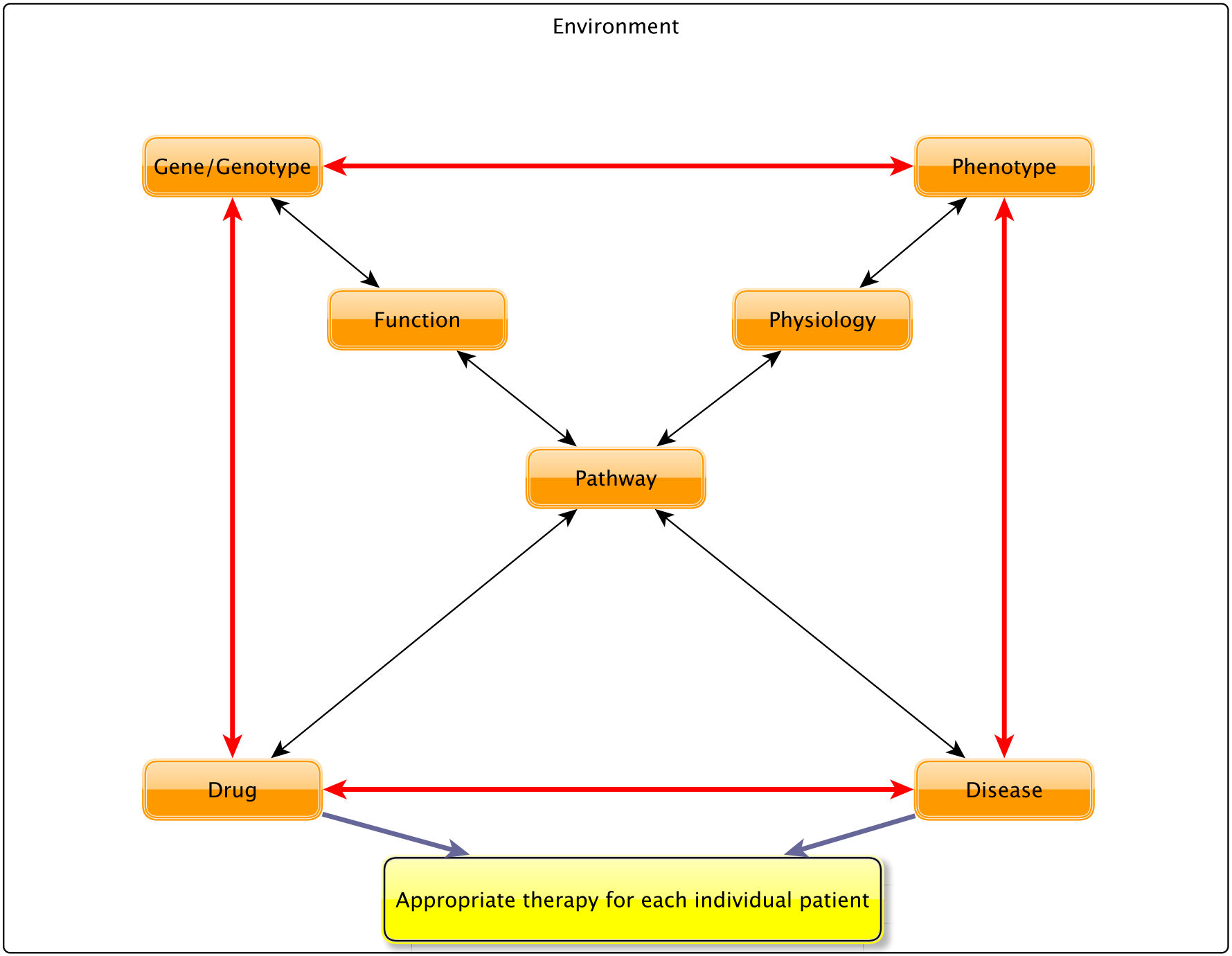
Ontologies & Biomedical Research

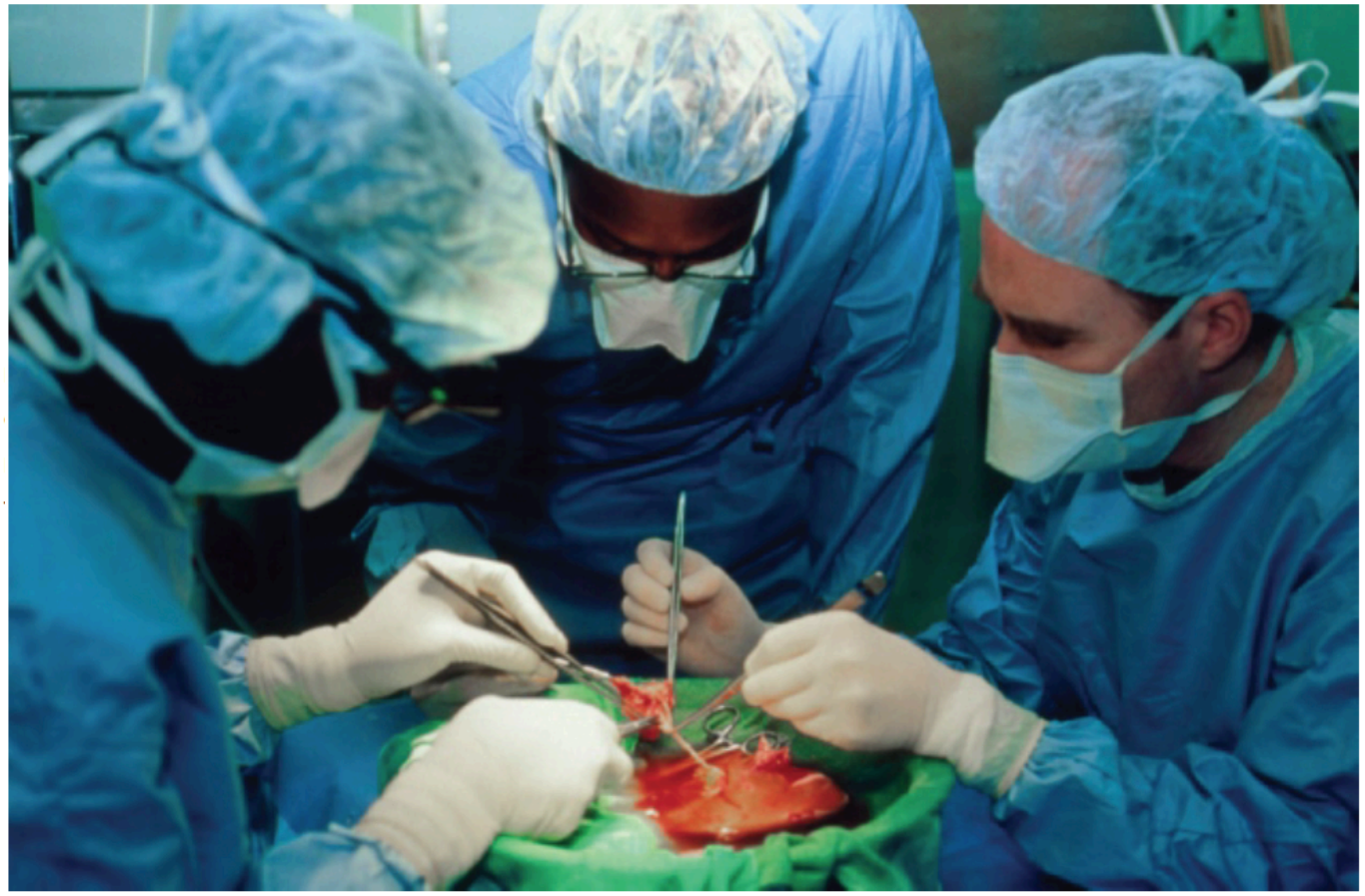

Environment
Gene/Genotype
Phenotype
Function
Physiology
Pathway
Drug
Disease
Appropriate therapy for each individual patient

Data-crunch highlights potential transplant drugs

Widely prescribed statin could have alternative applications.

Monya Baker

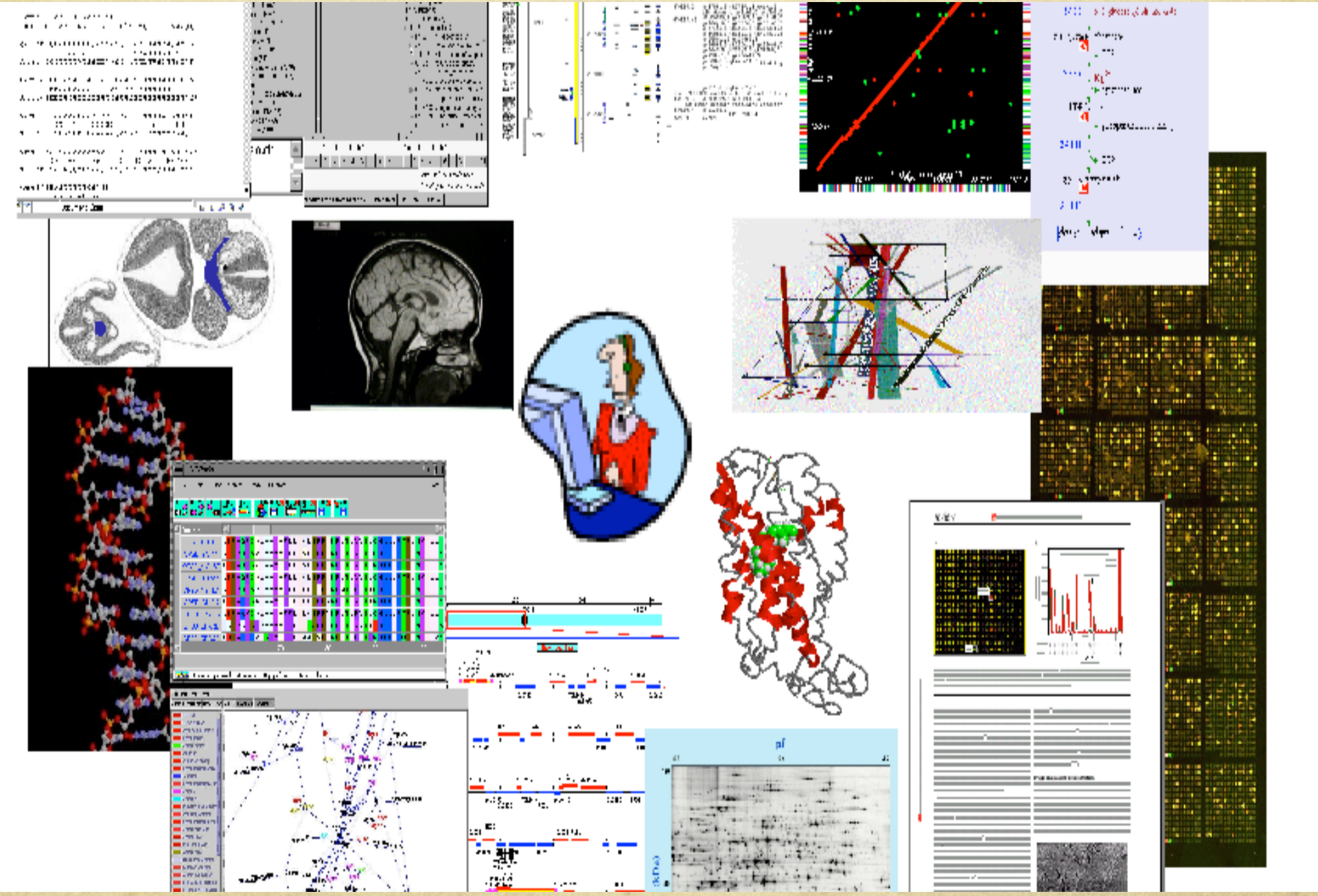
14 October 2013



> ""It took me about thirty minutes. Honestly, it is scary how easy it seems now, in retrospect.""

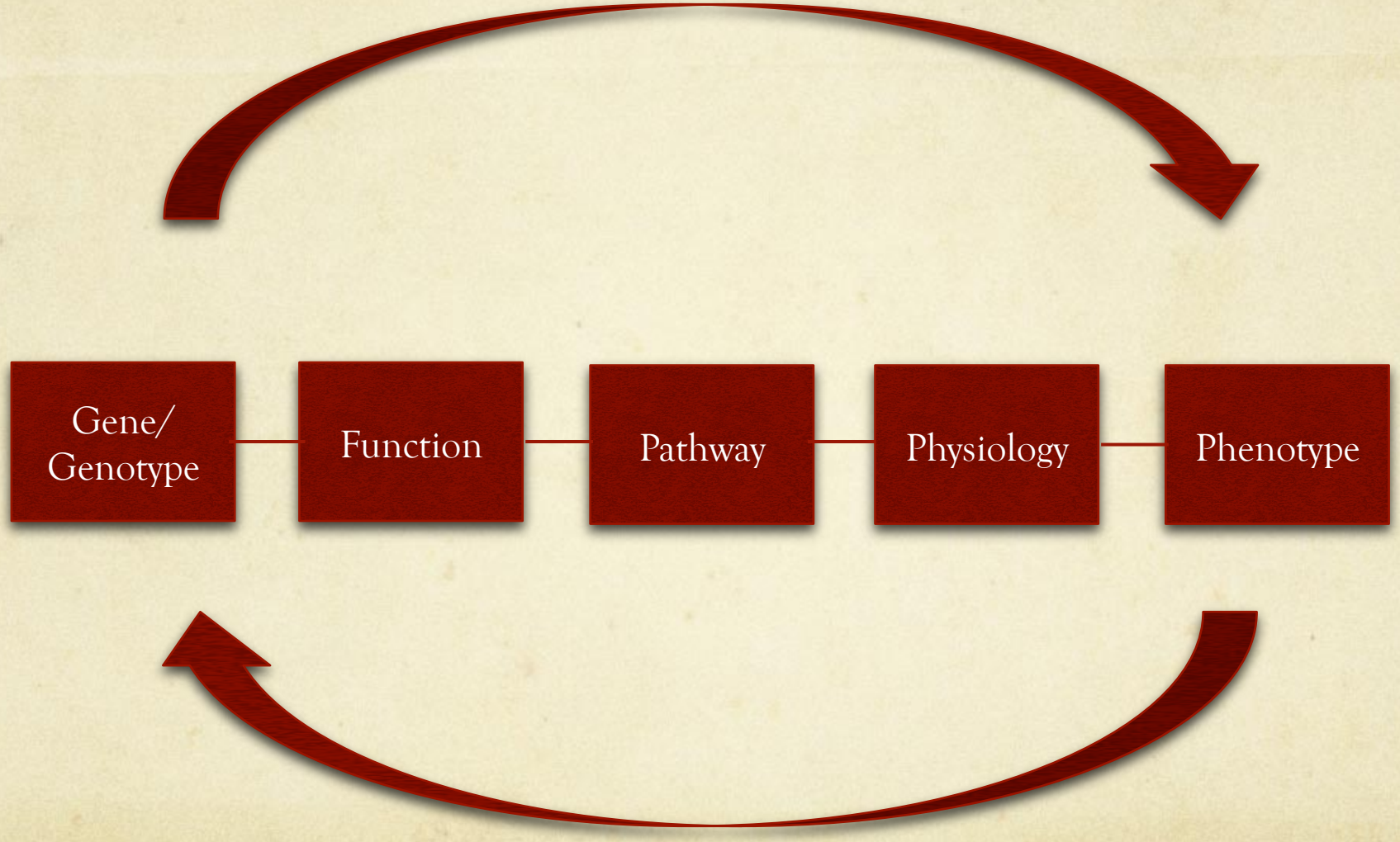
"This is a good story, and there is some promise for future directions," Suthanthiran adds. "It will be nice to see these drugs evaluated in a prospective clinical trial."

