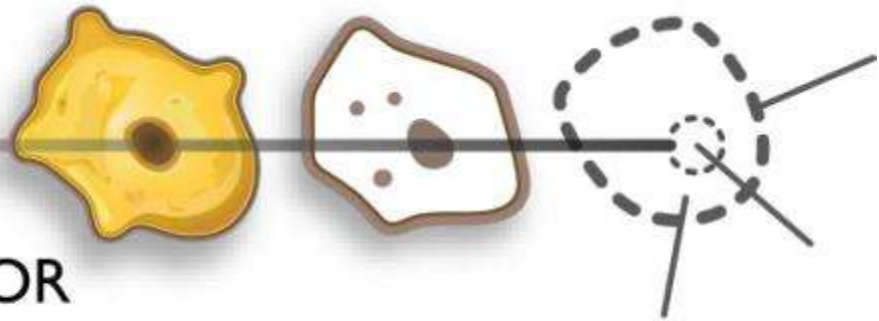


# Annotation Analytics for Gene and Protein functions

Nigam Shah, MBBS, PhD  
nigam@stanford.edu



NATIONAL CENTER FOR

**BIOMEDICAL ONTOLOGY**

# Annotation service

Process textual metadata to automatically tag text with as many ontology terms as possible.

GDS Summary			
Accession:	GDS906 <a href="#">View Expression (GEO profiles)</a>		
Title:	Bladder smooth muscle cell response to mechanical stretch		
DataSet type:	gene expression array-based (RNA / in situ oligonucleotide)		
Summary:	Expression profiling of cultured bladder smooth muscle cells subjected to repetitive mechanical stimulation for 4 hours. Chronic overdistension results in bladder wall thickening, associated with loss of muscle contractility. Results identify genes whose expression is altered by mechanical stimuli.		
Platform:	GPL96: Affymetrix GeneChip Human Genome U133 Array Set HU-U133A		
Citations:	Adam RM, Eaton SH, Estrada C, Nimgaonkar A et al. Mechanical stretch is a highly selective regulator of gene expression in human bladder smooth muscle cells. <i>Physiol Genomics</i> 2004 Dec 15;20(1):36-44. PMID: 15467014		
Sample organism:	Homo sapiens	Platform organism:	Homo sapiens
Feature count:	22283	Value type:	count
Series:	GSE1595	Series published:	07/25/2004
Last GDS update:	12/20/2004		

