - Comprehensive
 - Disease annotation specific
 - Cell annotation specific
- Applications
 - Diseases
 - cell

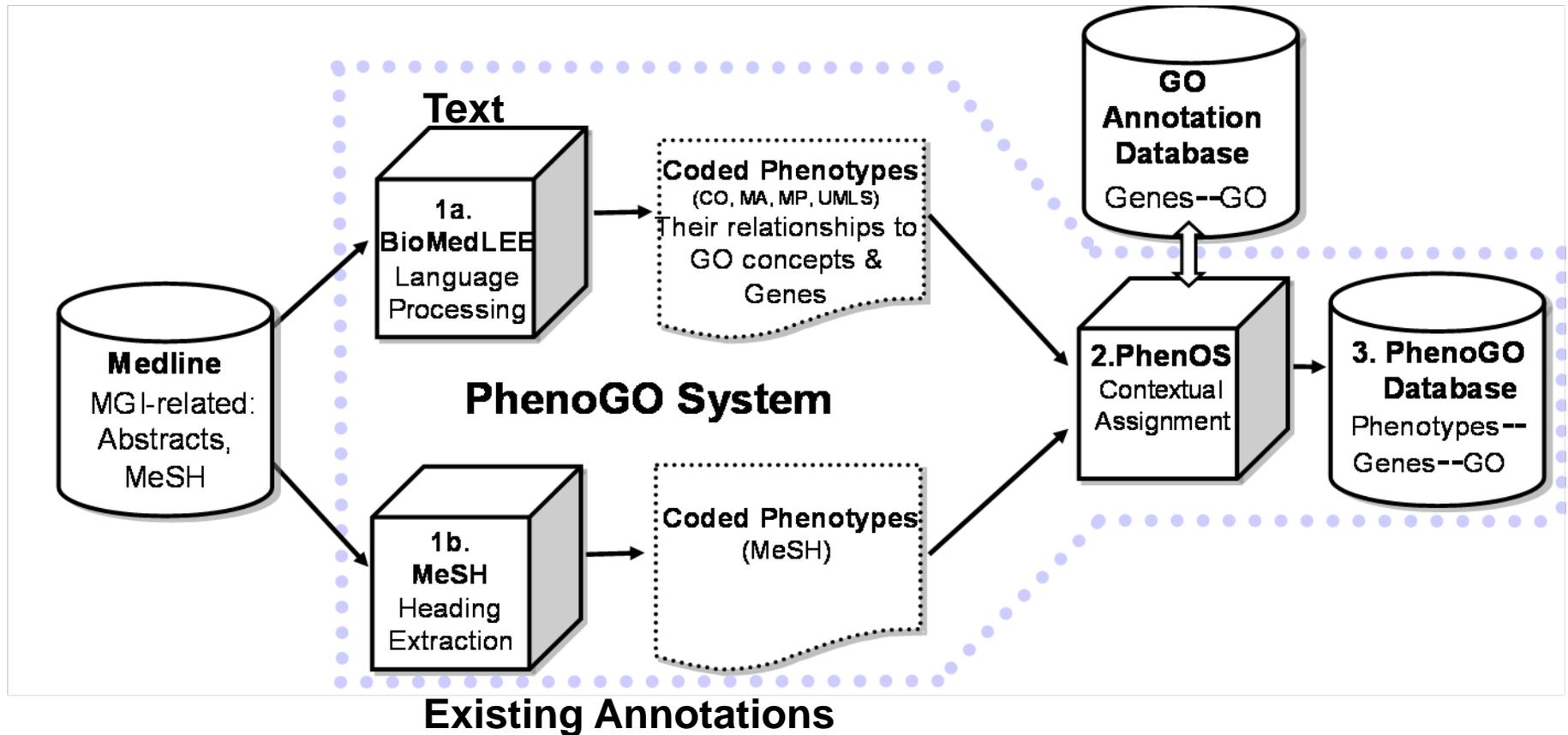


● ● ● | Introduction/Previous Work

- The PhenoGO database – adding context to Gene Ontology Annotations
- Open standards
 - Standardized ontologies and coding systems
 - Genes: MGI, SwissProt, UniGene, etc...
 - Phenotypes/contexts: Mammalian Phenotype Ontology, UMLS, Cell Type Ontology, MeSH, etc...
- 2006
 - 260,049 gene-GO-cell&anatomy context annotations
 - Specifically focused on the mouse

Lussier Y, Borlawsky T, Rappaport D, Liu Y, Friedman C. PhenoGO: assigning phenotypic context to gene ontology annotations with natural language processing. *Pac Symp Biocomput.* 2006;:64-75.

The PhenoGO Encoding Pipeline



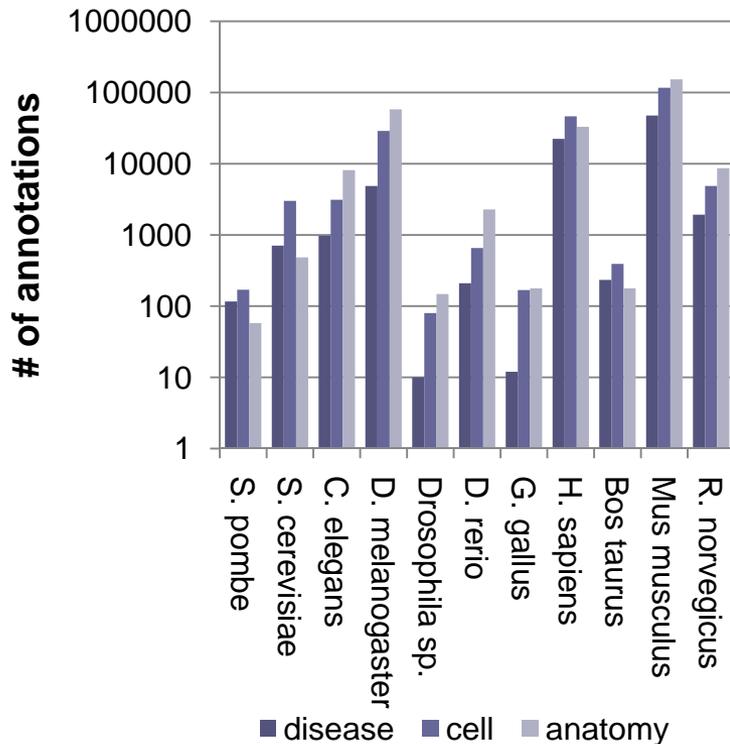


Database Expansion

- Addition of ten additional species
- Addition of diseases and disease-associated clinical phenotypes
- Database now includes:
 - *Schizosaccharomyces pombe*
 - *Saccharomyces cerevisiae*
 - *Caenorhabditis elegans*
 - *Drosophila melanogaster*
 - *Drosophila sp.*
 - *Danio rerio*
 - *Gallus gallus*
 - *Homo sapiens*
 - *Bos taurus*
 - *Mus musculus*
 - *Rattus norvegicus*

Database Contents

- 658,041 distinct PMID-gene-GO-context annotations

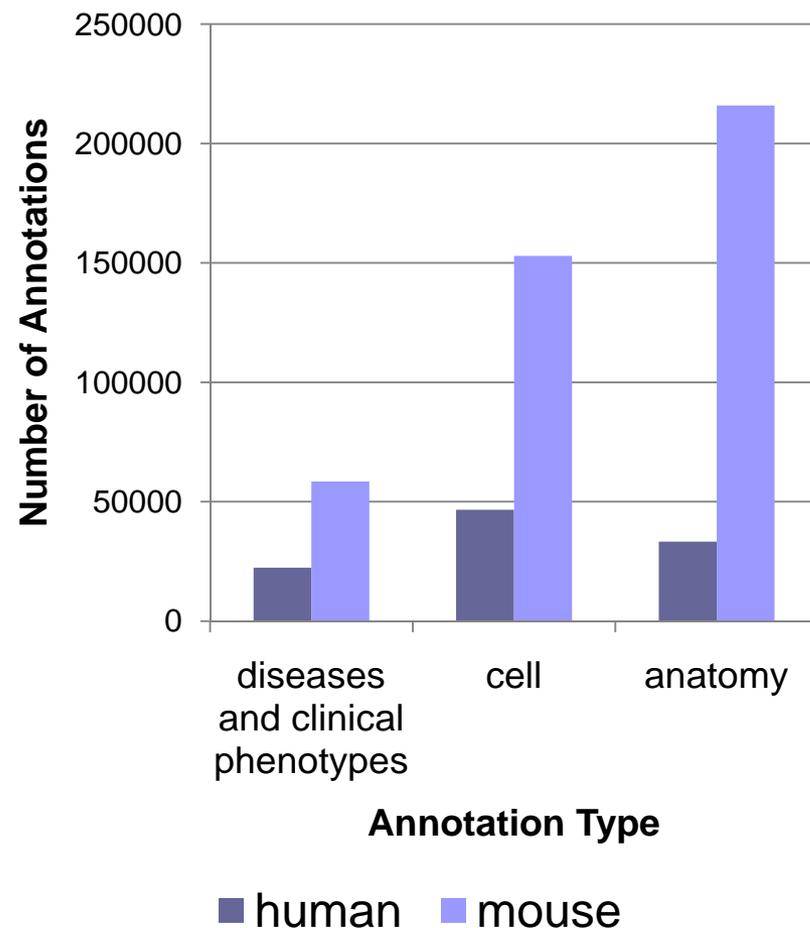


Taxon	Name	# Annotations
4896	Schizosaccharomyces pombe	344
4932	Saccharomyces cerevisiae	4,192
6239	Caenorhabditis elegans	12,212
7227	Drosophila melanogaster	91,782
7242	Drosophila sp.	238
7955	Danio rerio	3,142
9031	Gallus gallus	358
9606	Homo sapiens	102,262
9913	Bos taurus	804
10090	Mus musculus	427,275
10116	Rattus norvegicus	15,432
Total		658,041



