The 16th KJC Bioinformatics Symposium
Sokendai, Japan, August 29-31

Building a Knowledgebase for Precision Medicine

Lei Liu

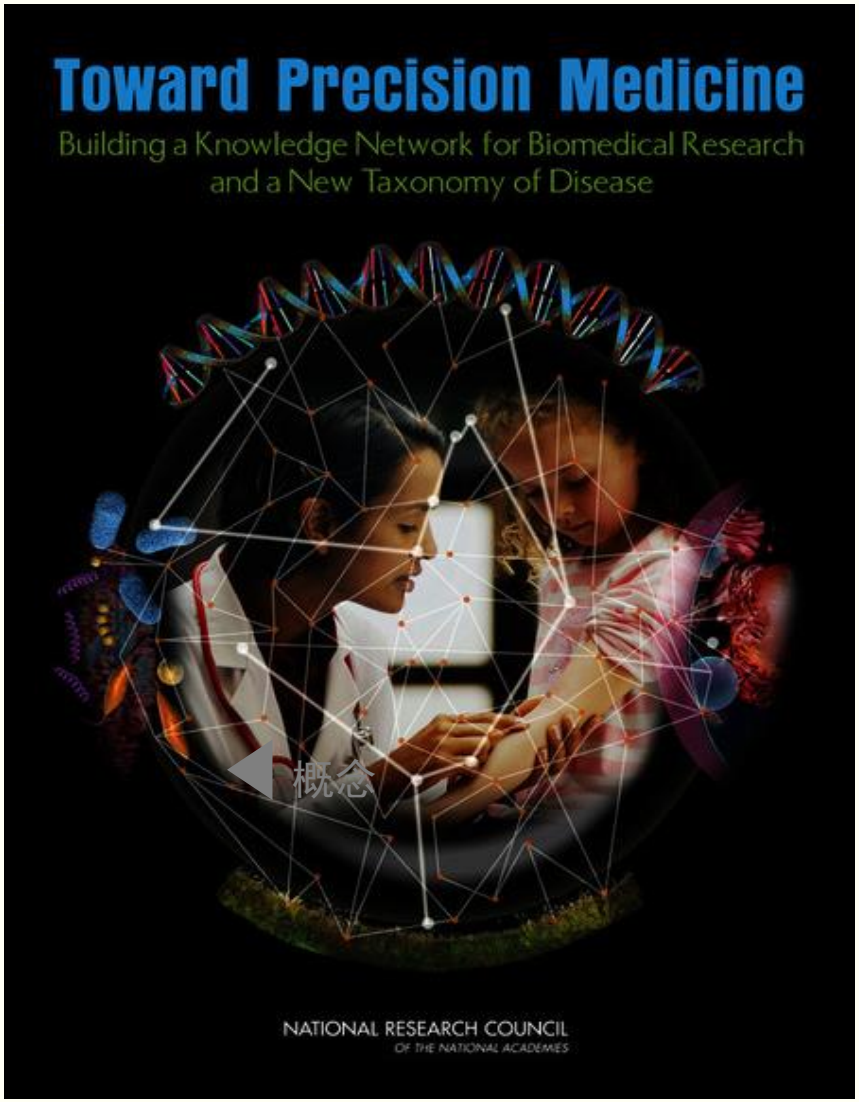
Institute of Biomedical Sciences

Fudan University

2018.8.30

Precision Medicine

NRC：《**Toward Precision Medicine**: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease》 November， 2011



Toward Precision Medicine

- 10 types of cancer
- Type 2 diabetes
- Rare diseases

Biospecimen

- Cohorts
- Clinical trials
- Biobank

Precision Medicine:

- Diagnosis
- Prognosis
- Therapy
- Monitoring
- Prevention

Phenotype

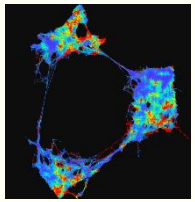
- EHR/EMR
- Clinical information
- Survival data
- Human Phenomics Project

Omics

- Genomics
- Transcriptomics
- Proteomics
- Metabolomics

Big Data

- Data quality
- Standardization
- Data mining
- SOPs



China Precision Medicine Initiative

Cohort Studies

- Genotypes
- Phenotypes
- Environment
- Life Styles

...

Omics Analysis

- Genomics
- Proteomics
- Metabolomics

...

Large Samples, Big Data

Knowledgebase

Heterogeneous
Huge Data Sets

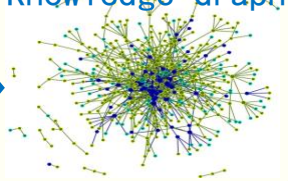
Ontologies

Literatures

Databases

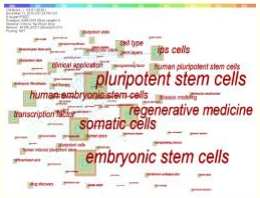
Data Curation

Knowledge Graph



Applications in Precision Medicine

Research

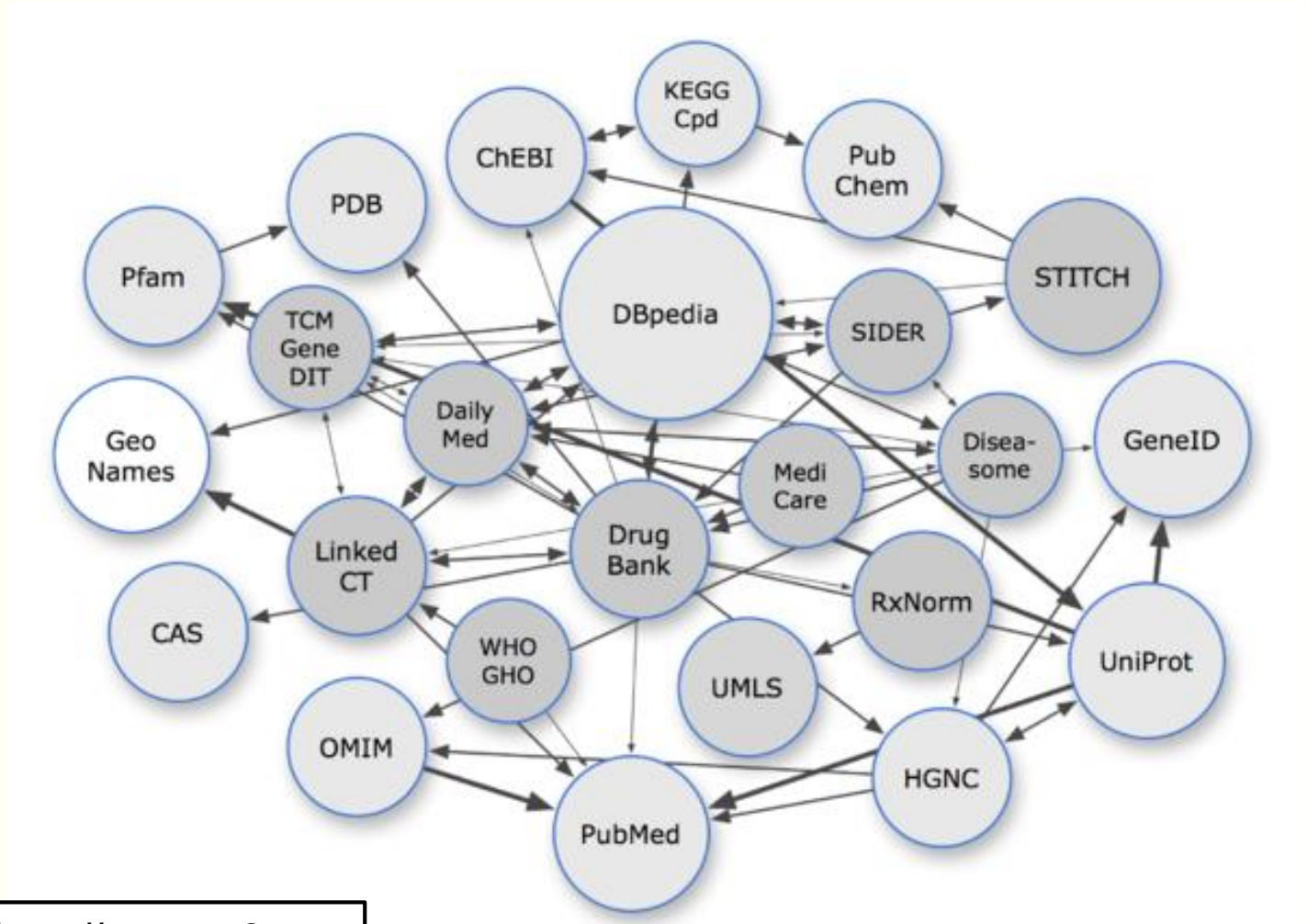


Search Engine
Data Analysis
Discovery
...

Clinical

- Diagnosis
- Prognosis
- Treatment
- ...

Linked Open Data in Biomedince



Source: http://www.w3.org

National Key R&D Plan Project:

Knowledgebase of Precision Medicine for Disease Studies

Lead Institute：Fudan University

Lead PI：Lei Liu

Participants：

Duration：2016-2020

Institute of Medical Informatics, Chinese Academy of Medicine

Chinese Military Academy of Medicine

National Center for Protein Sciences

Shanghai Institute of Life Sciences, Chinese Academy of Science

Beijing Institute of Genomics, Chinese Academy of Science

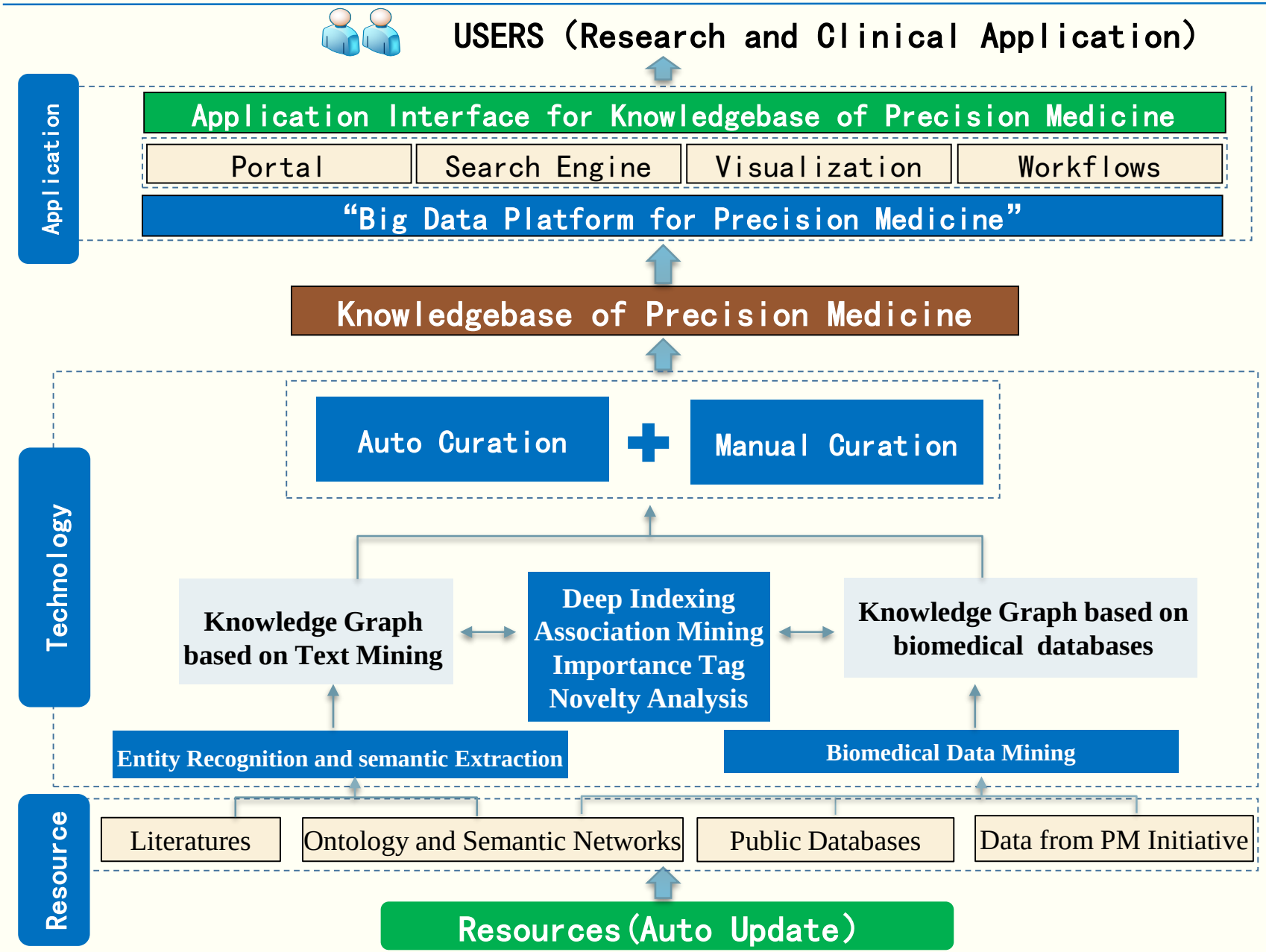
Harbin Institute of Technology

Zhejiang University

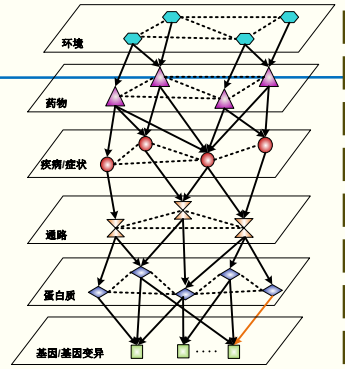
Dalian University of Technology

Shanghai Center for Bioinformation Technology

Implementation Schema



PM Ontology and Semantic Networks



面向恶性肿瘤、代谢系统疾病、呼吸系统疾病、心脑血管疾病等重大疾病

- reformulated to
- tradename of
- has precise ingredient
- ingredient of
- ……

Drug

- inhibitor
- partial agonist
- activator
- stimulator
- reducer
- ……

- inhibitor
- substrate
- encoded by
- ……

- Gene has Mutation
- Mutation has Size
- ……

Genome

- transcribe
- regulate

Proteome

- treats
- disrupts
- ……

- positively regulates
- negatively regulates
- Mutation related to Disease

Symptom & Disease

Environment

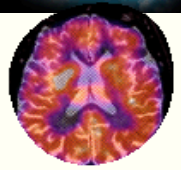
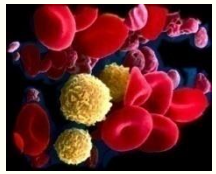
- is associated with
- interacts with
- has part
- ……

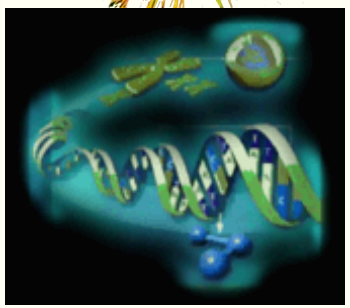
- is a
- affects
- complicates
- co-occurs with

Ontologies at Different Levels

Biology

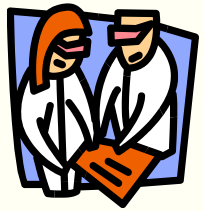
Length Metric
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}

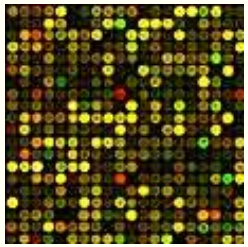






ICD10(疾病）
SNOMED(疾病)
NCBI 分类系统
SNOMED (器官)

哺乳动物的表型
SNOMED (形态学)
ATCC (细胞株)
Cell Ontology细胞本体（细胞类型）
Gene Ontology基因本体（亚细胞）
Gene Nomenclature 基因命名
Quaternary code 四进制码