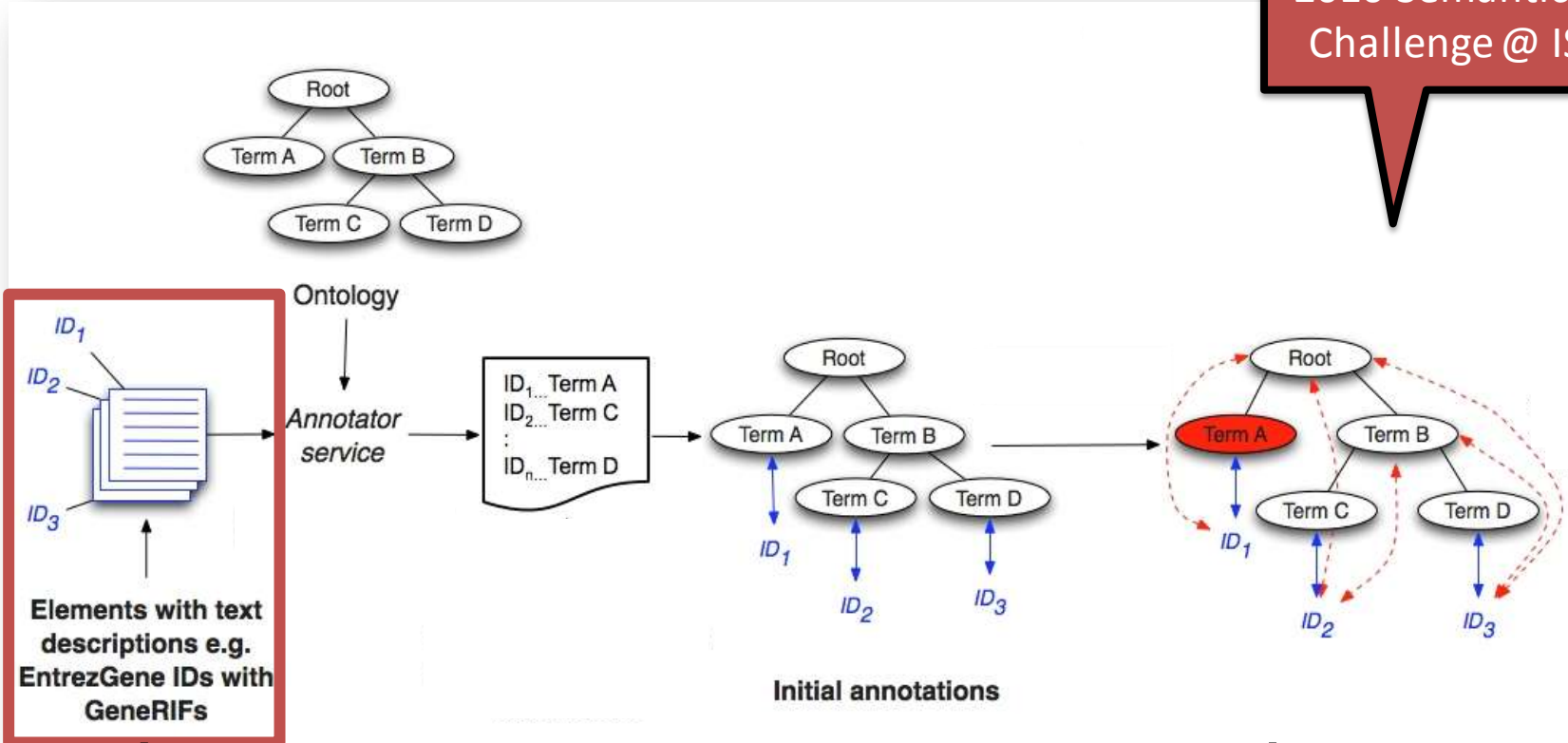


107 million calls,  
~1000 GB data

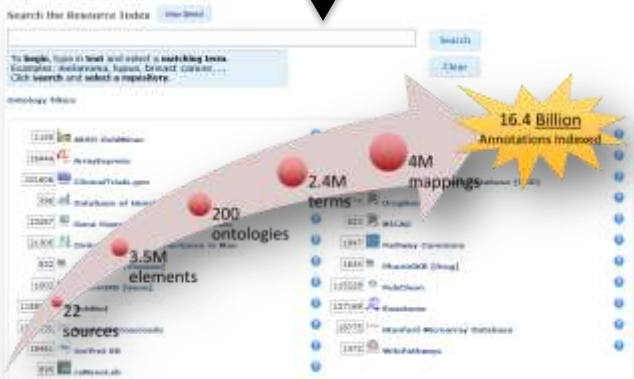
Expression, Expression of bladder, bladder, smooth, bladder muscle, muscle, smooth muscle, cells, mechanical, mechanical stimulation, stimulation, Chronic, results, bladder overdistension, associated, associated with, with, loss, genes, altered