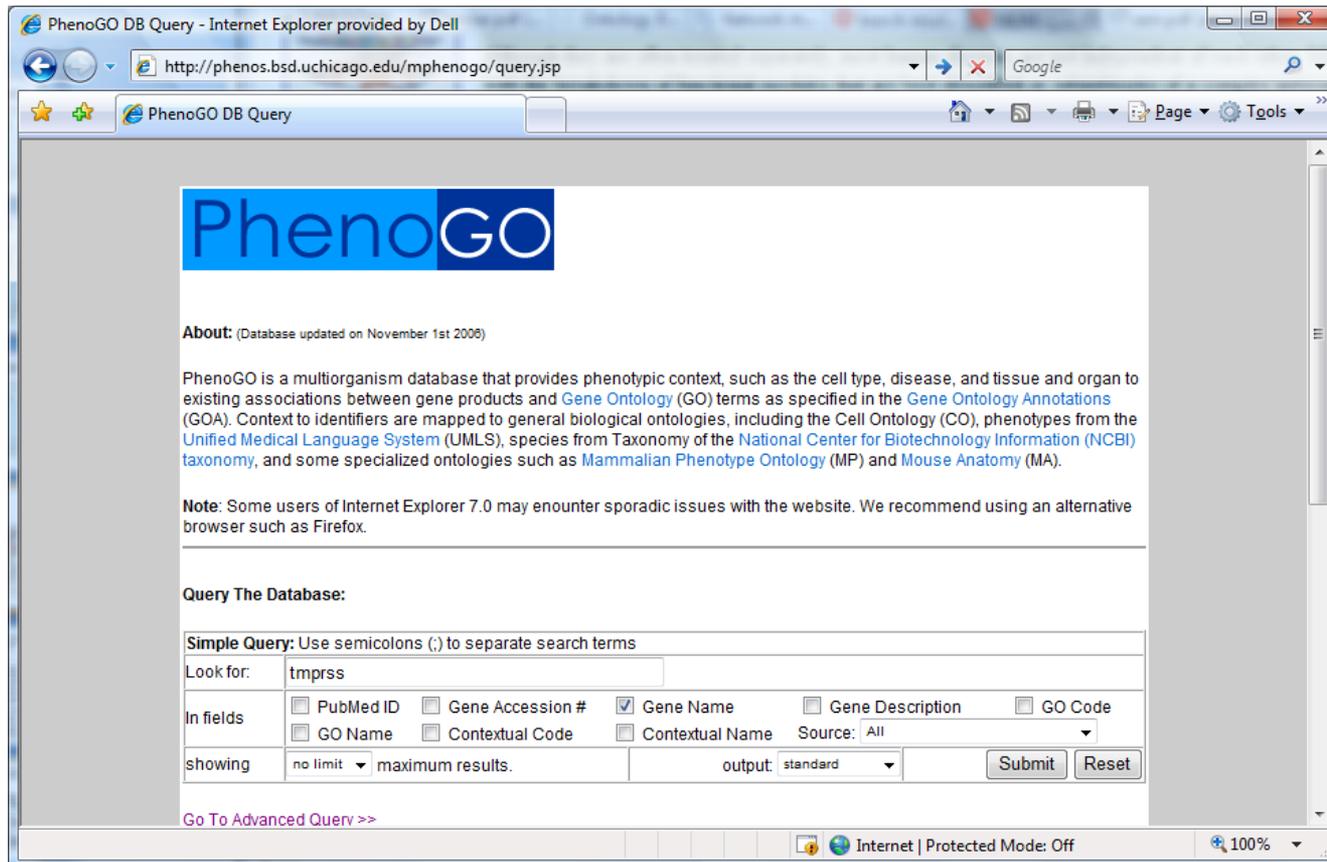
Diseases and Disorders

- 78,947 total annotations associated with diseases and disorders
 - 69,899 Human and mouse annotations
 - 3,209 distinct disease and disorders of 6,084 distinct contexts



Web Query: Basic

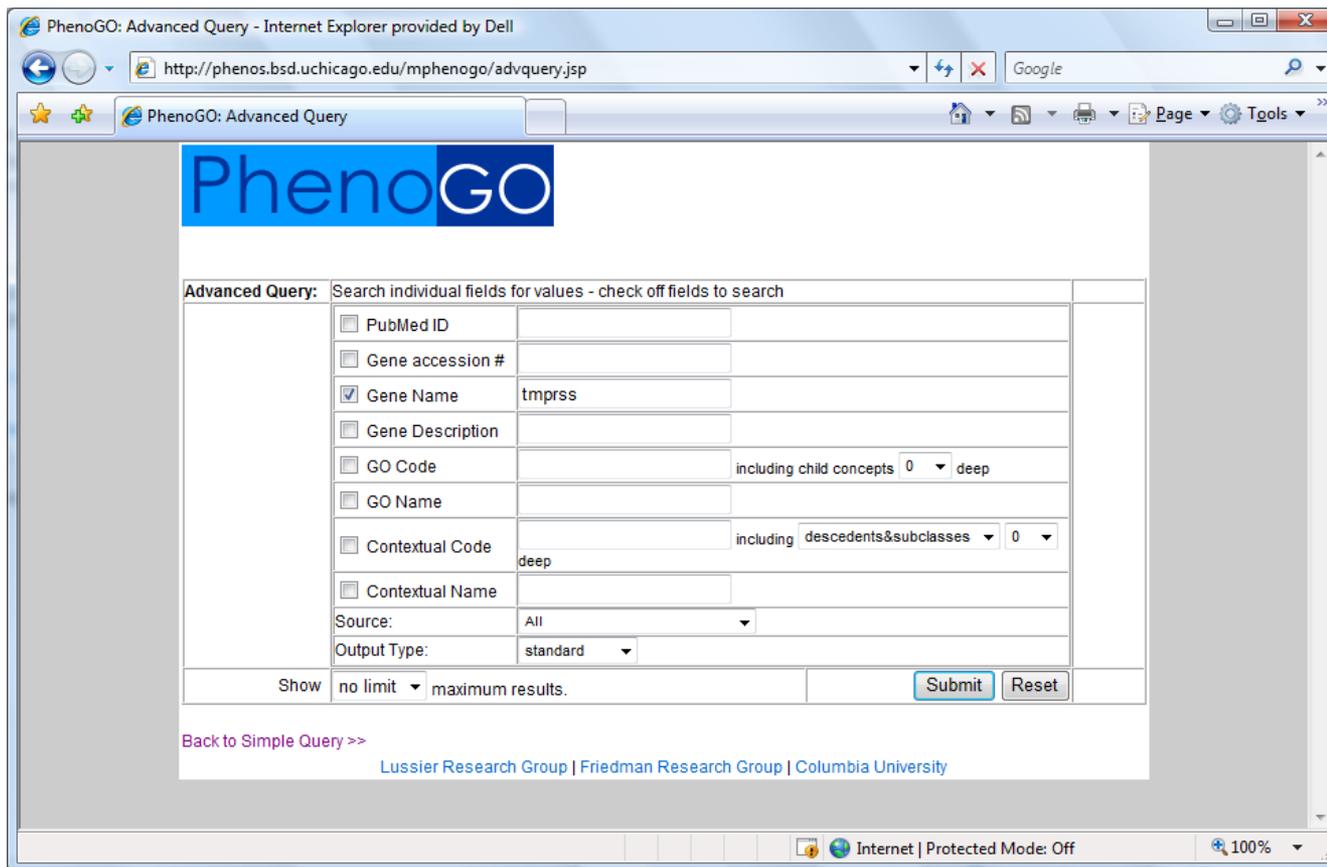
- Runs a query analogous to an SQL OR for the search terms



The screenshot shows the PhenoGO DB Query web interface in Internet Explorer. The browser window title is "PhenoGO DB Query - Internet Explorer provided by Dell". The address bar shows the URL "http://phenos.bsd.uchicago.edu/mphenogo/query.jsp". The page features the PhenoGO logo at the top. Below the logo, there is an "About:" section with the text "(Database updated on November 1st 2008)". The main content area contains a "Query The Database:" section. This section includes a "Simple Query:" instruction: "Use semicolons (;) to separate search terms". The "Look for:" field contains the text "tmprss". Below this, there are several checkboxes for "In fields": "PubMed ID", "Gene Accession #", "Gene Name" (checked), "Gene Description", "GO Code", "GO Name", "Contextual Code", and "Contextual Name". The "Source:" dropdown menu is set to "All". The "showing" section has a "no limit" dropdown and "maximum results." text. The "output:" dropdown menu is set to "standard". There are "Submit" and "Reset" buttons. At the bottom of the page, there is a link "Go To Advanced Query >>". The browser status bar at the bottom shows "Internet | Protected Mode: Off" and "100%".