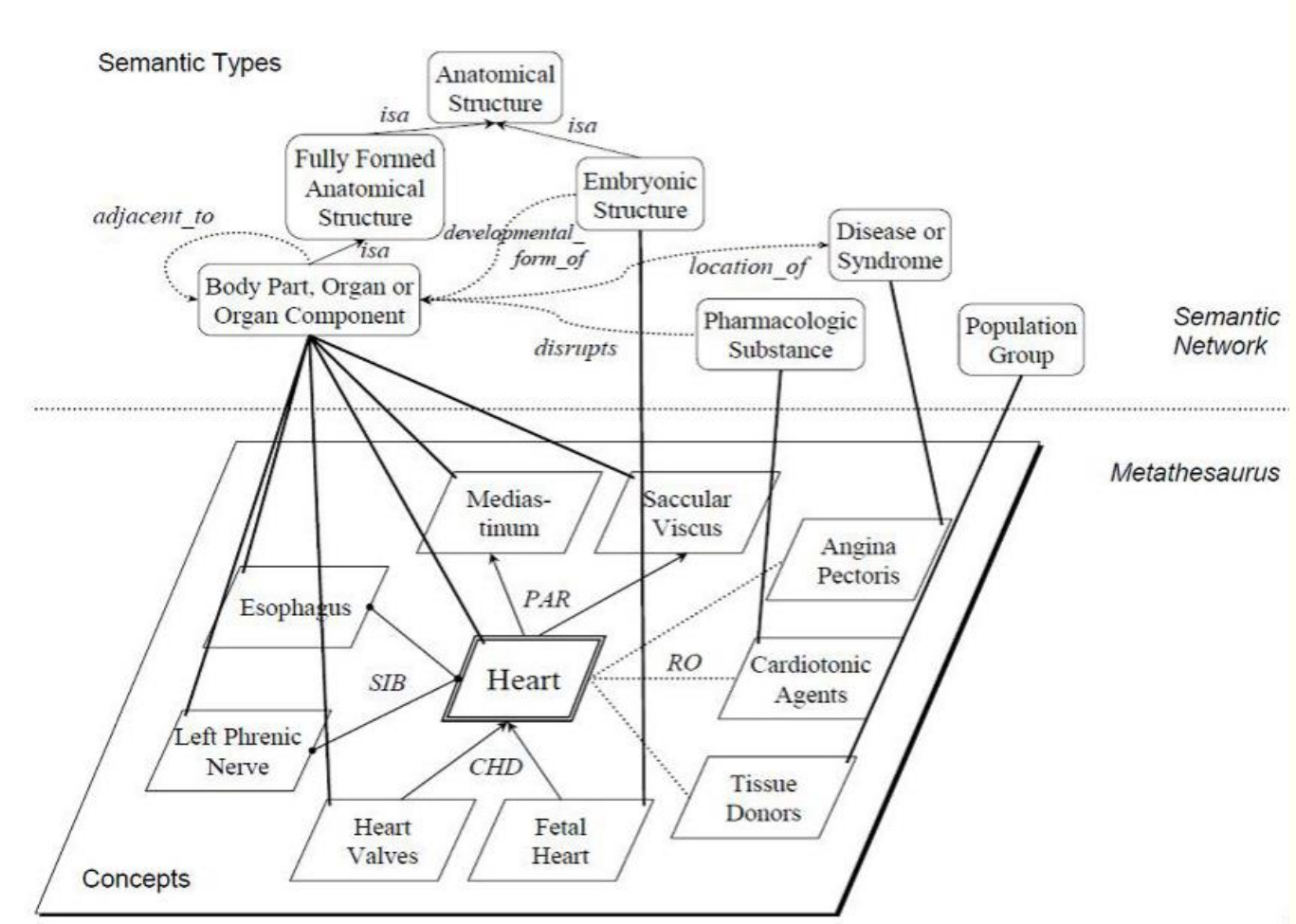



Technology

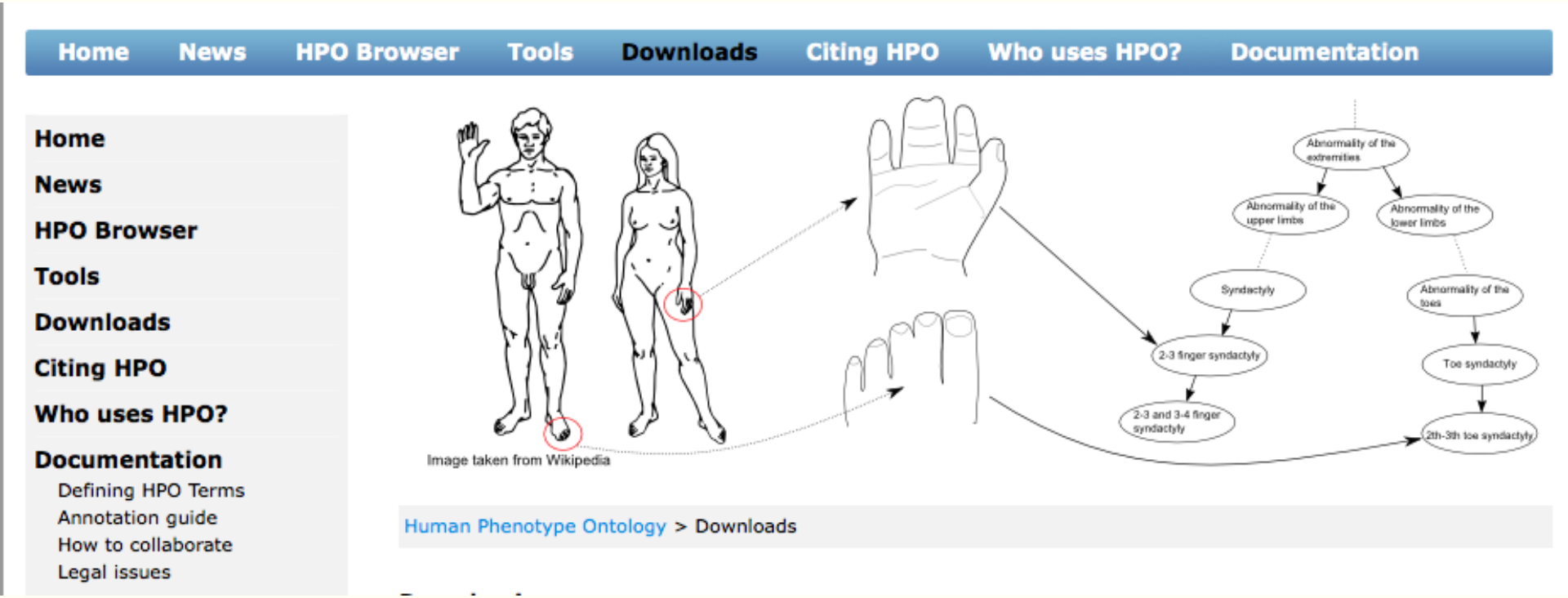
Population
Individual
Organ
Tissue
cell
subcellular
Virus
DNA
Bases

UMLS Semantic Network

- **UMLS** = Unified Medical Language System (NLM)
- Composed of:
  - Metathesaurus
  - Semantic Network
  - Lexicon
- Contains approximately 5 million codes representing 1 million concepts derived from 100 source terminologies



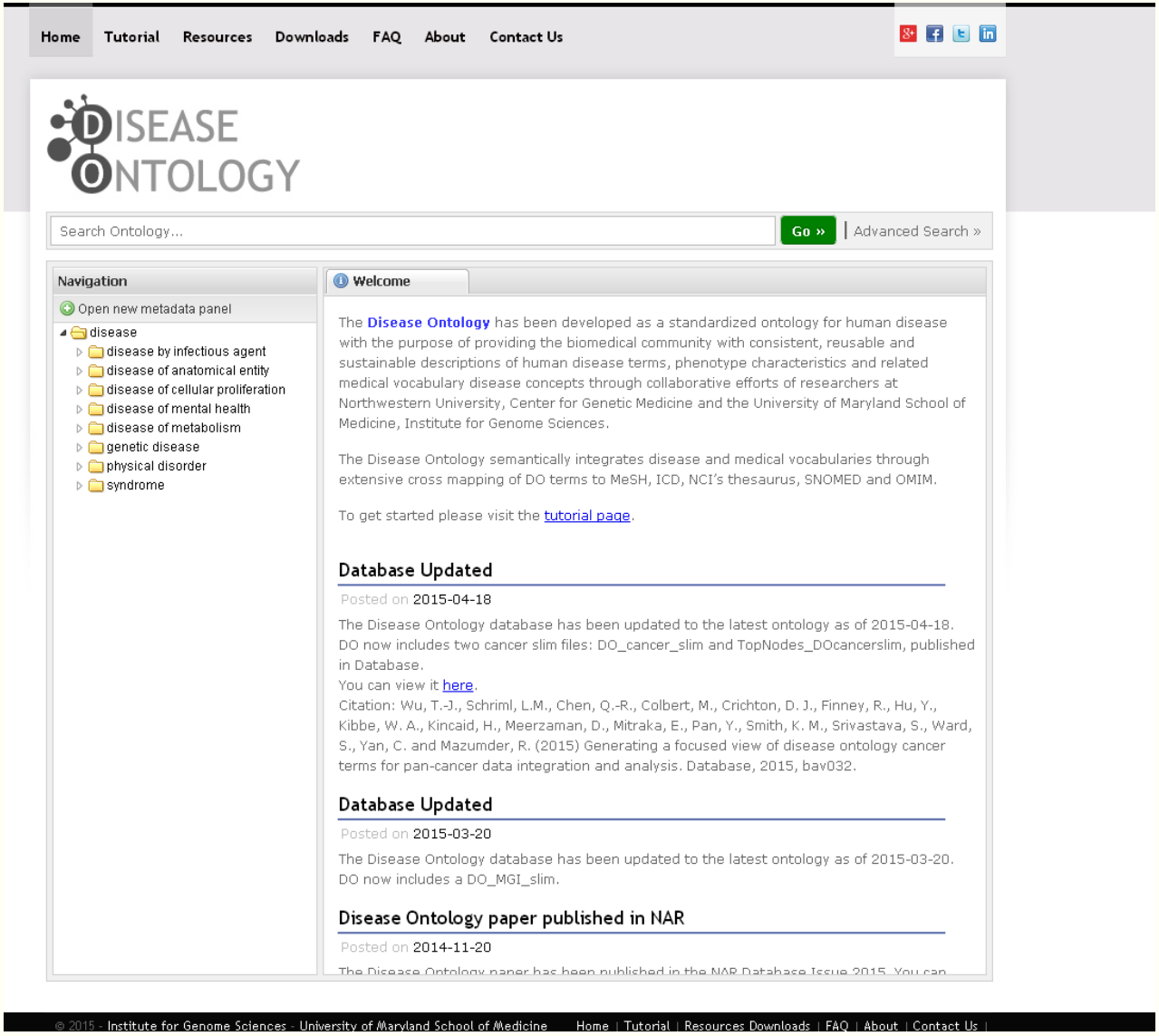
HPO: human phenotype ontology



Structure between phenotype
Relationship between phenotype and disease

http://www.human-phenotype-ontology.org/

DO: Disease ontology



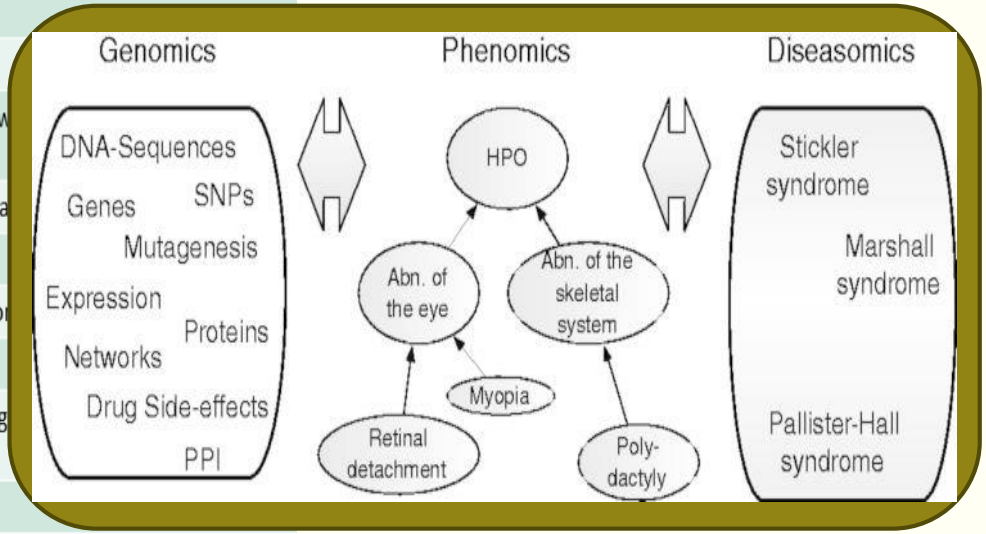
http://disease-ontology.org/

Gene ontology

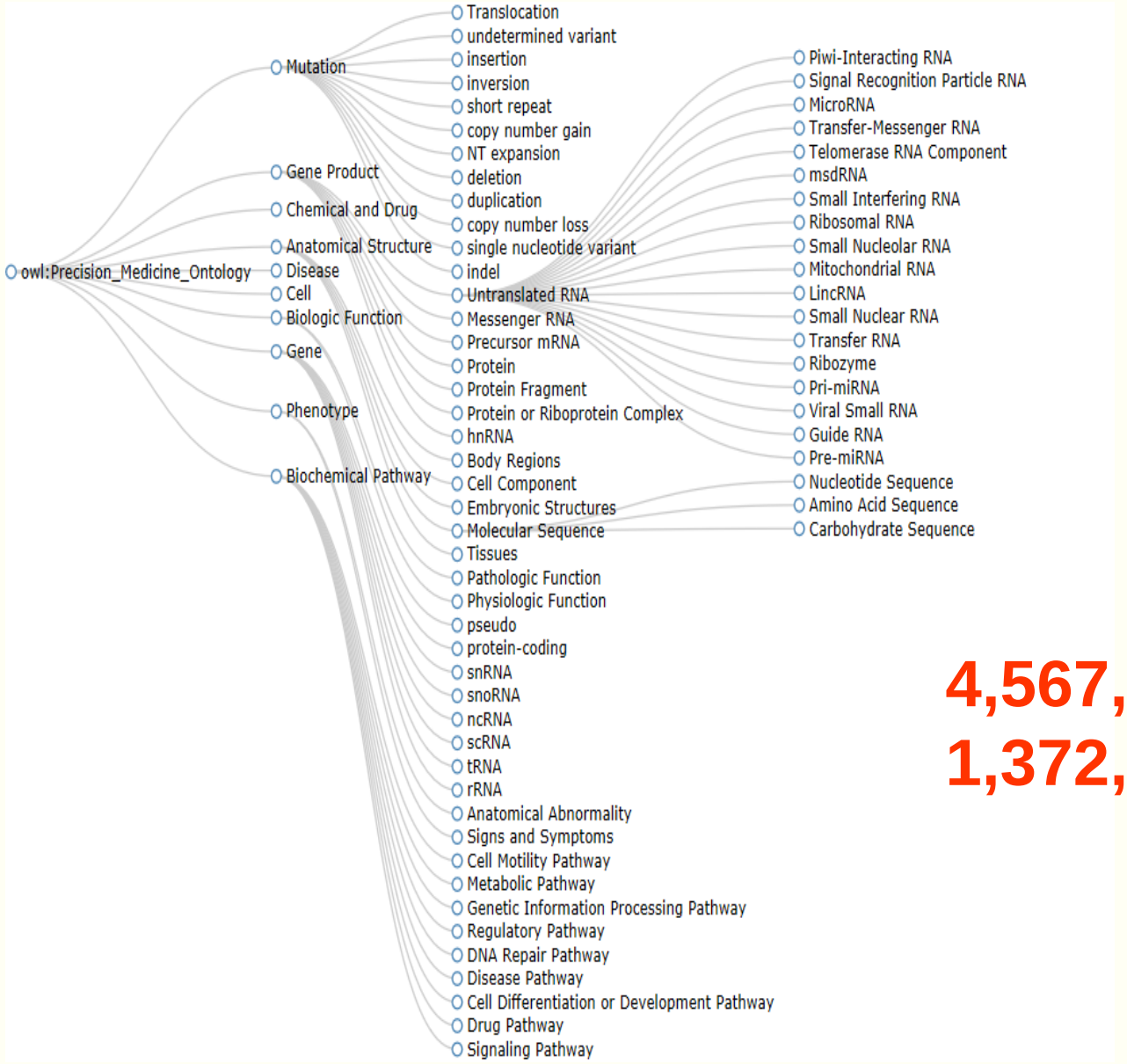
Human phenotype ontology

Disease ontology

Relationship	Definition
Gene has Mutation	A mutation occurs in or near a gene, usually at a given position.
Patient/Cohort has Mutation	A patient or cohort has a specific genetic variation.
Mutation related to Disease	A mutation is associated with (or causes) a disease.
Mutation has Size	Indicates the number or frequency of mutations.
Disease related to Gene	A disease is associated with a gene—that is, a gene (w causes a disease.
Disease related to Body Part	A disease may occur in a body part, or have a body pa
Patient has Age	A patient has a given age
Cohort has Age	A summary age for a cohort. Often listed as a mean or
Patient/Cohort has Gender	A patient or cohort is male or female.
Patient/Cohort has Geographic Location	A patient or cohort has a given ethnicity or lives in a g
Patient/Cohort has Disease	A patient or cohort has a disease.
Cohort has Size	The size of a cohort group



PMOV2.0——2018.05



58 ontologies, thesaurus, controlled vocabularies

SNOMED CT

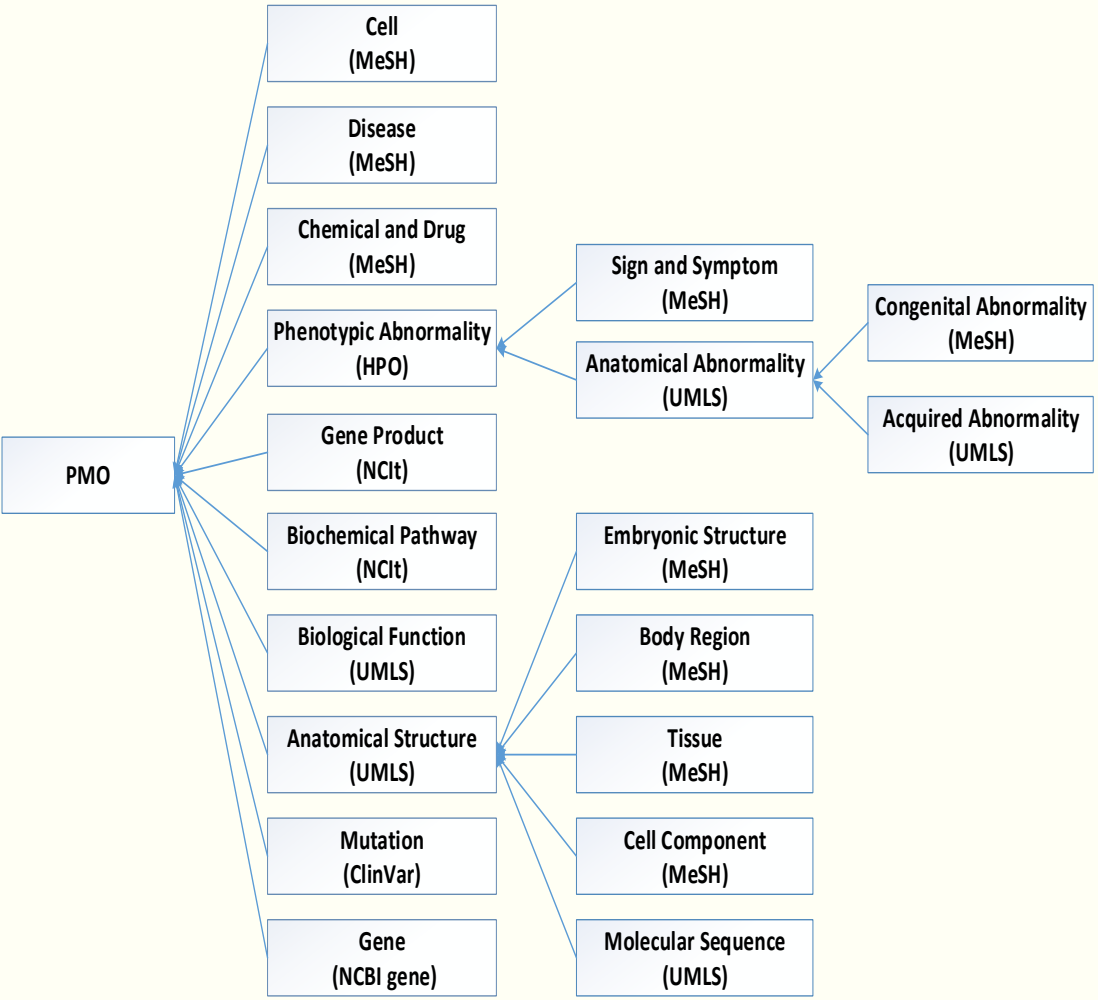
ICD-10 Version:2016

NCI thesaurus

MeSH Medical Subject Headings

4,567,208 terms;
1,372,967 concepts

PMO: sources and statistics



NUMBER OF CLASSES	15621
NUMBER OF INSTANCES	707892
NUMBER OF PROPERTIES	93
MAXIMUM DEPTH	13
MAXIMUM NUMBER OF CHILDREN	102
AVERAGE NUMBER OF CHILDREN	4
CLASSES WITH A SINGLE CHILD	1702
CLASSES WITH MORE THAN 25 CHILDREN	83