# Resource index

Won 1<sup>st</sup> prize at the 2010 Semantic Web Challenge @ ISWC

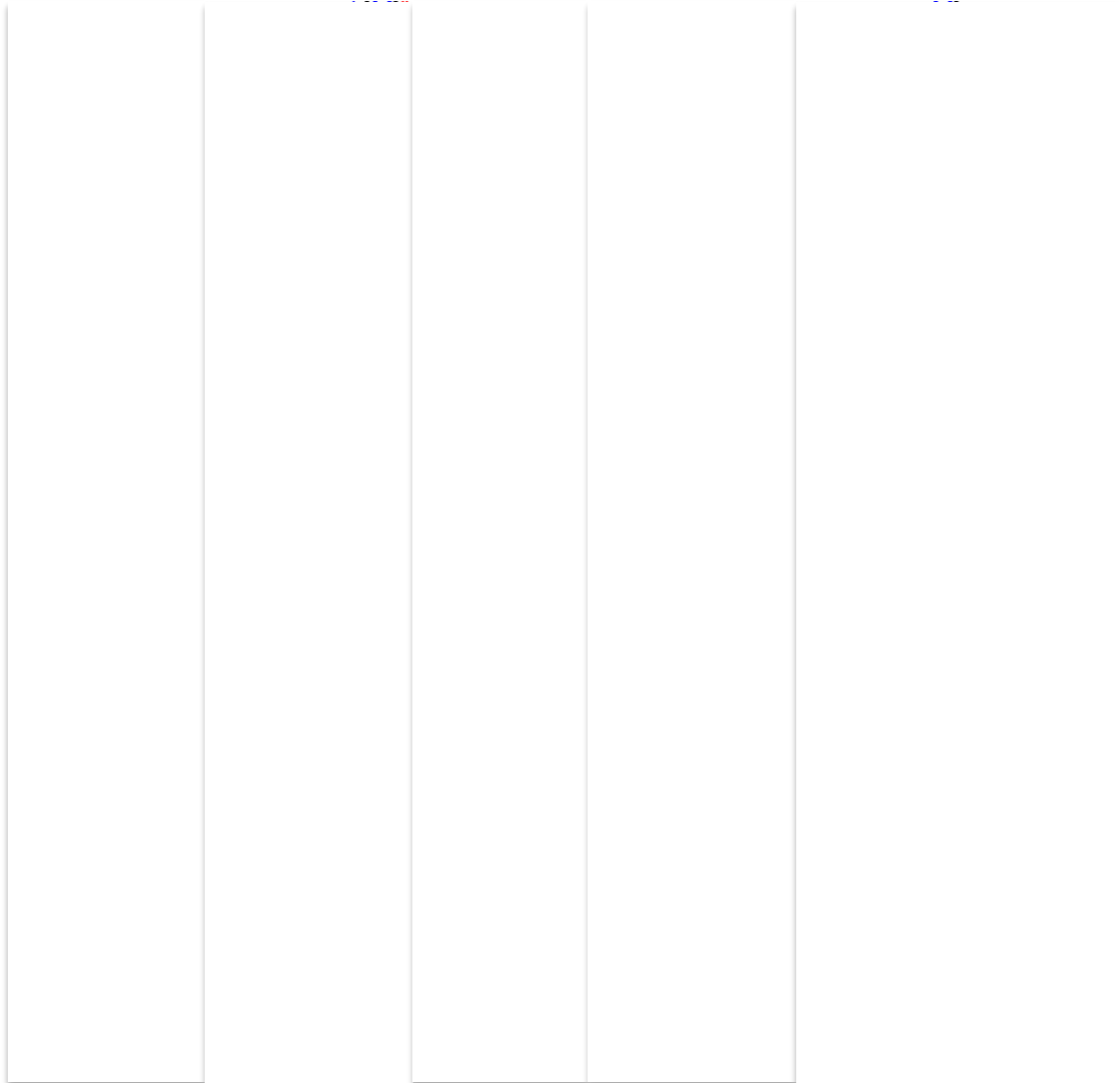


- Pubmed Abstracts
- Adverse Events (AERS)
- GEO
- ⋮
- Clinical Trials
- Drug Bank