Web Query: Advanced

- Runs a query analogous to an SQL AND for the search terms



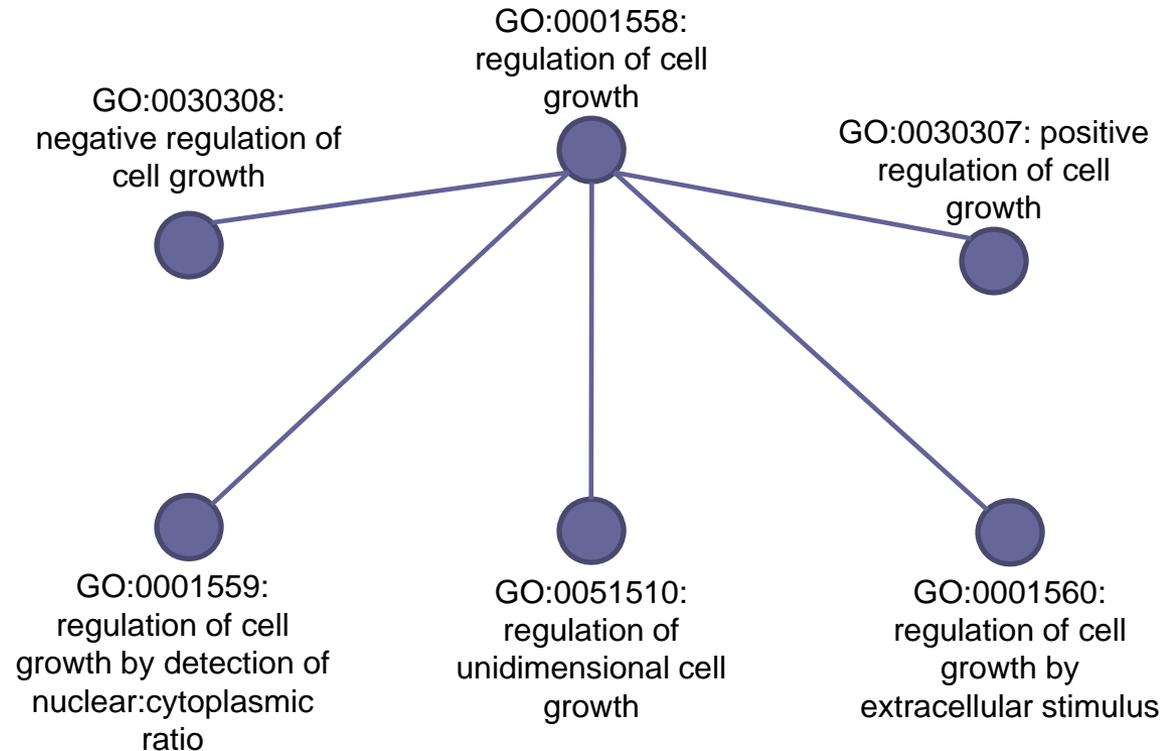
The screenshot shows the PhenoGO Advanced Query interface in Internet Explorer. The browser title is "PhenoGO: Advanced Query - Internet Explorer provided by Dell" and the address bar shows "http://phenos.bsd.uchicago.edu/mphenogo/advquery.jsp". The page features the PhenoGO logo at the top. Below the logo is a form titled "Advanced Query: Search individual fields for values - check off fields to search". The form includes several input fields and checkboxes:

Field	Value	Options
PubMed ID		
Gene accession #		
<input checked="" type="checkbox"/> Gene Name	tmprss	
Gene Description		
GO Code		including child concepts 0 deep
GO Name		
Contextual Code		including descendents&subclasses 0 deep
Contextual Name		
Source:	All	
Output Type:	standard	

At the bottom of the form, there is a "Show" section with "no limit" selected for "maximum results." and "Submit" and "Reset" buttons. Below the form, there is a link "Back to Simple Query >>" and a footer "Lussier Research Group | Friedman Research Group | Columbia University". The browser status bar at the bottom indicates "Internet | Protected Mode: Off" and "100%".

Hierarchical Query

- Query for more specific descendent terms
- Uses a number of ancestor-descendent tables to build a SQL query



Result Formats

o HTML-formatted

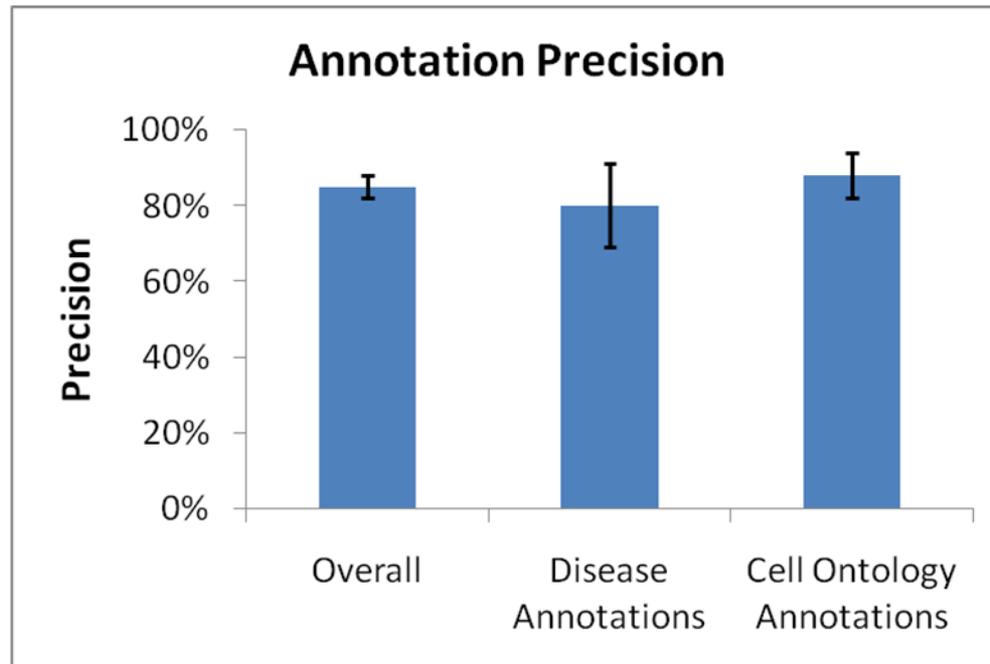
result #	pmid	gene code	gene name	gene description	GO code	GO name	context code	context name	taxon	source
1	10786627	MGI:1353645	Tmprss8	transmembrane protease, serine 8 (intestinal)	GO:0008236	serine-type peptidase activity	UMLS:C0020885	ileum	10090	MGI
2	10786627	MGI:1353645	Tmprss8	transmembrane protease, serine 8 (intestinal)	GO:0008236	serine-type peptidase activity	UMLS:C0021852	small intestine	10090	MGI
3	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0006508	proteolysis and peptidolysis	UMLS:C0015965	fetus	10090	MGI
4	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0004252	serine-type endopeptidase activity	UMLS:C0015965	fetus	10090	MGI
5	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0005886	plasma membrane	UMLS:C0015965	fetus	10090	MGI
6	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0006508	proteolysis and peptidolysis	UMLS:C0023884	liver	10090	MGI
7	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0004252	serine-type endopeptidase activity	UMLS:C0023884	liver	10090	MGI
8	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0005886	plasma membrane	UMLS:C0023884	liver	10090	MGI
9	12393794	MGI:2155445	Tmprss3	transmembrane protease, serine 3	GO:0006883	sodium ion homeostasis	UMLS:C0029207	organ of corti structure	10090	MGI
10	12393794	MGI:2155445	Tmprss3	transmembrane protease, serine 3	GO:0017080	sodium channel regulator activity	UMLS:C0029207	organ of corti structure	10090	MGI
11	12149247	MGI:1919003	Tmprss6	transmembrane serine protease 6	GO:0006508	proteolysis and peptidolysis	UMLS:C0023884	liver	10090	MGI

o Tab-delimited text

```
pmid gene code gene name gene description GO code GO name context taxon source
10786627 MGI:1353645 Tmprss8 transmembrane protease, serine 8 (intestinal) GO:0008236 serine-type peptidase activity UMLS:C0020885 ileum 10090 MGI
10786627 MGI:1353645 Tmprss8 transmembrane protease, serine 8 (intestinal) GO:0008236 serine-type peptidase activity UMLS:C0021852 small intestine 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0006508 proteolysis and peptidolysis UMLS:C0015965 fetus 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0004252 serine-type endopeptidase activity UMLS:C0015965 fetus 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0005886 plasma membrane UMLS:C0015965 fetus 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0006508 proteolysis and peptidolysis UMLS:C0023884 liver 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0004252 serine-type endopeptidase activity UMLS:C0023884 liver 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0005886 plasma membrane UMLS:C0023884 liver 10090 MGI
12393794 MGI:2155445 Tmprss3 transmembrane protease, serine 3 GO:0006883 sodium ion homeostasis UMLS:C0029207 organ of corti structure 10090 MGI
12393794 MGI:2155445 Tmprss3 transmembrane protease, serine 3 GO:0017080 sodium channel regulator activity UMLS:C0029207 organ of corti structure 10090 MGI
12149247 MGI:1919003 Tmprss6 transmembrane serine protease 6 GO:0006508 proteolysis and peptidolysis UMLS:C0023884 liver 10090 MGI
```

Evaluations

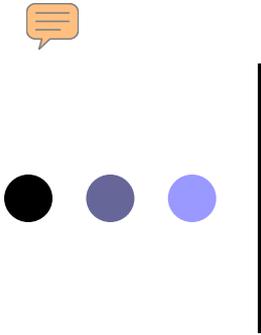
- Comprehensive Evaluation
- Cell Ontology-specific Evaluation
- Disease Annotations-specific Evaluation





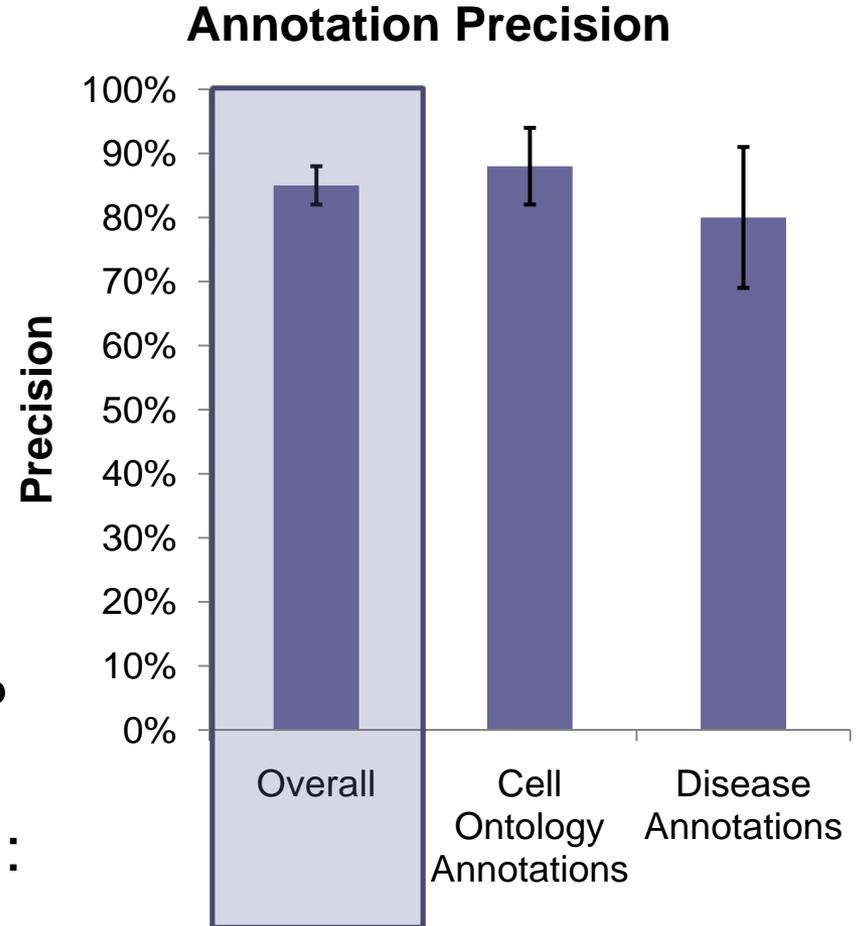
Evaluation Methodology

- Precision:
 - A set of randomly extracted entries was extracted and each was evaluated for correctness manually
- Recall
 - Randomly extracted sentences were presented to evaluators and each was checked to exist in the database if it made sense