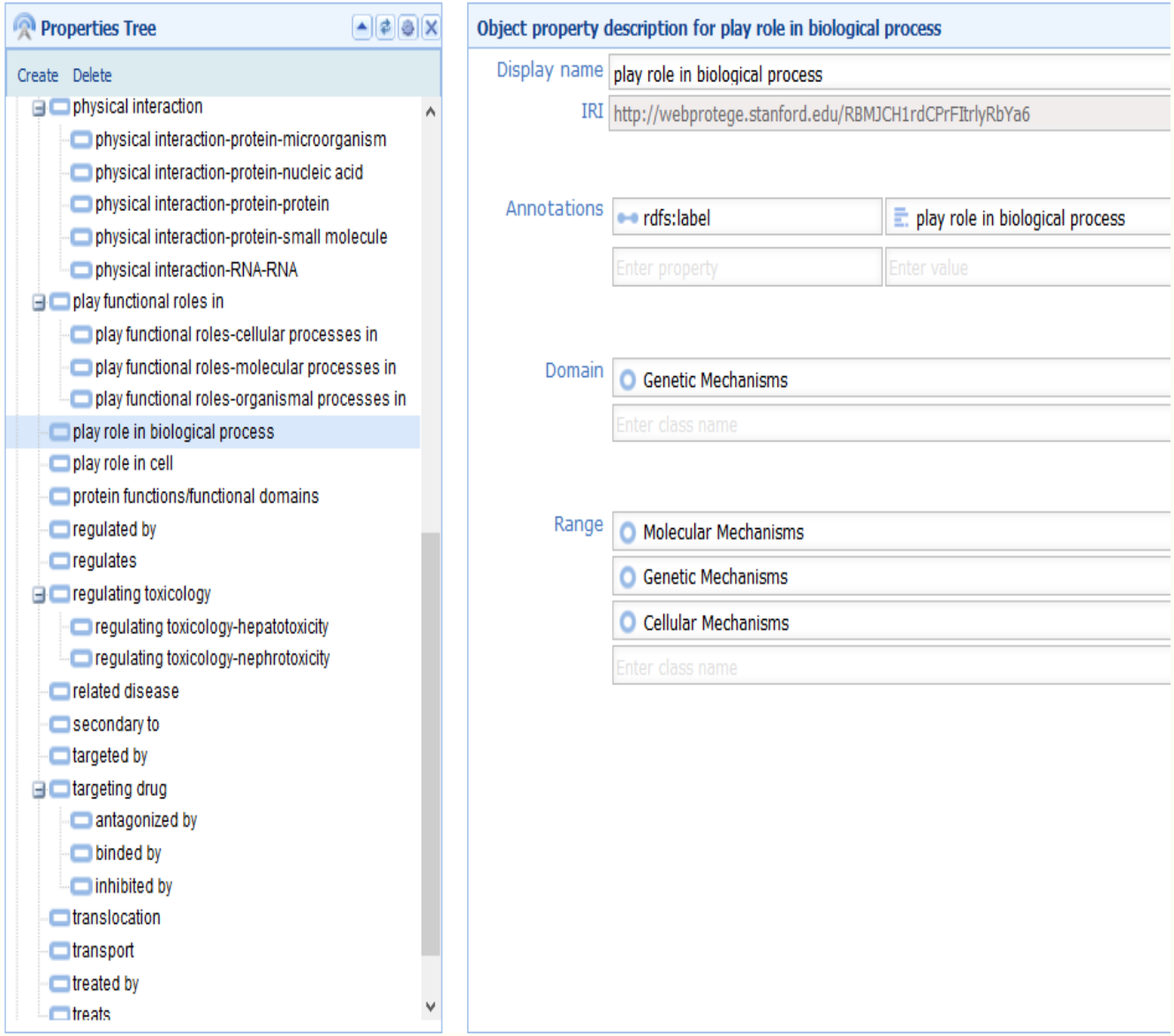
Concept types related to genetic mutations

Concept	Definition
Gene	A segment of DNA that codes for a protein
Mutation	A mutation is an alteration (deletion, insertion, substitution) of nucleotides (DNA, RNA) or amino acids (Protein)
Body part	An organ or anatomical location in a person.
Disease	An abnormal condition affecting the body of an organism.
Patient	An individual with a disease
Cohort	A group of people; specifically any group or population of people that may be assigned a disease or characteristic.
Size	A number indicating the number of people in a *cohort*, or the number/frequency of a *mutation*.
Age	A number or range indicating how old a person/group of people is.
Gender	Terms indicating whether someone is male or female
Geographical location	Terms indicating where a person/group of people comes from, either based on ethnic origin or where they live.
……	……

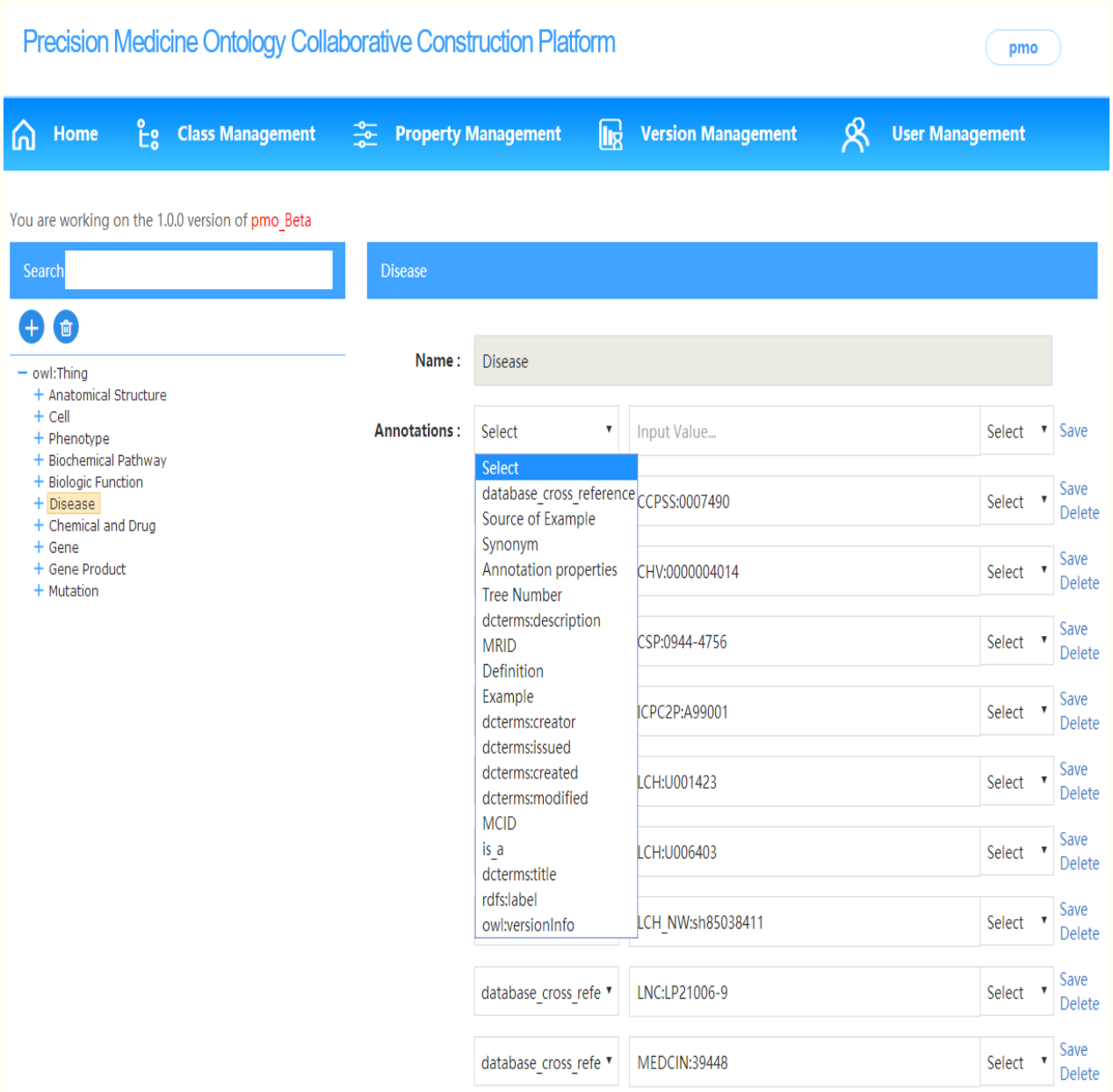
- Verspoor, K., A. Jimeno Yepes, L. Cavedon, T. McIntosh, A. Herten-Crabb, Z. Thomas and J. P. Plazzer (2013). "Annotating the biomedical literature for the human variome." Database (Oxford) **2013: bat019.**

Relationship	Definition
Gene has Mutation	A mutation occurs in or near a gene, usually at a given position.
Patient/Cohort has Mutation	A patient or cohort has a specific genetic variation.
Mutation related to Disease	A mutation is associated with (or causes) a disease.
Mutation has Size	Indicates the number or frequency of mutations.
Disease related to Gene	A disease is associated with a gene—that is, a gene (when mutated) is linked to, or causes a disease.
Disease related to Body Part	A disease may occur in a body part, or have a body part in its name.
Patient has Age	A patient has a given age
Cohort has Age	A summary age for a cohort. Often listed as a mean or an age limit.
Patient/Cohort has Gender	A patient or cohort is male or female.
Patient/Cohort has Geographic Location	A patient or cohort has a given ethnicity or lives in a given place.
Patient/Cohort has Disease	A patient or cohort has a disease.
Cohort has Size	The size of a cohort group
……	……