Nature doi:10.1038/nature.2013.13944



Reverse Genetics

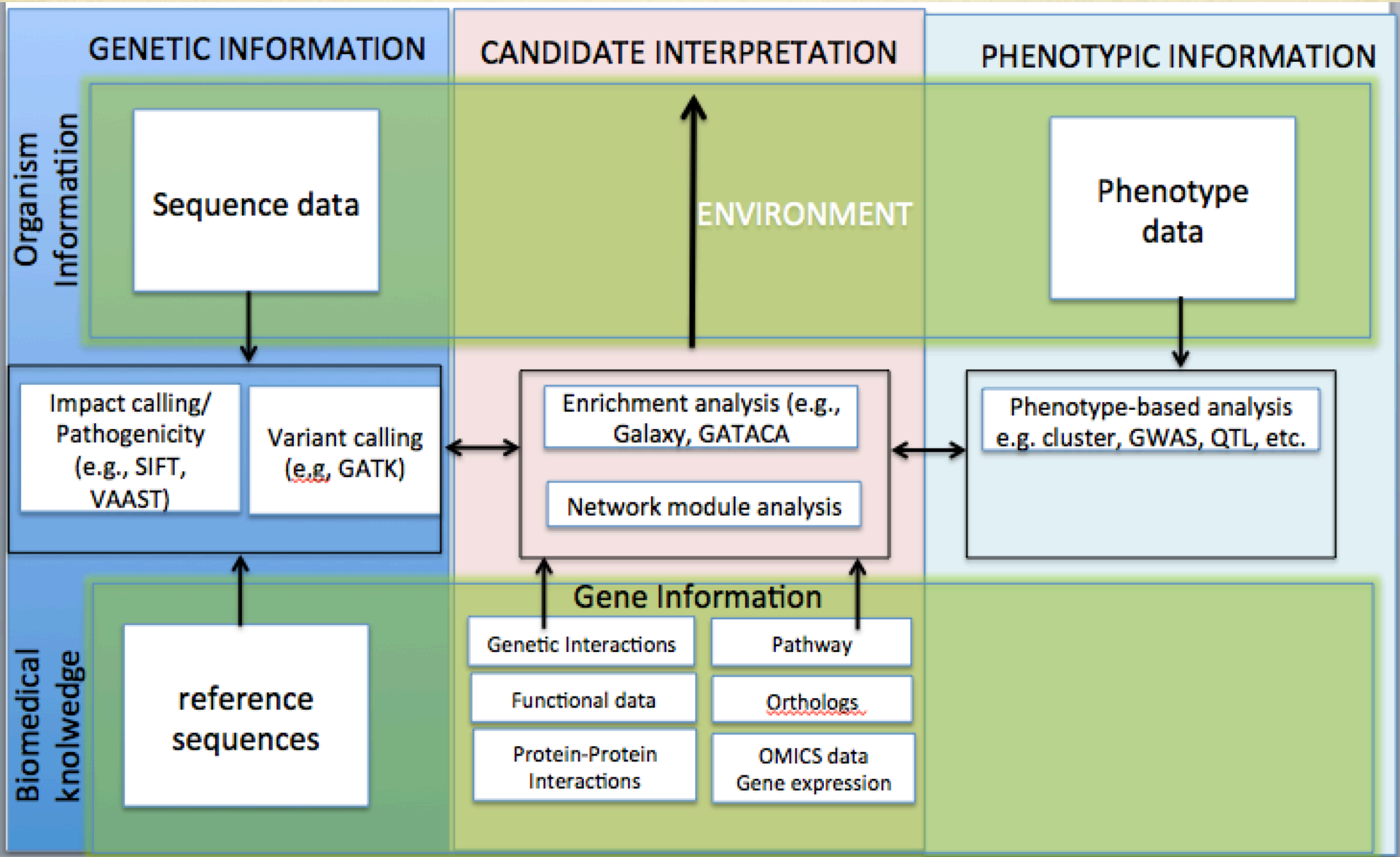
Functional Analysis



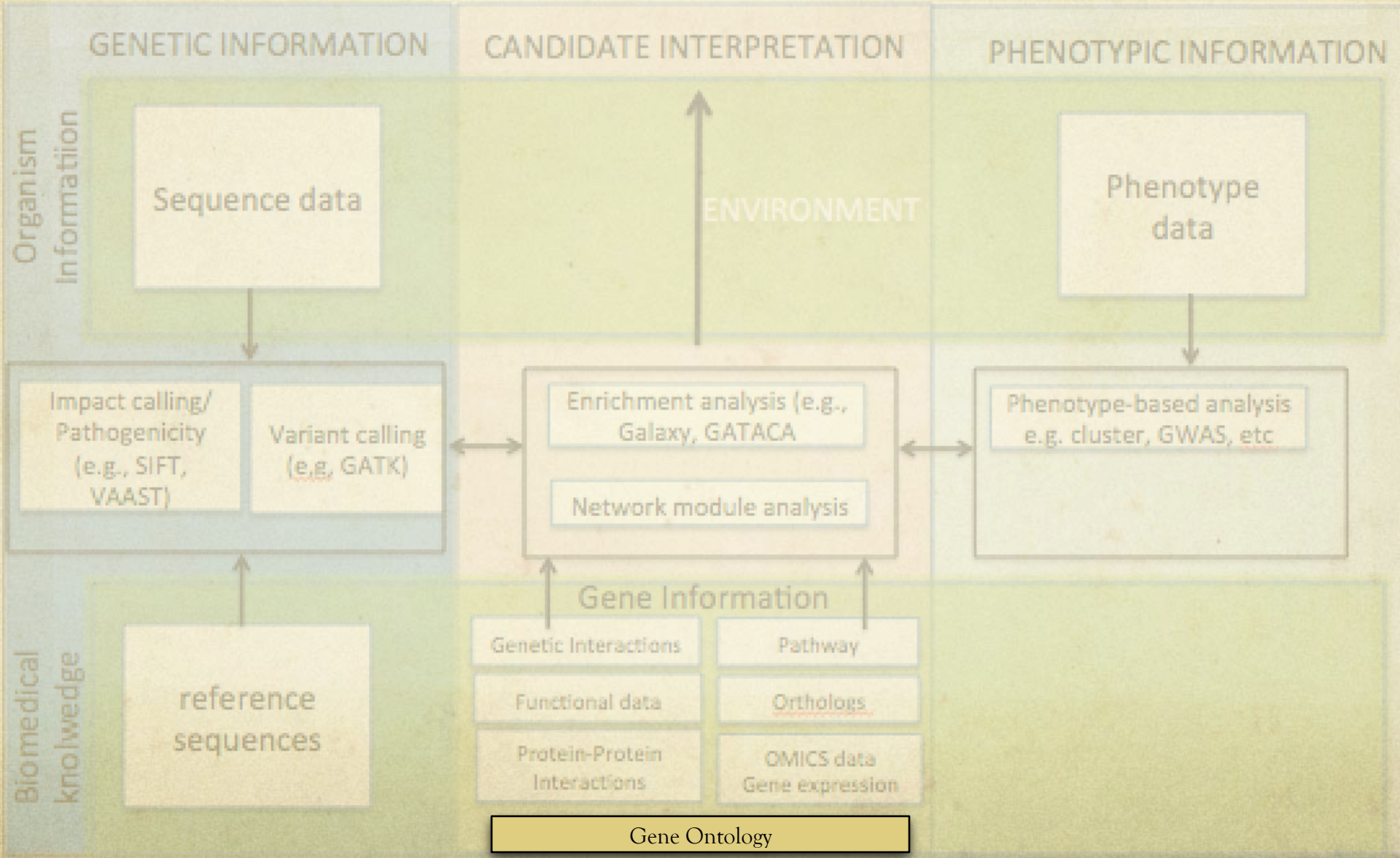
Forward Genetics

Positional Cloning

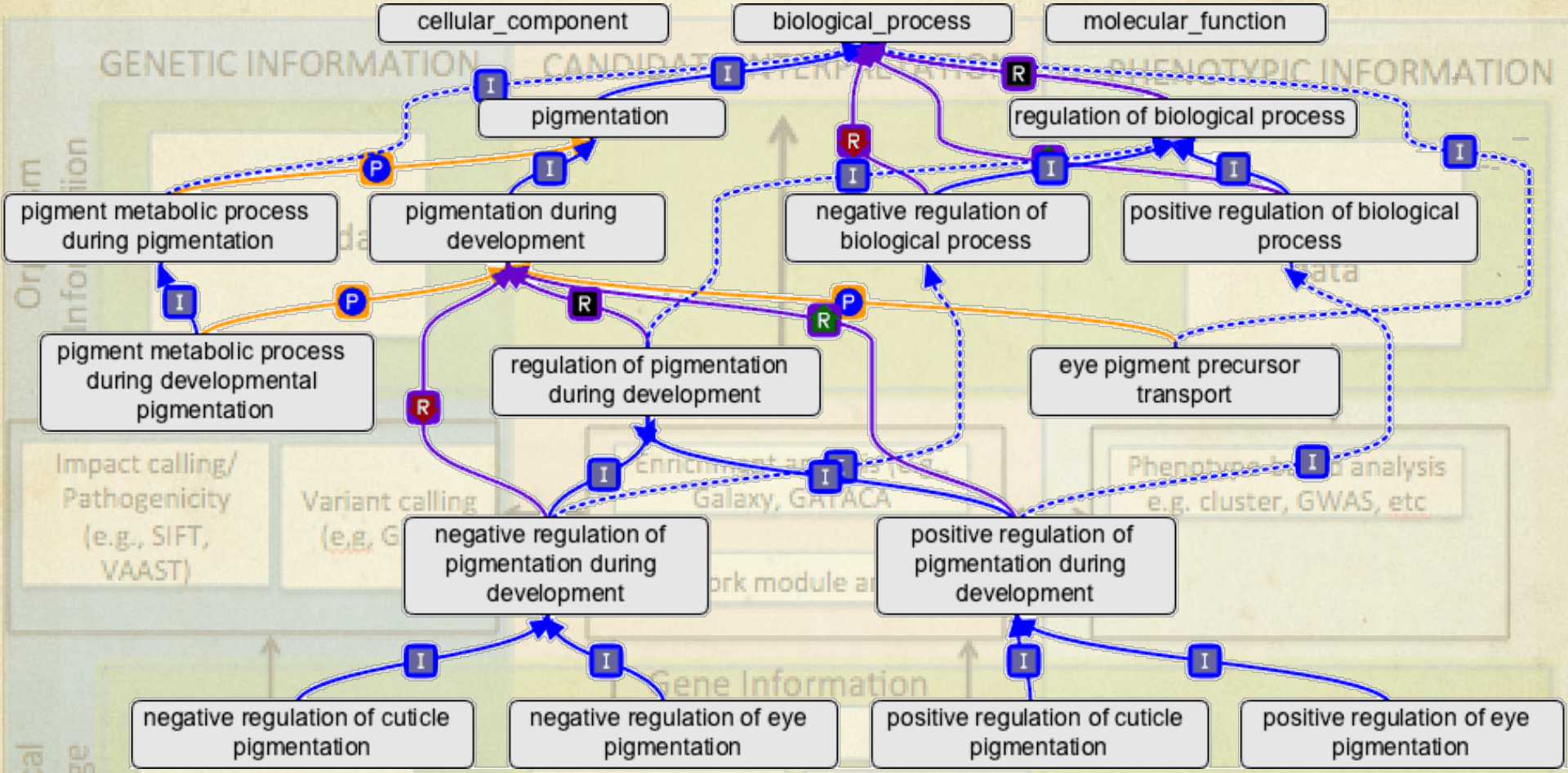
Candidate gene prioritization



Gene Ontology



Gene Ontology

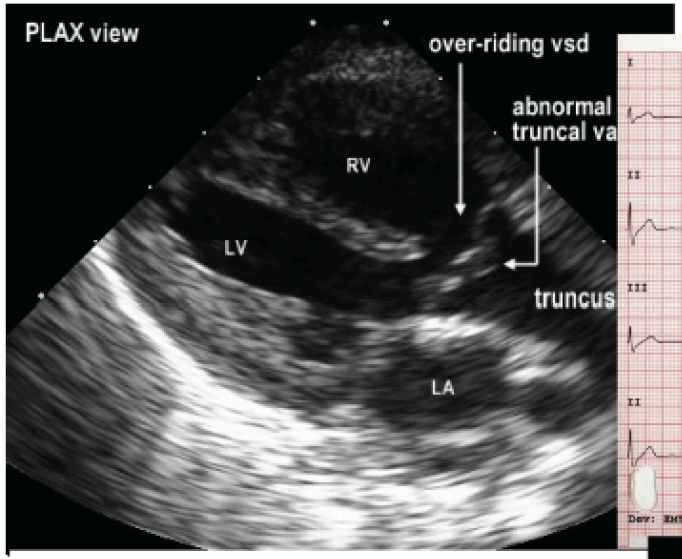


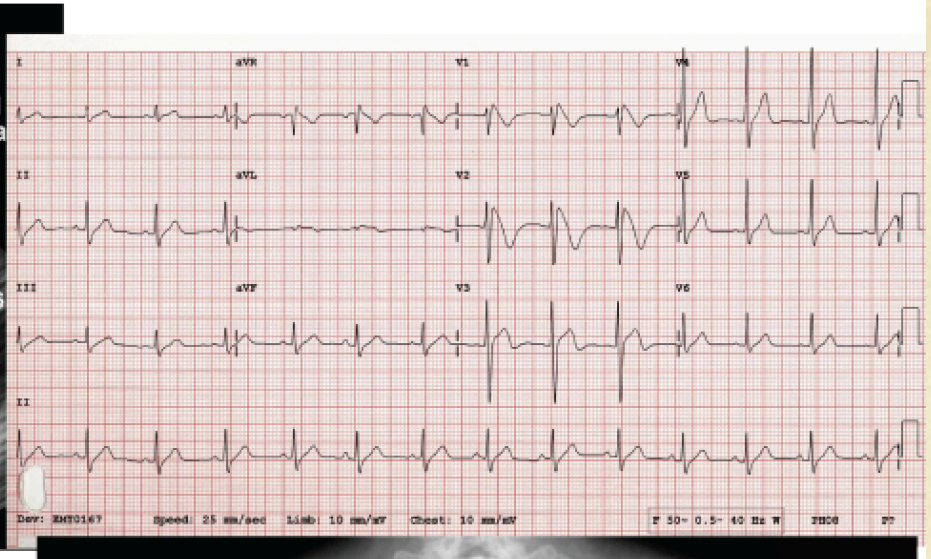
Gene Ontology

How much data?

- Our ability to identify causative variants/variants of interest *depends* on the layer of biological knowledge
- More genetic data will *increase our ability* to prioritise gene candidates (GWAS, QTL, etc.)
- More phenotype data will *alter our potential* for revealing gene candidates

Phenotypes in the clinic



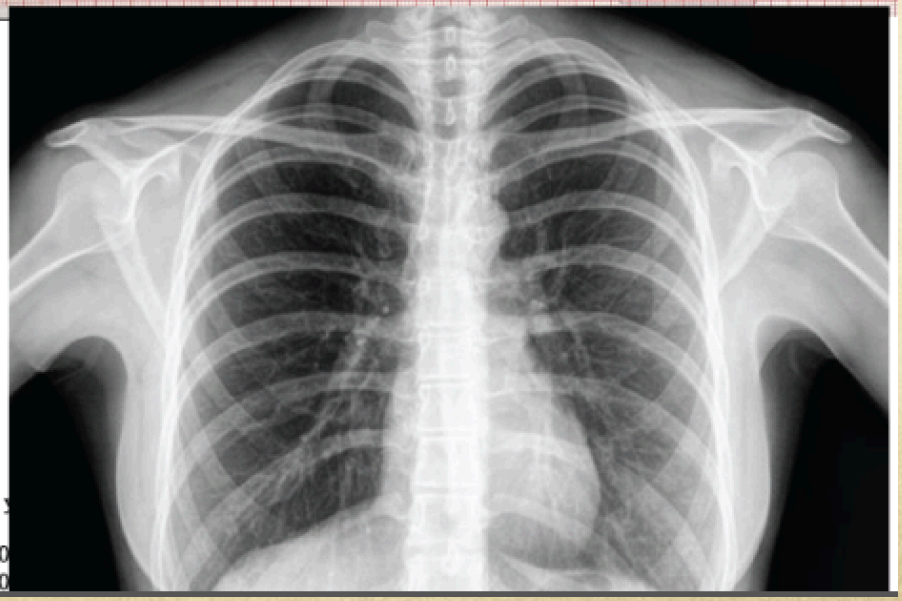


Complete Blood Count:

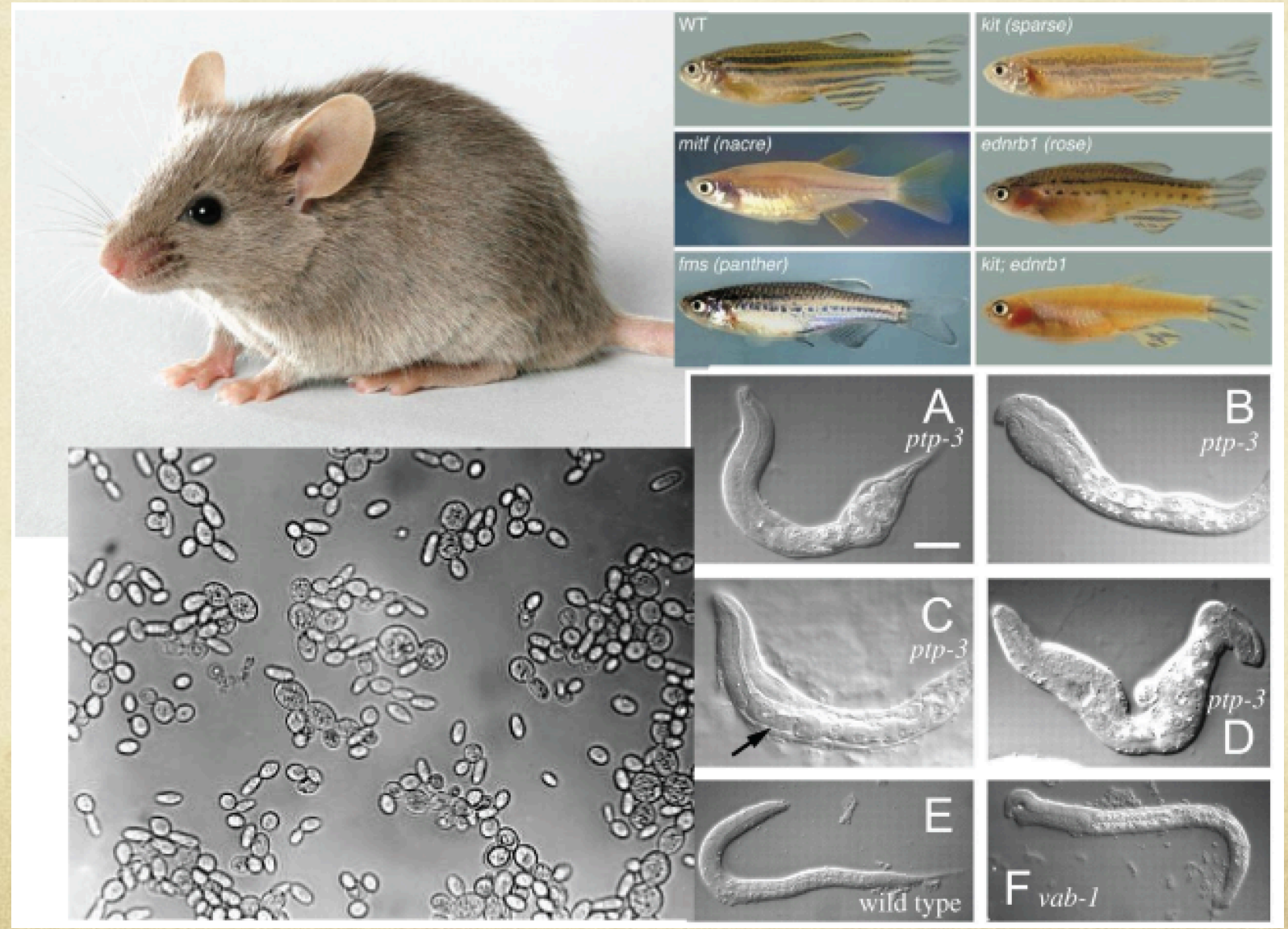
	Patient Value	Normal Range 2 years – 6 years
WBC	8.4 x 10^9 / L	(5.0 – 17.0)
RBC	2.77 x 10^{12} / L	(3.90 – 5.30)
Hgb	7.5 g/dl	(11.5 – 13.5)
Hct	21.8 %	(34.0 – 40.0)
MCV	78.6 fl	(75.0 – 87.0)
MCH	26.9 pg	(25.0 – 31.0)
MCHC	34.2 gm/dl	(31.0 – 36.0)
RDW	17.3 %	(11.5 – 15.0)
PLT	192 x 10^9 / L	(150 – 450)

Differential:

	Absolute	Normal Range Number	2 years – 6 y
Neutrophils	43 %	(3.61)	(1.50 – 8.50
Bands	6 %	(0.50)	(0.00 – 1.00



Animal Model Phenotypes



Plant Phenotypes

Comparative Phenomics

Power of the Phenotype

The meaningful cross species and across domain translation of phenotype is essential → phenotype-driven gene function discovery and comparative pathobiology

Goal - "A platform for facilitating mutual understanding and interoperability of phenotype information *across*

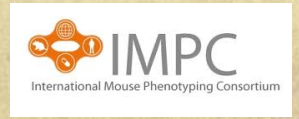
- **species**,
- **domains of knowledge**,

and amongst people and machines" .....

PATO today

PATO is now being used as a community standard for phenotype description

- many consortia (e.g. Phenoscape, The Virtual Human Physiology project (VPH), IMPC, BIRN, NIF)
- most of the major model organism databases, (e.g. example Flybase, Dictybase, Wormbase, Zfin, Mouse genome database (MGD))
- international projects



PATO's Semantic Framework

- Conceptual Layer
- Semantic Components Layer
- Unification Layer
- Integration Layer

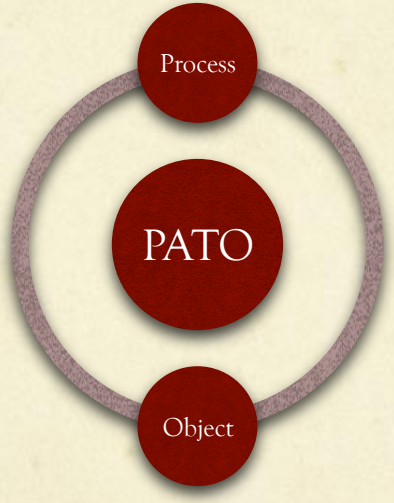
PATO Conceptual Layer

EQ

EQ Model

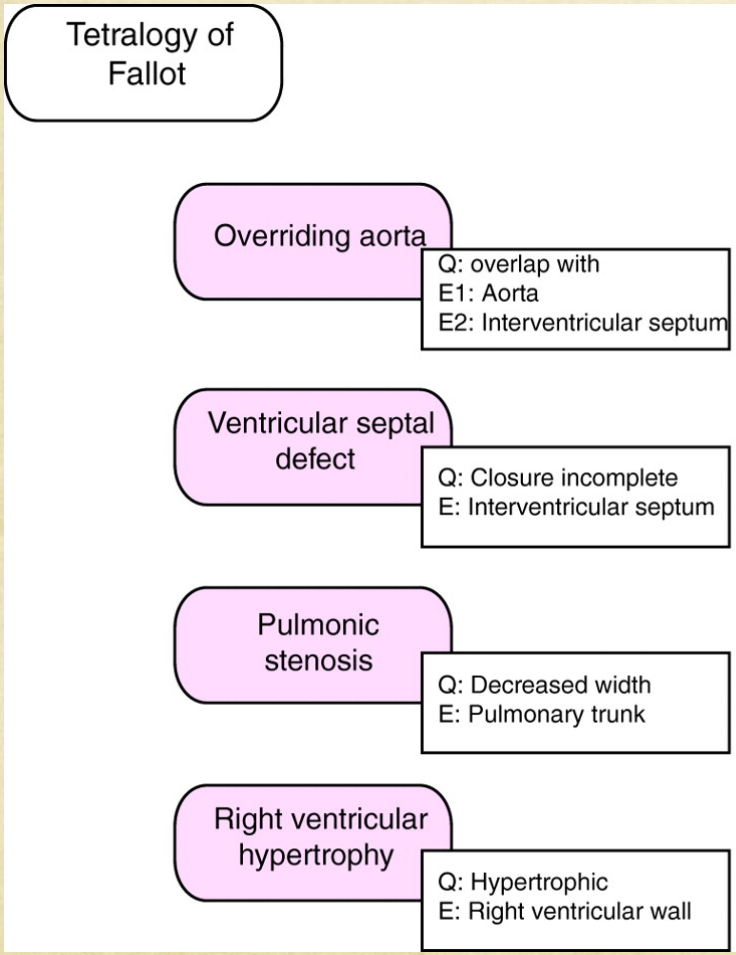
link Entities (E) from GO, CheBI, FMA etc. to Qualities (Q) from PATO →

EQ statements



Genome Biol. 2005;6(1):R8.
Using ontologies to describe mouse phenotypes.
Gkoutos GV, Green EC, Mallon AM, Hancock JM, Davidson D

Endophenotype disaggregation of the tetralogy of Fallot (OMIM:187500)



[Term]id: HP:0001636 ! Tetralogy of Fallot
intersection_of: PATO:0000001 ! Quality
intersection_of: has_part HP:0002623 ! Overriding aorta
intersection_of: has_part HP:0001629 ! Ventricular septal defect
intersection_of: has_part HP:0001642 ! Pulmonic stenosis
intersection_of: has_part HP:0001667 ! Right ventricular hypertrophy

[Term]id: HP:0002623 ! Overriding aorta
intersection_of: PATO:0001590 ! overlap with
intersection_of: inheres_in FMA:3734 ! Aorta
intersection_of: towards FMA:7135 ! Membranous interventricular septum

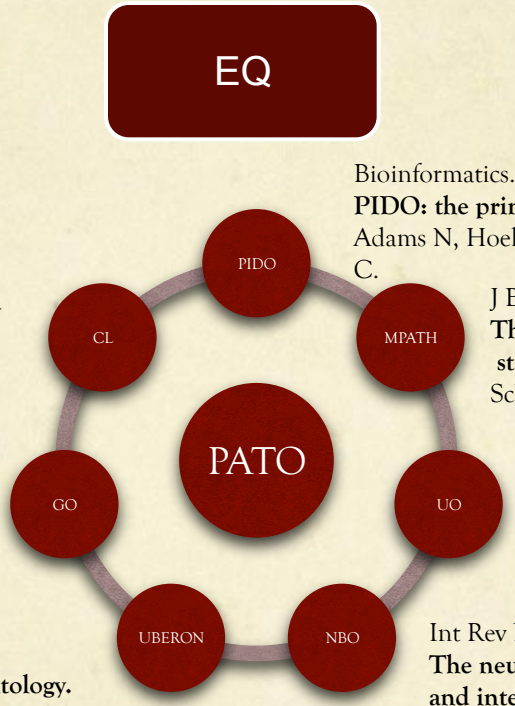
[Term]id: HP:0001629 ! Ventricular septal defect
intersection_of: PATO:0000609 ! closure incomplete
intersection_of: inheres_in FMA:7133 ! Interventricular septum

[Term]id: HP:0001642 ! Pulmonic stenosis
intersection_of: PATO:0000599 ! decreased width
intersection_of: inheres_in FMA:8615 ! Pulmonary arterial trunk

[Term]id: HP:0001667! Right ventricular hypertrophy
intersection_of: PATO:0000584! hypertrophic
intersection_of: inheres_in FMA:9533! Right ventricular wall

Semantic Components Layer

- Behavior
- Cell Phenotype
- Pathology
- UBERON
- Measurements
- etc



Bioinformatics. 2012 Jul 1;28(13):1783-9.
Semantic integration of physiology phenotypes with an application to the Cellular Phenotype Ontology.
Hoehndorf R, Harris MA, Herre H, Rustici G, Gkoutos GV.

Mamm Genome. 2013 Nov 1. [Epub ahead of print]
Analyzing gene expression data in mice with the Neuro Behavior Ontology.
Hoehndorf R, Hancock JM, Hardy NW, Mallon AM, Schofield PN, Gkoutos GV.

Genome Biol. 2012 Jan 31;13(1):R5.
Uberon, an integrative multi-species anatomy ontology.
Mungall CJ, Torniai C, Gkoutos GV, Lewis SE, Haendel MA.

Bioinformatics. 2011 Nov 15;27(22):3193-9.
PIDO: the primary immunodeficiency disease ontology.
Adams N, Hoehndorf R, Gkoutos GV, Hansen G, Hennig C.

J Biomed Semantics. 2013 Sep 13;4(1):18.
The mouse pathology ontology, MPATH; structure and applications
Schofield PN, Sundberg J, Sundberg B, McKerlie C, Gkoutos GV.

Database (Oxford). 2012 Oct 10;2012:bas033.
The Units Ontology: a tool for integrating units of measurement in science.
Gkoutos GV, Schofield PN, Hoehndorf R.

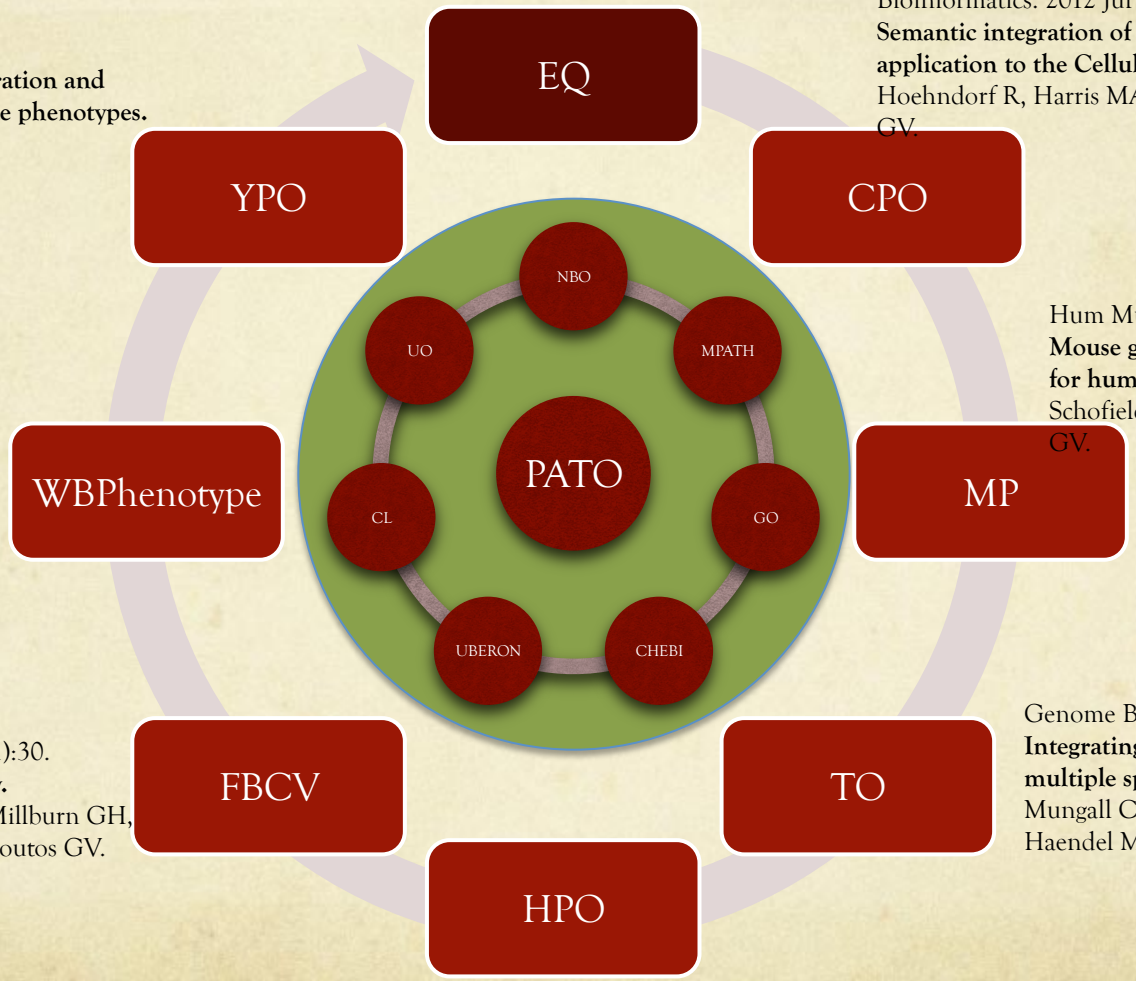
Int Rev Neurobiol. 2012;103:69-87.
The neurobehavior ontology: an ontology for annotation and integration of behavior and behavioral phenotypes.
Gkoutos GV, Schofield PN, Hoehndorf R.