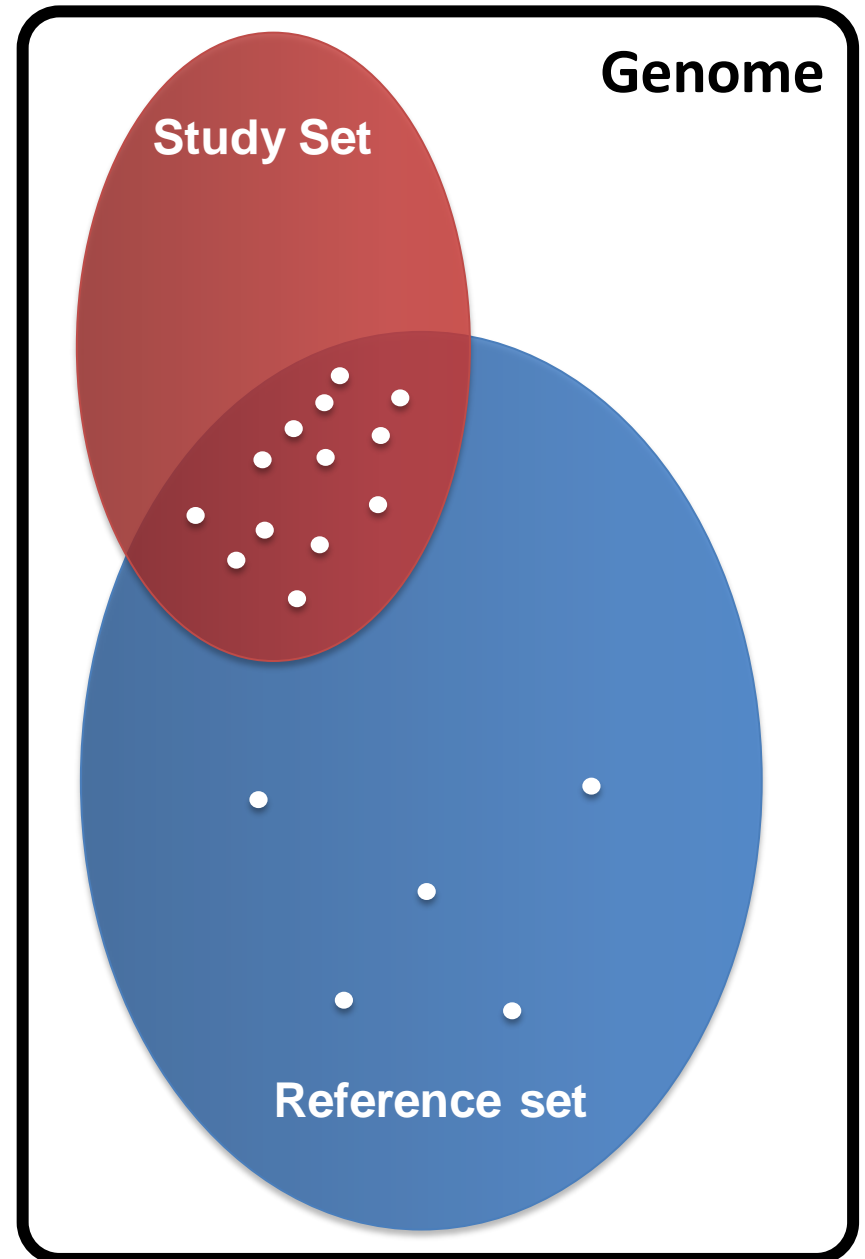
# Understanding the genome

- Units of study range in length from 'whole chromosome' to 'single nucleotide'
- E.g. three copies of Chr. 21 → Down's syndrome
- The focus is on finding the functional associations of strings in the genome