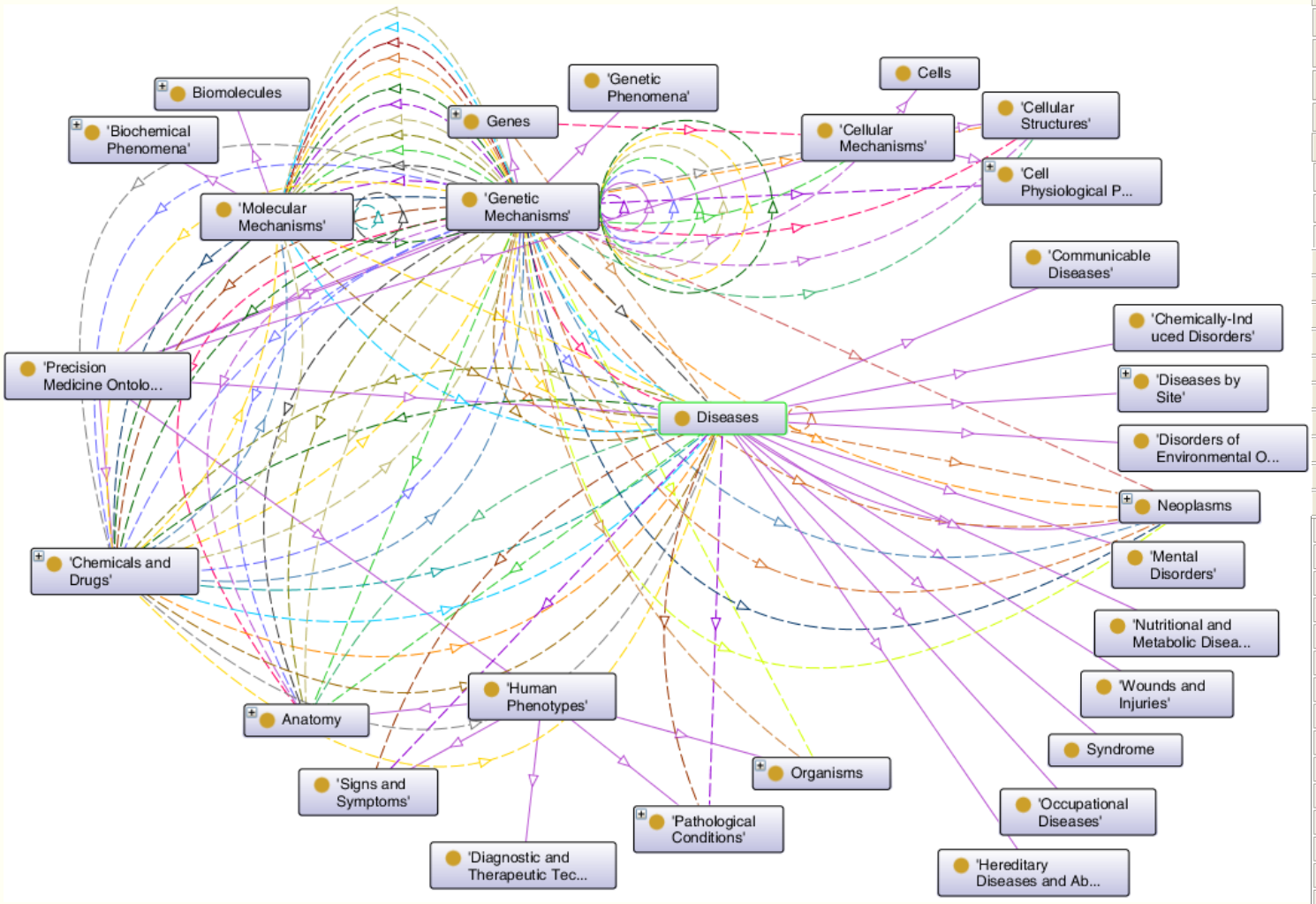
Semantic Relationships in PMO



Precision Medicine Ontology Collaborative Construction Platform

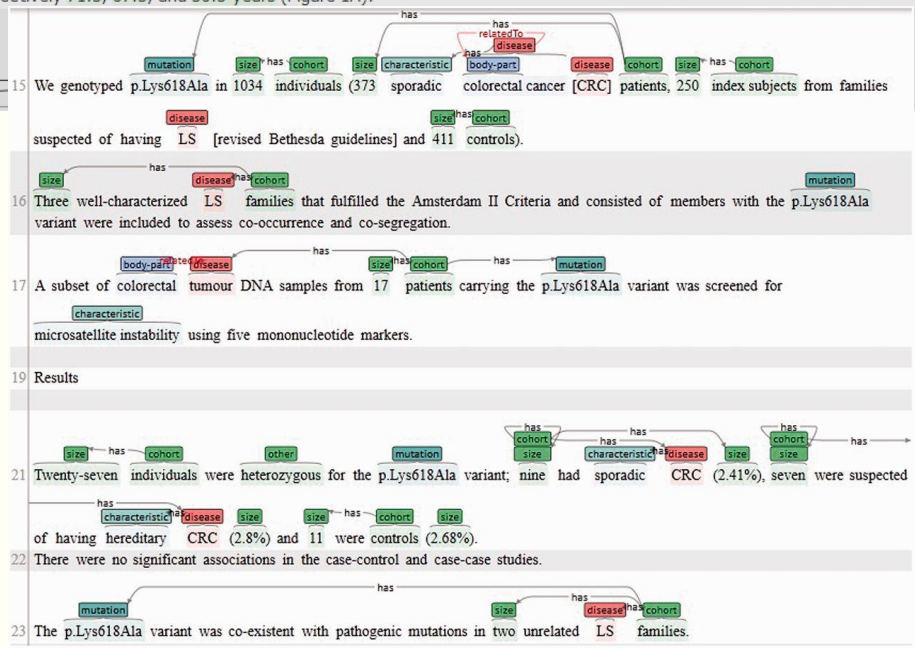


Complex Semantic Relationship



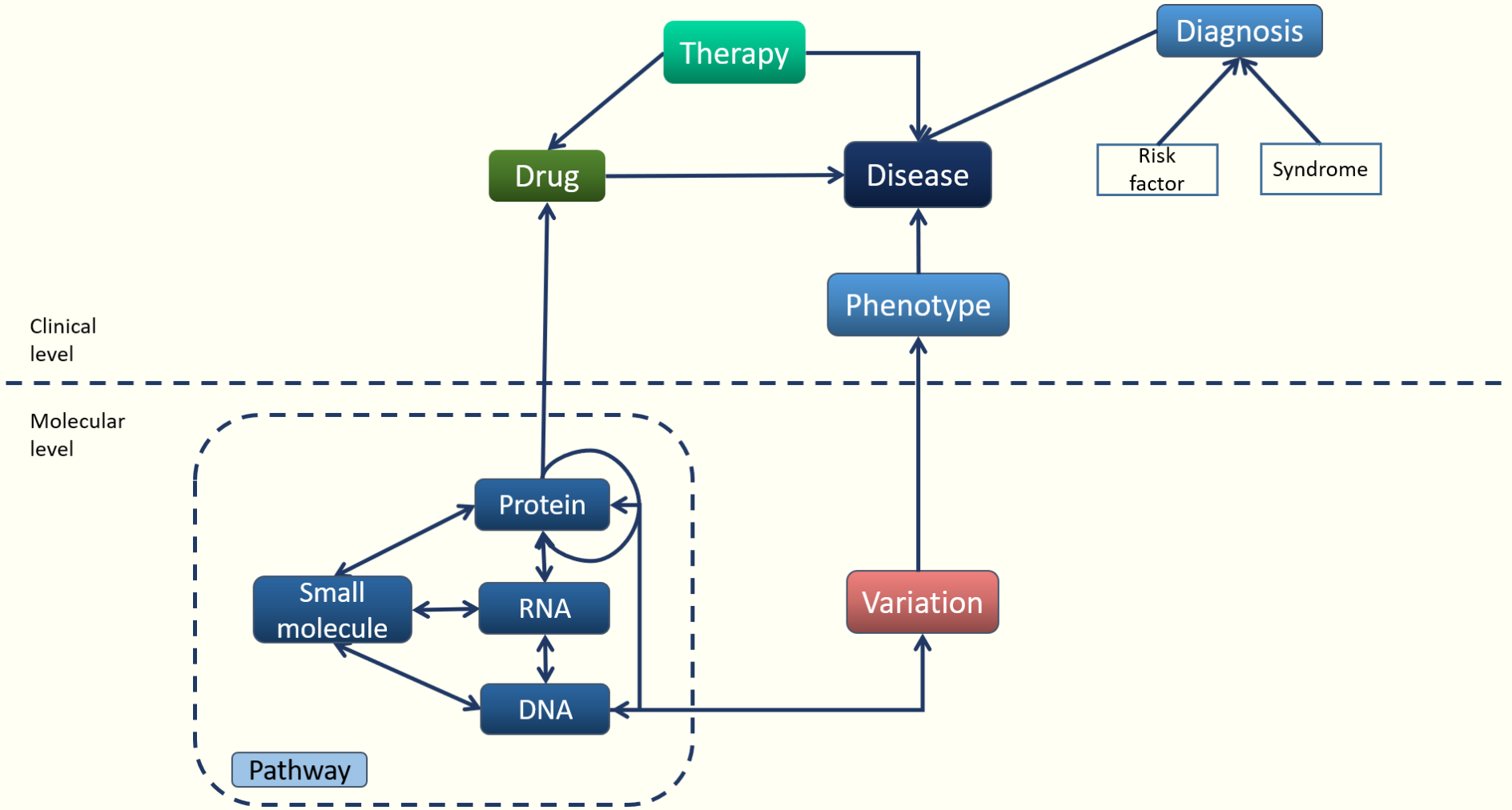
- 'antagonized by' (Domain>Range)
- 'binded by' (Domain>Range)
- 'express and localize' (Domain>Range)
- 'express and localize-biofluid' (Domain>Range)
- 'express and localize-cell or tissue expression' (Domain>Range)
- 'express and localize-subcellular location' (Domain>Range)
- 'express in disease' (Domain>Range)
- 'expression regulated by' (Domain>Range)
- 'has clinical finding' (Domain>Range)
- 'has finding site' (Domain>Range)
- 'has molecular function' (Domain>Range)
- 'has subcellular location' (Domain>Range)
- 'inhibited by' (Domain>Range)
- 'is a member of' (Domain>Range)
- 'is biomarker of' (Domain>Range)
- 'play role in cell' (Domain>Range)
- 'protein functions/functional domains' (Domain>Range)
- 'regulated by' (Domain>Range)
- 'regulating toxicology' (Domain>Range)
- 'regulating toxicology-hepatotoxicity' (Domain>Range)
- 'regulating toxicology-nephrotoxicity' (Domain>Range)
- 'related disease' (Domain>Range)
- 'targeted by' (Domain>Range)
- 'targeting drug' (Domain>Range)
- 'treated by' (Domain>Range)
- 'is biomarker-diagnosis of' (Domain>Range)
- 'is biomarker-disease progression of' (Domain>Range)
- 'is biomarker-prognosis of' (Domain>Range)
- 'is biomarker-response to therapy of' (Domain>Range)
- 'modificated and regulated by' (Domain>Range)
- 'phosphorylation regulated by' (Domain>Range)
- 'physical interaction' (Domain>Range)
- 'physical interaction-protein-microorganism' (Domain>Range)
- 'physical interaction-protein-nucleic acid' (Domain>Range)
- 'physical interaction-protein-protein' (Domain>Range)
- 'physical interaction-protein-small molecule' (Domain>Range)
- 'physical interaction-RNA-RNA' (Domain>Range)
- 'play functional roles in' (Domain>Range)
- 'play functional roles-cellular processes in' (Domain>Range)
- 'play functional roles-molecular processes in' (Domain>Range)
- 'play functional roles-organismal processes in' (Domain>Range)
- 'play role in biological process' (Domain>Range)
- binding (Domain>Range)
- decreases (Domain>Range)

complex semantic relationships embedded in text

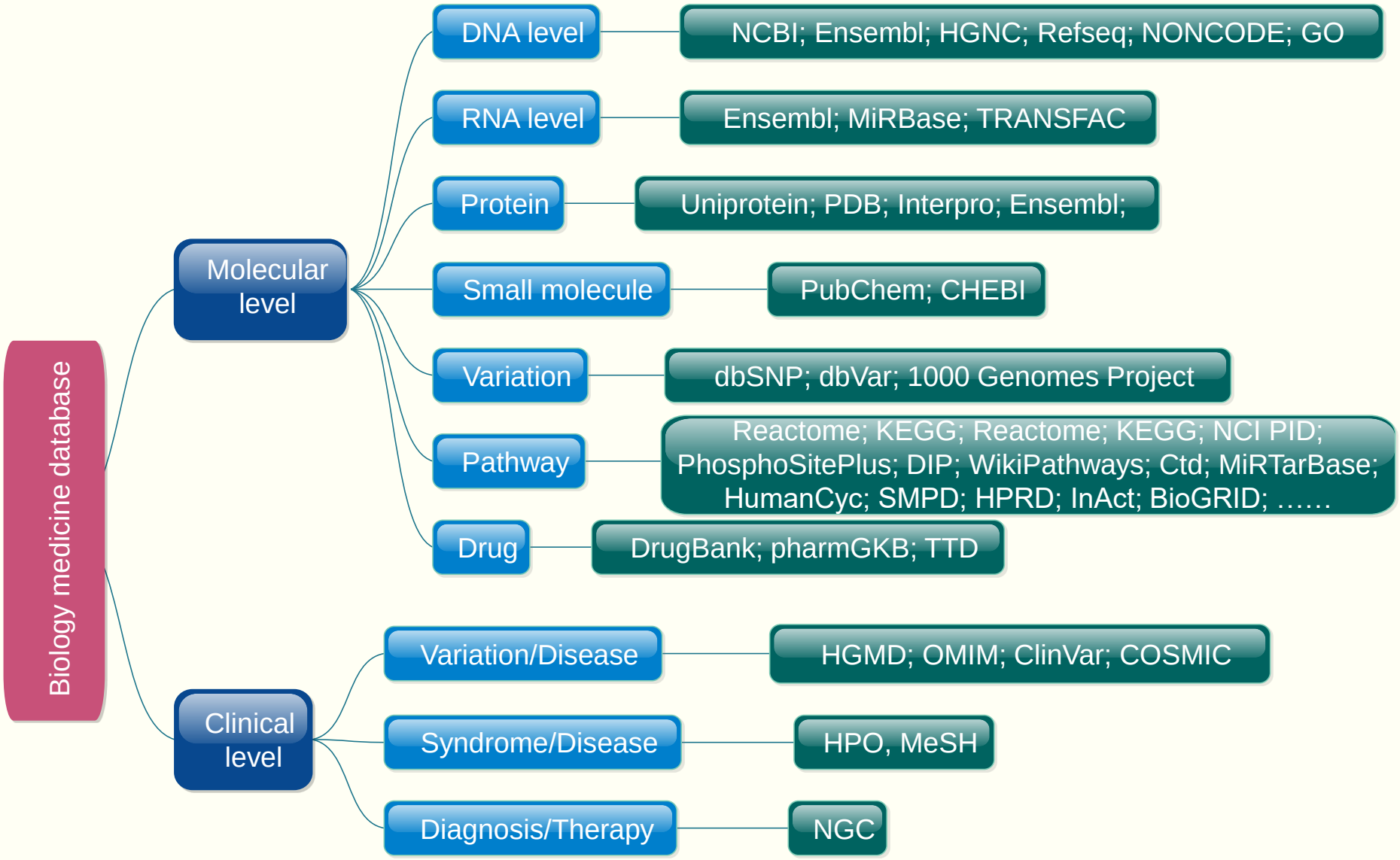


http://www.opennicta.com.au/home/health/variome

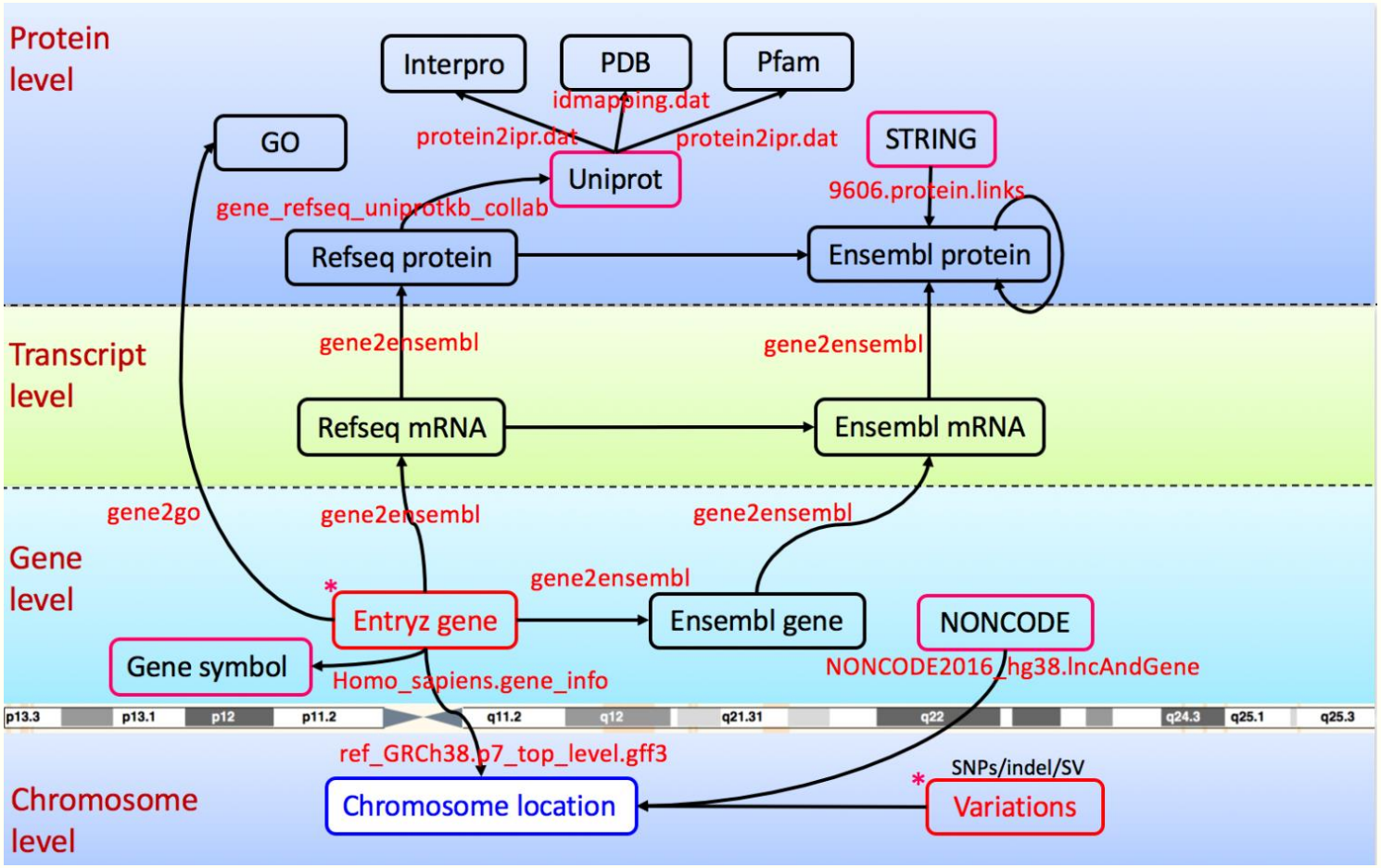
Top Design of Knowledge Integration



Integration of Databases



Genomics Data Integration



gene	Protein coding	20,656
	Protein no-coding	38,943

Variation Integration

Variation ⟷ Disease

50 YEARS OMIM
Human Genetics Knowledge for the World

HGMD®

COSMIC
Catalogue Of Somatic Mutations In Cancer

ClinVar

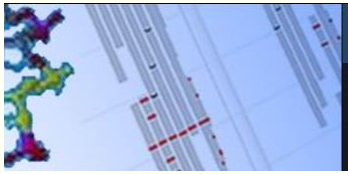
ClinVar aggregates information about g

- 10,725 Diseases
- 18,022 genes

Variation

dbVar

dbVar is NCBI's database of genomic struc
element insertions, translocations, and cor



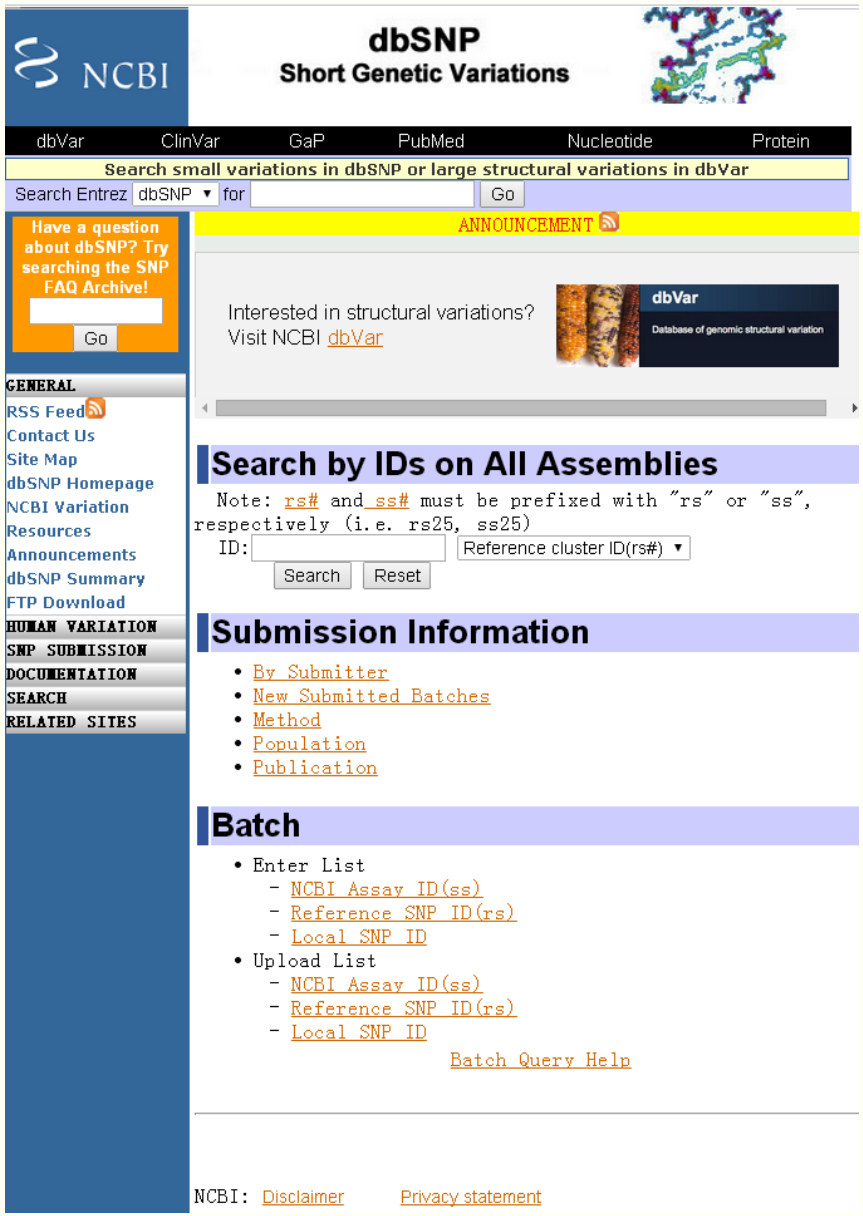
dbSNP

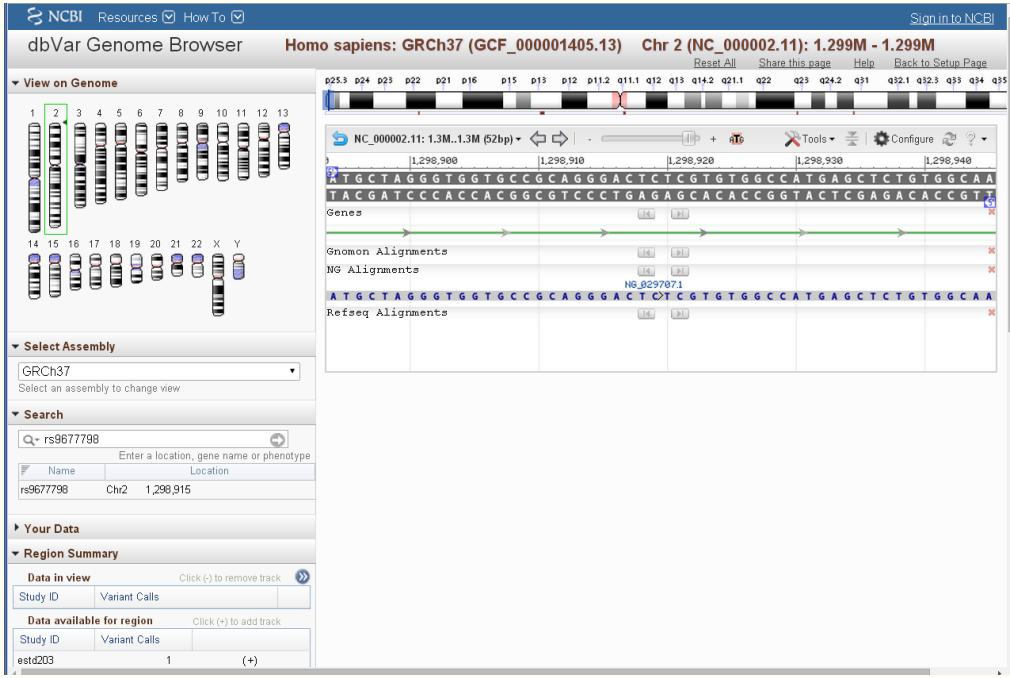
Database of single nucleotide polymorp
insertions/deletions, microsatellites, an