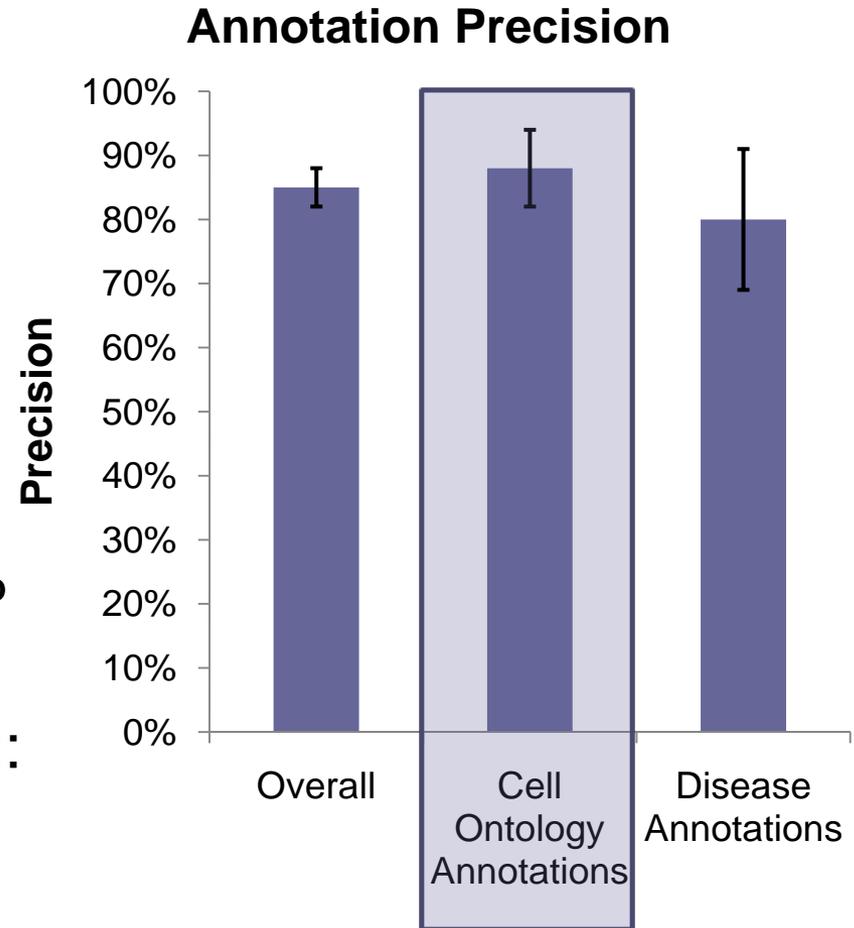
Comprehensive Evaluation

- 2 Experienced Evaluators with biological backgrounds
- n = 300 phenotypic annotations evaluated independently
- Precision: 85% (95% CI: 82%-89%)
- Recall: 76% (95% CI: 69-83%)



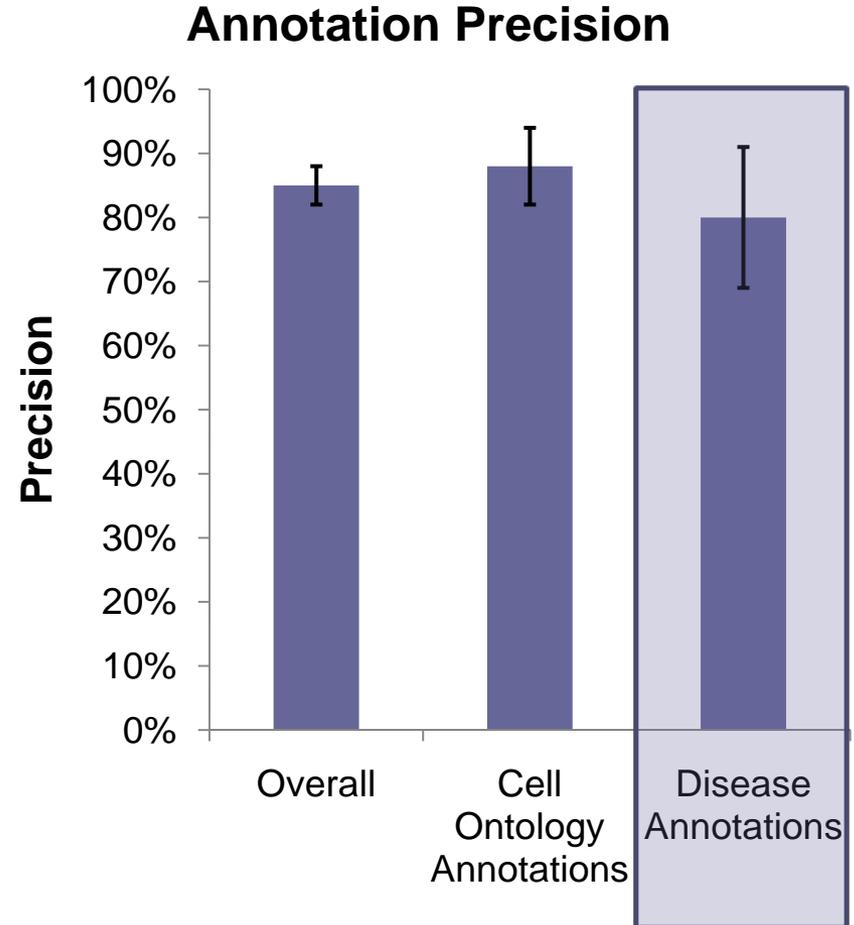
Cell Ontology-Specific Evaluation

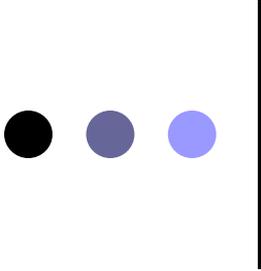
- 2 Experienced Evaluators with biological backgrounds
- $n = 50$ cell ontology-specific annotations
- Precision: 88% (95% CI: 82%-94%)
- Recall: 79% (95% CI: 69%-89%)



Disease-Specific Evaluation

- 2 Experienced Evaluators with biological background
- n = 50 disease and disease-associated clinical phenotypes evaluated independently
- Precision : 80% (95% CI:69%-90%)





Applications

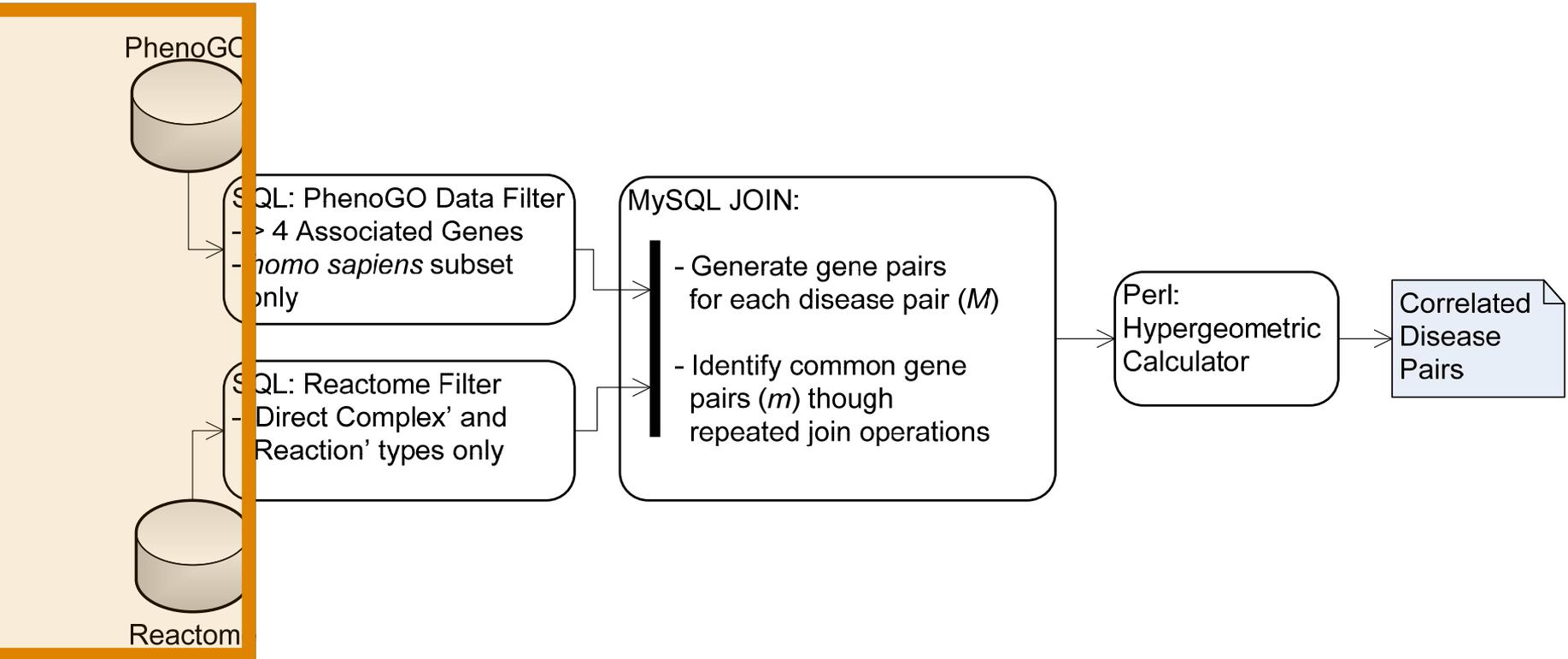
- Gene-Phenotype annotations play a central role in the study of the networks associated with human disease
- Approaches
 - Manual annotation
 - Krauthammer M, et al., *PNAS* (2004)
 - Goh, et al., *PNAS* (2007)
 - NLP
 - Lage, et al., *Nature Biotech.* (2007)