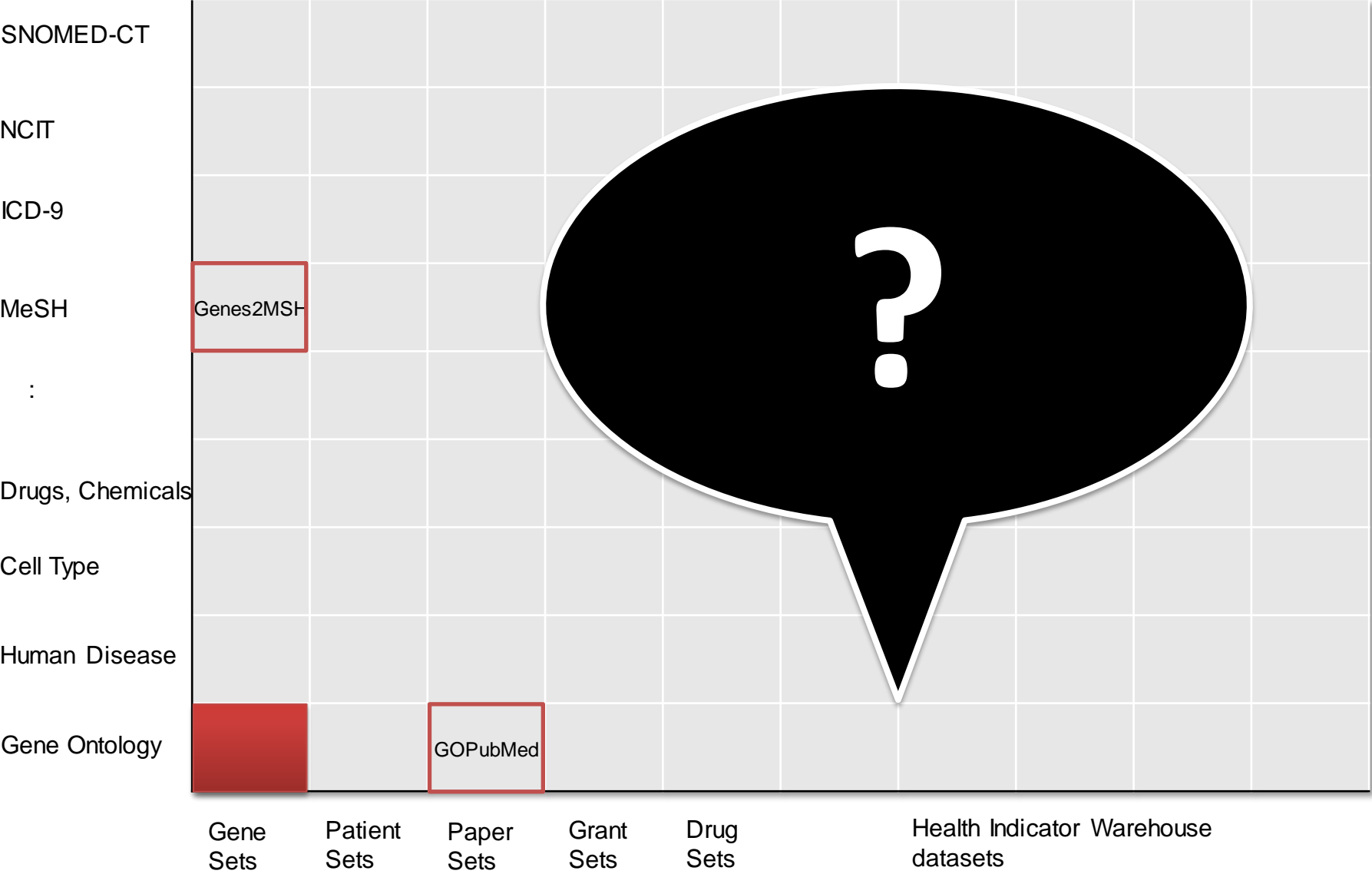


# Generic GO based analysis routine

- Get annotations for each gene in a set
- Count the occurrence of each annotation term in the study set
- Count the occurrence of that term in some reference set (whole genome?)
- P-value for how surprising their overlap is.