Mutation →38,836 Genes

Known SNP database

dbSNP



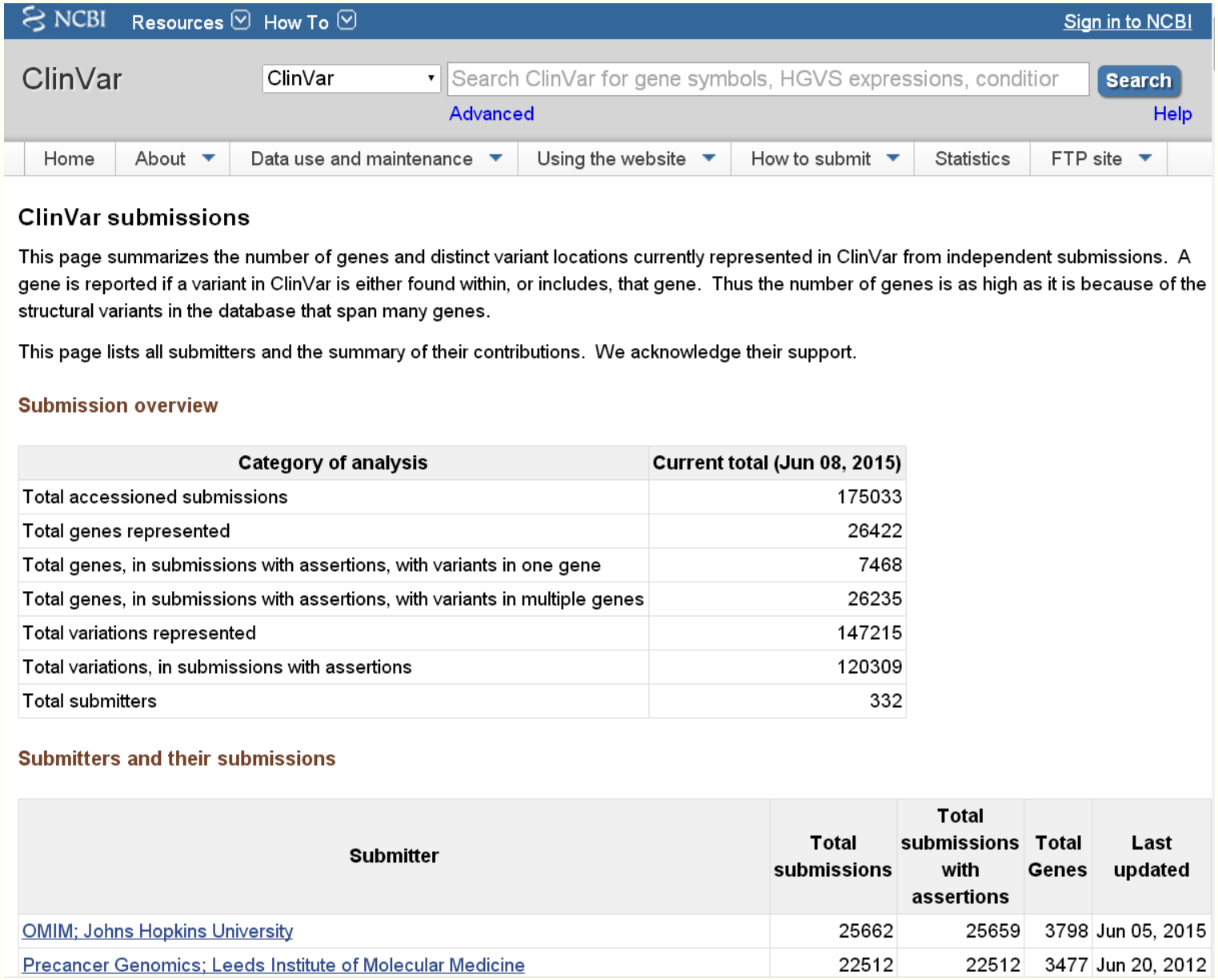


dbSNP build 141：61,060,456

http://www.ncbi.nlm.nih.gov/SNP/

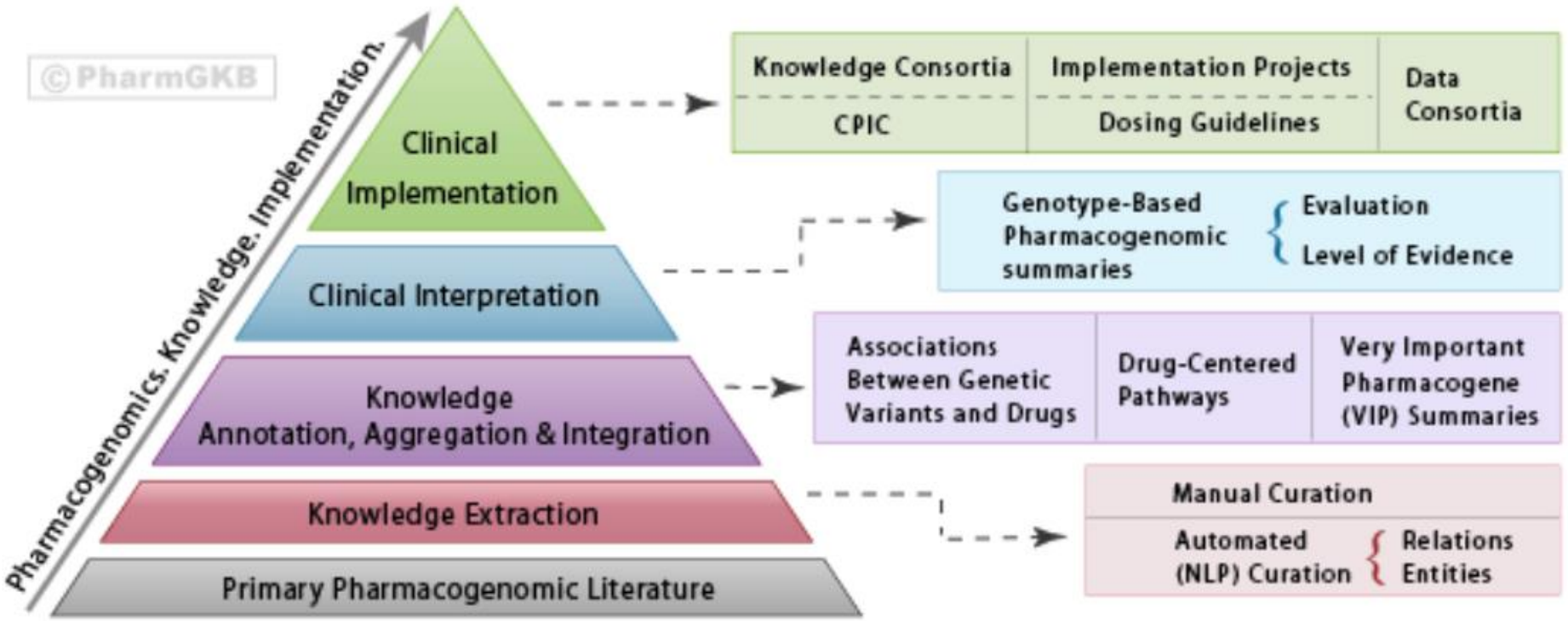
Known association

ClinVar



http://www.ncbi.nlm.nih.gov/clinvar/

The PharmGKB Knowledge Pyramid



PharmGKB提供以下信息：

VA： Variant Annotations
PW： Drug-Centered Pathway
VIP： Very Important Pharmacogene Summaries
CA：Clinical Annotations
DG：Pharmacogenomics-Based Drug-Dosing Guidelines
DL：Drug Labels with Pharmacogenomic Information

文章发表情况：

Total Publications: 284

- Pathway Publications: 38
- Guideline Publications: 23
- VIP Publications: 32
- Other Publications: 191

Downloads

In addition to the PharmGKB website, we are pleased to make PharmGKB data and knowledge available as downloads. We have found that it is often critical to check with our curators at feedback@pharmgkb.org before embarking on a large project using these data, to be sure that the files and data we make available are being interpreted correctly. PharmGKB generally does NOT need to be a co-author on such analyses; we just want to make sure that there is a correct understanding of our data before lots of resources are spent.

Examples of papers that have been written by others using PharmGKB information

Primary PharmGKB Data

Versions of external data sources used in these files

- Genes: genes.zip (3 MB)
- RSID mapping: rsid.zip (151 MB)
- Drugs: drugs.zip (456 KB)
- Diseases: diseases.zip (312 KB)
- Pathways in BioPax format: pathways-biopax.zip (484 KB)
- Pathways in tsv format: pathways-tsv.zip (67 KB)
- Dosing Guidelines in JSON format: dosingGuidelines.json.zip (414 KB)
- Drug Label summary in tsv format: drugLabels.zip (6 KB)

Variant and Clinical Annotations Data

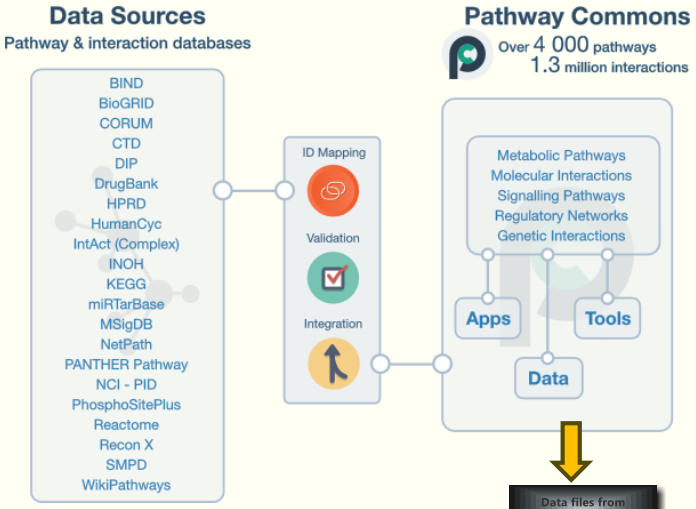
To access PharmGKB's variant and clinical annotations data please submit a Variant and Clinical Annotations Data Request.

Variant, Gene and Drug Relationship Data

To access PharmGKB's variant, gene and drug relationship data please submit a Variant, Gene and Drug Relationship Data Request.

http://www.pharmgkb.org/

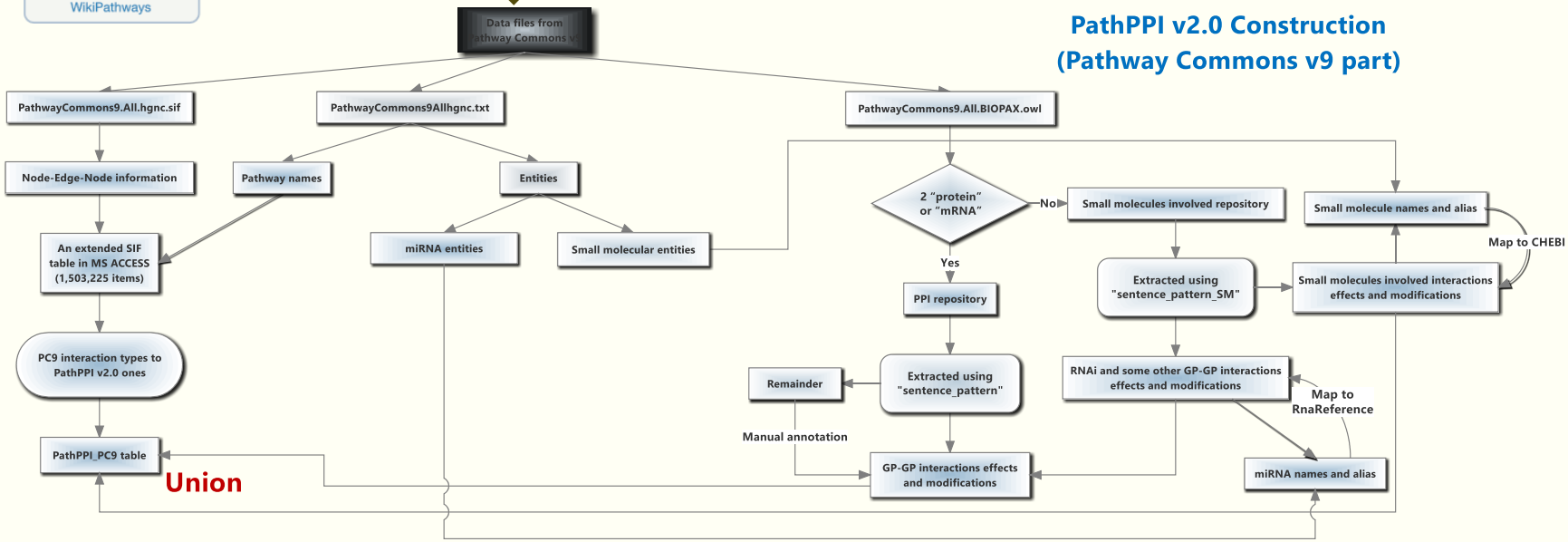
Network Integration



Pathway Commons v9, BioPAX L3 →

PathPPI v2.0

PathPPI v2.0 Construction
(Pathway Commons v9 part)



Network Integration

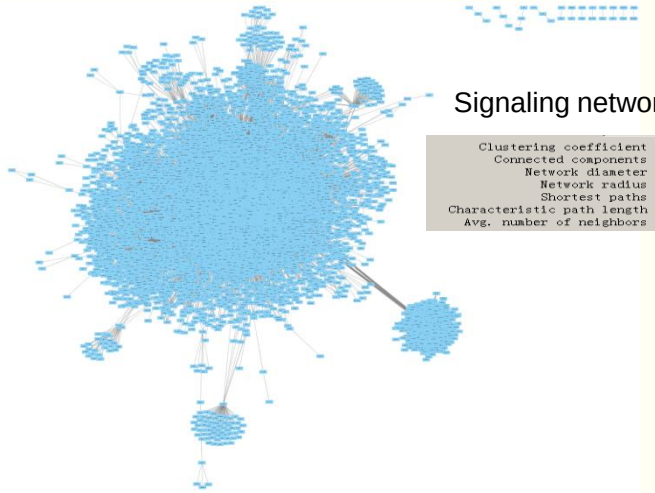
Pathway Type	Edge Type	Directionality	Edge No.
Signaling Pathway	SR: Signaling Regulation	Directed	172,765
	ER: Expression Regulation	Directed	122,786
	CAI: Complex Assembly Interaction	Undirected	177,227
	TR: Transport Regulation	Directed	7,296
	TRc: Transport Regulation chemical	Directed	3,285
	ca: x chemical affects P	Directed	469,519
	RNAi: RNAinterference	Directed	317,556
TechPPI	TechPPI: Technical Protein-Protein Interaction	Undirected	316,437
Metabolic Pathway	sp: metabolic reaction $s \xrightarrow{E} p \xrightarrow{F}$	Directed	14,428
	sE: metabolic reaction $s \xrightarrow{E} p \xrightarrow{F}$	Directed	22,480
	Ep: metabolic reaction $s \xrightarrow{E} p \xrightarrow{F}$	Directed	21,334
	EE: metabolic reaction $s \xrightarrow{E} p \xrightarrow{F}$	Directed	154,975
	rw: x reacts with y	Undirected	3,912

Network

Source Database	PathPPI v1.0	PathPPI v2.0
Reactome	2012.03.14	v56
NCI PID	2012.03.17	2015.17.27
NetPath	2012.04.14	2012.04.14
INOH	2010.01.31	2011.03.22 (v4.0)
KEGG	2009.10.12	2011.07
BioCarta	2010.08.13	
SPIKE	2011.03.22	
PhosphoSitePlus		2016.03.15
HumanCyc		v19.5
PANTHER Pathway		v3.4
TRANSFAC		v7.4
miRTarBase		v4.5
DrugBank		v4.3
Recon X		v2.03
Ctd		2015.06.03
SMPD		v2.0
WikiPathways		2015.09.29
HPRD	R9	R9
IntAct	2013.10.22	2016.02.16
BioGRID	R3.2.105	R3.44.135
DIP	2013.07.07	2016.01.14
MINT	2013.03.26	
BIND		2010.12.15
CORUM		2012.02.17