Unification Layer

J Biomed Semantics. 2012 Sep 21;3
Ontology-based cross-species integration and analysis of Saccharomyces cerevisiae phenotypes.
Gkoutos GV, Hoehndorf R.

Bioinformatics. 2012 Jul 1;28(13):1783-9.
Semantic integration of physiology phenotypes with an application to the Cellular Phenotype Ontology.
Hoehndorf R, Harris MA, Herre H, Rustici G, Gkoutos GV.

Provision of PATO based equivalence definitions

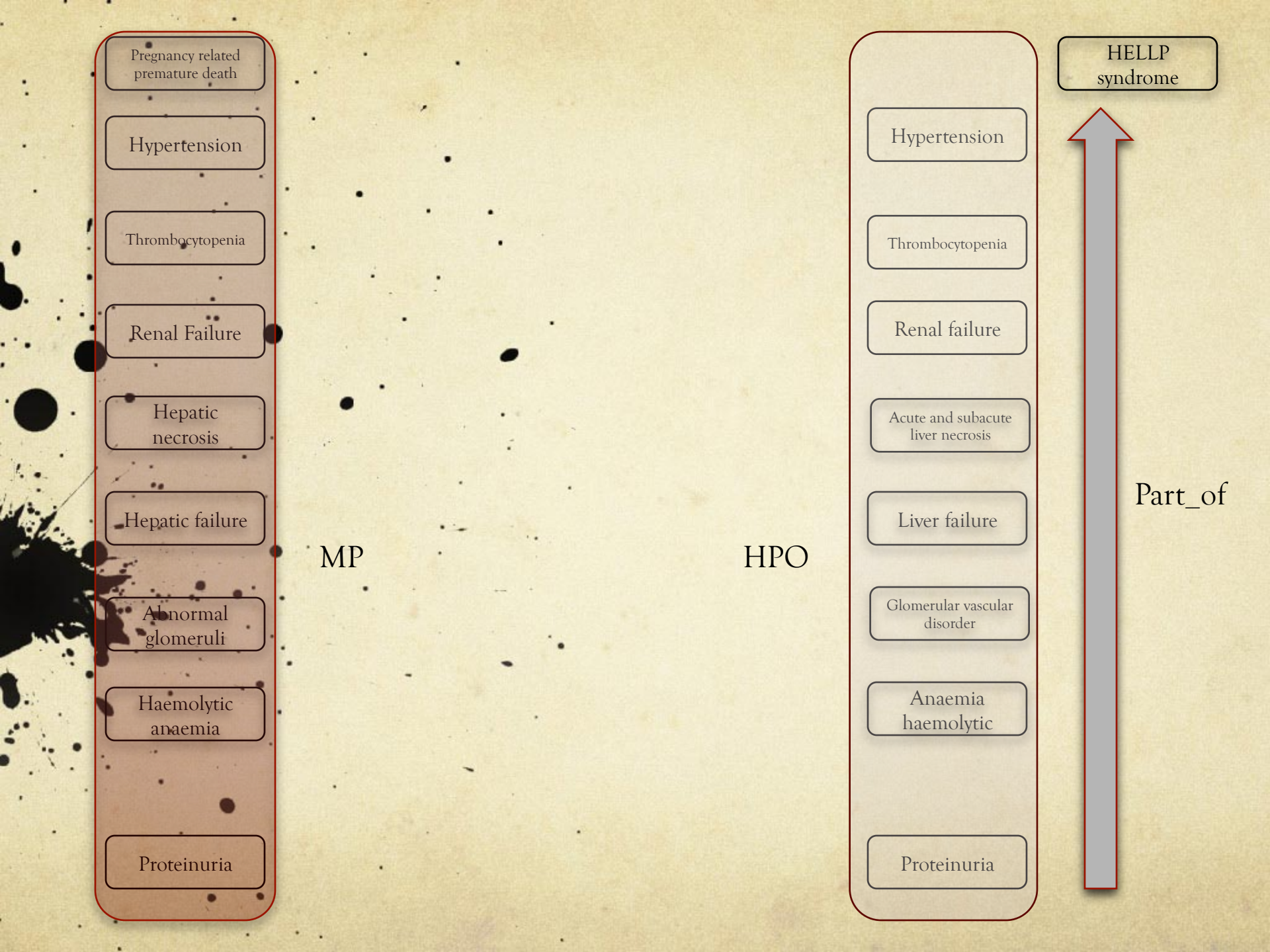


Hum Mutat. 2012 May;33(5):826-36.
Mouse genetic and phenotypic resources for human genetics.
Schofield PN, Hoehndorf R, Gkoutos GV.

J Biomed Semantics. 2013 Oct 18;4(1):30.
The Drosophila phenotype ontology.
Osumi-Sutherland D, Marygold SJ, Millburn GH, Stefancsik R, Falls K, Brown NH, Gkoutos GV.

Genome Biol. 2010 Jan 8;11(1):R2.
Integrating phenotype ontologies across multiple species.
Mungall CJ, Gkoutos GV, Smith CL, Haendel MA, Lewis SE, Ashburner M.

Conf Proc IEEE Eng Med Biol Soc. 2009;2009:7069-72.
Entity/quality-based logical definitions for the human skeletal phenome using PATO.
Gkoutos GV, Mungall C, Dolken S, Ashburner M, Lewis S, Hancock J, Kohler S, Robinson PN.

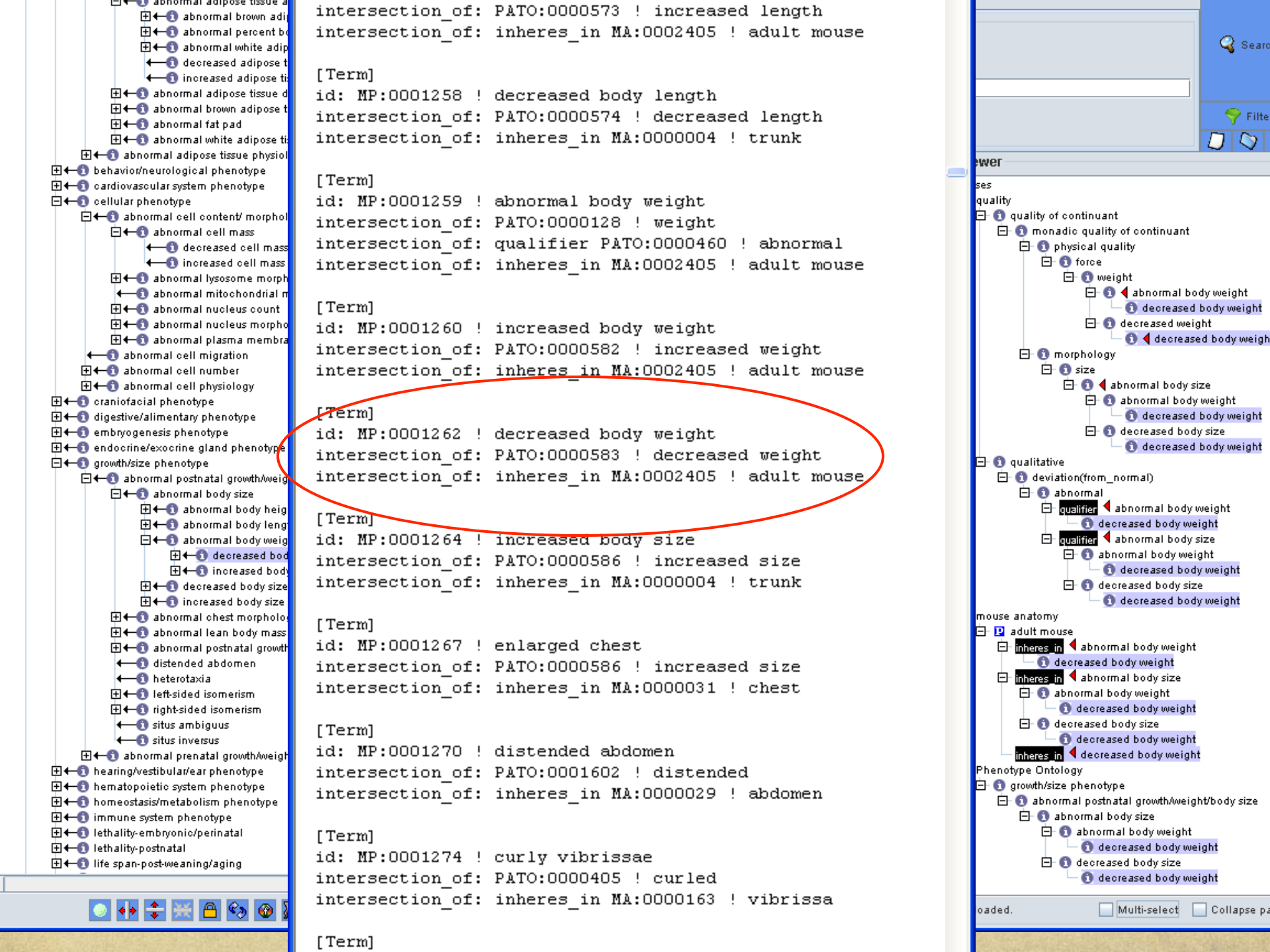

Pregnancy related premature death
Hypertension
Thrombocytopenia
Renal Failure
Hepatic necrosis
Hepatic failure
Abnormal glomeruli
Haemolytic anaemia
Proteinuria
MP
HPO
Hypertension
Thrombocytopenia
Renal failure
Acute and subacute liver necrosis
Liver failure
Glomerular vascular disorder
Anaemia haemolytic
Proteinuria
HELLP syndrome
Part_of

PATO-based definitions

Aristotelian definitions (genus-differentia)

A <**Q**> *which* **inheres_in** an <**E**>

[Term]
id: MP:0005325
name: abnormal renal glomerulus morphology
namespace: mammalian_phenotype_xp
Synonym: abnormal glumeruli morphology
Synonym: abnormal malpighian tuft morphology
def: " lower than normal average weight "[]
is_a: MP:0002827 ! abnormal renal corpuscle morphology
intersection_of: PATO:0000051 ! morphology
intersection_of: qualifier PATO:0000460 ! abnormal
intersection_of: inheres_in MA:0001657 ! glomerulus


intersection_of: PATO:0000573 ! increased length
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001258 ! decreased body length
intersection_of: PATO:0000574 ! decreased length
intersection_of: inheres_in MA:0000004 ! trunk
[Term]
id: MP:0001259 ! abnormal body weight
intersection_of: PATO:0000128 ! weight
intersection_of: qualifier PATO:0000460 ! abnormal
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001260 ! increased body weight
intersection_of: PATO:0000582 ! increased weight
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001262 ! decreased body weight
intersection_of: PATO:0000583 ! decreased weight
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001264 ! increased body size
intersection_of: PATO:0000586 ! increased size
intersection_of: inheres_in MA:0000004 ! trunk
[Term]
id: MP:0001267 ! enlarged chest
intersection_of: PATO:0000586 ! increased size
intersection_of: inheres_in MA:0000031 ! chest
[Term]
id: MP:0001270 ! distended abdomen
intersection_of: PATO:0001602 ! distended
intersection_of: inheres_in MA:0000029 ! abdomen
[Term]
id: MP:0001274 ! curly vibrissae
intersection_of: PATO:0000405 ! curled
intersection_of: inheres_in MA:0000163 ! vibrissa
[Term]
abnormal adipose tissue physiol
behavior/neurological phenotype
cardiovascular system phenotype
cellular phenotype
abnormal cell content/ morphol
abnormal cell mass
decreased cell mass
increased cell mass
abnormal lysosome morph
abnormal mitochondrial m
abnormal nucleus count
abnormal nucleus morpho
abnormal plasma membra
abnormal cell migration
abnormal cell number
abnormal cell physiology
craniofacial phenotype
digestive/alimentary phenotype
embryogenesis phenotype
endocrine/exocrine gland phenotype
growth/size phenotype
abnormal postnatal growth/weig
abnormal body size
abnormal body heig
abnormal body leng
abnormal body weig
decreased bod
increased body
decreased body size
increased body size
abnormal chest morpholo
abnormal lean body mass
abnormal postnatal growth
distended abdomen
heterotaxia
left-sided isomerism
right-sided isomerism
situs ambiguus
situs inversus
abnormal prenatal growth/weigh
hearing/vestibular/ear phenotype
hematopoietic system phenotype
homeostasis/metabolism phenotype
immune system phenotype
lethality-embryonic/perinatal
lethality-postnatal
life span-post-weaning/aging
Search
Filte
ewer
ses
quality
quality of continuant
monadic quality of continuant
physical quality
force
weight
abnormal body weight
decreased body weight
decreased weight
decreased body weigh
morphology
size
abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
qualitative
deviation(from_normal)
abnormal
qualifier abnormal body weight
decreased body weight
qualifier abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
mouse anatomy
P adult mouse
inheres_in abnormal body weight
decreased body weight
inheres_in abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
inheres_in decreased body weight
Phenotype Ontology
growth/size phenotype
abnormal postnatal growth/weight/body size
abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
oaded.
Multi-select
Collapse pa

MP - PATO based definitions

MP term	MP Definition	Entity	Quality
cataract MP:0001304	complete or partial opacity of the lens	lens MA:0000275 FMA:58241	opaque PATO:0000963
jaundice MP:0000611	clinical manifestation of hyperbilirubinemia, with deposition of bile pigments in the skin, resulting in yellowish staining of the skin and mucous membranes	skin MA:0000151 FMA:7163	yellow PATO:0000324
		skin mucous gland MA:0000148 mucous gland FMA:62888	yellow PATO:0000324
		pigment accumulation in tissues GO:0043480	yellow PATO:0000324
		pigment accumulation in tissues GO:0043480	mislocalized PATO:0000628

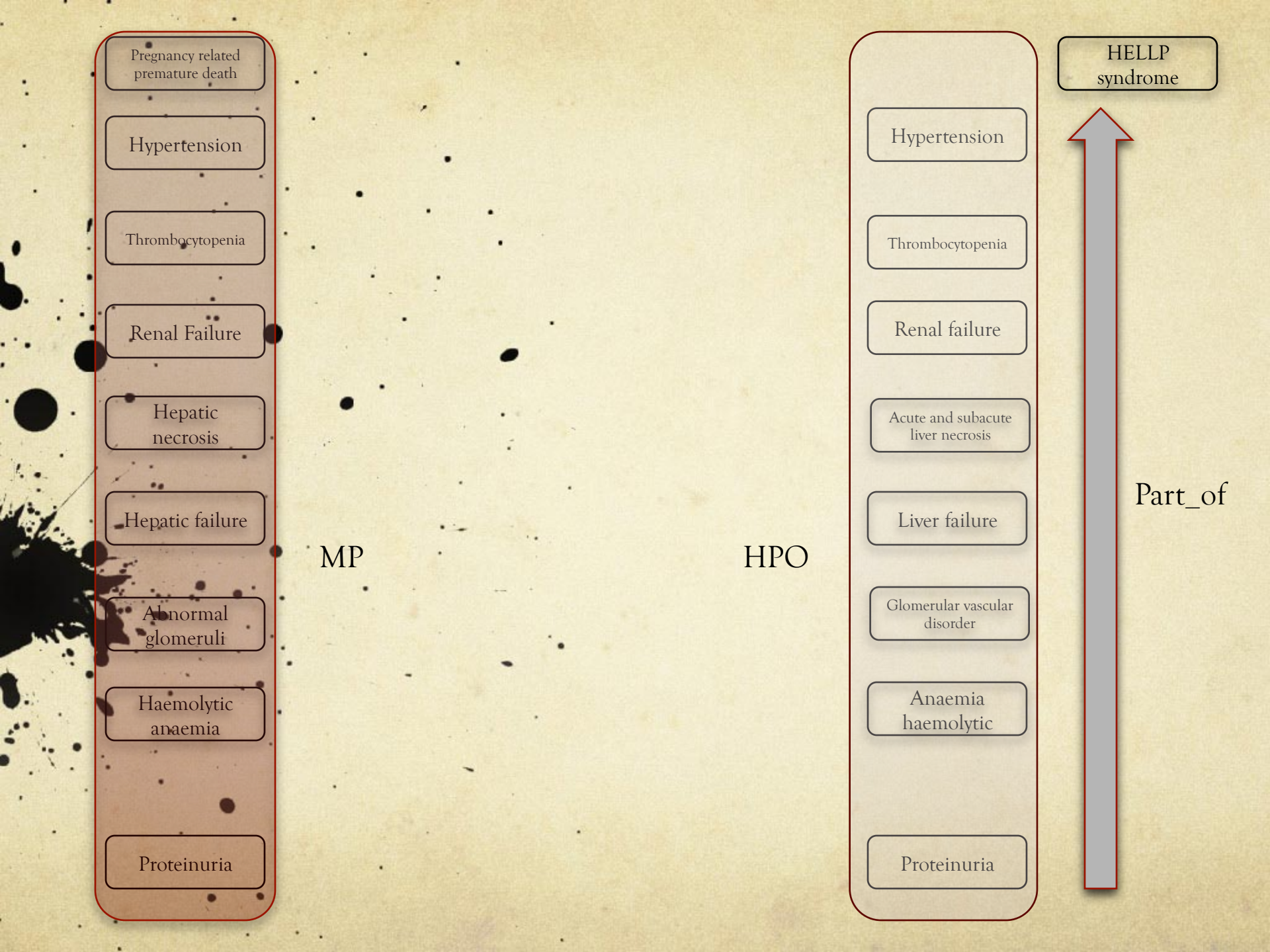
HPO-PATO based definitions

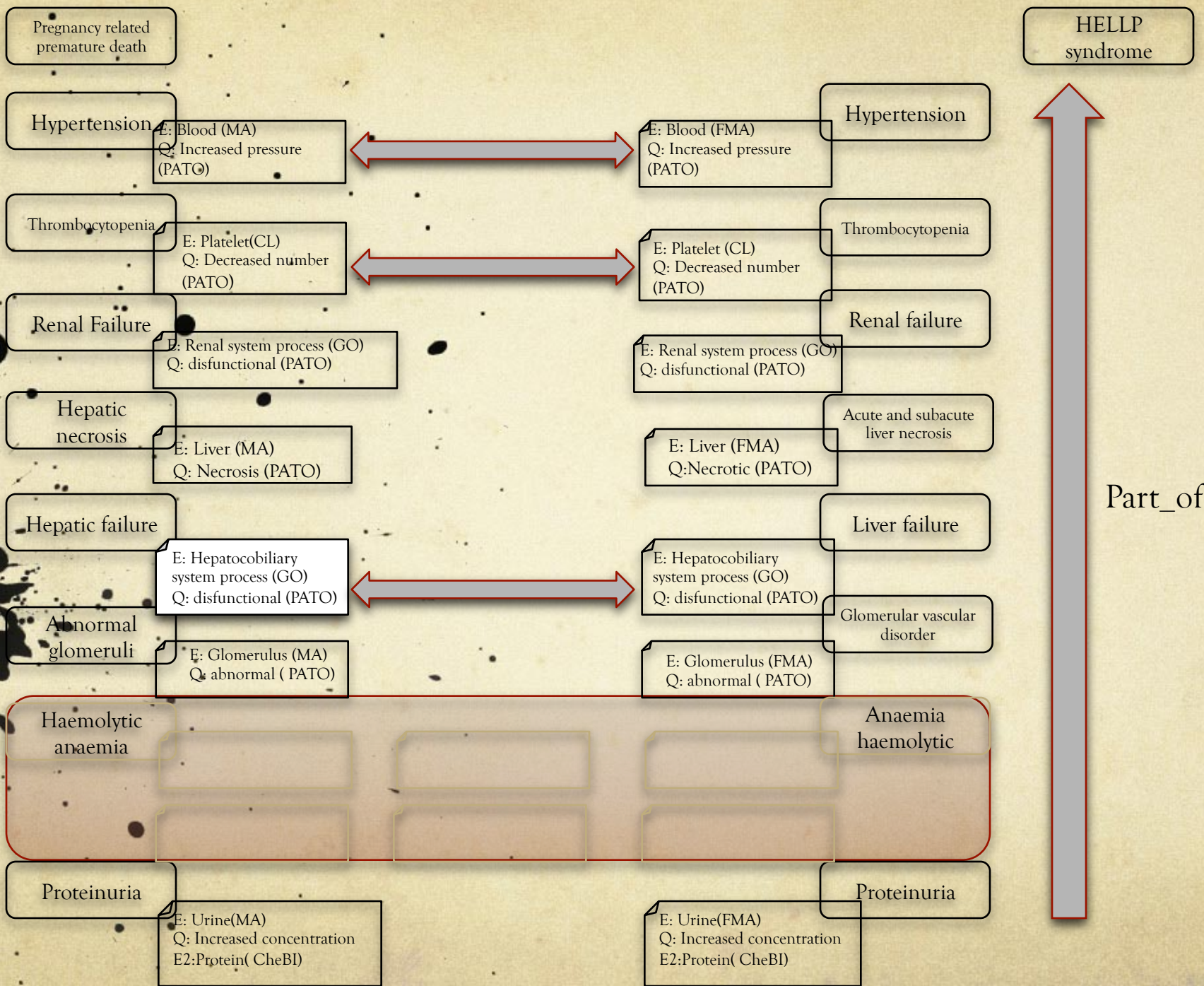
- OBO format

```
[Term]
id: HP:0004349 ! Reduced bone mineral density
intersection_of: PATO:0001790 ! decreased density
intersection_of: inheres_in FMA:30317 ! bone
```

- OWL format

```
Class: Hypoglycemia
EquivalentTo:
'decreased concentration'
and towards some 'glucose'
and inheres_in some 'portion of blood'
and qualifier some 'abnormal'
```