Source: Barabási AL., **N Engl J Med.** 2007 Jul 26;357(4):404-7

Methodology

Lussier PSB 12:76-87(2007)

- **Reactome:** Protein-Protein Interaction network
- **PhenoGO:** Phenotype-genotype association



Accrual Standardization

Integration

Discovery

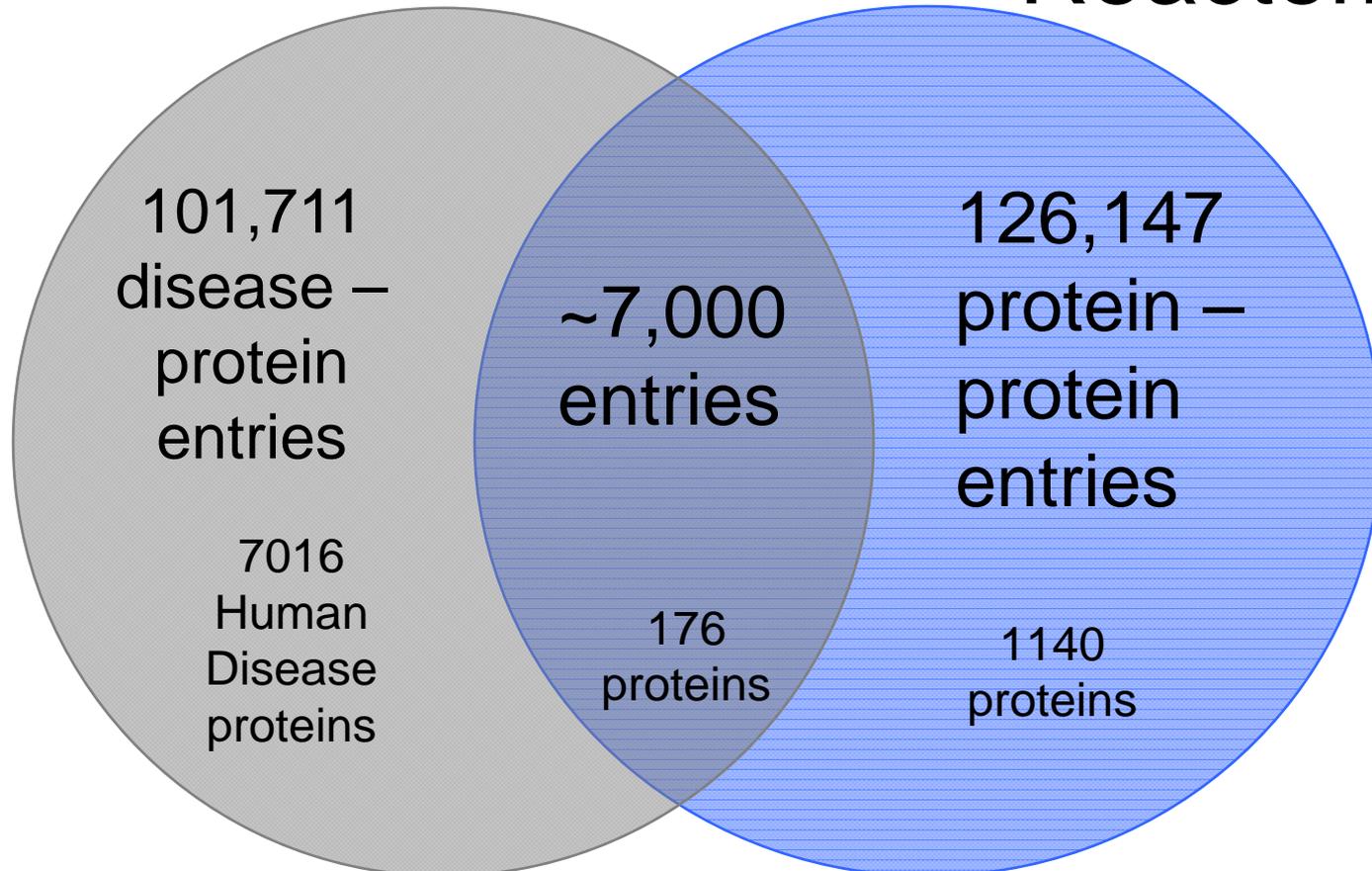


What do we have data on?

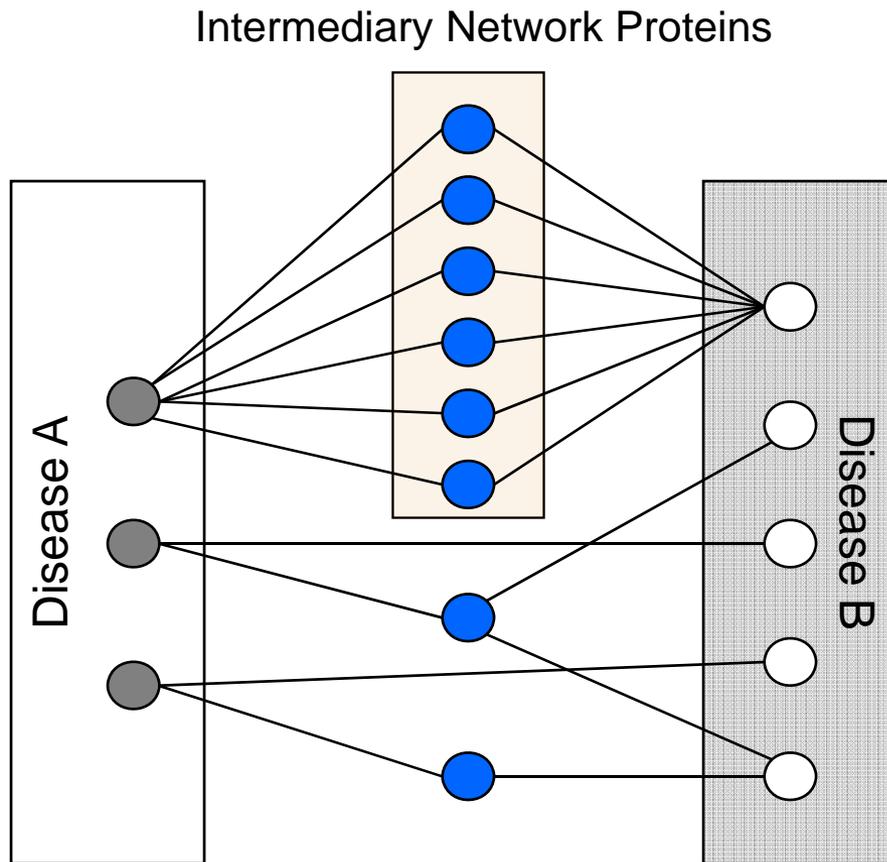
(Lussier. Pac Symp Biocomp 2007)

PhenoGO

Reactome



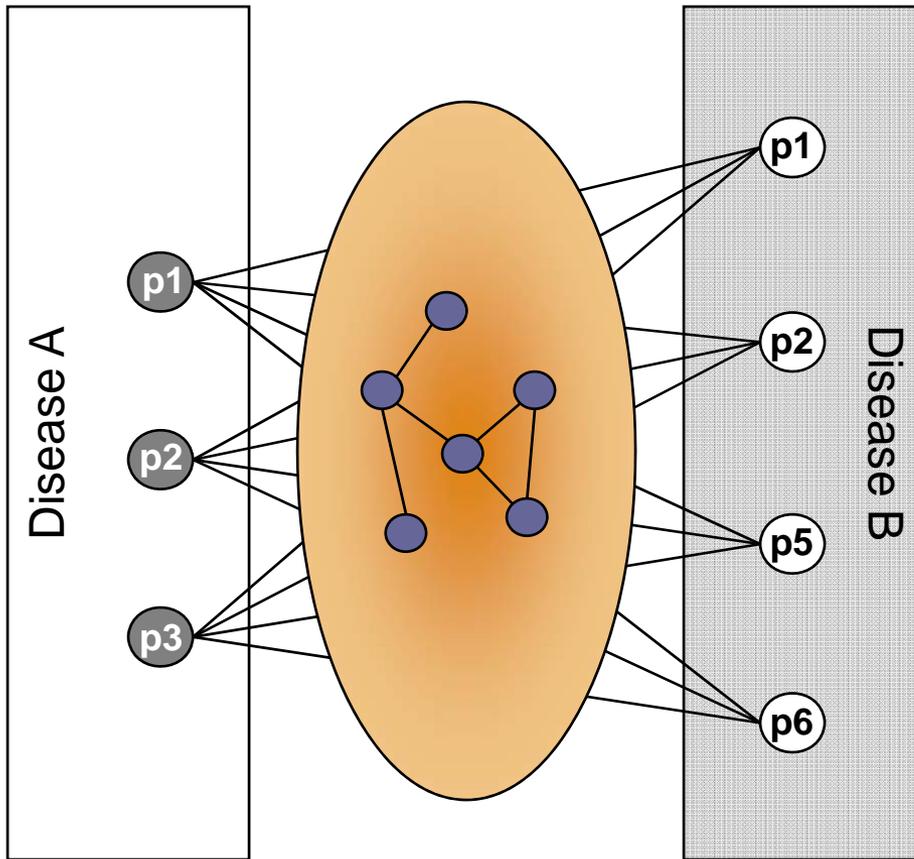
Highly connected hubs create biases when counting nodes or edges between disease proteins