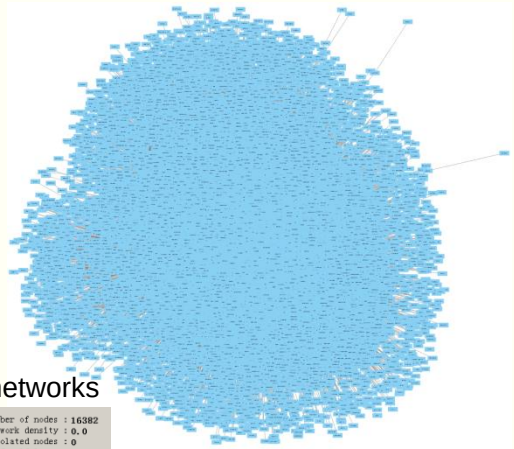
26 Modification

Modification	v1.0	v2.0	Modification	v1.0	v2.0
Acetylation	●	●	ADP-ribosylation		●
Botinylation	●		Acylation		●
Decanoylation	●		Alkylation		●
Dimethylation	●		Amination		●
Farnesylation	●	●	Carbamoylation		●
Fucosylation	●		Carboxylation		●
Galactosylation	●		Ethylation		●
Geranylgeranylation	●		Geranoylation		●
Glucosylation	●		Glucuronidation		●
Glycosylation	●	●	Glutathionylation		●
Glycylation	●		Glycation		●
Hydroxylation	●	●	N-glycosylation		●
Lipoylation	●	●	O-glycosylation		●
Methylation	●	●	Nitrosation		●
Myristoylation	●		Prenylation		●
Octanoylation	●		Ribosylation		●
Palmitoylation	●		Sulfation		●
Phosphopantetheine	●				
Phosphorylation	●	●			
Sumoylation	●	●			
Trimethylation	●				
Ubiquitination	●	●			



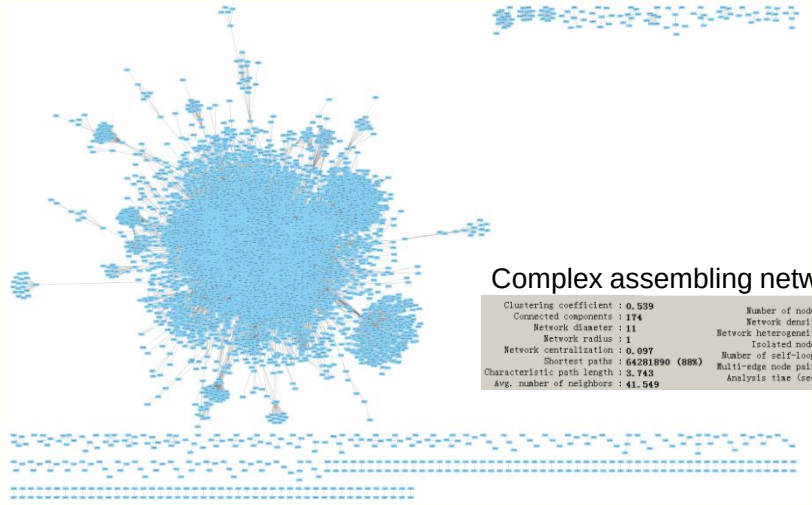
Signaling networks

Clustering coefficient	0.105	Number of nodes	8911
Connected components	13	Network density	0.0
Network diameter	10	Isolated nodes	0
Network radius	1	Number of self-loops	0
Shortest paths	31576996 (39%)	Multi-edge node pairs	7855
Characteristic path length	3.860	Analysis time (sec)	13.015
Avg. number of neighbors	37.013		



Expression regulatory networks

Clustering coefficient	0.201	Number of nodes	16382
Connected components	1	Network density	0.0
Network diameter	10	Isolated nodes	0
Network radius	1	Number of self-loops	0
Shortest paths	11916671 (7%)	Multi-edge node pairs	524
Characteristic path length	3.632	Analysis time (sec)	12.774
Avg. number of neighbors	14.926		

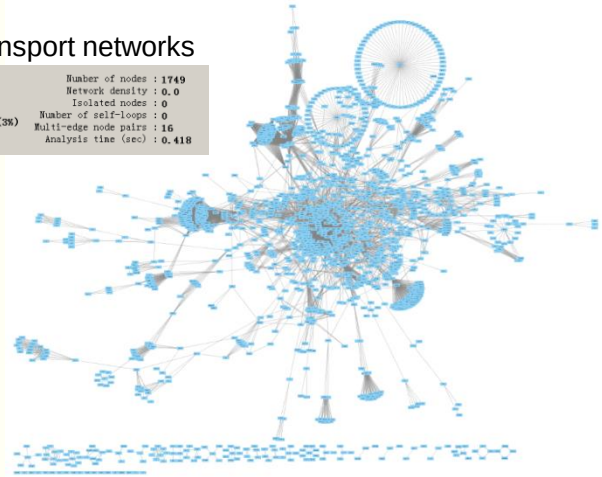


Complex assembling networks

Clustering coefficient	0.539	Number of nodes	8531
Connected components	174	Network density	0.005
Network diameter	11	Network heterogeneity	1.580
Network radius	1	Isolated nodes	0
Network centralization	0.097	Number of self-loops	0
Shortest paths	64281890 (88%)	Multi-edge node pairs	0
Characteristic path length	3.743	Analysis time (sec)	19.294
Avg. number of neighbors	41.549		

Transport networks

Clustering coefficient	0.025	Number of nodes	1749
Connected components	34	Network density	0.0
Network diameter	12	Isolated nodes	0
Network radius	1	Number of self-loops	0
Shortest paths	93773 (3%)	Multi-edge node pairs	16
Characteristic path length	4.626	Analysis time (sec)	0.418
Avg. number of neighbors	8.325		



Knowledge Graph

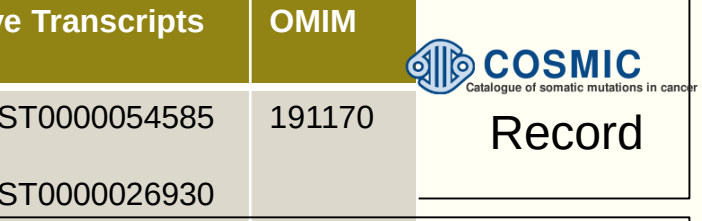
Gene Name	Gene Id	Alternative Transcripts	OMIM
TP53	COSG501	TP53_ENST00000545858, TP53_ENST00000269305	191170

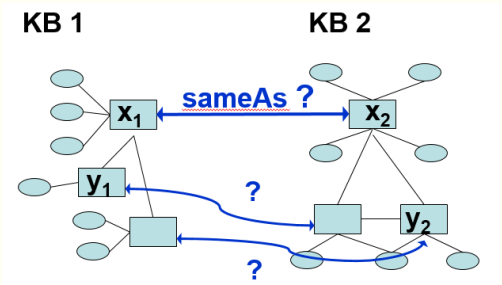
COSMIC Catalogue of somatic mutations in cancer

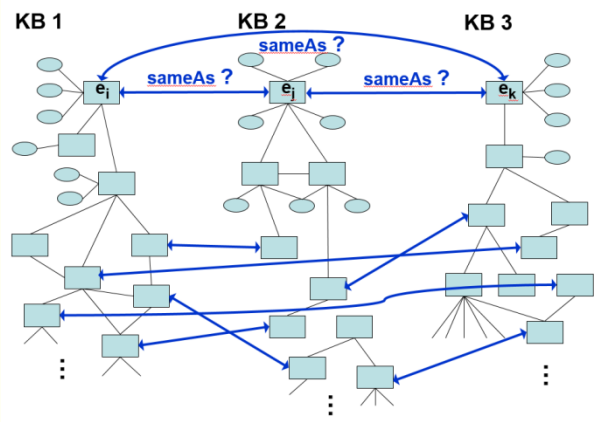
Record

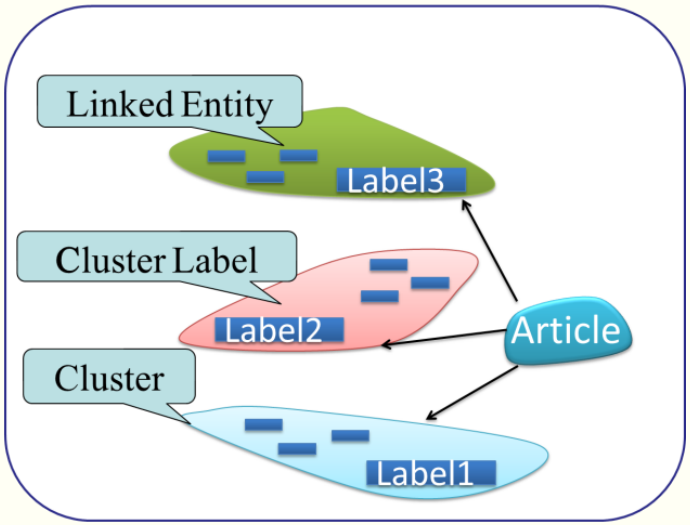
Subject	Predicate	Object
COSG501	Gene name	TP53
COSG501	Aternative Transcripts	TP53_ENST00000545858
COSG501	Aternative Transcripts	TP53_ENST00000269305
COSG501	OMIM	**OMIM_TP53**