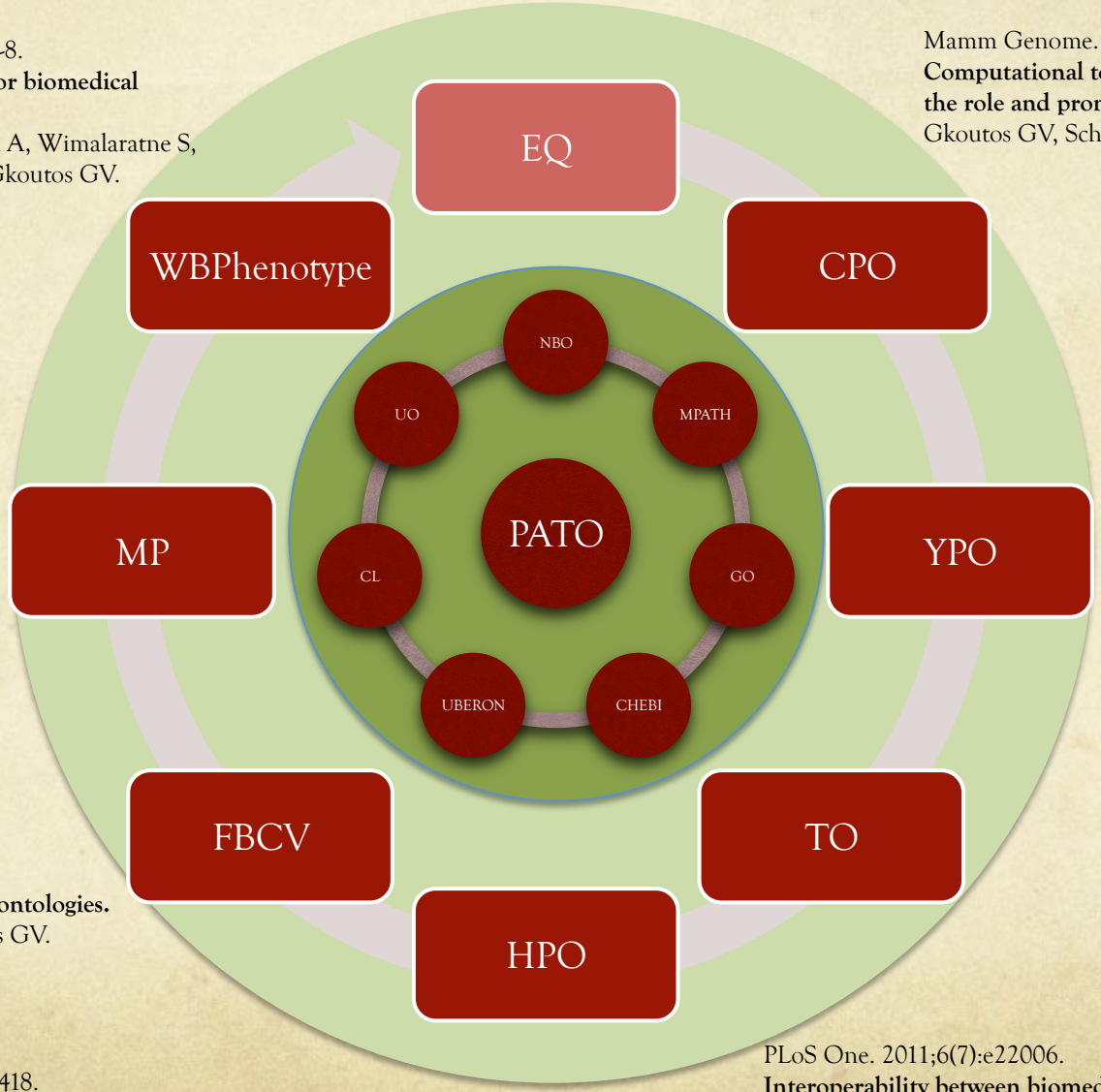

Pregnancy related premature death
Hypertension
Thrombocytopenia
Renal Failure
Hepatic necrosis
Hepatic failure
Abnormal glomeruli
Haemolytic anaemia
Proteinuria
MP
HPO
Hypertension
Thrombocytopenia
Renal failure
Acute and subacute liver necrosis
Liver failure
Glomerular vascular disorder
Anaemia haemolytic
Proteinuria
HELLP syndrome
Part_of


Pregnancy related premature death
HELLP syndrome
Hypertension
E: Blood (MA)
Q: Increased pressure (PATO)
Hypertension
E: Blood (FMA)
Q: Increased pressure (PATO)
Thrombocytopenia
E: Platelet(CL)
Q: Decreased number (PATO)
Thrombocytopenia
E: Platelet (CL)
Q: Decreased number (PATO)
Renal Failure
E: Renal system process (GO)
Q: disfunctional (PATO)
Renal failure
E: Renal system process (GO)
Q: disfunctional (PATO)
Hepatic necrosis
E: Liver (MA)
Q: Necrosis (PATO)
Acute and subacute liver necrosis
E: Liver (FMA)
Q:Necrotic (PATO)
Hepatic failure
E: Hepatocobiliary system process (GO)
Q: disfunctional (PATO)
Liver failure
E: Hepatocobiliary system process (GO)
Q: disfunctional (PATO)
Abnormal glomeruli
E: Glomerulus (MA)
Q: abnormal (PATO)
Glomerular vascular disorder
E: Glomerulus (FMA)
Q: abnormal (PATO)
Haemolytic anaemia
Anaemia haemolytic
Proteinuria
E: Urine(MA)
Q: Increased concentration
E2:Protein(CheBI)
Proteinuria
E: Urine(FMA)
Q: Increased concentration
E2:Protein(CheBI)
Part_of

Formalisation Layer

Bioinformatics. 2011 Apr 1;27(7):1001-8.
A common layer of interoperability for biomedical ontologies based on OWL EL.
Hoehndorf R, Dumontier M, Oellrich A, Wimalaratne S, Rebholz-Schuhmann D, Schofield P, Gkoutos GV.

Mamm Genome. 2012 Oct;23(9-10):669-79.
Computational tools for comparative phenomics: the role and promise of ontologies.
Gkoutos GV, Schofield PN, Hoehndorf R.

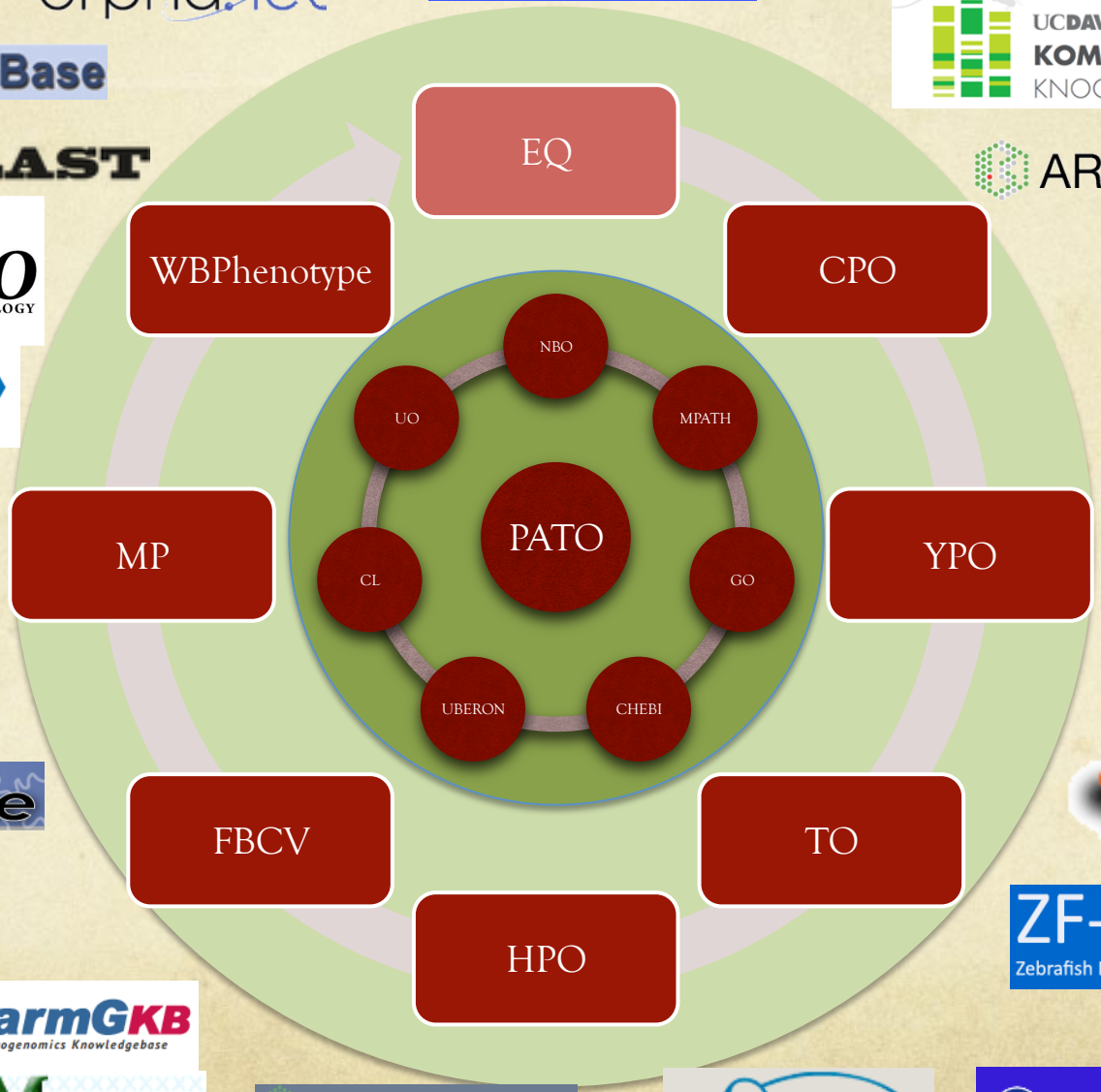


Brief Bioinform. 2012 Sep 8.
Evaluation of research in biomedical ontologies.
Hoehndorf R, Dumontier M, Gkoutos GV.

BMC Bioinformatics. 2011 Oct 27;12:418.
Improving ontologies by automatic reasoning and evaluation of logical definitions.
Köhler S, Bauer S, Mungall CJ, Carletti G, Smith CL, Schofield P, Gkoutos GV, Robinson PN.

PLoS One. 2011;6(7):e22006.
Interoperability between biomedical ontologies through relation expansion, upper-level ontologies and automatic reasoning.
Hoehndorf R, Oellrich A, Schuhmann D, Schofield P, Gkoutos GV.

Integration Layer



Pathbase
European mutant mouse pathology database

orphanet

STITCH 3.1

DrugBank
Open Data Drug & Drug Target Database

dictyBase

UCDAVIS
KOMP Phenotyping Pilot
KNOCKOUT MOUSE PROJECT

PhenomeBLAST

ARRAYEXPRESS

RICORDO
INTEROPERABLE ANATOMY AND PHYSIOLOGY

EMPRESS

EUMODIC

QB

IMPC
International Mouse Phenotyping Consortium

ZF-HEALTH
Zebrafish Regulomics for Human Health

MGI

eumorphia

WormBase

SGD

PHENOSCAPE

ZF-HEALTH
Zebrafish Regulomics for Human Health

PharmGKB
The Pharmacogenomics Knowledgebase

PomBase

OMIM
Online Mendelian Inheritance in Man

SIDER 2
Side Effect Resource

europhenome

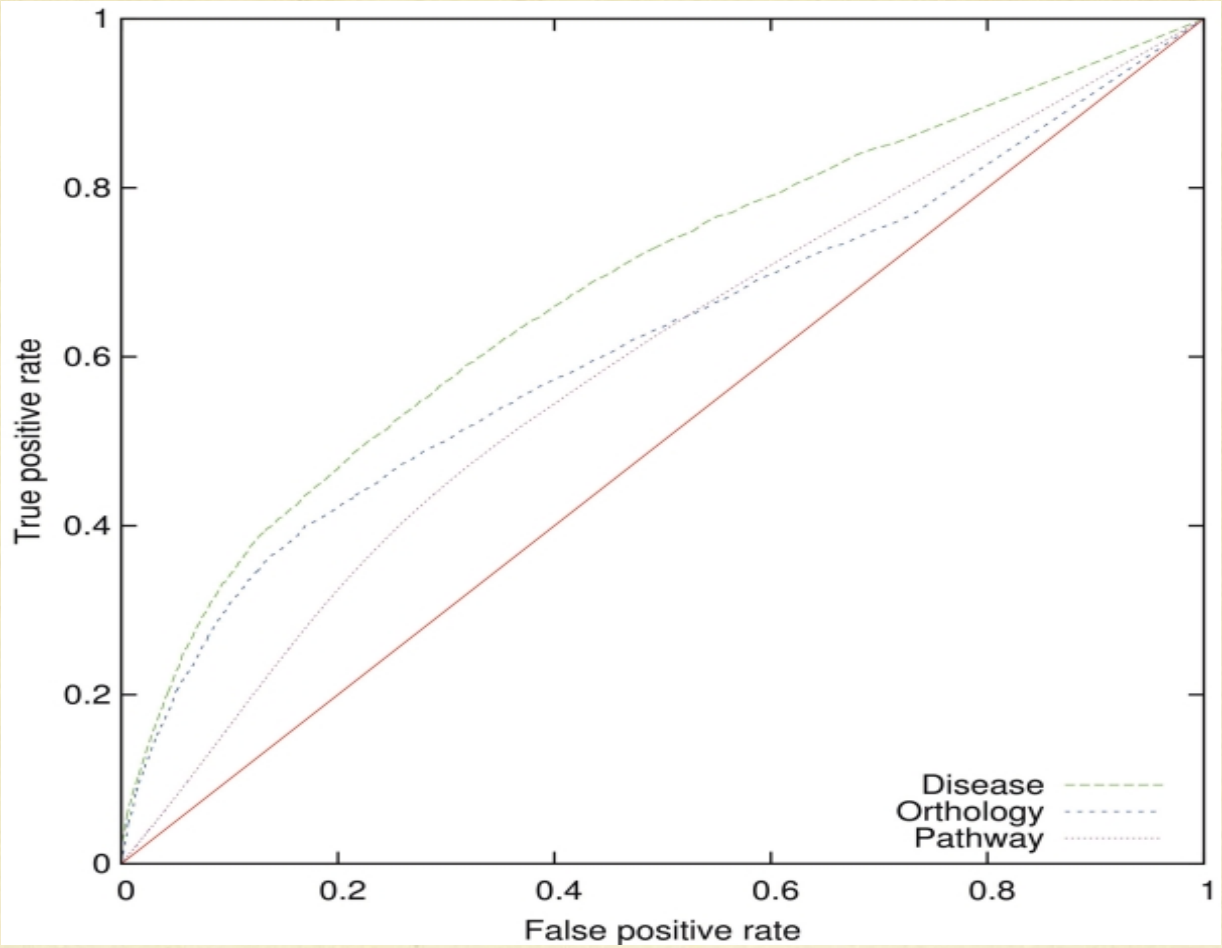
RIKEN

Cross Species Data Integration

Cross species integration framework

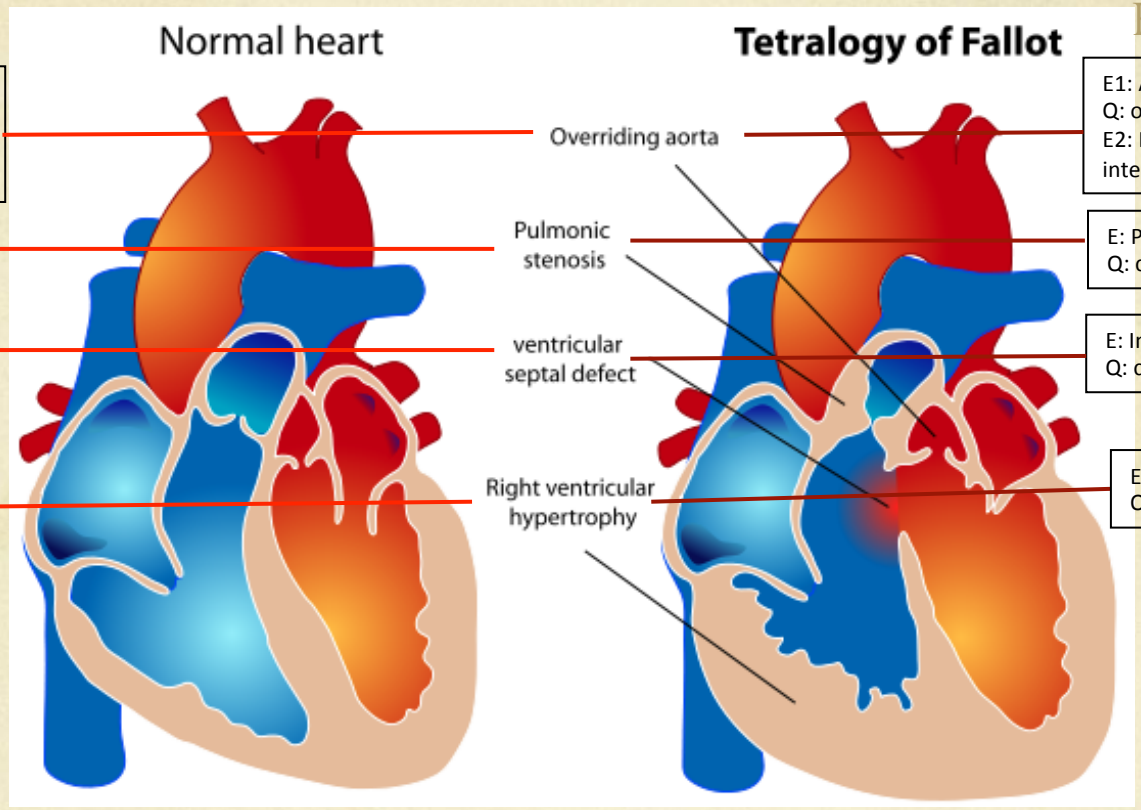
- A PATO-based cross species phenotype network based on experimental data from 5 model organisms yeast, fly, worm, fish and mouse and human disease phenotypes (OMIM)
- integration of anatomy and phenotype ontologies
  - more than 1.000,000 classes and 3,500,000 axioms
- PhenomeNET forms a network with more than 310.000 complex phenotype nodes representing complex phenotypes
- Semantic phenotype similarity - pairwise comparison of disease and animal phenotypes

PhenomeNet

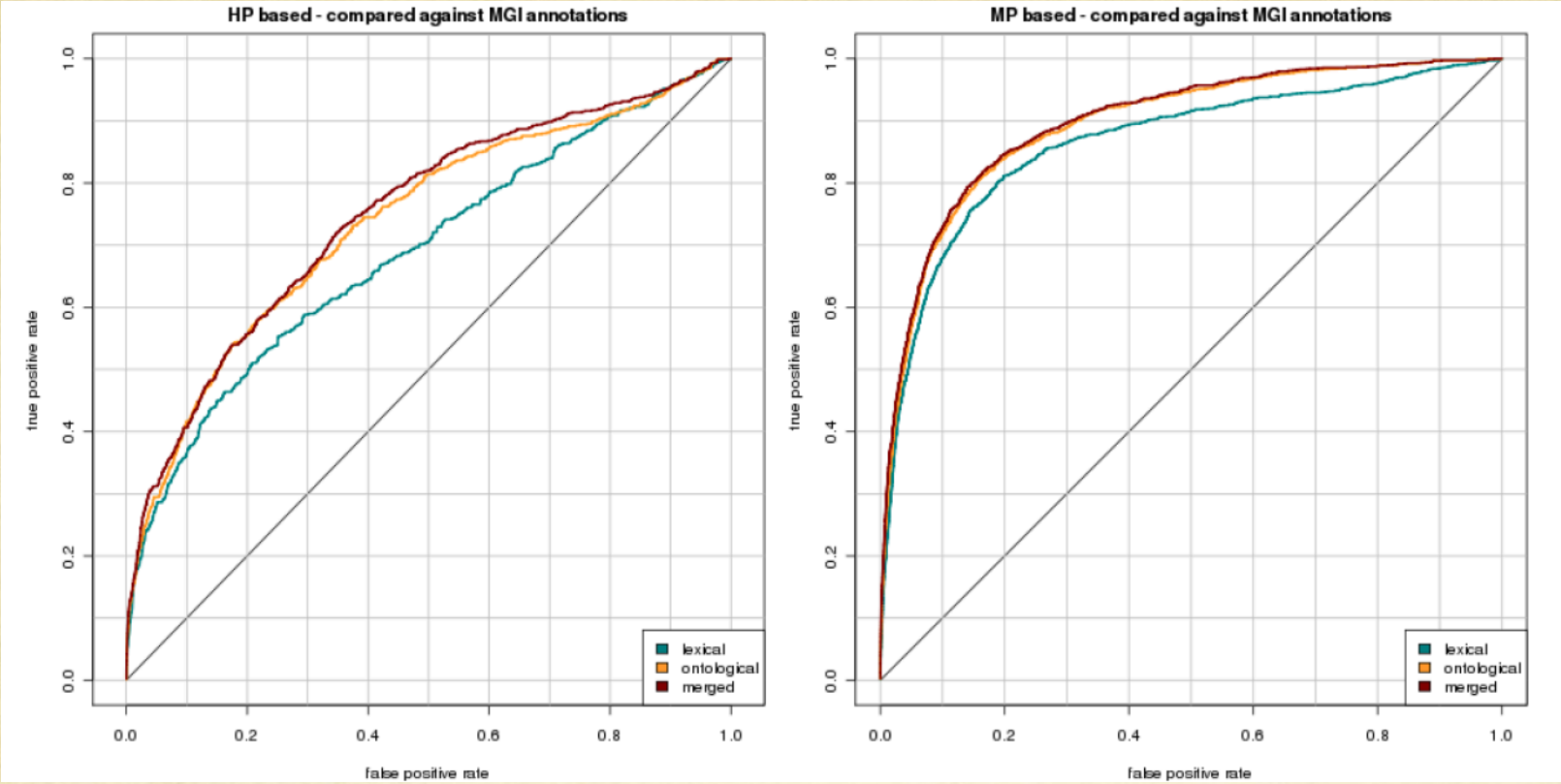


- quantitative evaluation based on predicting orthology, pathway, disease
- Receiver Operating Characteristic (ROC) Curve analysis
- Area Under Curve (AUC) = 0.7

Candidate disease gene prioritization



- Predict all known human and mouse disease genes
- Adam19 and Fgf15 mouse genes
- using zebrafish phenotypes - mammalian homologues of Cx36.7 and Nkx2.5 are involved in TOF



AUC = 0.9

- Enhance the network e.g.
  - Semantics e.g Behavior and pathology related phenotypes etc.
  - Methods e.g. text mining, machine learning etc.
  - Resources e.g. OrphaNet
- PhenomeNET now significantly outperforms previous phenotype-based approaches of predicting gene–disease associations
- Performance matches gene prioritization methods based on prior information about molecular causes of a disease


ClinVar
EMPReSS
PomBase
SGD
OMIM
Online Mendelian Inheritance in Man
UCDAVIS
KOMP Phenotyping Pilot
KNOCKOUT MOUSE PROJECT
dictyBase
ISCA
dbGAP
IRDiRC
Pathbase
European mutant mouse pathology database
PhenomeBLAST
ARRAYEXPRESS
REACTOME
europhenome
EUMODIC
MGI
PHENOSCAPE
orphanet
DECIPHER
v5.1
GRCh37
WormBase
ZF-HEALTH
Zebrafish Regulomics for Human Health
EMMA
RIKEN
IMPC
International Mouse Phenotyping Consortium
dbSNP

Ontologies &Biomedical Research

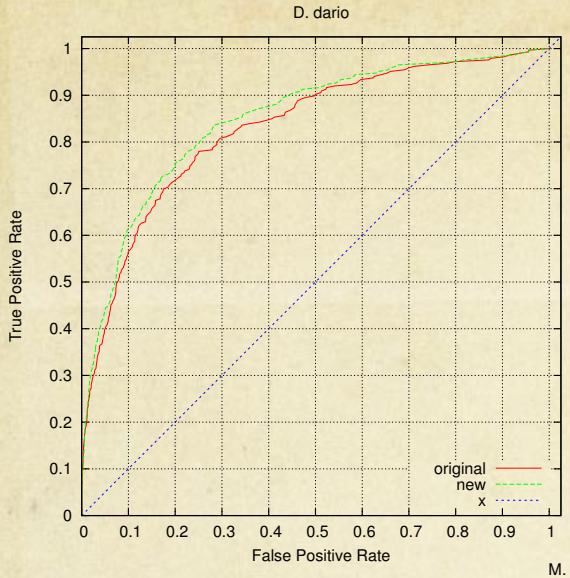
- Gene Function Determination
- Systematic Genome-Wide Phenotyping
- From Genotype to Phenotype
- Candidate Disease Gene Prioritisation
- Rare and Orphan Diseases
- Diagnostics Strategies to support Identification of Causative Genes
- Big Data
- Pharmacogenomics