- Nodes:
 - ● = disease A proteins
 - ○ = disease B proteins
 - ● = Network proteins
- Edges = prot-prot interac.
 - 6/8 shared interactions partners are due to only 2 endpoints
- Examples of Hubs:
 - RPS10 (118 partners)
 - ERCC2 (125 partners)

Measure of Shared Disease Proteins Networks:

Counting combinations of disease pairs is scalable



- For diseases A and B,
 - count unique connected disease-protein pairs regardless of:
 - the number of pathway(s) per pair
 - The number of intermediary proteins in the network



Discovery of Shared Phenotype-Associated Protein Interaction Subnetworks

o Integration of:

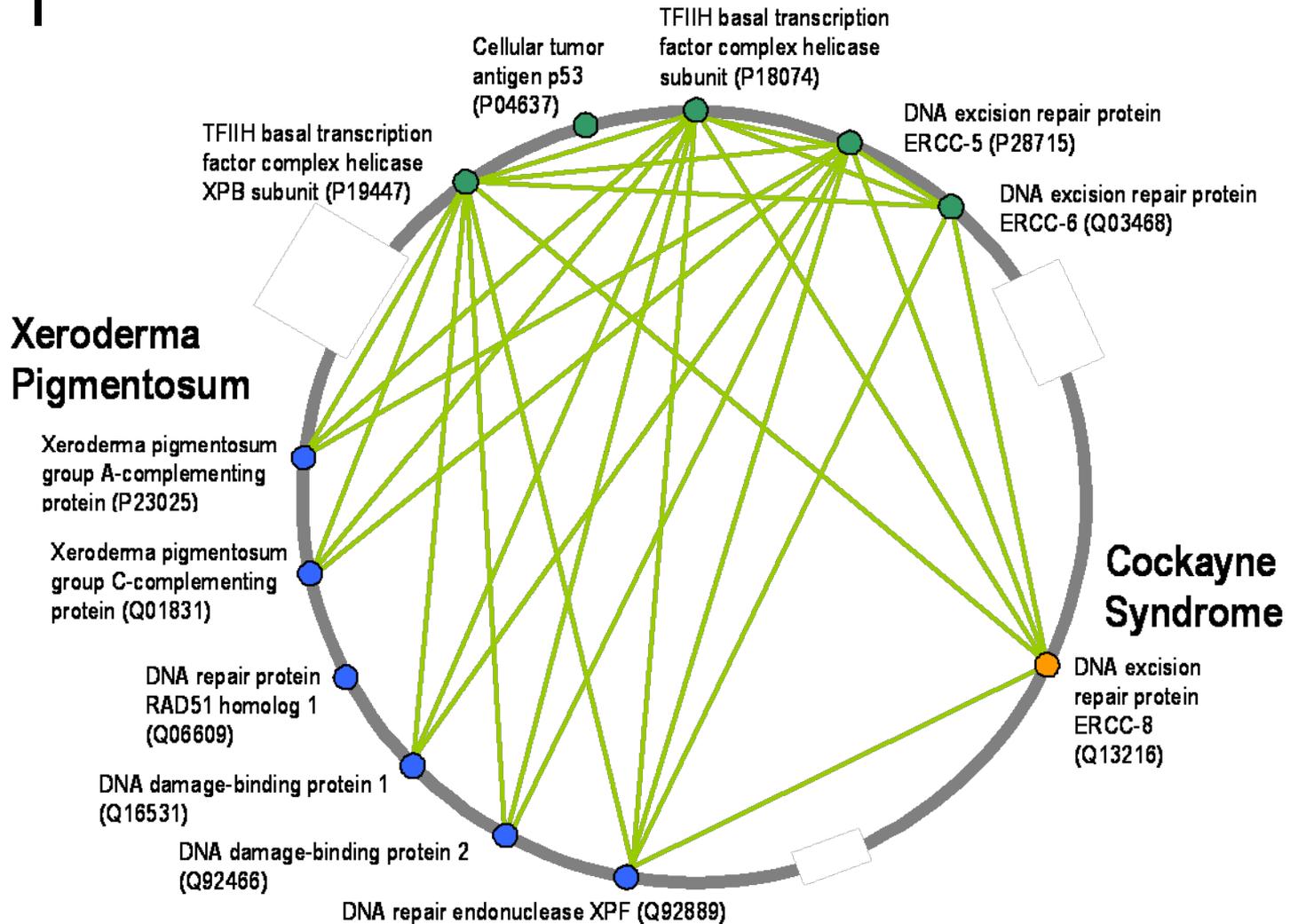
- The Reactome protein-protein interaction network
- Human-specific PhenoGO gene-disease associations

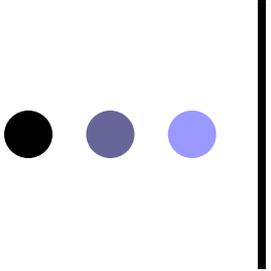
o Resulting in 11,703 disease-disease comparisons

UMLS ID	Disease 1	UMLS ID	Disease 2	Corrected pvalue
C0009207	Cockayne Syndrome	C0043346	Xeroderma Pigmentosum	8.5e-18
C0043346	Xeroderma Pigmentosum	C0085390	Li-Fraumeni Syndrome	4.9e-06
C0007001	Carbohydrate Metabolism, Inborn Errors	C0002514	Amino Acid Metabolism, Inborn Errors	6.2e-05
C0009404	Colorectal Neoplasms	C0950123	Genetic Diseases, Inborn	5.0e-05
C0085390	Li-Fraumeni Syndrome	C0009207	Cockayne Syndrome	1.9e-04

Sam L, Liu Y, Li J, Friedman C, Lussier YA. Discovery of protein interaction networks shared by diseases. *Pac Symp Biocomput.* 2007;;76-87.

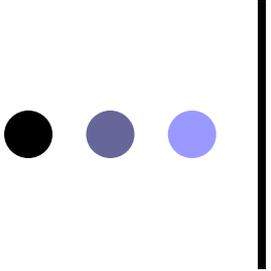
Example: Cockayne and XP



Top Results

UMLS ID	Disease 1	UMLS ID	Disease 2	P-PI (#)	pvalue
C0009207	Cockayne Syndrome	C0043346	Xeroderma Pigmentosum	38	7.3e-22
C0043346	Xeroderma Pigmentosum	C0085390	Li-Fraumeni Syndrome	24	6.7e-11
C0007001	Carbohydrate Metabolism, Inborn Errors	C0002514	Amino Acid Metabolism, Inborn Errors	9	8.3e-10
C0009404	Colorectal Neoplasms	C0950123	Genetic Diseases, Inborn	16	6.7e-10
C0085390	Li-Fraumeni Syndrome	C0009207	Cockayne Syndrome	16	2.7e-09
C0009404	Colorectal Neoplasms	C0015625	Fanconi's Anemia	8	1.5e-05
C0009404	Colorectal Neoplasms	C0085413	Polycystic Kidney, Autosomal Dominant	8	1.5e-05
C0024141	Lupus Erythematosus, Systemic	C0004364	Autoimmune Diseases	4	9.3e-05
C0024314	Lymphoproliferative Disorders	C0004364	Autoimmune Diseases	6	1.3e-04
C0024314	Lymphoproliferative Disorders	C0024141	Lupus Erythematosus, Systemic	6	1.3e-04



Proof of concept

- Fanconi Anemia and Breast Cancer Susceptibility. *Nature Genetics* 39(2);2007.

Results after multiple-testing adjustment

Dunn-Sidak correction

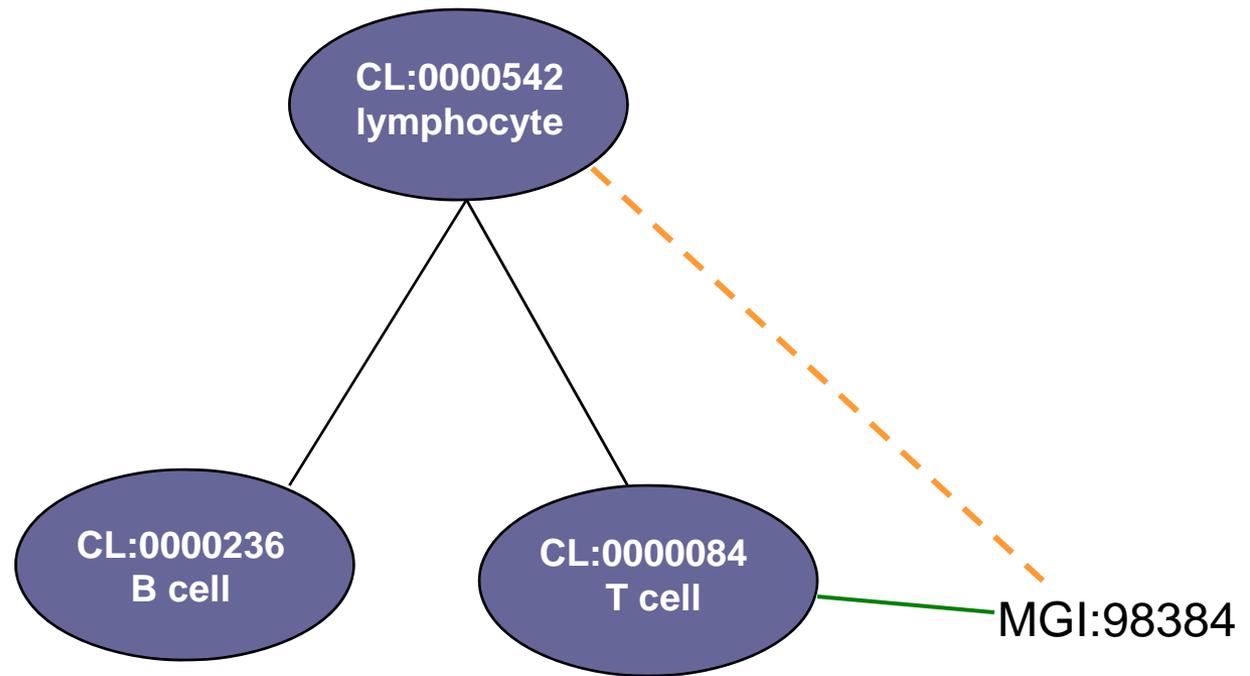
$$p' = 1 - (1 - p)^r$$

- **5 significant results**
- Bonferroni-type method
- Less severe, but still very conservative
- Study $r = 11,703$ comparisons