RDF

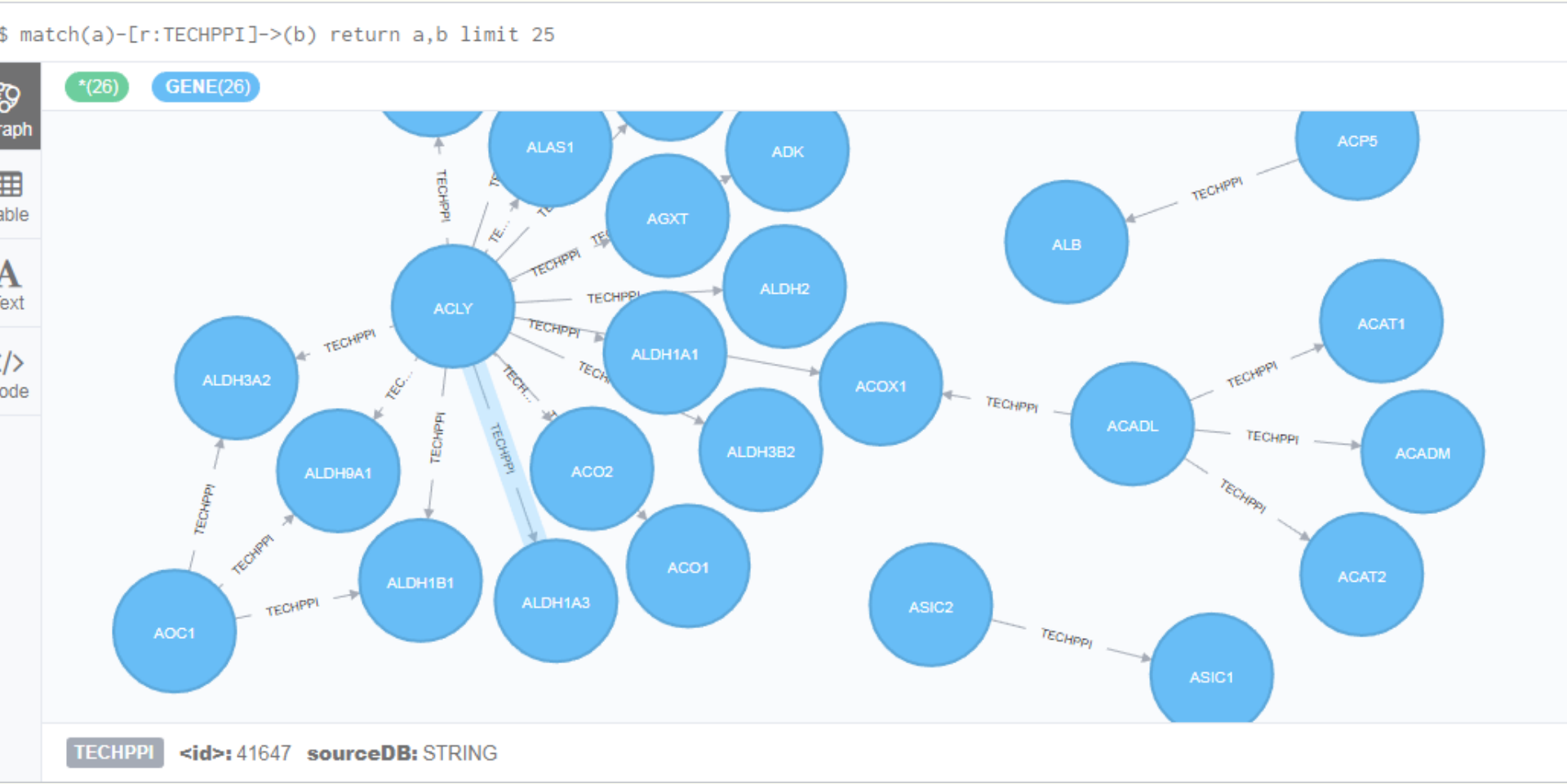




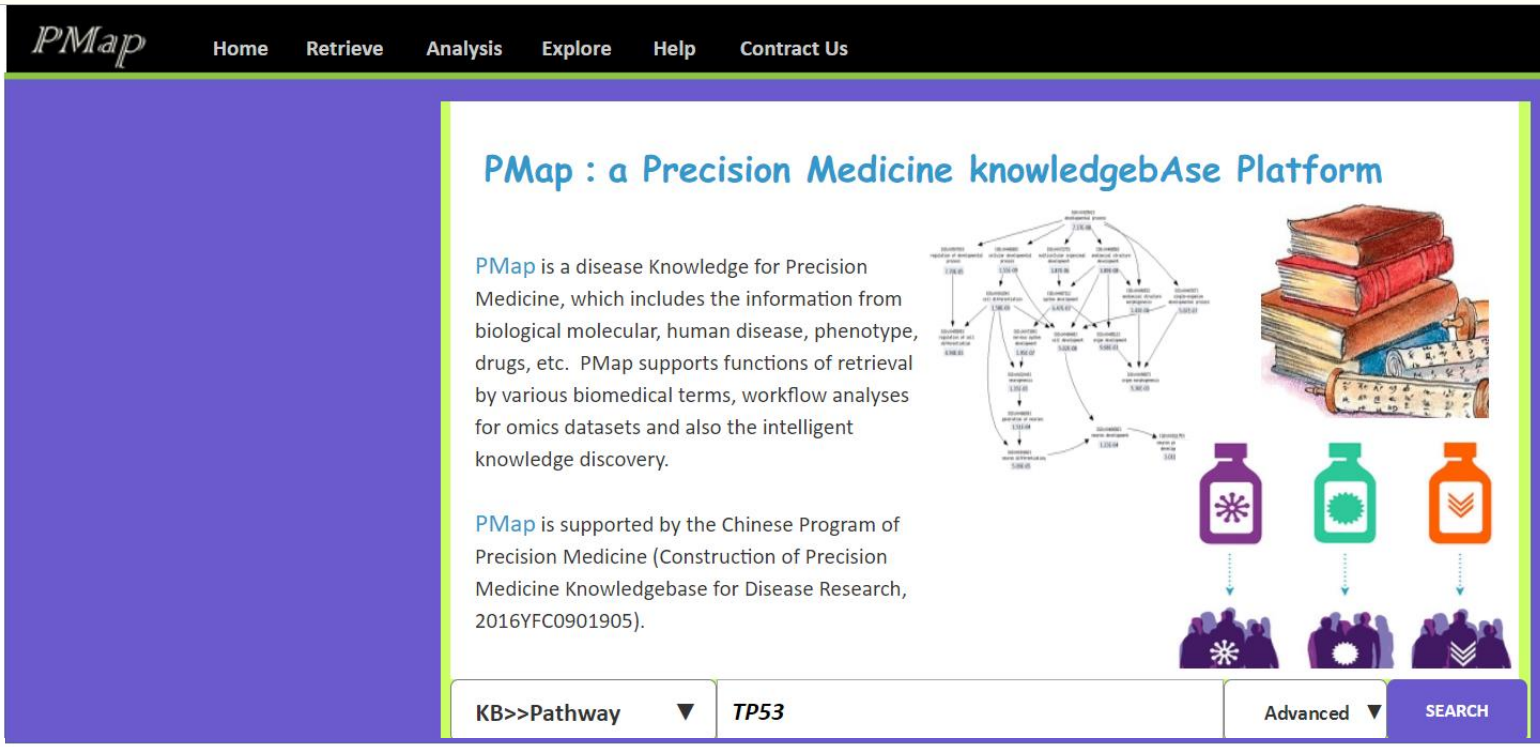


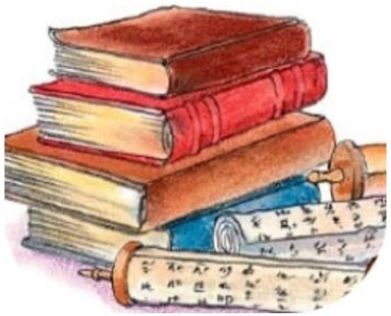


Visualization of Knowledge Graph

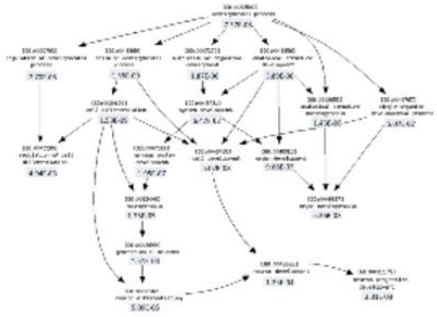


Tools for the Knowledgebase: PMap

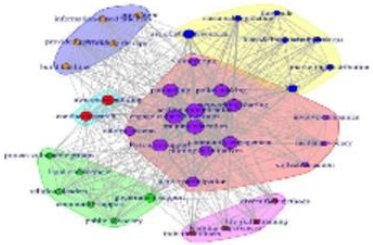




Literature

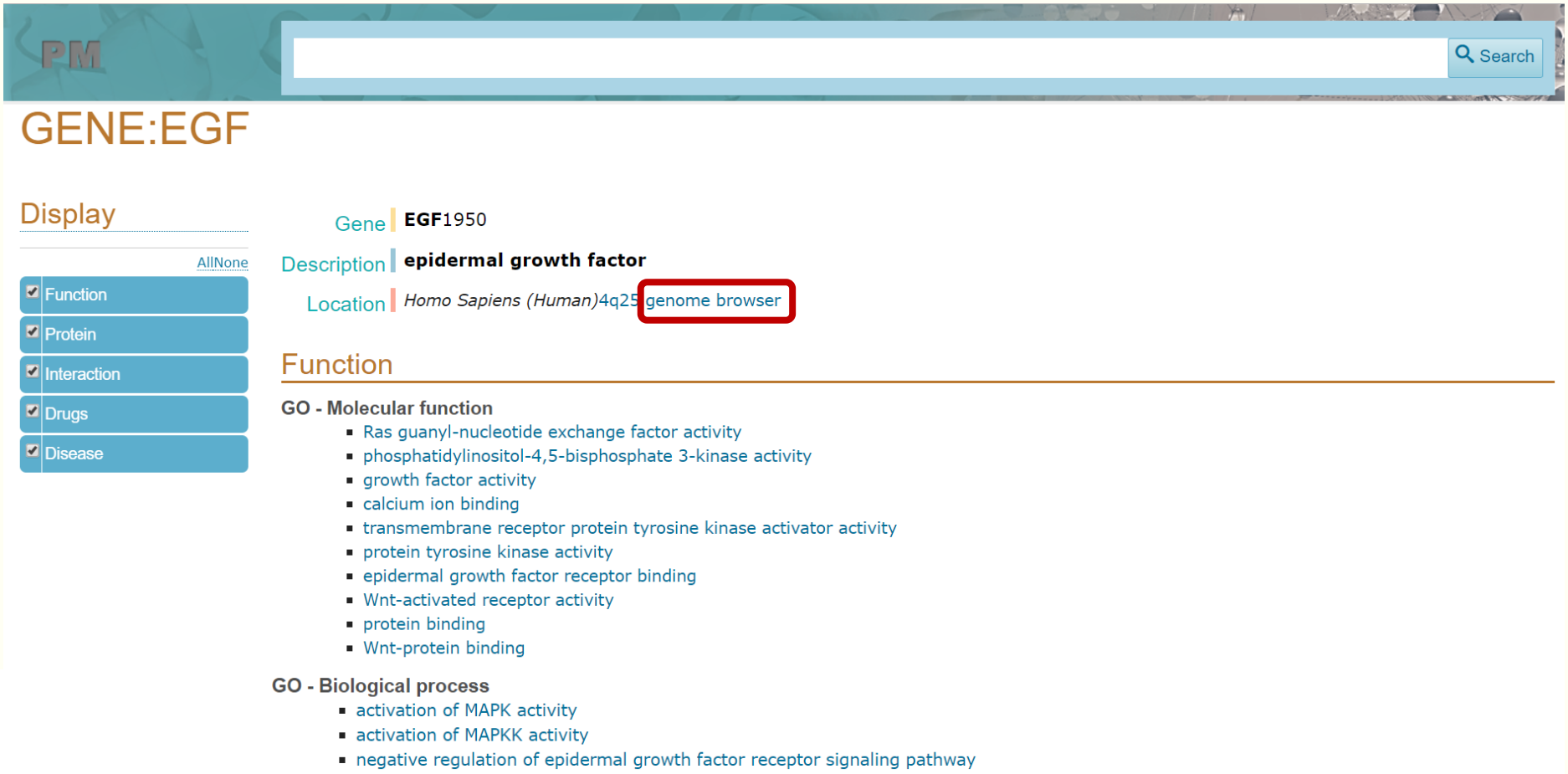


Ontology

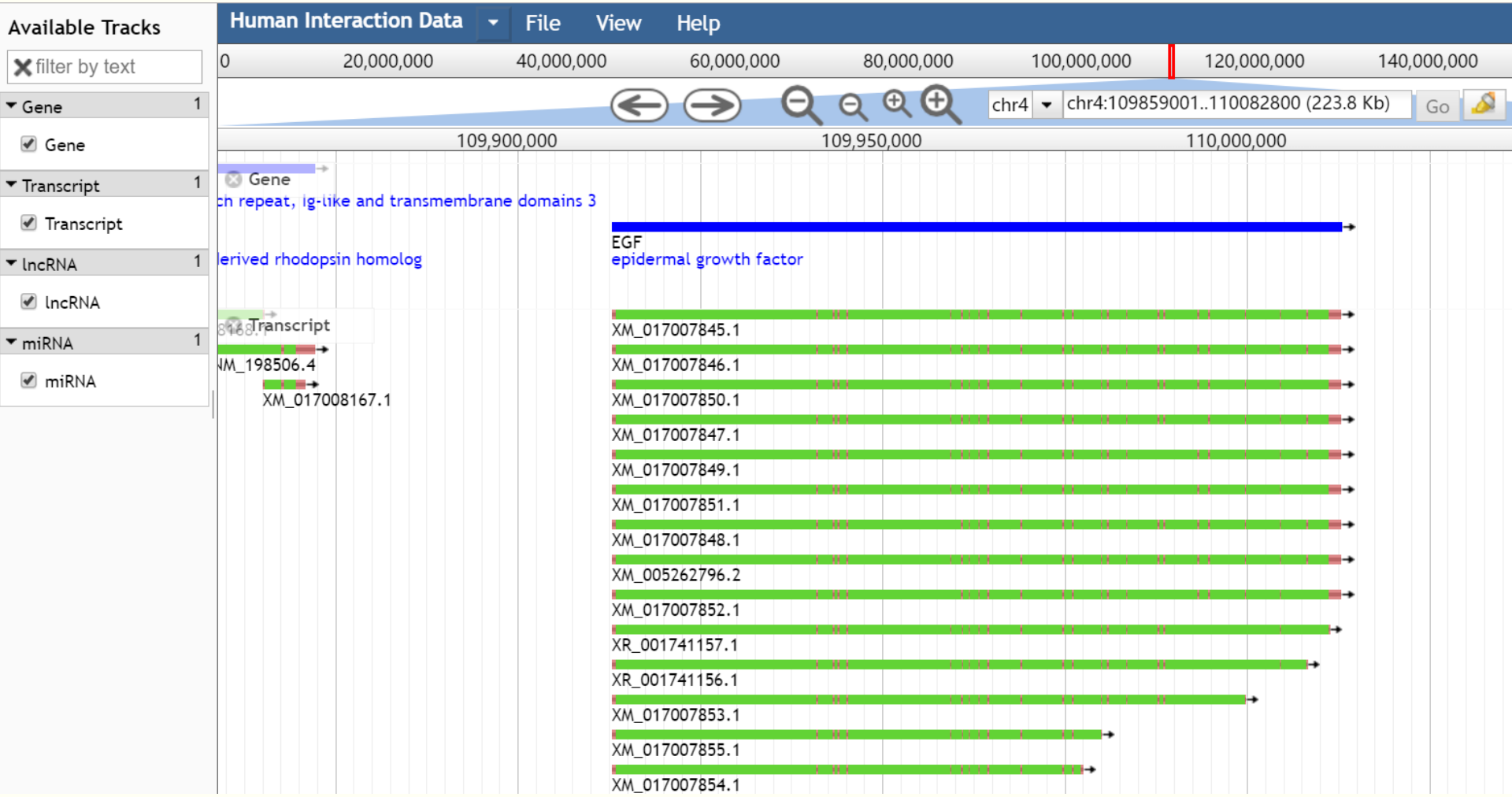


PMap-ODAS

Search Engine



Genome browser：EGF



EGF Related Drugs

Display

All None

- [x] Function
- [x] Protein
- [x] Interaction
- [x] Drugs
- [x] Disease

Drugs

- cetuximab Source:PharmGKB
- bevacizumab Source:PharmGKB
- cholecalciferol Source:PharmGKB
- gefitinib Source:PharmGKB
- oxaliplatin Source:PharmGKB
- fluorouracil Source:PharmGKB
- irinotecan Source:PharmGKB
- ethanol Source:PharmGKB
- capecitabine Source:PharmGKB
- dexamethasone Source:PharmGKB
- docetaxel Source:PharmGKB
- panitumumab Source:PharmGKB
- Fatty acid derivatives Source:PharmGKB
- Insulins And Analogues Source:PharmGKB
- estrogens Source:PharmGKB
- glucocorticoids Source:PharmGKB

Disease

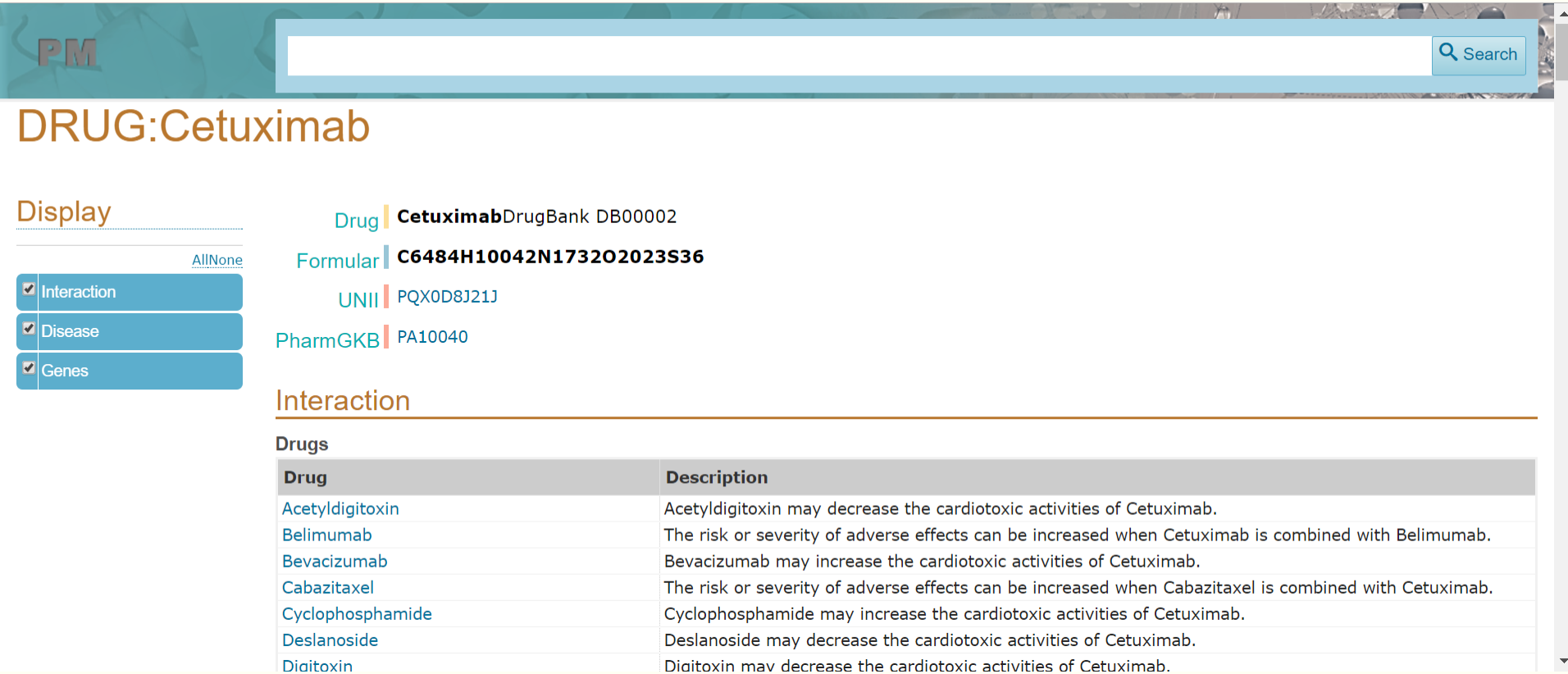
OMIM

OMIM	Hypomagnesemia 4, renal, 611718 (3)

COSMIC

Show 10 entries

Drug-Drug Interaction



Cetuximab Targets and Treatments

Trastuzumab	Trastuzumab may increase the cardiotoxic activities of Cetuximab.

Foods

Diseases

- Cetuximab, used in combination with irinotecan, is indicated for the treatment of EGFR-expressing, metastatic colorectal carcinoma in patients who are refractory to irinotecan-based chemotherapy. Cetuximab administered as a single agent is indicated for t Source:drugbank
- Colorectal cancer Source:TTD

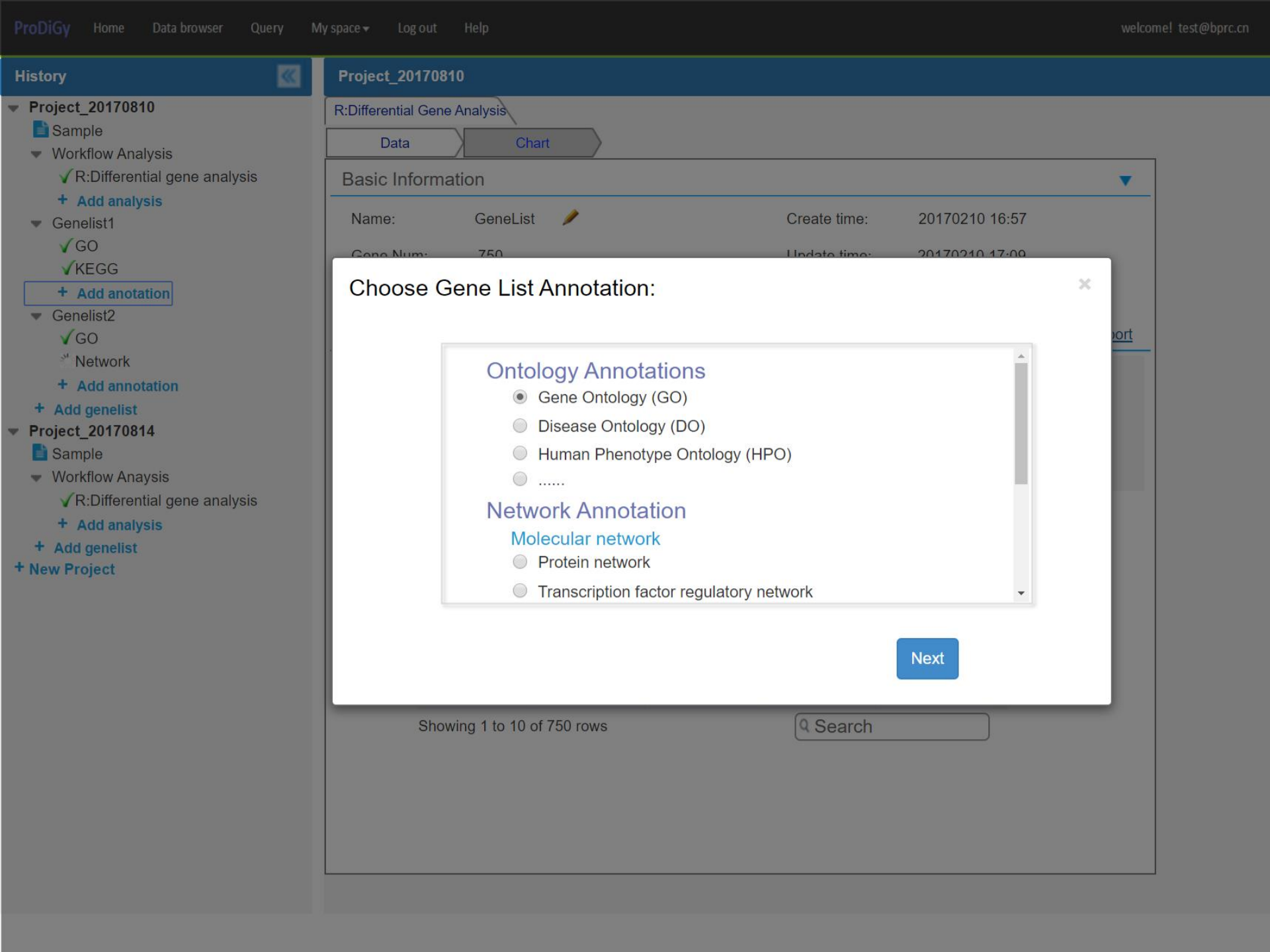
Genes

- Epidermal growth factor receptor Source:drugbank
- Low affinity immunoglobulin gamma Fc region receptor III-B Source:drugbank
- Complement C1r subcomponent Source:drugbank
- Complement C1q subcomponent subunit A Source:drugbank
- Complement C1q subcomponent subunit B Source:drugbank
- Complement C1q subcomponent subunit C Source:drugbank
- Low affinity immunoglobulin gamma Fc region receptor III-A Source:drugbank
- Complement C1s subcomponent Source:drugbank
- High affinity immunoglobulin gamma Fc receptor I Source:drugbank
- Low affinity immunoglobulin gamma Fc region receptor II-a Source:drugbank
- Low affinity immunoglobulin gamma Fc region receptor II-b Source:drugbank
- Low affinity immunoglobulin gamma Fc region receptor II-c Source:drugbank
- cyclin-dependent kinase 2 Source:PharmGKB
- cyclin-dependent kinase 4 Source:PharmGKB

Display

All None

- [x] Interaction
- [x] Disease
- [x] Genes


ProDiGy
Home
Data browser
Query
My space
Log out
Help
welcome! test@bprc.cn
History
Project_20170810
Sample
Workflow Analysis
R:Differential gene analysis
Add analysis
Genelist1
GO
KEGG
Add anotation
Genelist2
GO
Network
Add annotation
Add genelist
Project_20170814
Sample
Workflow Anaysis
R:Differential gene analysis
Add analysis
Add genelist
New Project
Project_20170810
R:Differential Gene Analysis
Data
Chart
Basic Information
Name:
GeneList
Create time:
20170210 16:57
Choose Gene List Annotation:
Ontology Annotations
Gene Ontology (GO)
Disease Ontology (DO)
Human Phenotype Ontology (HPO)
.....
Network Annotation
Molecular network
Protein network
Transcription factor regulatory network
Next
Showing 1 to 10 of 750 rows
Search

Ontology Annotations

- (•) Gene Ontology (GO)
- ( ) Disease Ontology (DO)
- ( ) Human Phenotype Ontology (HPO)
- ( ) .....

Network Annotation

Molecular network

- ( ) Protein network
- ( ) Transcription factor regulatory network
- ( ) Phosphorylation network
- ( ) Ubiquitination network
- ( ) .....