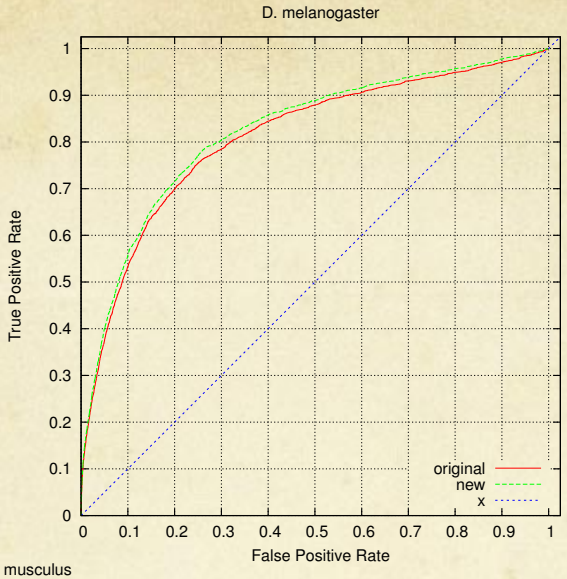
Gene function determination

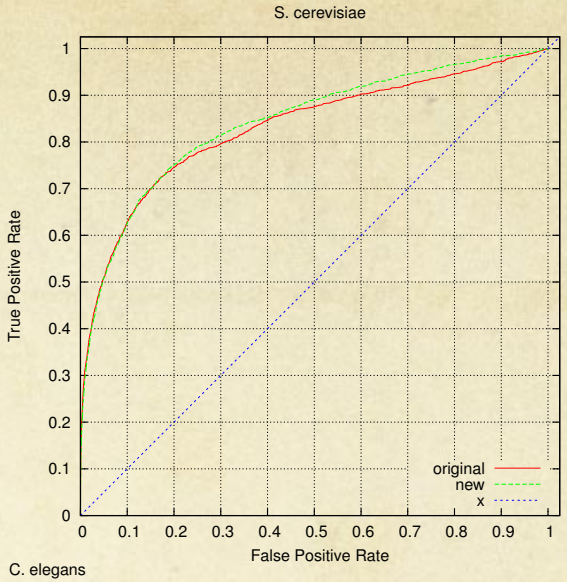
- phenotype data are commonly annotated, automatically, semi-automatically and manually, with terms from species-specific phenotype ontologies,

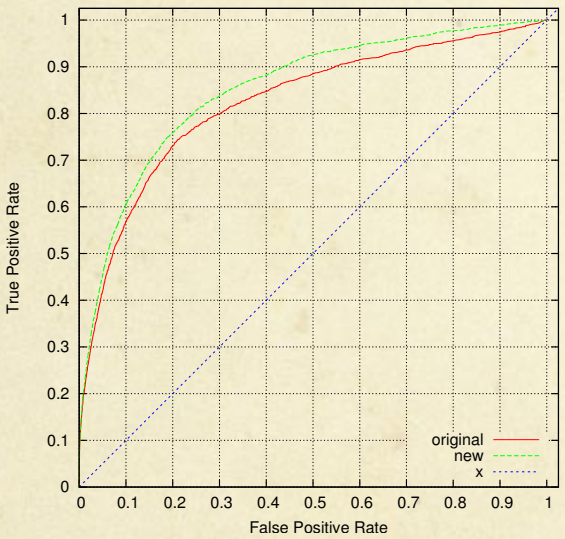
- utilize phenotype ontologies → infer the functions (impaired given a phenotype observation)

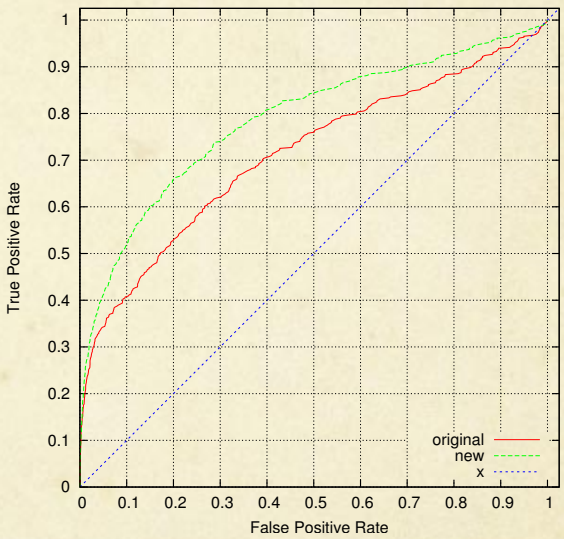
- phenotype data from 5 different species → evaluation
  - manually for biological correctness
  - predict genetic interactions (semantic similarity between genes)











Species	Inferred	p-value
Fish	1717	0.24
Yeast	14047	0.162
Fly	1633	0.041
Worm	9221	0.03
Mouse	15693	0.00007

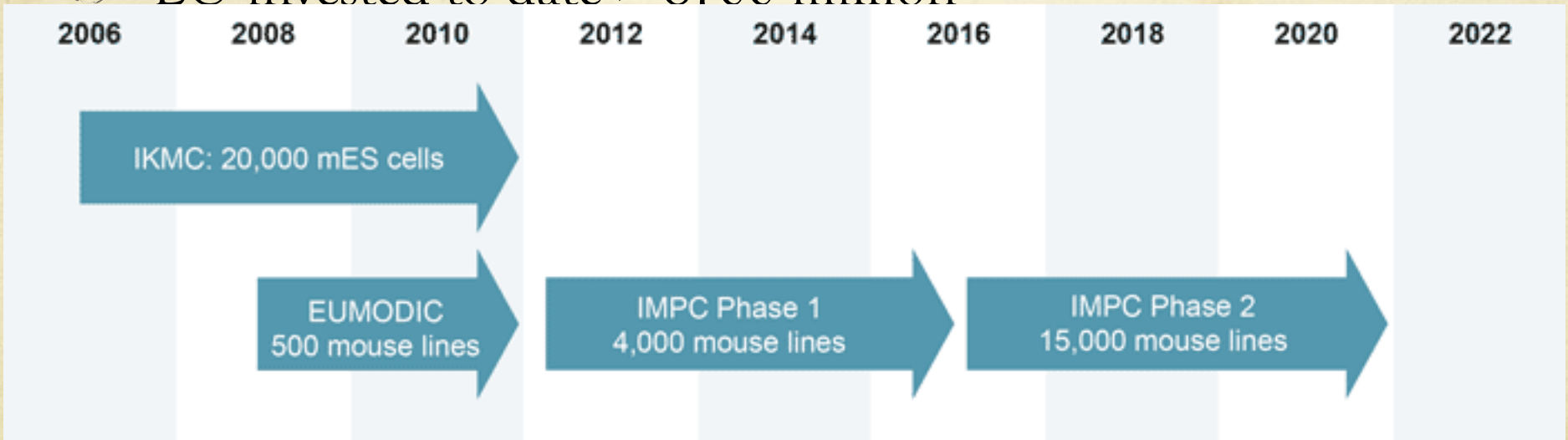
Systematic Genome-Wide Phenotyping

The promise of animal models

- Forward and reverse genetics (e.g. Collaborative Cross panel, IKMC/IMPC)

  → **understanding of gene function by taking a pan-genomic and pan-phenomic approach**

- EU invested to date > €700 million



- IMPC - systematic, agnostic phenotyping of the mouse genome

Gene-Disease associations based on minimal phenotype information

- The nature of a phenotyping pipeline is breadth whilst depth will rely on secondary and tertiary phenotyping carried out by domain experts
- Results to date has not revealed any significant associations
- Challenge – select genes based on primary screens to look at for secondary phenotyping
- Aim - platform for the identification of possible gene disease associations based on minimal phenotype information

Nsun2

Gene Details

Marker Name(s):	NOL1/NOP2/Sun domain family member 2 view this gene in MGI
Marker Type:	protein coding gene
Synonyms:	D13Wsu123e, Misu
Location:	Chr13:69672624-69774658(+)

Related Human Conditions (from OMIM) – ***no related Human Condition***

More Information

Data provided by Mouse Genome Informatics (MGI), Ensembl

Hyperactivity
Glucose homoeostasis
Decreased body fat
Decreased grip strength
Decreased body weight
Increased erythrocytes
Decreased blood lipids
Abnormal skeletal morphology and mineralisation
Cataracts
Abnormal cornea

WTSI Phenotyping

MP Ontology Based Heatmap

Allele Name	Colony Prefix	adipose tissue	behavior/neurological	cardiovascular system	cellular	craniofacial	digestive/alimentary	embryogenesis	endocrine/exocrine gland	growth/size	hearing/vestibular/ear	hematopoietic system	homeostasis/metabolism	immune system	integument	limbs/digits/tail	liver/biliary system	mortality/aging	muscle	nervous system	other	pigmentation	renal/urinary system	reproductive system	respiratory system	skeleton	taste/olfaction	tumorigenesis	vision/eye
Nsun2$^{tm1a(EUCOMM)Wtsi}$	MBKW	■	■	■	■	■			■	■	■	■	■	■	■	■		■		■		■	■	■		■			■

Legend: No Raw Data | No Significant Annotations | Significant Annotation Present | Link to a test report page

PhenomeBrowser

Related genotypes and diseases for MENTAL RETARDATION, AUTOSOMAL RECESSIVE 2

Contents

1. Related OMIM diseases and genes (with similarity value > 0.1)
2. Related mouse genotypes/genes (with similarity value > 0.1)
3. Related yeast genotypes/genes (with similarity value > 0.1)
4. Related worm genotypes/genes (with similarity value > 0.1)
5. Related fly genotypes/genes (with similarity value > 0.1)
6. Related zebrafish genotypes/genes (with similarity value > 0.1)

New query

Related OMIM diseases for MENTAL RETARDATION, AUTOSOMAL RECESSIVE 2

Rank	Name (ID)	Similarity		
1	MENTAL RETARDATION, AUTOSOMAL RECESSIVE 2 (OMIM:607417)	1	Show phenotype	explore
2	HYPERLEXIA (OMIM:238350)	0.423264395686282	Show phenotype	explore
3	SPECIFIC LANGUAGE IMPAIRMENT 2 (OMIM:606712)	0.323633032411548	Show phenotype	explore
4	CEREBROCORTICAL DEGENERATION OF INFANCY (OMIM:213950)	0.283242266968077	Show phenotype	explore
5	MICROCEPHALY WITH SPASTIC QUADRIPLEGIA (OMIM:251280)	0.262401084926546	Show phenotype	explore
6	INDOLYLACROYL GLYCINURIA WITH MENTAL RETARDATION (OMIM:243050)	0.256622755335947	Show phenotype	explore
7	AMINOADIPIC ACIDURIA (OMIM:204750)	0.252725770031264	Show phenotype	explore
8	CYSTEINE PEPTIDURIA (OMIM:219550)	0.246945469322116	Show phenotype	explore
9	PSEUDOURIDINURIA AND MENTAL DEFECT (OMIM:264500)	0.246945469322116	Show phenotype	explore
10	GLUTATHIONURIA (OMIM:231950)	0.246945469322116	Show phenotype	explore
11	MENTAL RETARDATION, X-LINKED 72 (OMIM:300271)	0.244120270515175	Show phenotype	explore

Nsun2 is the pathogenic gene at MRT5 on 5p15

%611091

MENTAL RETARDATION, AUTOSOMAL RECESSIVE 5; MRT5

HGNC Approved Gene Symbol: MRT5

Cytogenetic location: 5p15-p14 *Genomic coordinates (GRCh37):* 5:0 - 28,900,000 (from NCBI)

Gene Phenotype Relationships

Location	Phenotype	Phenotype MIM number
5p15-p14	Mental retardation, autosomal recessive, 5	611091

Phenotypic Series

TEXT

Clinical Features

Najmabadi et al. (2007) reported a large consanguineous Iranian family in which 3 individuals had nonsyndromic moderate to severe mental retardation.

Mapping

By linkage analysis in a large consanguineous Iranian family in which 3 individuals had nonsyndromic moderate to severe mental retardation, Najmabadi et al. (2007) identified a candidate locus on chromosome 5p, termed MRT5, with a maximum lod score of 3.0. Haplotype analysis delineated a 5.6-Mb candidate region between SNPs rs1824938 and rs60701.

By homozygosity mapping of a consanguineous Iranian family in which 2 patients had moderate nonsyndromic mental retardation, Kuss et al. (2011) found linkage to a 6.8-Mb region on chromosome 5p between SNPs rs2008927 and rs60701 (lod score of 3.9).

Abstract/Session Information for Program Number 1100W

Session Information

Session Title: Molecular Basis of Mendelian Disorders **Session Type:** Poster

Session Location: Exhibit Hall, Level 2, Convention Center **Session Time:** Wed 10:00AM-7:00PM

Abstract Information

Program Number: 1100W **Presentation Time:** Wed, Oct 12, 2011, 3:00PM-4:00PM

Keywords: Molecular Basis of Mendelian Disorders, KW011 - BRAIN/NERVOUS SYSTEM, KW106 - MENTAL RETARDATION, KW113 - MODEL ORGANISMS, KW123 - NEUROGENETICS, KW109 - METHYLATION

Abstract Content

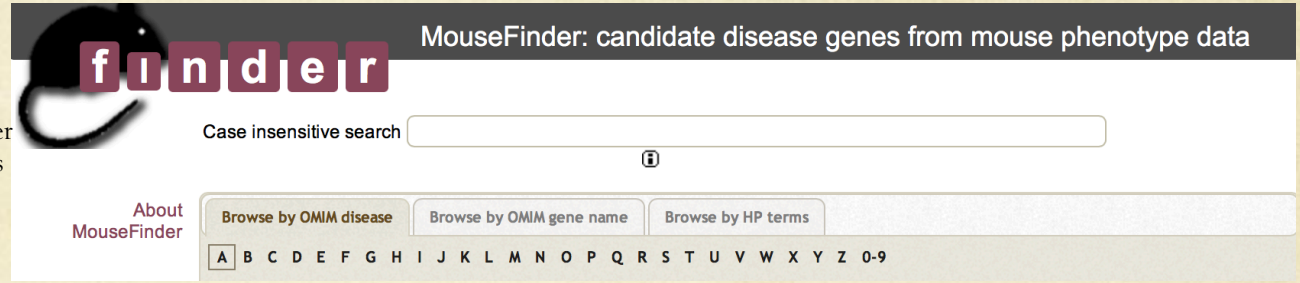
Mutations in the NSUN2 gene cause autosomal recessive intellectual disability in Middle Eastern populations with elevated frequency. *L. Abbasi Moheb[1], S. Mertel[2], L. Nouri Vahid[3], K. Kahrizi[3], A. Tzschach[1], D. Wieczorek[4], M. Garshasbi[1], S. Cirak[5], SS. Abedini[3], H. Najmabadi[3], HH. Ropers[1], S. Sigrist[2], AW. Kuss[1,6]* 1) Human Molec Gen, Max Planck Inst Molec Gen, Berlin, Germany; 2) Genetics, Institute for Biology, Freie Universität Berlin, Berlin, Germany; 3) Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran; 4) Institut für Humangenetik, Universitätsklinikum , Essen, Germany; 5) The Dubowitz Neuromuscular Centre, University College London Institute of Child Health London, UK; 6) Current address: Institute for Human Genetics, Interfaculty Institute for Genetics and Functional Genomics, Ernst Moritz Arndt University, Greifswald, Germany.

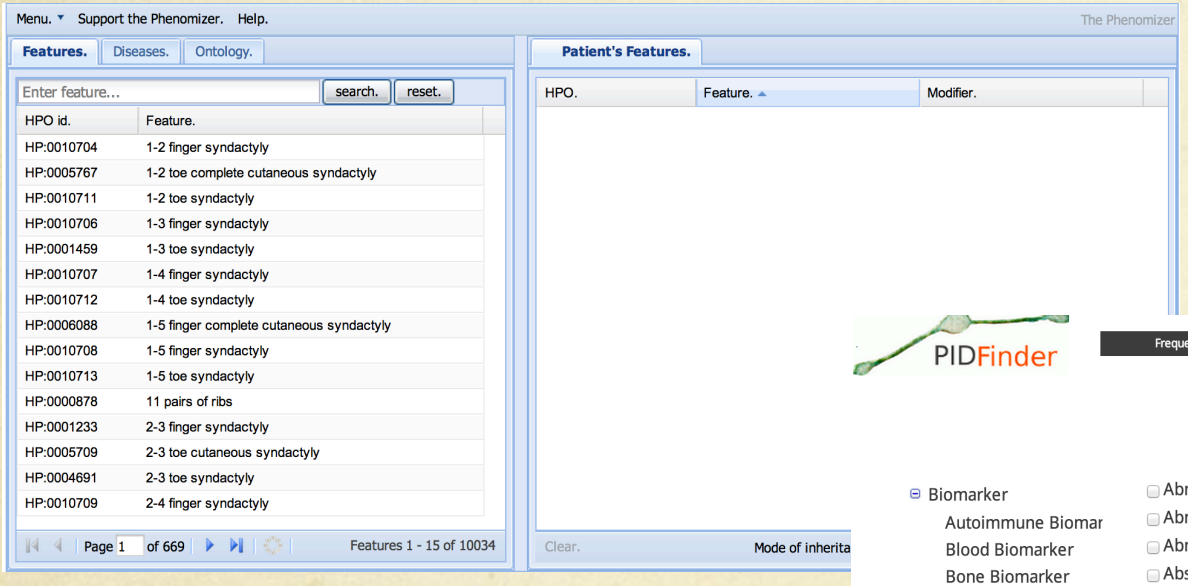
During the course of our investigations into the autosomal recessive causes of intellectual disability (ARID) we have previously identified numerous new loci for this condition.Interestingly, so far no more than six hotspot loci for unspecific or non-syndromic autosomal recessive intellectual disability (NS-ARID) have been identified (Kuss et al. 2011, Human Genetics Vol.129, 141-148). In this study we now resolved the underlying gene defect of MRT 5 and report three deleterious mutations in NSUN2. These were found in two independent consanguineous Iranian families and one Turkish family with several patients suffering from non-syndromic ARID. NSUN2 encodes a methyltransferase, which catalyzes the intron-dependent formation of 5-methylcytosine at C34 of tRNA-leu(CAA). Two of the observed changes were nonsense mutations (p.Q227X and p.Q372X), which cause a complete loss of NSUN2 transcripts in the patients. In the third family we found an intronic exchange of an adenosine for a guanine 11 nucleotides upstream of exon 6. This change causes exon 6 to be skipped during splicing and results in the loss of the main transcript. Hence all mutations lead to a loss of NSUN2 protein function in homozygous mutation carriers and thus in all likelihood cause the patient phenotype. In order to gain further evidence for an involvement of NSUN2 in cognitive functions, we studied fruit fly mutants that lack the NSUN2 ortholog. These experiments revealed a marked learning impairment in mutant flies, which clearly underscores the relevance of NSUN2 in higher brain functions.

- 3 patients in consanguineous Iranian and Turkish families
- Non-syndromic ARID
- Complete loss of Nsun2 transcripts
  - Two nonsense mutations
  - Intronic exchange of A for G 11 nt upstream of exon 6. Cause exon 6 skipping and loss of main transcript.

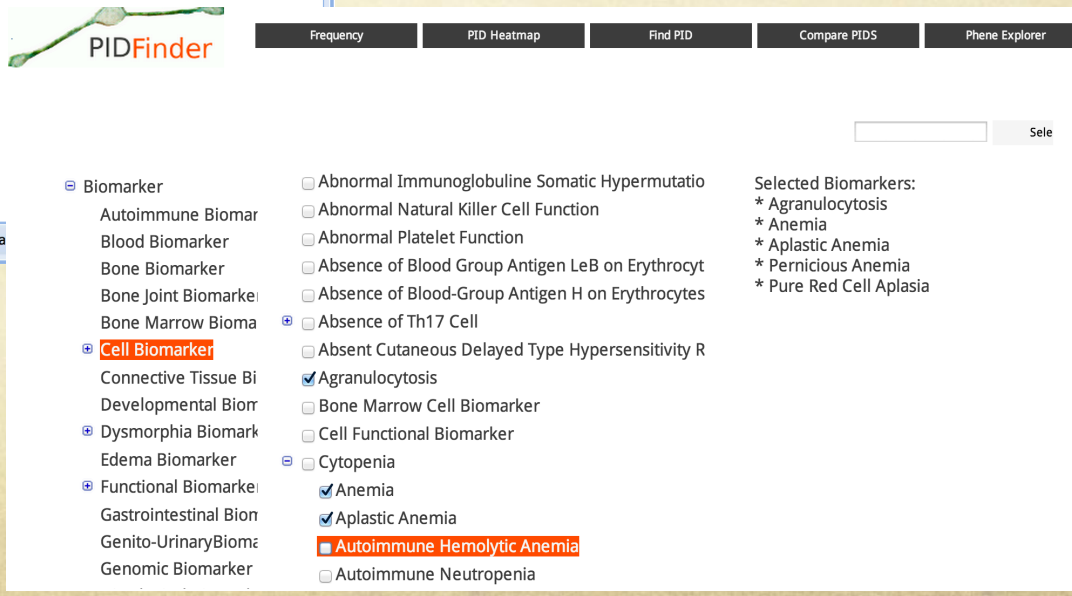
Diagnostics applications

Hum Mutat. 2012 May;33(5):858-66.
MouseFinder: Candidate disease genes from mouse phenotype data.
Chen CK, Mungall CJ, Gkoutos GV, Doelken SC, Köhler S, Ruef BJ, Smith C, Westerfield M, Robinson PN, Lewis SE, Schofield PN, Smedley D.





Nucleic Acids Res. 2013 Nov 11.
The Human Phenotype Ontology project: linking Molecular biology and disease through phenotype data.
Köhler S, *et al*

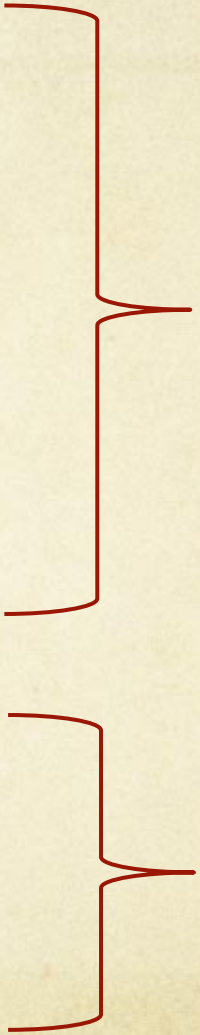


Bioinformatics. 2011 Nov 15;27(22):3193-9.
PIDO: the primary immunodeficiency disease ontology.
Adams N, Hoehndorf R, Gkoutos GV, Hansen G, Hennig C.