Permutation Resampling

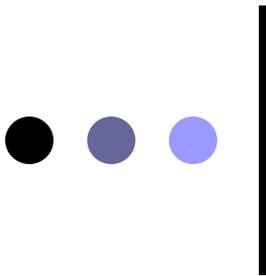
- **162 significant results**
- **Uncorrected $p < .092$ for 5% error rate**
- How much does the p-value depend on which gene-phenotype associations exist?
- Build background sample from randomized gene-phenotype data
- 1000 iterations

Assign gene/GO using Cell Ontology's structure



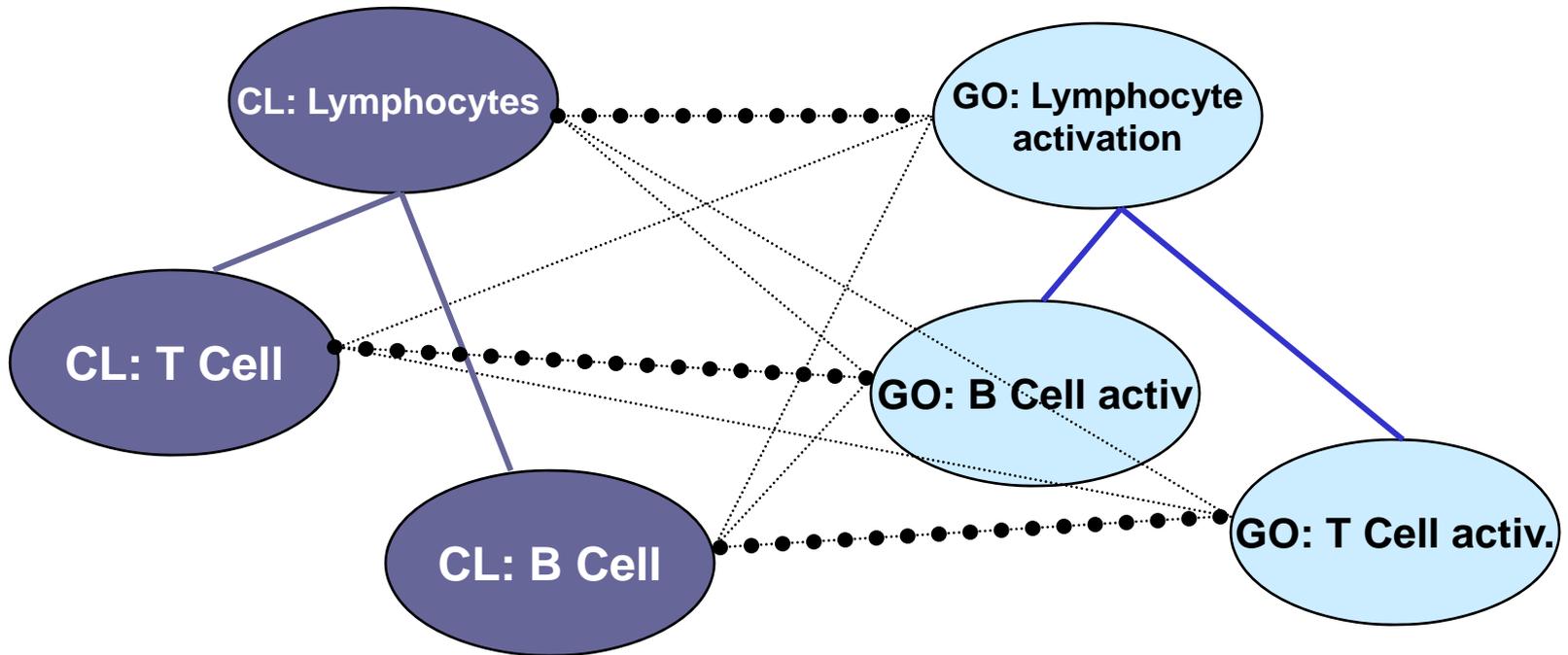
Existing relation in PhenoGO: **MGI:98384—CL:0000084**

Inferred relation: **MGI:98384—CL:0000542**



Cell Ontology

Gene Ontology



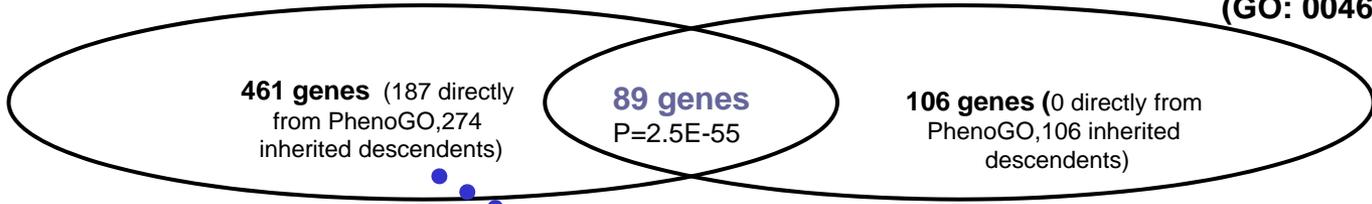
Evaluate relationships between GO and CO
identify shared genes between each pair of node,
and calculate probability of the distribution

Genetic Scalar Analysis of Phenotypes (geneSAP - unpublished) (mouse)



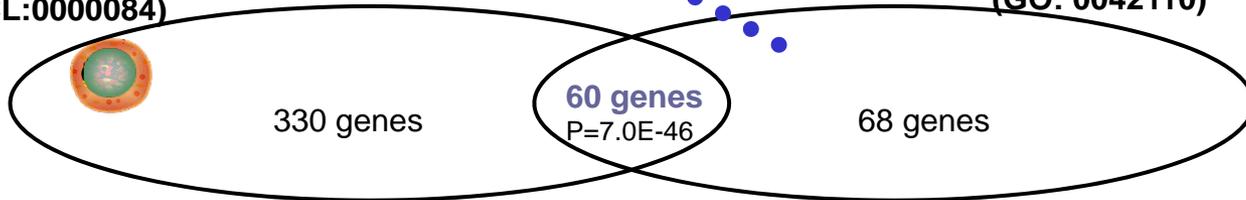
Lymphocytes
(CL:0000542)

Lymphocyte activation
(GO: 0046649)



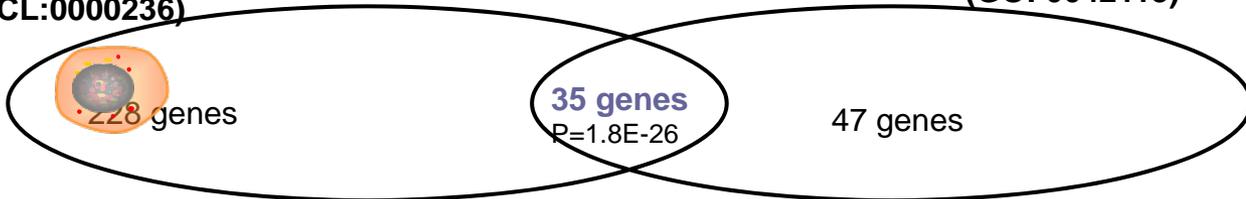
T-cell
(CL:0000084)

T cell activation
(GO: 0042110)