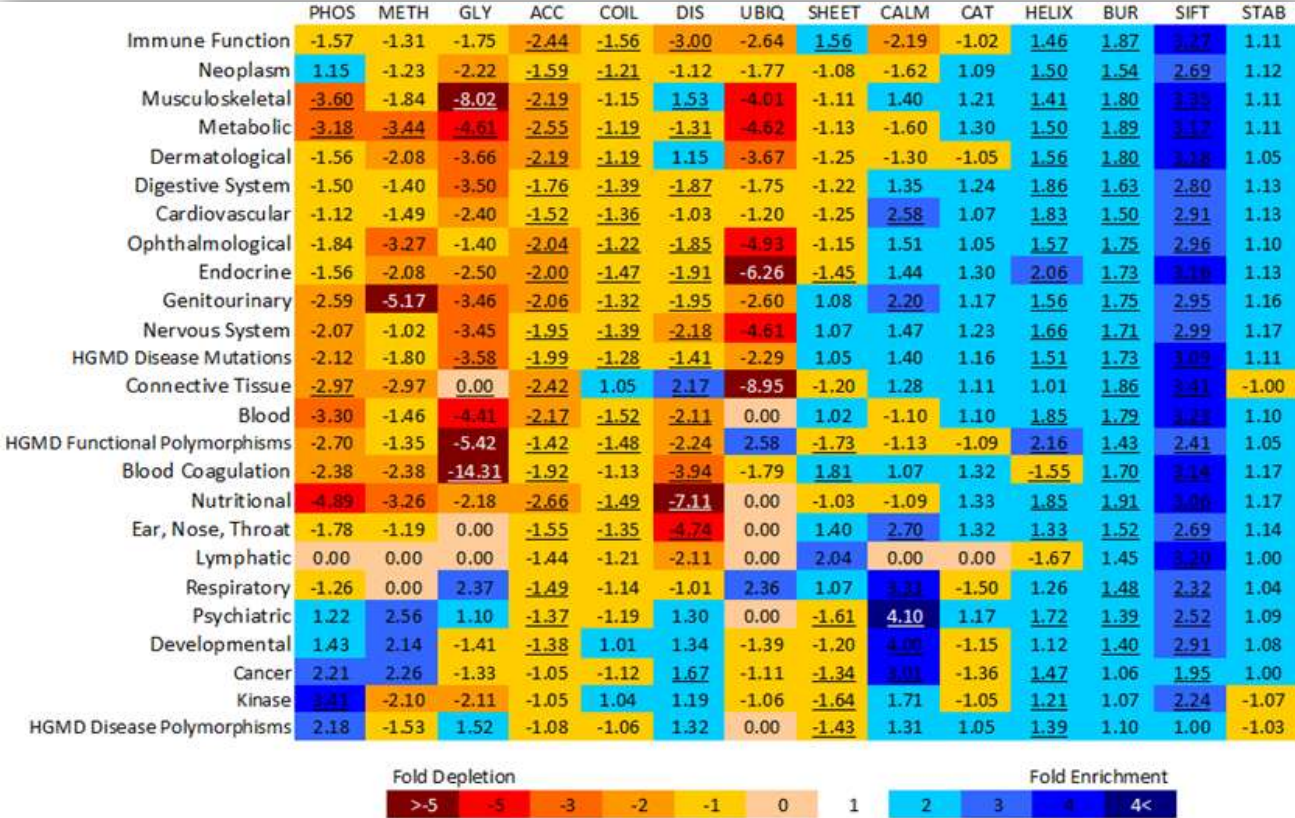
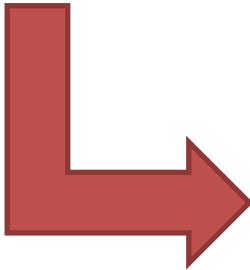