From Genotype to Phenotype

From Genotype to Phenotype

- similar function
  - GO-based semantic similarity
- similar cellular location
  - GO-based semantic similarity
- similar protein interactions/pathways
  - distance in PPI network
- similar (gene, protein) sequence
  - Smith–Waterman distance
- expression in similar organ systems
  - semantic similarity over anatomy ontology
- similar phenotypes
  - semantic similarity over phenotype ontology



Existing measures

Additional measures

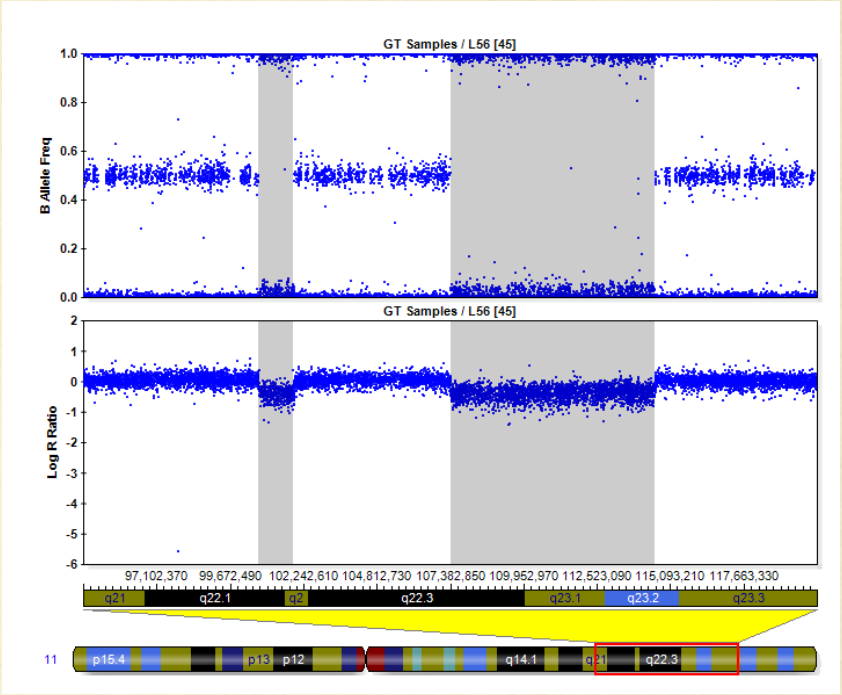
Predict the phenotypes resulting from single gene knockouts in mice

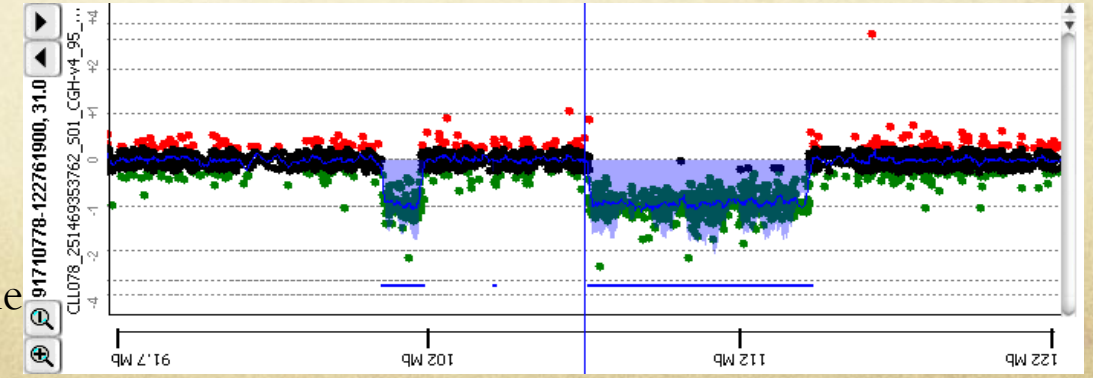
- training data (genotype)
  - 15,000 mouse models with single gene mutations
- SVM to combine similarity scores
- work in progress: 0.70 AUC for predicting gene–disease associations using PhenomeNET

Candidate Disease Gene Prioritisation

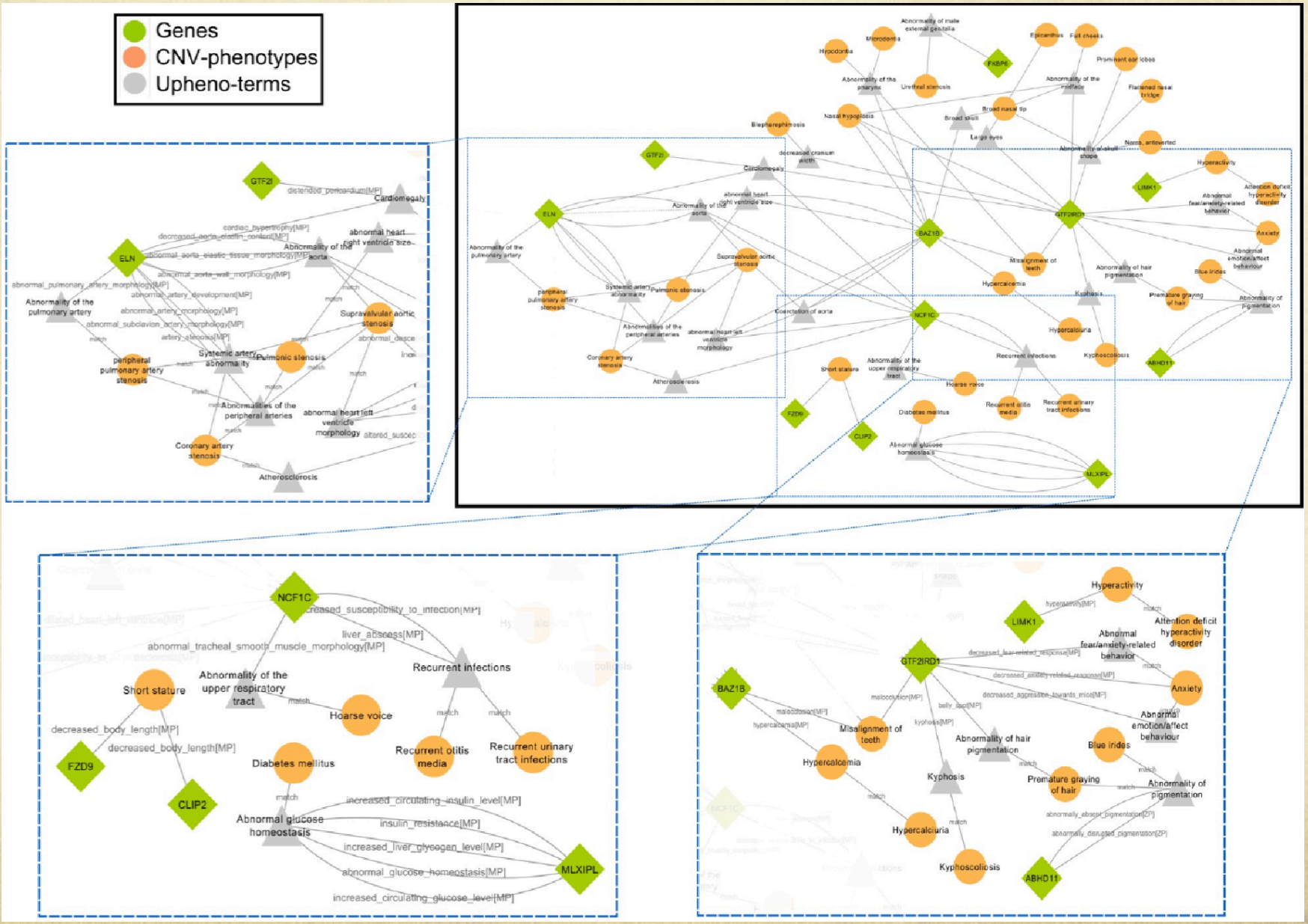
Copy Number Variation (CNV)

- Common in human genomes
  - 35% of the human genome demonstrates evidence of coverage by CNVs
  - 300,000 SNPs from each of three European population isolates, spanning from Northern to Southern Europe,
  - detected 4016 CNVs in 1964 individuals, clustering into 743 CNVRs.
- Associated with common traits via SNP analysis
- Associated with recurrent syndrome



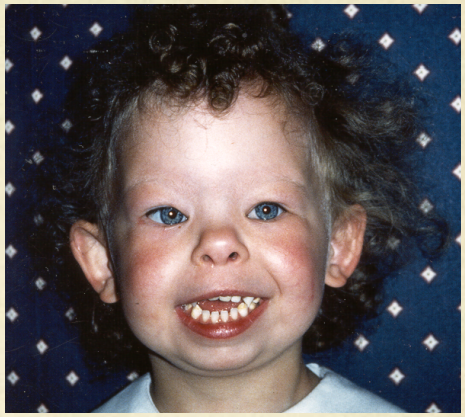


Semantic Modelling of 27 CNVs



Genotype-phenotype correlations in pathogenic CNVs

- 879 potential gene candidates - 430 of which were not previously reported in the literature



- evaluated against two patients suffering from *Williams-Beuren Syndrome (WBS)* with rare microdeletions seen in clinic
  - analysis generated a profile of 32 phenotypic abnormalities connected to 11 candidate genes through 39 associations.

- basis for understanding previously uninterpretable genotype-phenotype correlations in pathogenic CNVs

Dis Model Mech. 2013 Mar;6(2):358-72.
Phenotypic overlap in the contribution of individual genes to CNV pathogenicity revealed by cross-species computational analysis of single-gene mutations in humans, mice and zebrafish.
Doelken SC, Köhler S, Mungall CJ, Gkoutos GV, Ruef BJ, Smith C, Smedley D, Bauer S, Klopocki E, Schofield PN, Westerfield M, Robinson PN, Lewis SE.

Rare and Orphan Diseases

Rare and orphan diseases

Number of Entries:

Prefix	Autosomal	X Linked	Y Linked	Mitochondrial	Totals
* Gene description	12,750	627	48	35	13,460
+ Gene and phenotype, combined	250	14	0	2	266
# Phenotype description, molecular basis known	2,836	240	4	28	3,108
% Phenotype description or locus, molecular basis unknown	1,628	135	5	0	1,768
Other, mainly phenotypes with suspected mendelian basis	1,819	130	2	0	1,951
Totals	19,283	1,146	59	65	20,553

- at least 4000 diseases without known molecular basis
- disease-gene identification methods have a limited focal range
- necessary to suggest possible causative genes

Bassoe Syndrome

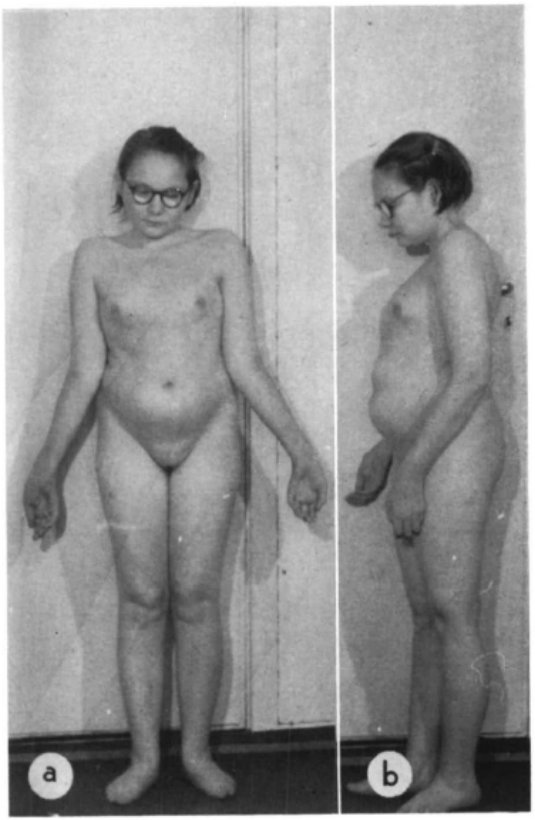


FIG. 2. *Case 1.* **a.** Showing cubitus valgus, elongated extremities and lack of pubic hair. **b.** Showing protruding abdomen, thoracic kyphosis and underdeveloped breasts.

FAMILIAL CONGENITAL MUSCULAR DYSTROPHY WITH GONADAL DYSGENESIS

HANS H. BASSÖE, M.D.*

Hammerfest Hospital, Hammerfest, Norway
(Medical Section—Head, H. Schartum-Hansen)

IN A family living in a small isolated village in Finnmark county, Norway, we have observed 7 persons suffering from congenital muscular dystrophy. Their symptoms were similar to those seen in congenital amyotonia (Oppenheim). Several children in the family had died very early in life; 3 others had been still born. In the third generation (III, Fig. 1) the child mortality was 33.3 per cent, whereas the average child mortality in Finnmark in the period 1926–1930 was 9.17 per cent. In the same generation, 2 siblings (Cases 1 and 2) were affected. In addition to the muscular dystrophy there was gonadal dysgenesis—ovarian agenesis and testicular

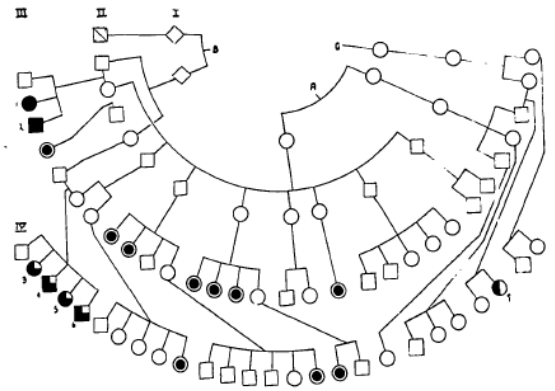


FIG. 1. *Family history.* Upright white squares and circles indicate normal men and women; slanted white squares indicate unknown sex. Black squares and circles indicate men with Klinefelter's syndrome and women with ovarian agenesis, both with cataract and muscular dystrophy. Black squares and circles with $\frac{1}{4}$ white area indicate men and women with muscular dystrophy, but without endocrine disorder. Black circle with $\frac{1}{2}$ white area indicates woman with muscular dystrophy and epicanthus. Black double circles indicate stillbirths.

Received for publication February 9, 1956.

* Present address: Medical Department B, University Clinic of Bergen, Bergen, Norway.

Bassoe Syndrome

organ system	Orphanet	mouse models (MGI)	additional mouse phenotypes reported in literature
skeletal	kyphosis, hypertensible joints, cubitus valgus	abnormal spine curvature, lordosis	kyphosis [24], kypholordosis [25], spinal defects [26]
muscular	amyotrophy, hypotonia, muscle hypotrophy	abnormal muscle morphology	muscle hypotrophy [27], muscle wasting [27]
behavioural	abnormal gait, amimia	abnormal gait, hypoactivity, tremors	failure to thrive [25], ataxia [24], defects in presynaptic function [27]
visual	cataract, strabismus	nuclear cataracts, microphthalmia	cataracts [26]
reproductive	testicular atrophy, hypogonadism, hypogenitalism, abnormal ovaries, reduced fertility	testicular atrophy, male infertility	decreased testicular weight [28], testicular degeneration [26,28], increased apoptosis of postmeiotic spermatids [28], oligospermia [28], decreased fertility [26,29], reduced sperm count and motility [26,29], ovarian abnormalities [29]

HIP1 candidate gene for Bassoe syndrome

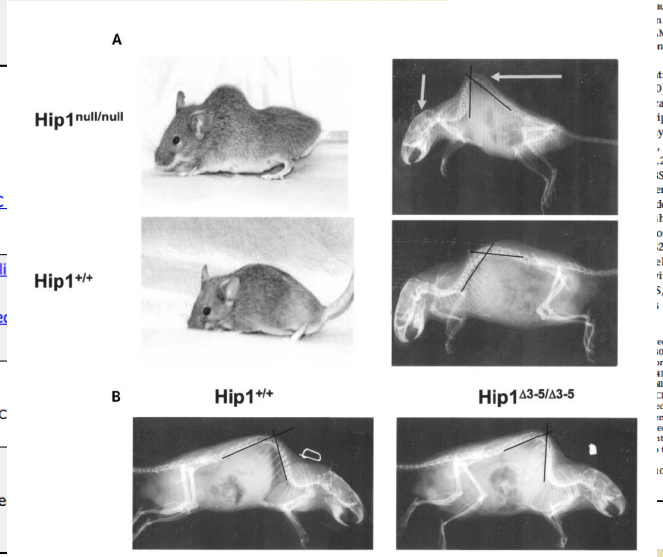
MGI

About Help FAQ

Search ▾ Download ▾ More Resources ▾ Submit Data Find Mice (IMSR) Analysis Tools

Hip1
Gene Detail

Symbol	**Hip1**
Name	**huntingtin interacting protein 1**
ID	MGI:1099804
Synonyms	2610109B09Rik, A930014B11Rik, E130315I21Rik, HIP-1, MGC:27616
Feature Type	protein coding gene
Genetic Map	Chromosome 5 75.18 cM, cytoband F-G2 Detailed Genetic Map ± 1 cM Mapping data(13)
Sequence Map	Chr5:135406523-135545122 bp, - strand From Ensembl annotation of GRCm38 Get FASTA 138600 bp ± 0 kb flank VEGA Genome Browser | Ensembl Genome Browser | UCSC
Mammalian homology	human; rat; cattle; chimpanzee; dog, domestic (Mammali Comparative Map (Mouse/Human Hip1 ± 2 cM) Protein SuperFamily: huntingtin-interacting protein 1-relate Gene Tree: Hip1
NEW Human ortholog	HIP1 huntingtin interacting protein 1 NCBI Gene ID 3092 Human Synonyms: ILWEQ Human Chr7:75163409-75368279 bp, - strand Reference GRC Human Diseases Associated with Human HIP1 (1)
Alleles and phenotypes	All alleles(32) : Targeted(8) Gene trapped(24) Homozygous mutants may exhibit axial skeleton defe whereas others did not.
Gene Ontology (GO) classifications	All GO classifications: (29 annotations) Process activation of cysteine-type endopeptidase activity involved in apoptotic proc Component clathrin-coated vesicle, cytoplasm, ... Function actin binding, clathrin binding, ... External Resources: FuncBase
Expression	Literature Summary: (3 records)



REPORT

Recurrent Distal 7q11.23 Deletion Including *HIP1* and *YWHAG* Identified in Patients with Intellectual Disabilities, Epilepsy, and Neurobehavioral Problems

Melissa B. Ramocki,[1,2,16] Magdalena Bartnik,[3,16] Przemyslaw Szafranski,[4,16] Katarzyna E. Kołodziejska,[4] Zhilian Xia,[4] Jaclyn Bravo,[4] G. Steve Miller,[5,6] Diana L. Rodriguez,[1,2] Charles A. Williams,[7] Patricia I. Bader,[8] Elżbieta Szczepanik,[9] Tomasz Mazurczak,[9] Dorota Antczak-Marach,[9] James G. Coldwell,[5] Cigdem I. Akman,[1,2] Karen McAlmon,[10,11,12] Melinda P. Cohen,[13] James McGrath,[14] Elizabeth Roeder,[15] Jennifer Mueller,[7] Sung-Hae L. Kang,[4] Carlos A. Bacino,[4] Ankita Patel,[4] Ewa Bocian,[3] Chad A. Shaw,[4] Sau Wai Cheung,[4] Tadeusz Mazurczak,[3] and Paweł Stankiewicz[3,4,*]

We report 26 individuals from ten unrelated families who exhibit variable expression and/or incomplete penetrance of epilepsy, learning difficulties, intellectual disabilities, and/or neurobehavioral abnormalities as a result of a heterozygous microdeletion distally adjacent to the Williams-Beuren syndrome region on chromosome 7q11.23. In six families with a common recurrent ~1.2 Mb deletion that includes the Huntingtin-interacting protein 1 (*HIP1*) and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma (*YWHAG*) genes and that is flanked by large complex low-copy repeats, we identified sites for nonallelic homologous recombination in two patients. There were no cases of this ~1.2 Mb distal 7q11.23 deletion copy number variant identified in over 20,000 control samples surveyed. Three individuals with smaller, nonrecurrent deletions (~180–500 kb) that include *HIP1* but not *YWHAG* suggest that deletion of *HIP1* is sufficient to cause neurological disease. Mice with targeted mutation in the *Hip1* gene (*Hip1*$^{-/-}$) develop a neurological phenotype characterized by failure to thrive, tremor, and gait ataxia. Overall, our data characterize a neurodevelopmental and

The American Journal of Human Genetics *87*, 857–865, December 10, 2010 857

Interface Focus. 2013 Apr 6;3(2):20120055.
An integrative, translational approach to understanding ra and orphan genetically based diseases.
Hoehndorf R, Schofield PN, Gkoutos GV.

orphaNews Europe
Newsletter of the European Union Committee of Experts on Rare Diseases
Summary
1 October 2013
print

PhenomeNet: a phenotype-based system to identify known candidate genes for diseases now integrates clinical signs from Orphanet

"integration and computational analysis of human disease and animal model phenotypes using PhenomeNet has the potential to reveal novel insights into the pathobiology underlying genetic diseases"

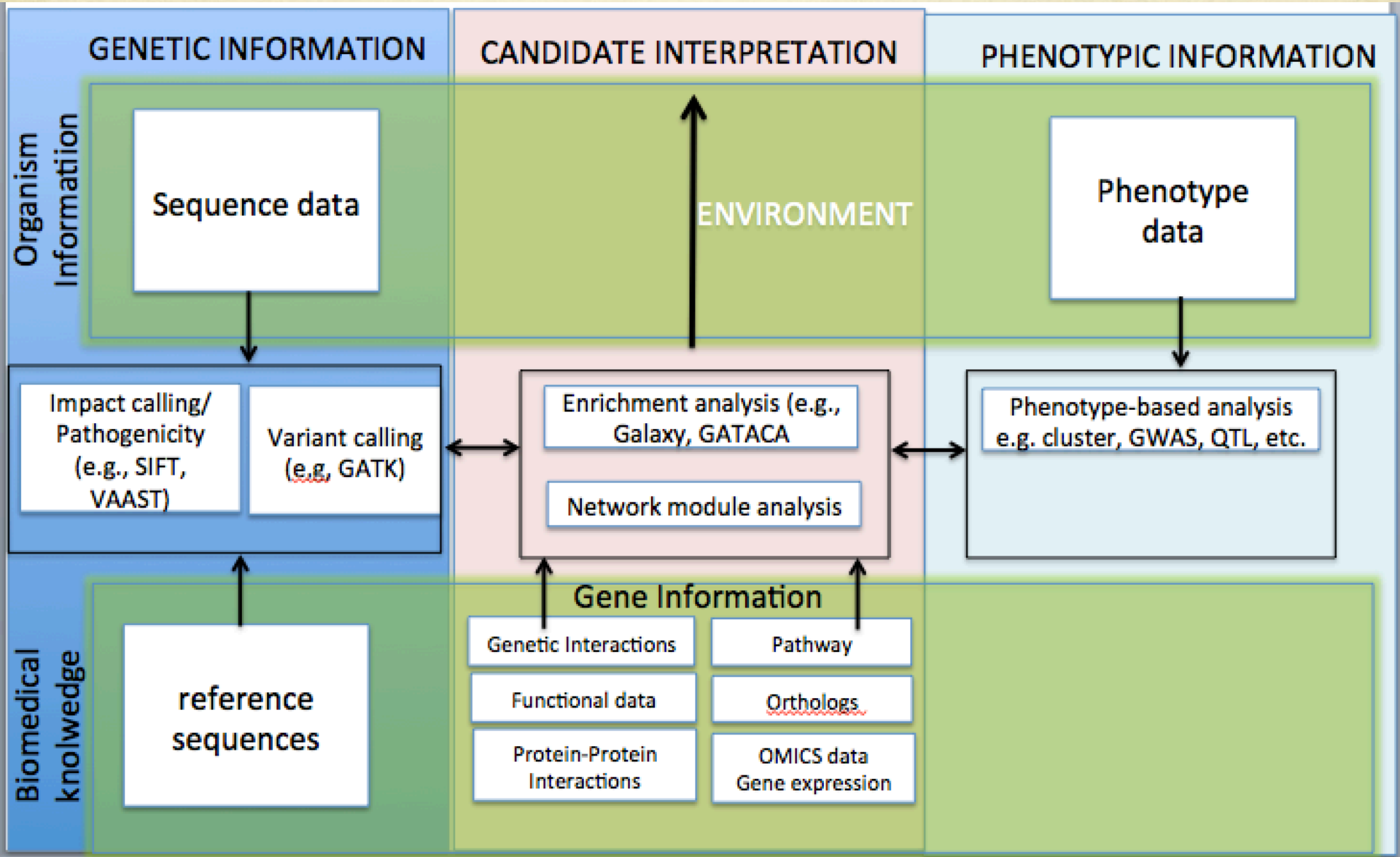
The Desminopathy Reporter

Making sense of missense, nonsense, and other vexsome gene mutations

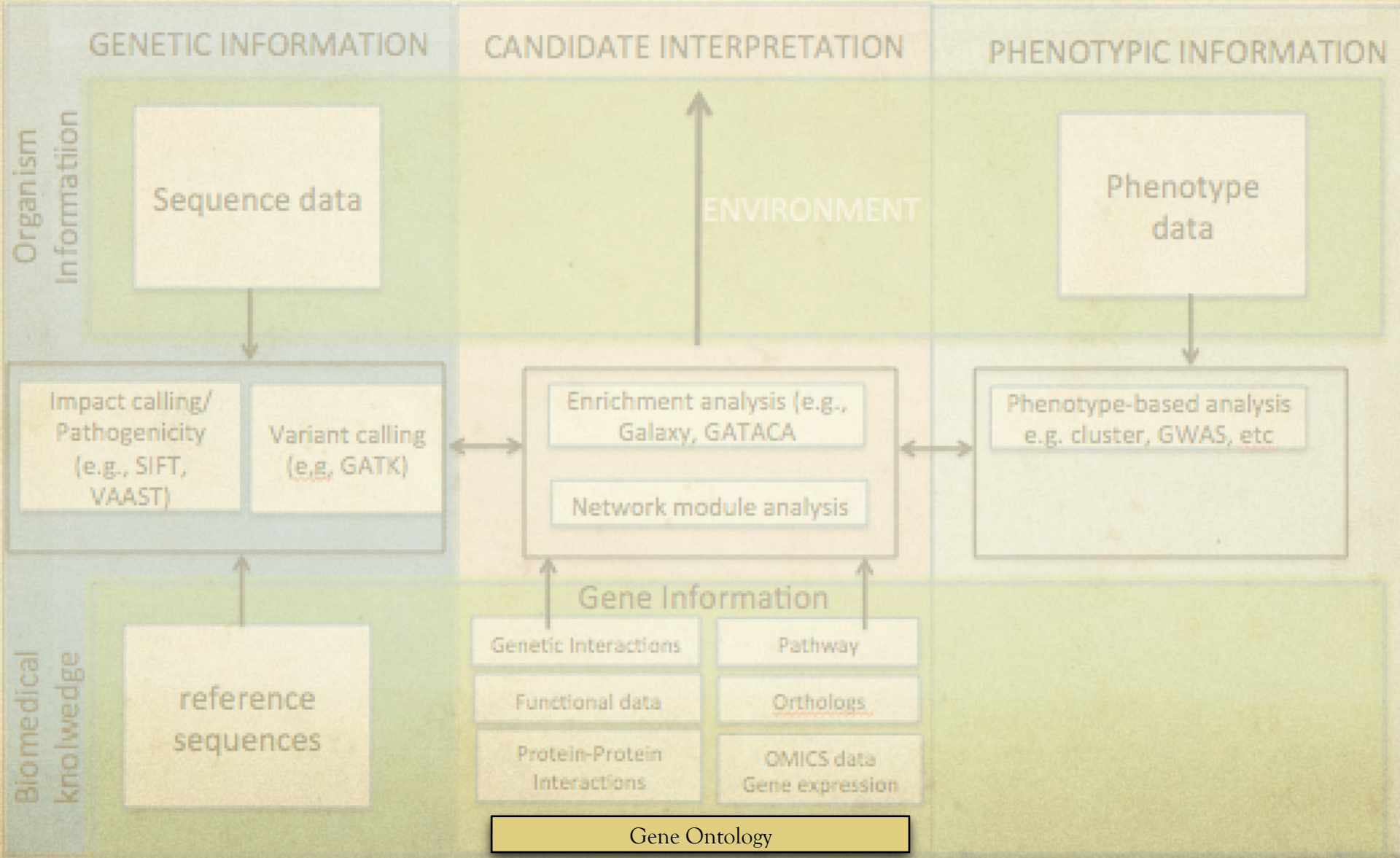
Research In Action

PhenomeNet is an integrated, searchable, cross-species phenotype network. It's also a wowsa! for "identifying candidate genes for genetic diseases [in humans] based on the similarity between a disease and animal model phenotypes." In an article