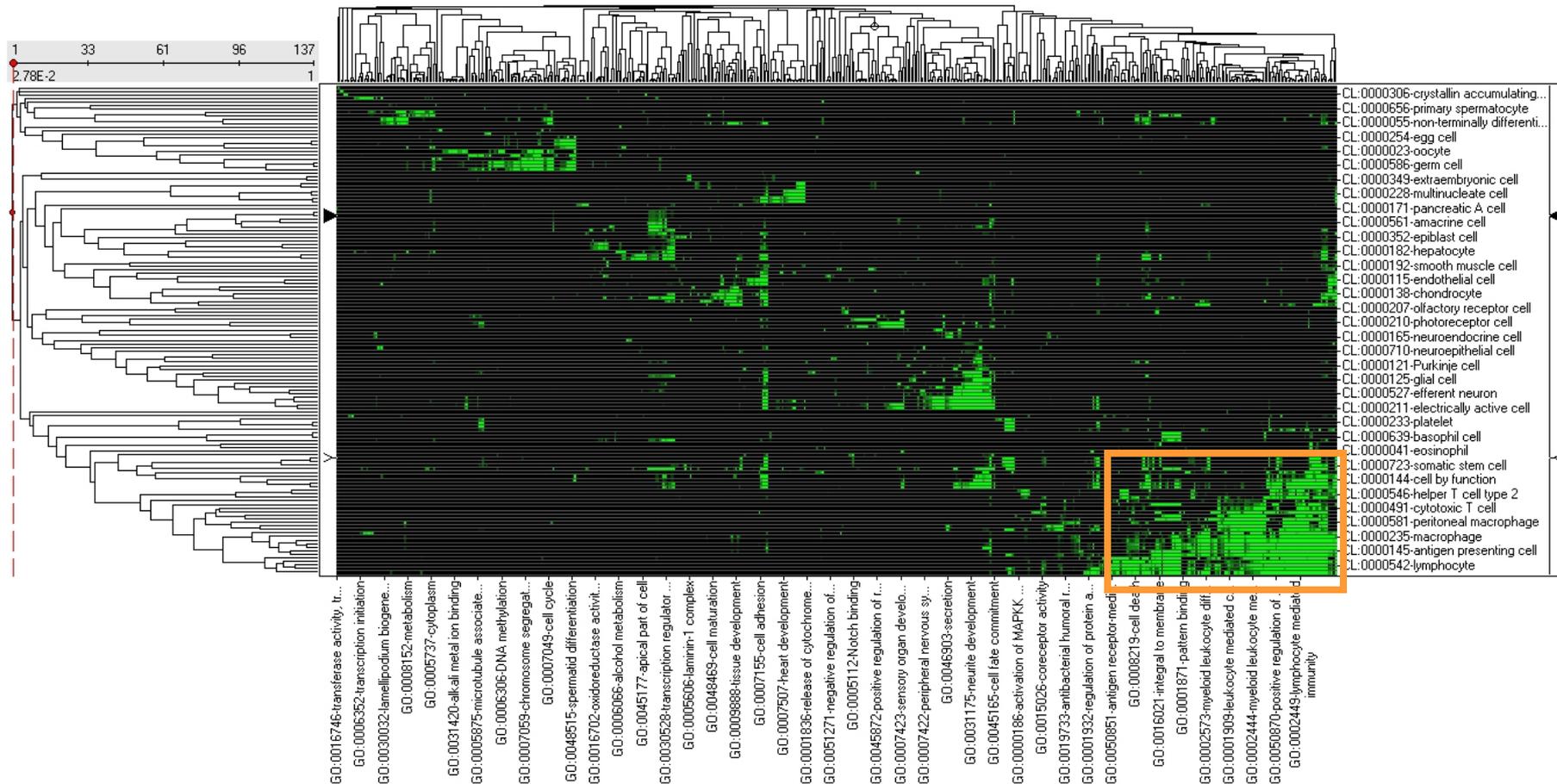
B-cell
(CL:0000236)

B cell activation
(GO: 0042113)



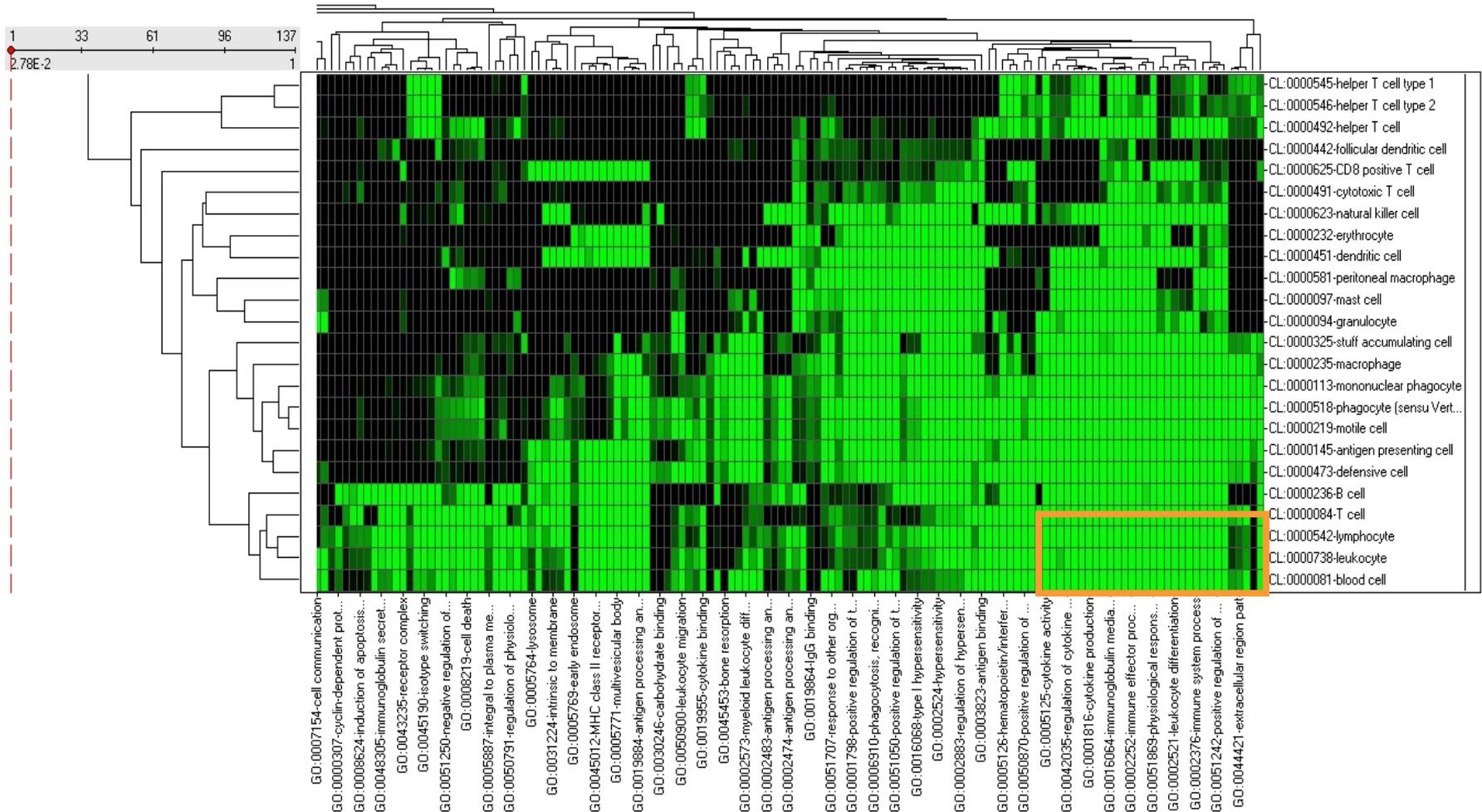
Clustering of mouse CO-GO

Hierarchical Clustering



Lymphocyte cluster in mouse

Hierarchical Clustering



GO:0002573-myeloid leukocyte diffr...
 GO:0002483-antigen processing an...
 GO:0002474-antigen processing an...
 GO:0019864-IgG binding
 GO:0051707-response to other org...
 GO:0001798-positive regulation of t...
 GO:0006910-phagocytosis, recogni...
 GO:0051050-positive regulation of t...
 GO:0016068-type I hypersensitivity
 GO:0002524-hypersensitivity
 GO:0002883-regulation of hypersen...
 GO:0003823-antigen binding
 GO:0005126-hematopoietin/interfer...
 GO:0050870-positive regulation of ...
 GO:0005125-cytokine activity
 GO:0042035-regulation of cytokine ...
 GO:0001816-cytokine production
 GO:0016064-immunoglobulin media...
 GO:0002252-immune effector proc...
 GO:0051869-physiological respons...
 GO:0002521-leukocyte differentiation
 GO:0002376-immune system process
 GO:0051242-positive regulation of ...
 GO:0044421-extracellular region part

-CL:0000581-peritoneal macrophage
 -CL:0000097-mast cell
 -CL:0000094-granulocyte
 -CL:0000325-stuff accumulating cell
 -CL:0000235-macrophage
 -CL:0000113-mononuclear phagocyte
 -CL:0000518-phagocyte (sensu Vert...
 -CL:0000219-motile cell
 -CL:0000145-antigen presenting cell
 -CL:0000473-defensive cell
 -CL:0000236-B cell
 -CL:0000084-T cell
 -CL:0000542-lymphocyte
 -CL:0000738-leukocyte
 -CL:0000081-blood cell



geneSAP Results

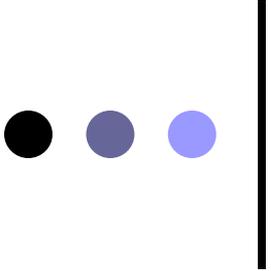
Paradigm (*subsumed microparadigms*)

o Mouse

- 1,030 (7,998)
GO-CL; $p < .05$
- CO: 136 (48)
- GO: 553 (480)
- ***Precision > 75%***

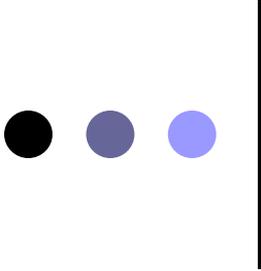
o Drosophila

- 400 (2,158)
GO-CL; $p < .05$
- CO: 41 (25)
- GO: 285 (307)
- ***Precision > 85%***



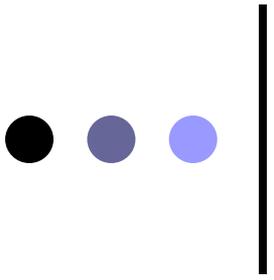
Discussion

- Proposed some solutions
- Needs for ontologies and more
 - Relate clinical and biological narratives to structured model with OBO Ontologies
 - Undefined relationships in OBO
 - Controlled relationships between OBO ontologies
 - ontology of relationships (S Lewis 2007)
 - Usefulness of higher order information models
 - Ternary gene-GO-CL, gene-GO-MP...
 - Quaternary?



Acknowledgements

- **University of Chicago**
 - Jianrong Li
 - Lee Sam (now at U of Michigan)
 - Tara Borlawsky (now at Ohio State)
 - Yang Liu
- **Columbia University**
 - Andrea Califano
 - Aris Floratos
 - Carol Friedman
- **Stanford University**
 - Mark Musen
- **University of Buffalo**
 - Barry Smith
- **NIH/NLM**
 - Olivier Bodenreider
- **Erci Newmann**
- **Grants**
 - NLM, 1K22 LM008308-01 (YL)
 - NLM, R01 LM007659-01 (CF)
 - NLM 1U54 CA121852-01A1 National Center: Multi-Scale Study of Cellular Networks



Questions?