# Annotation Analytics Landscape

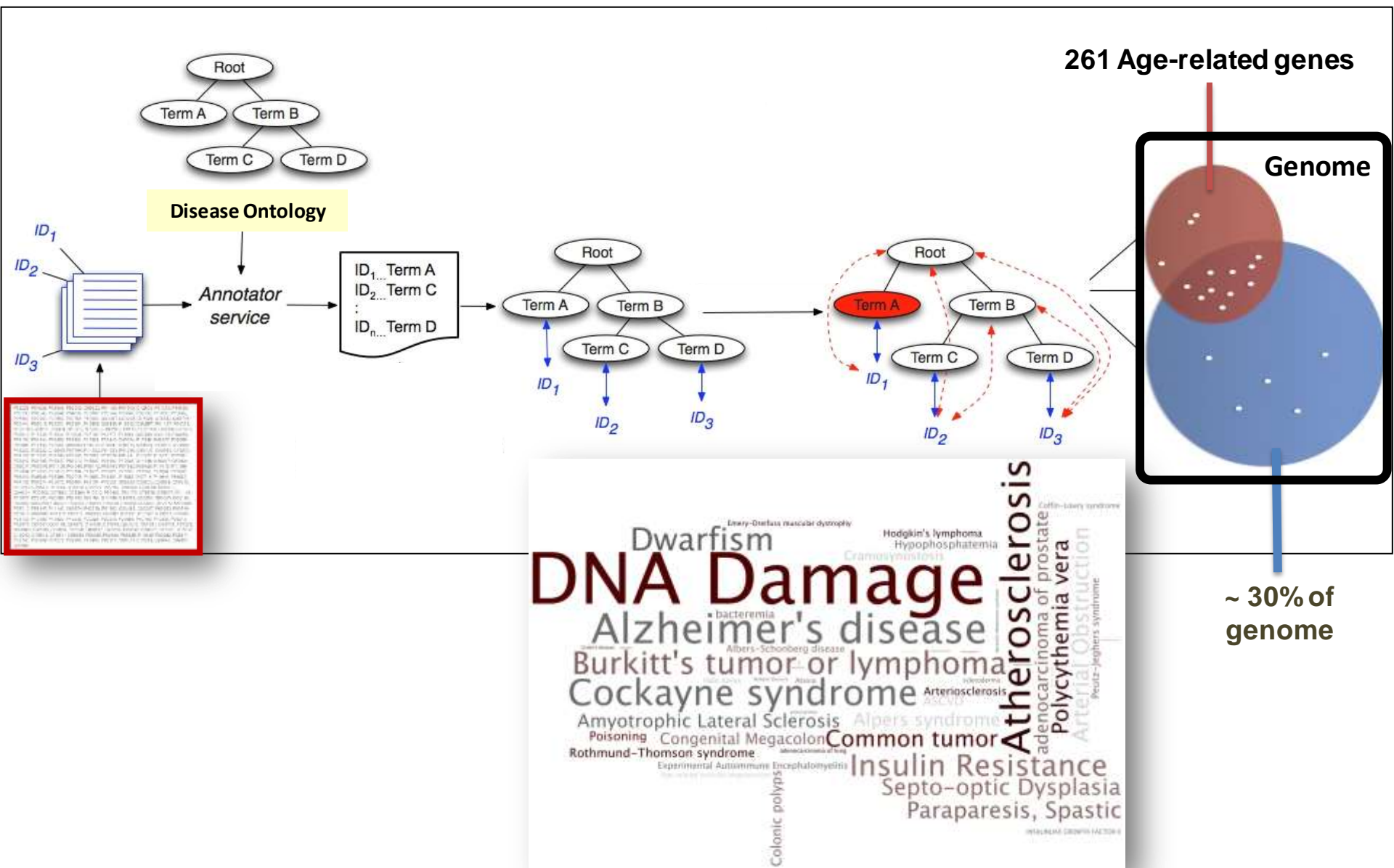


# Mutation enrichment

May have weak glycosidase activity towards glucuronylated steroids. However, it lacks essential active site Glu residues at positions 239 and 872, suggesting it may be inactive as a glycosidase in vivo. May be involved in the regulation of calcium and phosphorus homeostasis by inhibiting the synthesis of active vitamin D (By similarity). Essential factor for the specific interaction between FGF23 and FGFR1 (By similarity).  
 The Klotho peptide generated by cleavage of the membrane-bound isoform may be an anti-aging circulating hormone which would extend life span by inhibiting insulin/IGF 1 signaling (By similarity).