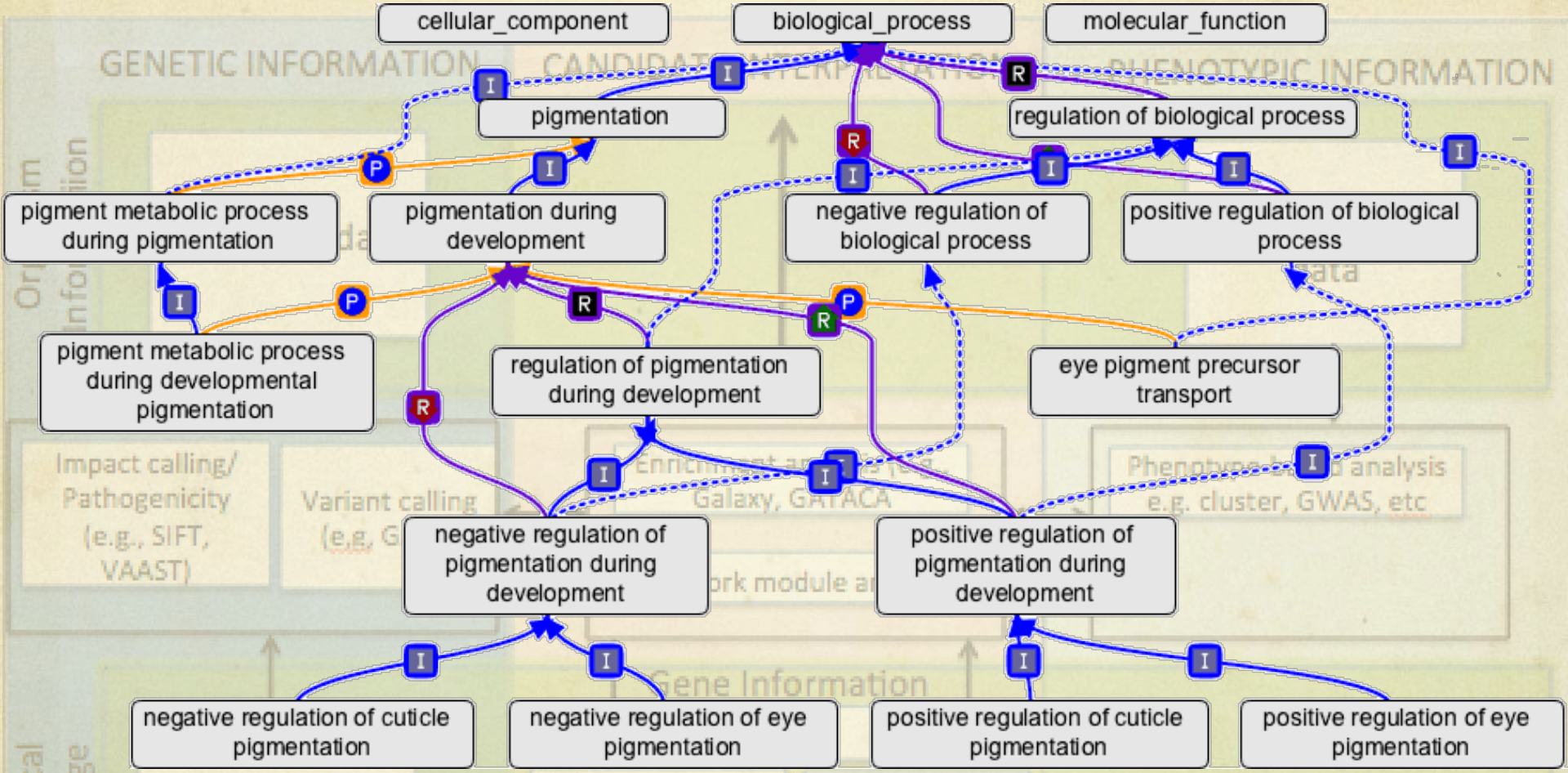
Candidate gene prioritization



Gene Ontology

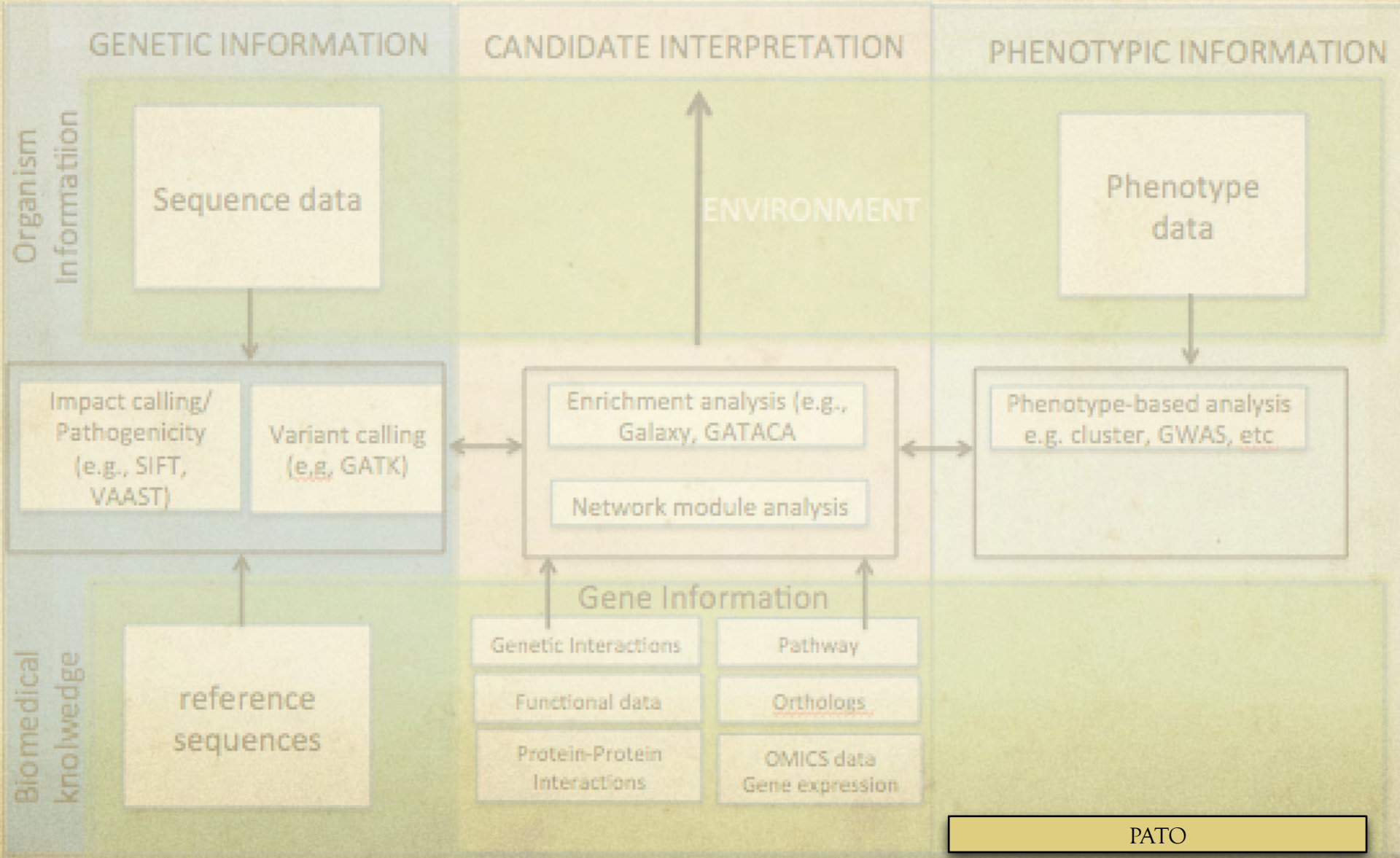


Gene Ontology



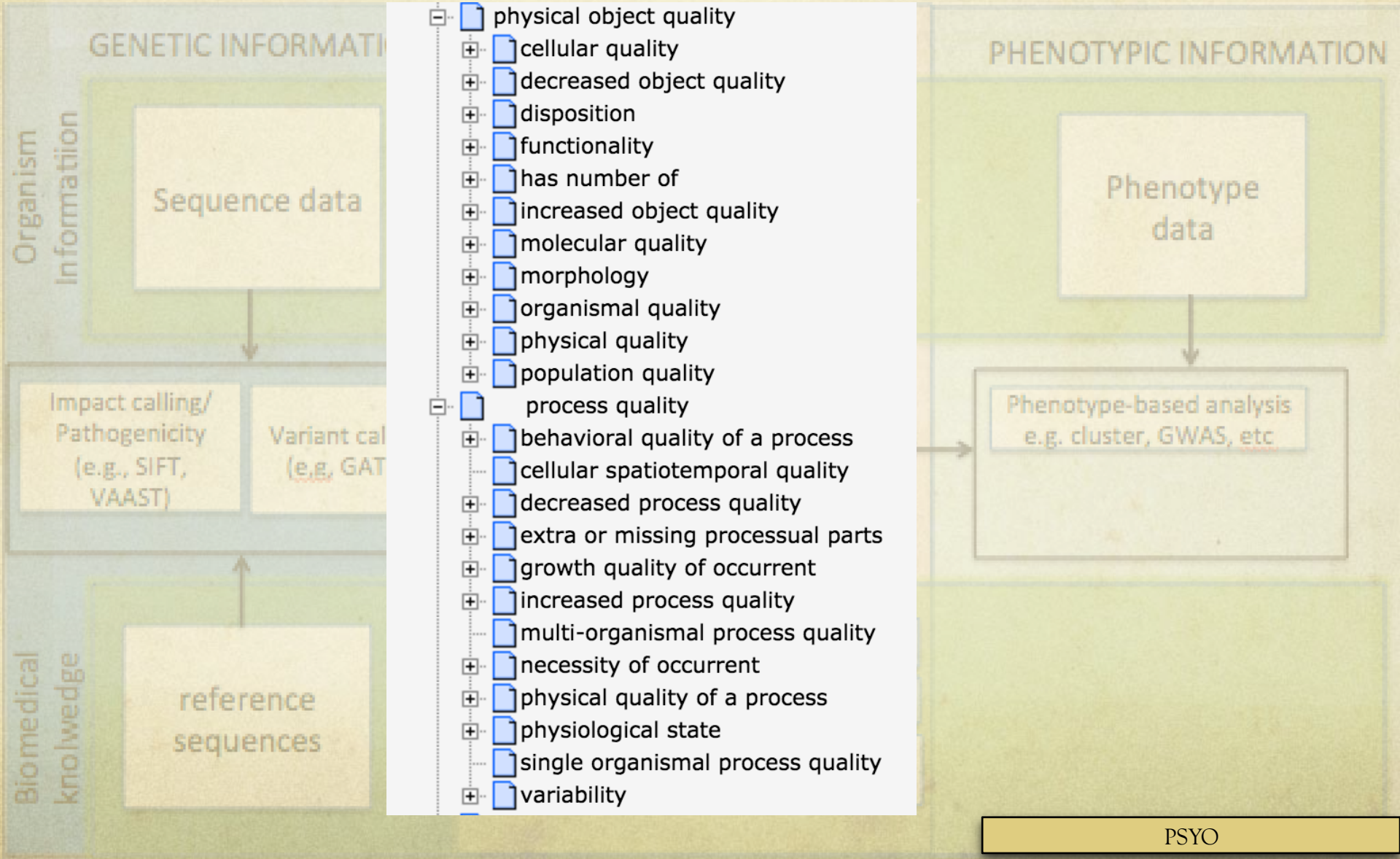
Gene Ontology

PATO


GENETIC INFORMATION
CANDIDATE INTERPRETATION
PHENOTYPIC INFORMATION
Organism Information
Sequence data
ENVIRONMENT
Phenotype data
Impact calling/ Pathogenicity (e.g., SIFT, VAAST)
Variant calling (e,g, GATK)
Enrichment analysis (e.g., Galaxy, GATACA
Network module analysis
Phenotype-based analysis e.g. cluster, GWAS, etc
Gene Information
Biomedical knolwedge
reference sequences
Genetic Interactions
Pathway
Functional data
Orthologs
Protein-Protein Interactions
OMICS data Gene expression

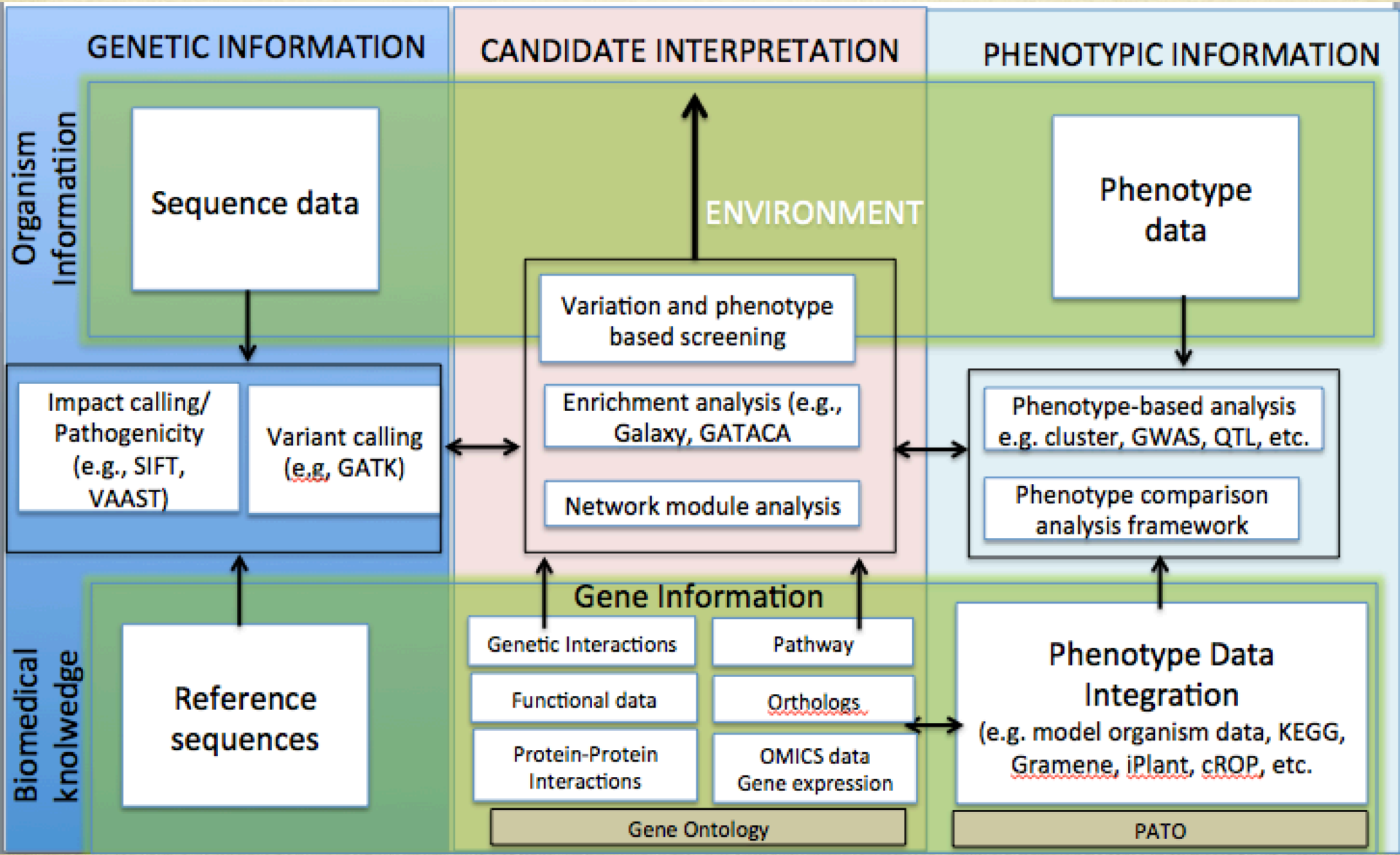
PATO

PATO


GENETIC INFORMATI
Organism Information
Sequence data
Impact calling/ Pathogenicity (e.g., SIFT, VAAST)
Variant cal (e.g, GAT
Biomedical knolwedge
reference sequences
PHENOTYPIC INFORMATION
Phenotype data
Phenotype-based analysis e.g. cluster, GWAS, etc
physical object quality
cellular quality
decreased object quality
disposition
functionality
has number of
increased object quality
molecular quality
morphology
organismal quality
physical quality
population quality
process quality
behavioral quality of a process
cellular spatiotemporal quality
decreased process quality
extra or missing processual parts
growth quality of occurrent
increased process quality
multi-organismal process quality
necessity of occurrent
physical quality of a process
physiological state
single organismal process quality
variability

PSYO

Candidate gene prioritization

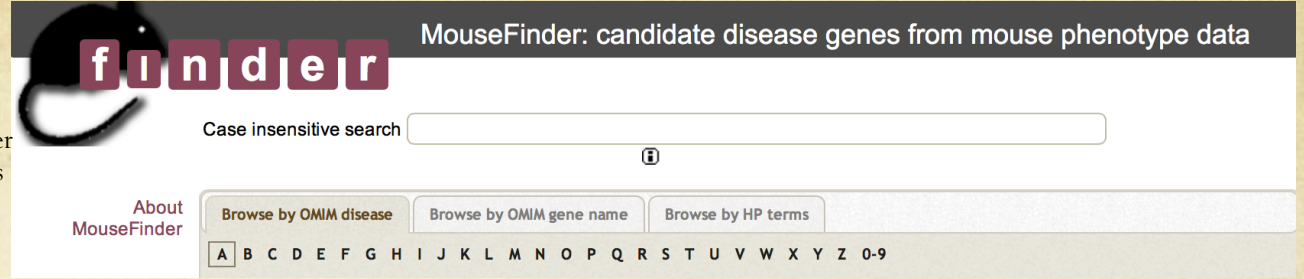


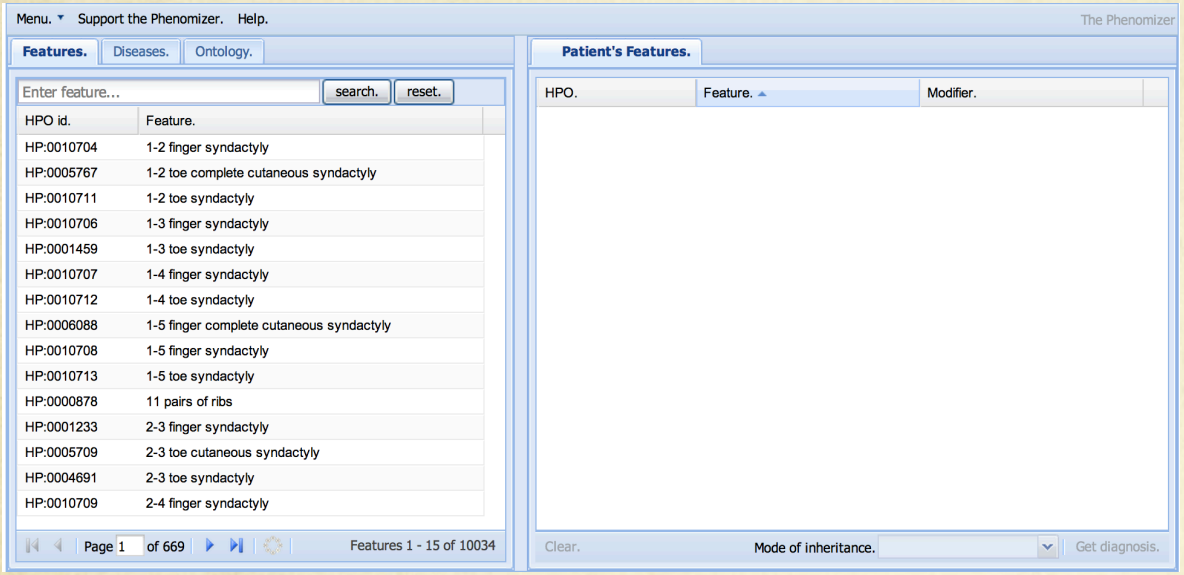
Diagnostics Strategies to support Identification of Causative Genes

Clinical diagnostics decision support systems

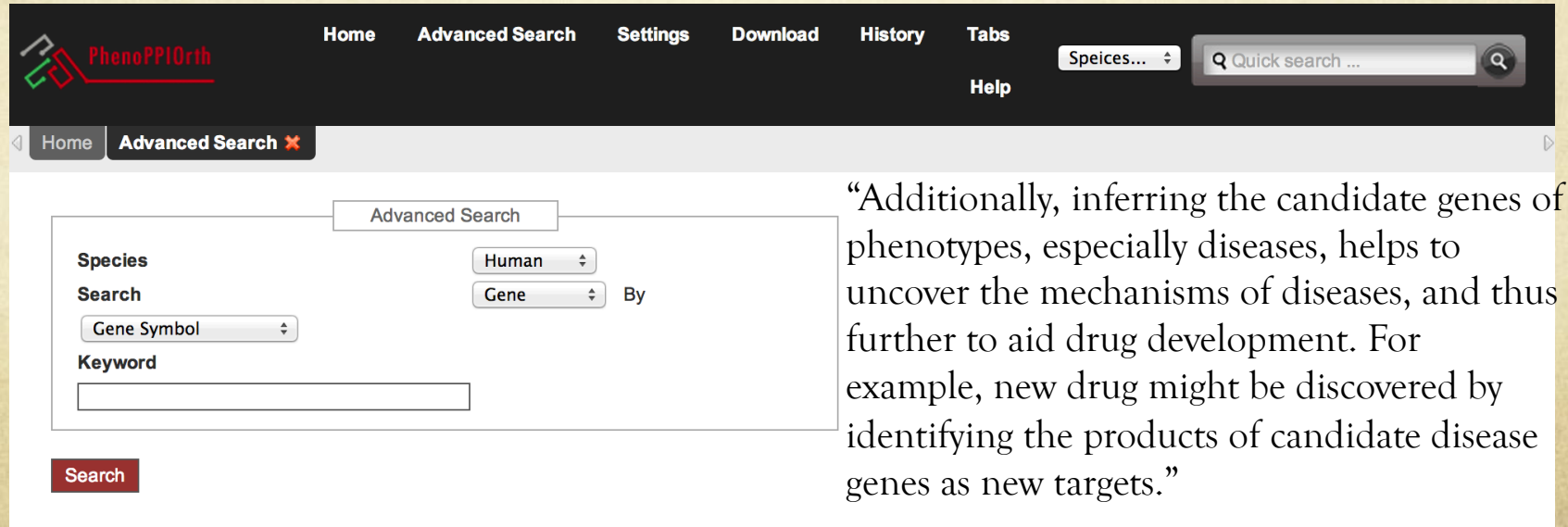
- “show me all genes involved in degenerative processes of the brain or heart for which no evidence of cerebellar degeneration is available in mouse models”
- “show me all genes associated with a particular process that are also associated with mental retardation”
- “prioritize these genes in order with their relevance to a particular set of phenotypes or a particular syndrome”

Hum Mutat. 2012 May;33(5):858-66.
MouseFinder: Candidate disease genes from mouse phenotype data.
Chen CK, Mungall CJ, Gkoutos GV, Doelken SC, Köhler S, Ruef BJ, Smith C, Westerfield M, Robinson PN, Lewis SE, Schofield PN, Smedley D.