Pathway

- ( ) KEGG analysis
- ( ) Reactome Analysis

Knowledge network

- ( ) .....

History

- Project_20170810
  - Sample
  - Workflow Analysis
    - R:Differential gene analysis
    - + Add analysis
  - Genelist1
    - GO
    - KEGG
    - + Add anotation
  - Genelist2
    - GO
    - Network
    - + Add annotation
  - + Add genelist
- Project_20170814
  - Sample
  - Workflow Anaysis
    - R:Differential gene analysis
    - + Add analysis
  - + Add genelist
- + New Project

Project_20170810

R: GeneOntology

Basic Information

Name: GOAnalysisResult

GO Version: ******版本

Input data: GeneList(750)

Create time: 20170210 16:57

Update time: 20170210 17:09

Size on disk: 5 kb

Table View | Tree View

Download

Mapped Gene Num.: 500

Sig. GO terms: 50

Category Display: Display GO Full Hierachy

PValue >= 0.05

GO terms view: table

Targets >= 2

Selected 30 GO terms, Export to tree view.

Sig. Depletec | Depleted | Enriched | Sig. Enriched

GO term ID		GeneRatio	BgRatio	PValue	p.adjust	Targets
Edit filters.		Edit filters.	Edit filters	Edit filte	Edit filt	Edit filters
GO:0005515	protein binding	372(76.23%)	10893(64.14%)	3.92E-09	3.92E-06	372
GO:0008009	chemokine activity	12(2.46%)	48(0.28%)	7.54E-09	3.92E-06	12
GO:0042379	chemokine receptor binding	12(2.46%)	58(0.34%)	7.43E-08	2.58E-05	12
GO:0003678	DNA helicase activity	11(2.25%)	50(0.29%)	1.34E-07	3.48E-05	11
GO:0043138	3'-5' DNA helicase activity	6(1.23%)	12(0.07%)	4.36E-07	9.07E-05	6
GO:0042802	identical protein binding	71(14.55%)	1376(8.10%)	9.10E-07	0.000152493	71
GO:0015631	tubulin binding	25(5.12%)	289(1.70%)	1.03E-06	0.000152493	25
GO:0008017	microtubule binding	21(4.30%)	217(1.28%)	1.26E-06	0.000164208	21
GO:0048248	CXCR3 chemokine receptor binding	4(0.82%)	5(0.03%)	3.29E-06	0.000356267	4
GO:0045236	CXCR chemokine receptor binding	6(1.23%)	16(0.09%)	3.42E-06	0.000356267	6
GO:0004386	helicase activity	16(3.28%)	150(0.88%)	6.73E-06	0.000586543	16
GO:0019901	protein kinase binding	35(7.17%)	541(3.19%)	7.03E-06	0.000586543	35
GO:0019900	kinase binding	38(7.79%)	612(3.60%)	7.32E-06	0.000586543	38
GO:0005524	ATP binding	71(14.55%)	1473(8.67%)	1.01E-05	0.000747546	71
GO:0003777	microtubule motor activity	11(2.25%)	77(0.45%)	1.19E-05	0.000793709	11
GO:0032559	adenyl ribonucleotide binding	72(14.75%)	1509(8.89%)	1.22E-05	0.000793709	72
GO:0030554	adenyl nucleotide binding	72(14.75%)	1520(8.95%)	1.56E-05	0.000954452	72
GO:0003688	DNA replication origin binding	5(1.02%)	13(0.08%)	2.04E-05	0.001180507	5
GO:0003682	chromatin binding	31(6.35%)	482(2.84%)	2.64E-05	0.001448643	31
GO:0043168	anion binding	110(22.54%)	2666(15.70%)	3.63E-05	0.001888793	110
GO:0003690	double-stranded DNA binding	41(8.40%)	738(4.35%)	4.36E-05	0.002162286	41
GO:0016787	hydrolase activity	105(21.52%)	2544(14.98%)	5.69E-05	0.002694465	105
GO:0048020	CCR chemokine receptor binding	7(1.43%)	36(0.21%)	6.28E-05	0.002841317	7
GO:0097367	carbohydrate derivative binding	93(19.06%)	2203(12.97%)	7.42E-05	0.003218342	93
GO:0035173	histone kinase activity	5(1.02%)	17(0.10%)	8.92E-05	0.003696812	5
GO:0061575	cyclin-dependent protein serine/thre	3(0.61%)	4(0.02%)	9.23E-05	0.003696812	3
GO:0005488	binding	439(89.96%)	14281(84.09%)	1.00E-04	0.003854447	439

Records per page: 10 Showing 1 to 10 of 1000 GO terms

First Previous 1 2 3 4 5 Next Last

History

- Project_20170810
  - Sample
  - Workflow Analysis
    - R:Differential gene analysis
    - + Add analysis
  - Genelist1
    - GO
    - KEGG
    - + Add anotation
  - Genelist2
    - GO
    - Network
    - + Add annotation
  - + Add genelist
- Project_20170814
  - Sample
  - Workflow Anaysis
    - R:Differential gene analysis
    - + Add analysis
  - + Add genelist
- + New Project

Project_20170810

R: GeneOntology

Basic Information

Name: GOAnalysisResult
GO Version: ******版本
Input data: GeneList(750)
Create time: 20170210 16:57
Update time: 20170210 17:09
Size on disk: 5 kb

Table View | Tree View | Download

Mapped Gene Num.: 500
Category Display: Display GO Full Hierachy
GO terms view: ontology
Sig. GO terms: 50
PValue >= 0.05
Targets >= 2

Selected 30 GO terms, Export to tree view.

Sig. Depletec | Depleted | Enriched | Sig. Enriched

GO term ID		GeneRatio	BgRatio	PValue	p.adjust	Targets
GO:0003674	Molecular_function	372(76.23%)	10893(64.14%)			
GO:0000988	Transcription Factor Activity, Protein Binding	12(2.46%)	48(0.28%)			
		12(2.46%)	58(0.34%)			
GO:0001071	Nucleic Acid Binding Transcription Factor Activity	11(2.25%)	50(0.29%)			
		6(1.23%)	12(0.07%)			
GO:0003677	DNA Binding	71(14.55%)	1376(8.10%)			
GO:0003723	RNA Binding	25(5.12%)	289(1.70%)	1.03E-06	0.000152493	25
GO:0003729	MRNA Binding	21(4.30%)	217(1.28%)	1.26E-06	0.000164208	21
GO:0003924	GTPase Activity	4(0.82%)	5(0.03%)	3.29E-06	0.000356267	4
GO:0004386	Helicase Activity	6(1.23%)	16(0.09%)	3.42E-06	0.000356267	6
GO:0004518	Nuclease Activity	16(3.28%)	150(0.88%)	6.73E-06	0.000586543	16
GO:0004871	Signal Transducer Activity	35(7.17%)	541(3.19%)	7.03E-06	0.000586543	35
GO:0005198	Structural Molecule Activity	38(7.79%)	612(3.60%)	7.32E-06	0.000586543	38
GO:0008092	Cytoskeletal Protein Binding	71(14.55%)	1473(8.67%)	1.01E-05	0.000747546	71
GO:0008134	Transcription Factor Binding	11(2.25%)	77(0.45%)	1.19E-05	0.000793709	11
GO:0008168	Methyltransferase Activity	72(14.75%)	1509(8.89%)	1.22E-05	0.000793709	72
GO:0008233	Peptidase Activity	72(14.75%)	1520(8.95%)	1.56E-05	0.000954452	72
GO:0008289	Lipid Binding	5(1.02%)	13(0.08%)	2.04E-05	0.001180507	5
	Protein Transporter	31(6.35%)	482(2.84%)	2.64E-05	0.001448643	31
		110(22.54%)	2666(15.70%)	3.63E-05	0.001888793	110
		41(8.40%)	738(4.35%)	4.36E-05	0.002162286	41
		105(21.52%)	2544(14.98%)	5.69E-05	0.002694465	105
		7(1.43%)	36(0.21%)	6.28E-05	0.002841317	7
		93(19.06%)	2203(12.97%)	7.42E-05	0.003218342	93
		5(1.02%)	17(0.10%)	8.92E-05	0.003696812	5
		3(0.61%)	4(0.02%)	9.23E-05	0.003696812	3
		439(89.96%)	14281(84.09%)	1.00E-04	0.003854447	439

Edit filters. Edit filters. Edit filters Edit filters Edit filters Edit filters

No Filter | Equal | Not Equal | Less | Greater

Records per page: 10 Showing 1 to 10 of 1000 GO terms

First Previous 1 2 3 4 5 Next Last

History

- Project_20170810
  - Sample
  - Workflow Analysis
    - R:Differential gene analysis
    - + Add analysis
  - Genelist1
    - GO
    - KEGG
    - + Add anotation
  - Genelist2
    - GO
    - Network
    - + Add annotation
  - + Add genelist
- Project_20170814
  - Sample
  - Workflow Anaysis
    - R:Differential gene analysis
    - + Add analysis
  - + Add genelist
- + New Project

Project_20170810

R: GeneOntology

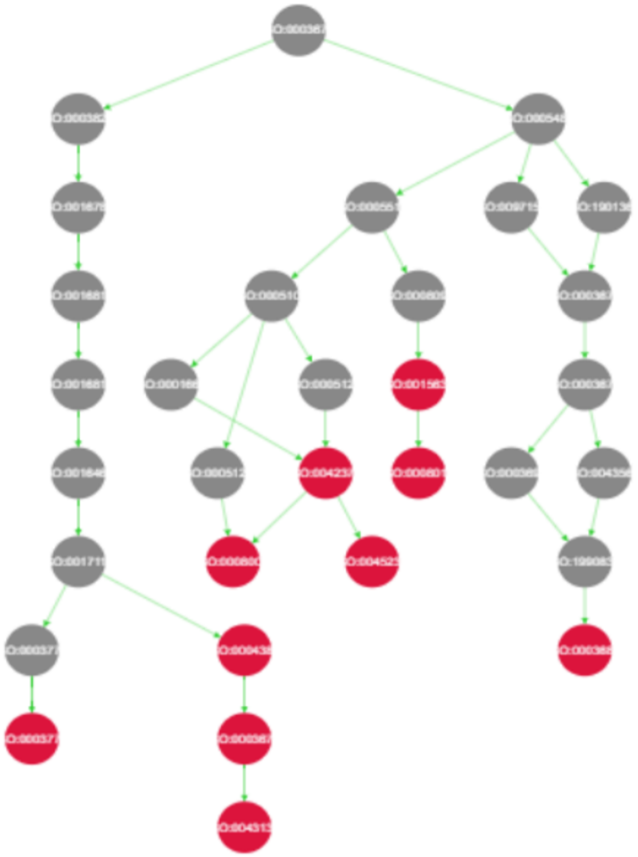
Basic Information

Name:	GOAnalysisResult	Create time:	20170210 16:57
GO Version:	******版本	Update time:	20170210 17:09
Input data:	GeneList(750)	Size on disk:	5 kb

Table View | Tree View

Export picture Download

Tree View of the Selected GO Terms



History

- Project_20170810
  - Sample
  - Workflow Analysis
    - R:Differential gene analysis
    - + Add analysis
  - Genelist1
    - GO
    - KEGG
    - + Add anotation
  - Genelist2
    - GO
    - Network
    - + Add annotation
  - + Add genelist
- Project_20170814
  - Sample
  - Workflow Anaysis
    - R:Differential gene analysis
    - + Add analysis
  - + Add genelist
- + New Project

Project_20170810

R: ProteinNetwork

Basic Information

Name: NetworkAnalysisResult

Sig. Networks: 50

Input data: GeneList(750)

Create time: 20170210 16:57

Update time: 20170210 17:09

Size on disk: 5 kb

Table View | Network View

Download

Sig. Networks: 50

Records per page: 10

700

Score >= 20

Targets >= 2

Default GO

Selected 30 GO terms, Export to Network view.

Network ID	Molecules in Network	Score	Targets
1	vANN, ^APOBR, ^ARPIN/C15orf38-AP3S2, ^CADM3, vCHCHD10, vENPP4, ^FOXA1, ^FOXA2, ^GATA3, vITGB4, vITGB7,^RCID1, ^KLF2,vLHFP	40	35
2	vADAD2, ^RPI, ^APP, ^CCDC149, vTOX2, vENPP4, ^FOXA1, ^FOXA2, ^GATA3, vITGB4, vITGB7,^RCID1, ^KLF2,vLHFP	39	34
3	vCCDC149, vTOX2, vENPP4, ^FOXA1, ^FOXA2, ^GATA3, vITGB4, vITGB7,^RCID1, ^KLF2,vLHFP	35	35
4	^CADM3, vCHCHD10, vENPP4, ^FOXA1,vANN, ^APOBR	34	32
5	^CADM3, vCHCHD10, vENPP4, ^FOXA1	30	31

Records per page: 10 Showing 1 to 10 of 50 rows

First Previous 1 2 3 4 5 Next Last

Targets mapped to this term:

Code	GO ID	GO Description	Targeted ID	Targeted Name	Gene Description
1.4	GO:0008009	chemokine activity	3627	A3GALT2	Alpha-1,3-galactosyltransferase 2 (EC 2.4.1.87) (Isoglobotriaosylceramide synthase) (iGb3 synthase) (iGb3S) [A3GALT2P] [IGBS3S]
1.4	GO:0008009	chemokine activity	10563	AADACL3	Arylacetamide deacetylase-like 3 (EC 3.1.1.-)
1.4	GO:0008009	chemokine activity	6373	AADACL4	Arylacetamide deacetylase-like 4 (EC 3.1.1.-)
1.4	GO:0008009	chemokine activity	4283	ABCA4	Retinal-specific ATP-binding cassette transporter (ATP-binding cassette sub-family A member 4) (RIM ABC transporter) (RIM protein) (RmP) (Stargardt disease protein) [ABCR]
1.4	GO:0008009	chemokine activity	6362	ABCB10	ATP-binding cassette sub-family B member 10, mitochondrial precursor (ATP-binding cassette transporter 10) (ABC transporter 10 protein) (Mitochondrial ATP-binding cassette 2) (M-ABC2)
1.4	GO:0008009	chemokine activity	6355	ABCD3	ATP-binding cassette sub-family D member 3 (70 kDa peroxisomal membrane protein) (PMP70) [PMP70] [PXMP1]
1.4	GO:0008009	chemokine activity	2921	ABL2	Abelson tyrosine-protein kinase 2 (EC 2.7.10.2) (Abelson murine leukemia viral oncogene homolog 2) (Abelson-related gene protein) (Tyrosine-protein kinase ARG) [ABLL] [ARG]
1.4	GO:0008009	chemokine activity	6364	ACADM	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial precursor (EC 1.3.8.7) (MCAD)
1.4	GO:0008009	chemokine activity	3576	ACAP3	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 3 (Centaurin-beta-5) (Cnt-b5) [CENTB5]
1.4	GO:0008009	chemokine activity	6352	ACBD3	Golgi resident protein GCP60 (Acyl-CoA-binding domain-containing protein 3) (Golgi complex-associated protein 1) (GOCAP1) (Golgi phosphoprotein 1) (GOLPH1) (PBR- and PKA-associated protein 7) (Peripheral benzodiazepine receptor-associated protein PAP7) [Contains: Golgi resident protein GCP60, N-terminally processed] [GCP60]
1.4	GO:0008009	chemokine activity	6347	ACBD6	Acyl-CoA-binding domain-containing protein 6
1.4	GO:0008009	chemokine activity	6351	ACKR1	Atypical chemokine receptor 1 (Duffy antigen/chemokine receptor) (Fy glycoprotein) (GpFy) (Glycoprotein D) (Plasmodium vivax receptor) (CD234 antigen) [DARC] [FY] [GPD]

History

- Project_20170810
  - Sample
  - Workflow Analysis
    - R:Differential gene analysis
    - + Add analysis
  - Genelist1
    - GO
    - KEGG
    - + Add anotation
  - Genelist2
    - GO
    - Network
    - + Add annotation
  - + Add genelist
- Project_20170814
  - Sample
  - Workflow Anaysis
    - R:Differential gene analysis
    - + Add analysis
  - + Add genelist
- + New Project

Project_20170810

R: ProteinNetwork

Basic Information

Name: NetworkAnalysisResult

Sig. Networks: 50

Input data: GeneList(750)

Create time: 20170210 16:57

Update time: 20170210 17:09

Size on disk: 5 kb

Table View | Network View

Export picture Download

Title: NetworkAnalysisChart

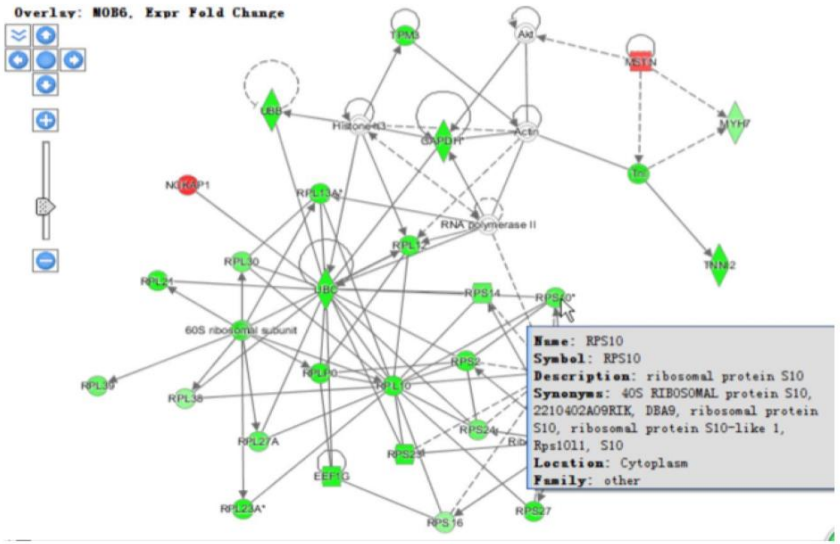
Layout: Force-Directed

Update time: 2017-02-10 17:09

Color plalettes: green-red

Border color: none

Default GO



Thank You!