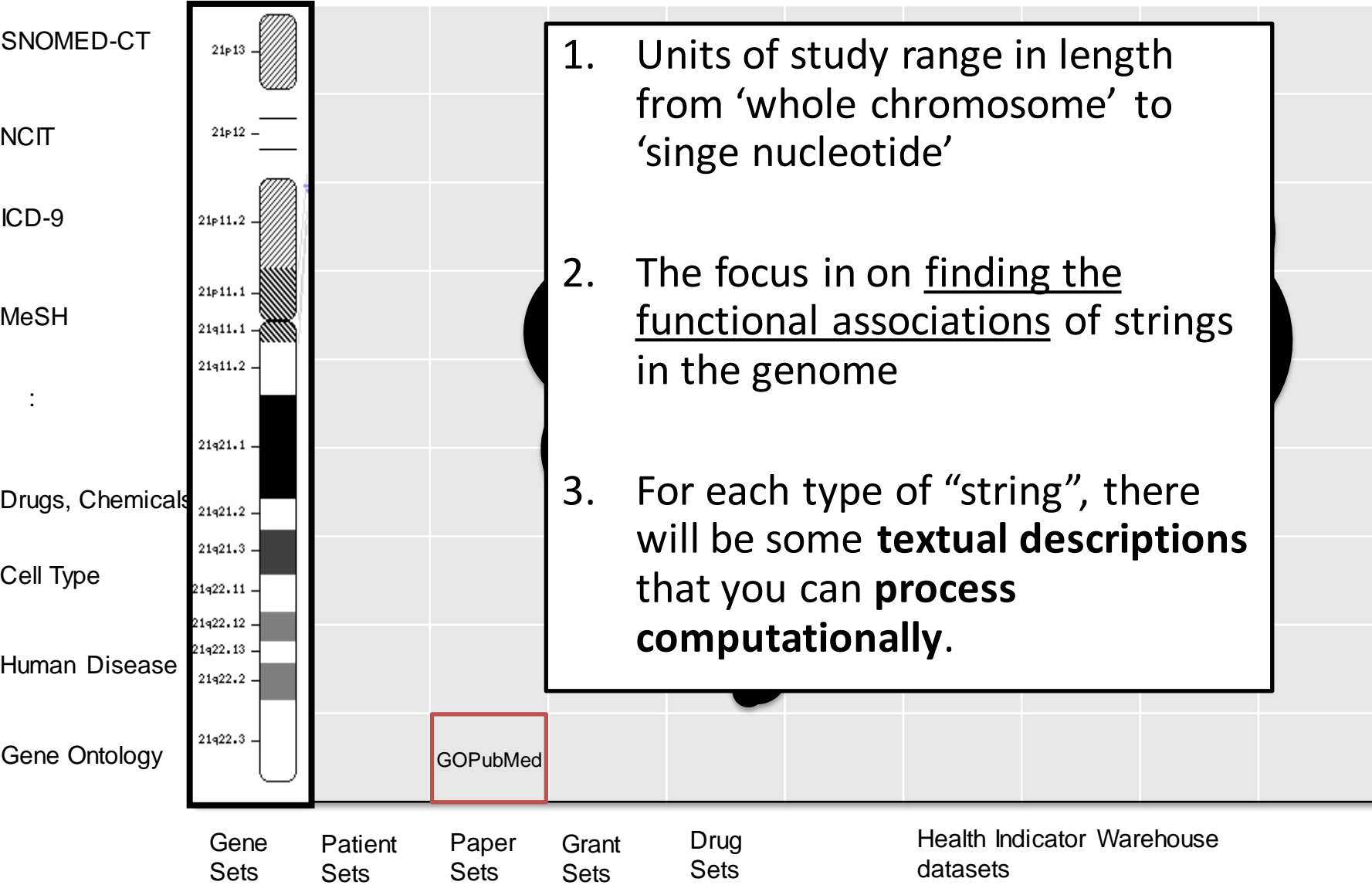


# Profiling a set of Aging genes





# Annotation Analytics Landscape



The team @  
[www.bioontology.org/project-team](http://www.bioontology.org/project-team)

NIH Roadmap grant U54 HG004028



STANFORD

SCHOOL OF MEDICINE