Nucleic Acids Res. 2013 Nov 11.
The Human Phenotype Ontology project: linking Molecular biology and disease through phenotype data.
Köhler S, *et al*



"Additionally, inferring the candidate genes of phenotypes, especially diseases, helps to uncover the mechanisms of diseases, and thus further to aid drug development. For example, new drug might be discovered by identifying the products of candidate disease genes as new targets."

Big Data

Can we apply the same approach across biology ?

Can we achieve the same type of phenotype data standardization for plants?

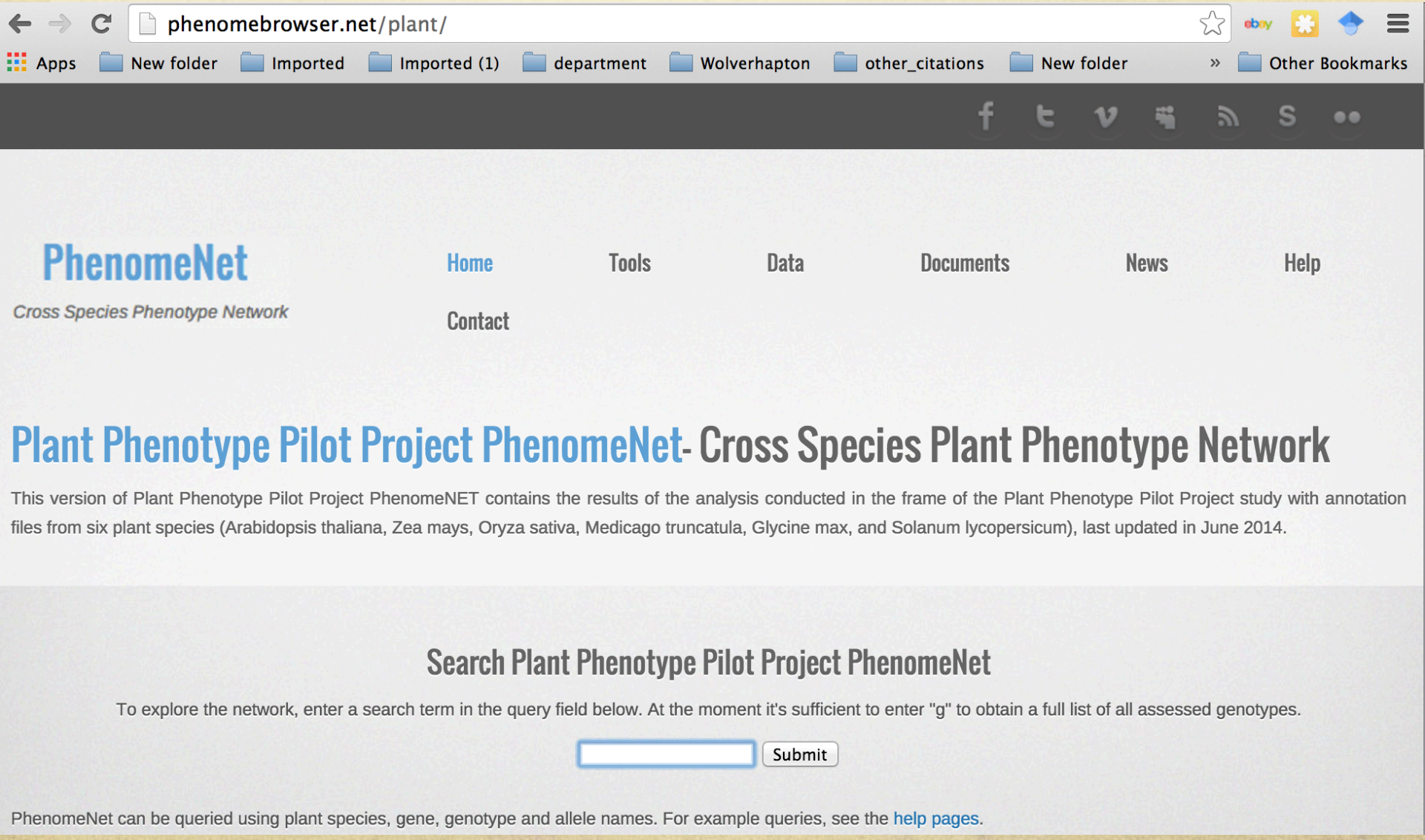
Description of Mutant Phenotype	Atomized Phenotype statements	Entity	Quality (PATO)
Dwarf with profuse slender tillers, small panicles	dwarf	PO: shoot system	decreased height
	profuse tillers	PO: whole plant	has extra parts of type (basal axillary shoot system)
	slender tillers	PO: basal axillary shoot system	slender
	small panicles	PO: inflorescence	decreased size
Delayed flowering; Reduction in total chlorophyll		GO: flowering	delayed
		ChEBI: chlorophyll	decreased concentration

Plant Phenotype Pilot Project (PPPP)

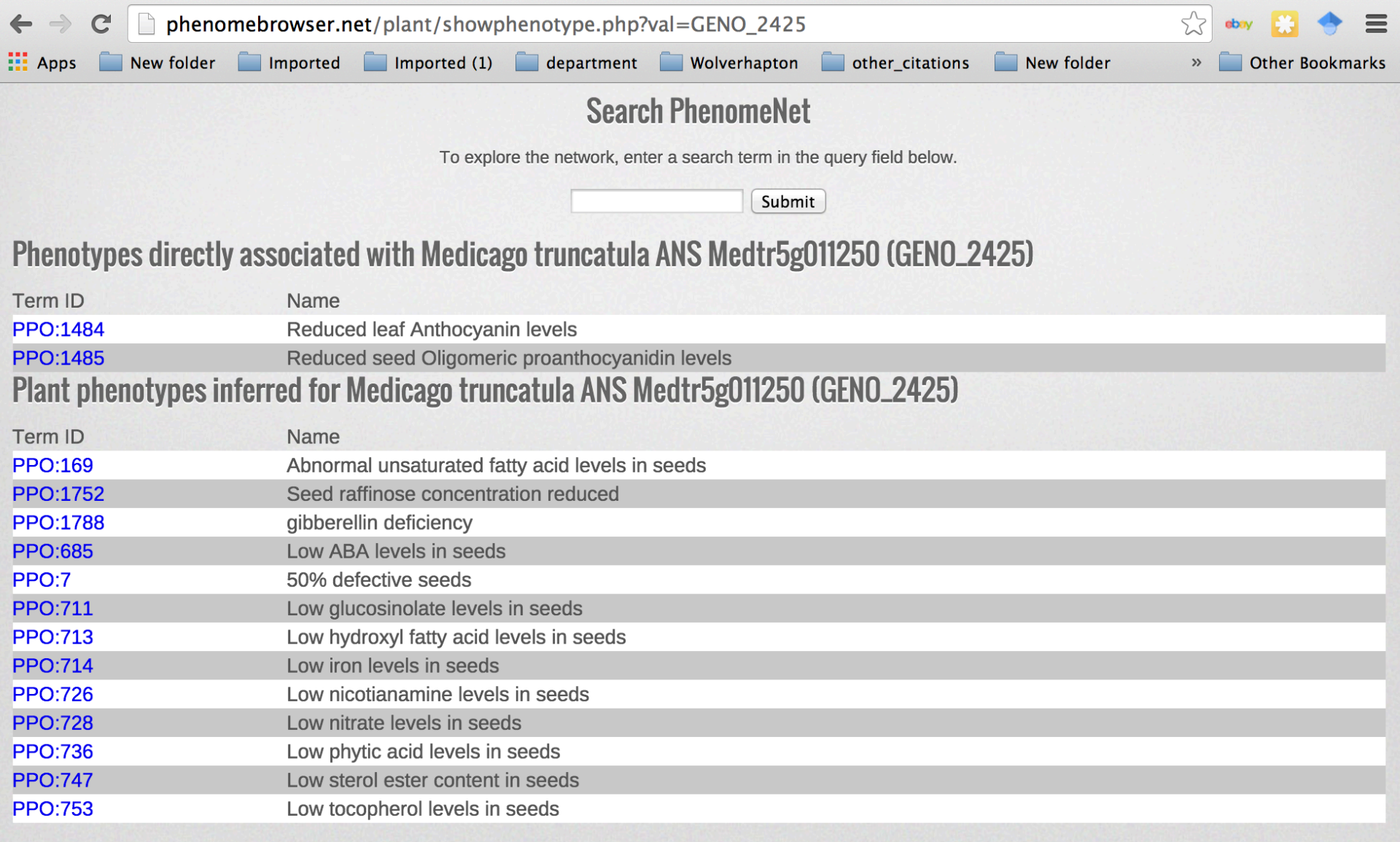
Species	#EQs (phenes)	#unique EQs across all genotypes	#genes	#genotypes	#phenotypes
Arabidopsis thaliana	5172	1260	2393	2393*	1385
Zea mays ssp *mays*	373	180	114	169	117
Oryza sativa L.	340	271	92	95	86
Solanum lycopersicum	269	174	72	128	90
Medicago truncatula	149	99	40	45	40
Glycine max	61	39	30	30*	24
Total	6364	2023	2741	2866	1742

* # genotypes equals # genes because no distinction between them was made for these species

Plant PhenomeNet



Medtr5g011250- leucoanthocyanidin dioxygenase-like protein



Search PhenomeNet

To explore the network, enter a search term in the query field below.

Submit

Related genotypes and diseases for Medicago truncatula ANS Medtr5g011250

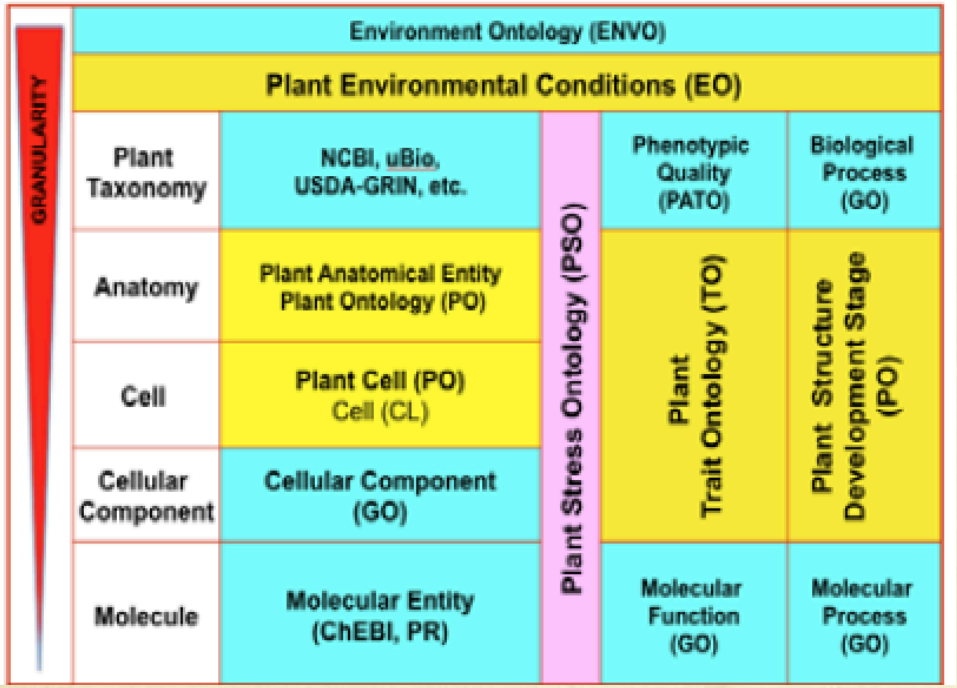
Contents

New query

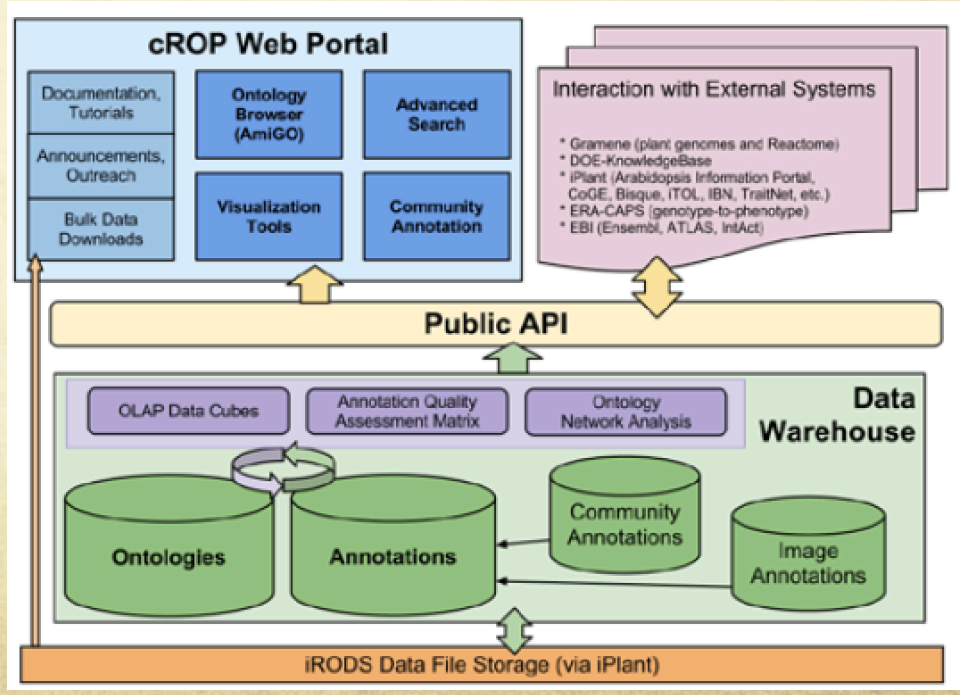
Related Plant phenotypes for Medicago truncatula ANS Medtr5g011250

Rank	Name (ID)	Similarity		
1	Medicago truncatula ANS Medtr5g011250 (GENO_2425)	1	Show	explore
2	Arabidopsis thaliana At5g60760 (GENO_2393)	0.9333333333	Show	explore
3	Arabidopsis thaliana AtNRT2.7 At5g14570 (GENO_319)	0.9333333333	Show	explore
4	Arabidopsis thaliana MYB28 At5g61420 (GENO_1461)	0.9333333333	Show	explore
5	Arabidopsis thaliana AtCBR At5g17770 (GENO_245)	0.9333333333	Show	explore
6	Arabidopsis thaliana VTE5 At5g04490 (GENO_2319)	0.9333333333	Show	explore
7	Glycine max RS1-2 Glyma06g18890.1 (GENO_2420)	0.875	Show	explore
8	Glycine max RS2 Glyma06g18890.1 (GENO_2421)	0.875	Show	explore
9	Arabidopsis thaliana YSL1 At4g24120 (GENO_2360)	0.8235294118	Show	explore
10	Solanum lycopersicum gamybl1 Solyc01g009070 gib-1 (GENO_2641)	0.8235294118	Show	explore
11	Arabidopsis thaliana NCED6 At3g24220 (GENO_1484)	0.7368421053	Show	explore
12	Arabidopsis thaliana PSAT1 At1g04010 (GENO_1787)	0.652173913	Show	explore
13	Arabidopsis thaliana ROD1 At3g15820 (GENO_1884)	0.1333333333	Show	explore

cROP - Common Reference Ontologies and Applications for Plant Biology



cROP-related ontologies



cROP portal

BIG DATA:
ontology-based heteregeneous, multi-modal data access

Aber-OWL:

- biological data
  - SPARQL endpoints
- literature
  - MEDLINE/PubMed
  - PubMed Central

[Tetralogy of Fallot with pulmonary atresia and major aortopulmonary collateral arteries].

Patients with **tetralogy** of **Fallot** (TOF), **pulmonary atresia** (PA), and major aortopulmonary collateral artery (MAPCA) usually need some staged surgical procedures. There is no clear consensus to the initial procedure and also the most proper initial procedure is different in each case. We report here a case of 6-year-old girl with TOF, PA, and MAPCA. We performed right ventricular outflow tract reconstruction (RVOTR) for the initial procedure because her pulmonary artery was extremely diminutive. Her pulmonary artery got good growth after the palliative surgery. Five years later she underwent complete repair after two other surgeries and 1 interventional embolization. RVOTR is a useful procedure for pulmonary artery growth, but it should be well considered that what size of RVOTR is needed in each case.

Imaging findings in uncorrected tetralogy of Fallot and pulmonary atresia with major aortopulmonary collateral arteries and septic embolism

Full text available.

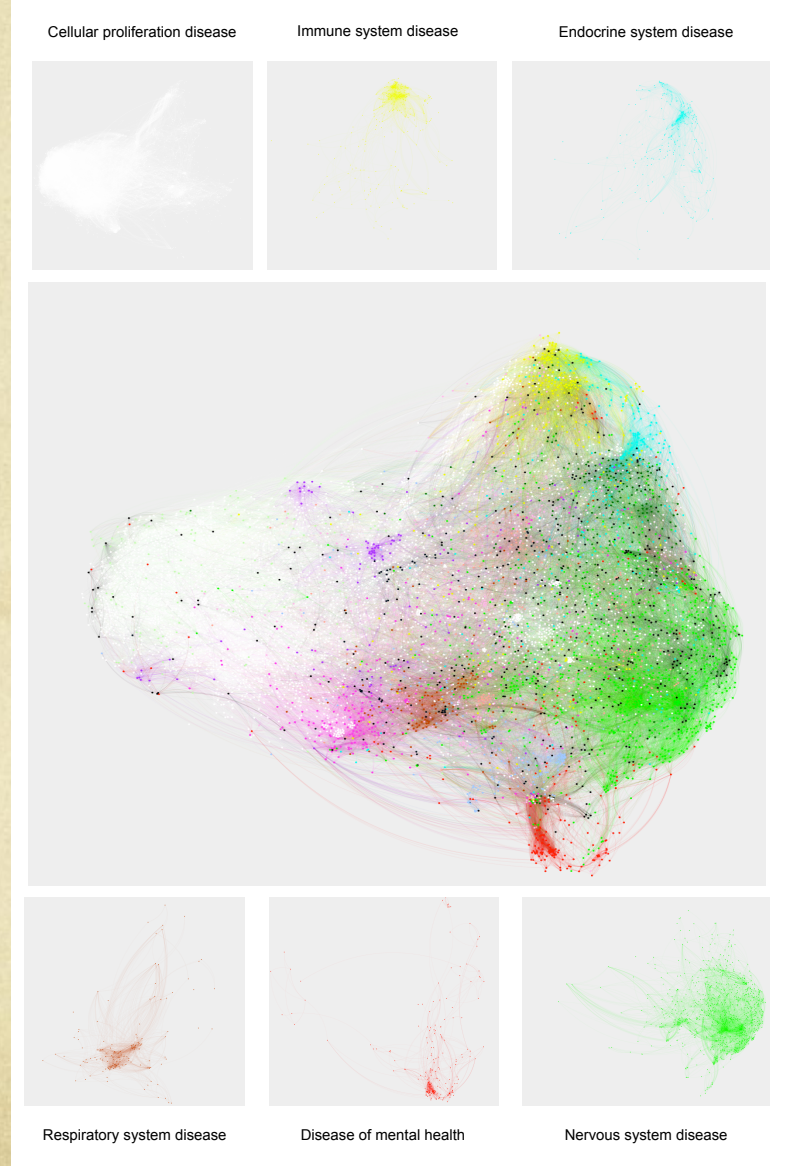
Tetralogy of **Fallot** (TOF) is one of the most common congenital heart malformations comprising a ventricular septal defect, right ventricular outflow tract obstruction, right ventricular hypertrophy, and overriding aorta. A rare variant includes pulmonary atresia and major aortopulmonary collateral arteries. Altered hemodynamics within the functional single-ventricle results in turbulent flow and predisposes to endocardial vegetation formation which may consequently lead to thromboembolic events. We present a rare case of an adult survivor of uncorrected TOF with pulmonary atresia.

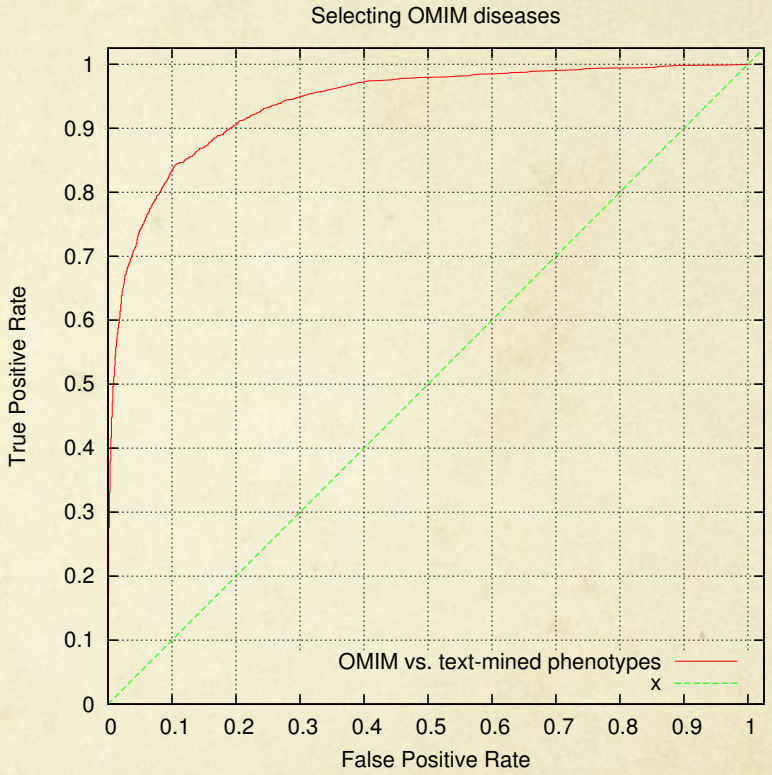
Early and late results of repair of tetralogy of Fallot with subarterial ventricular septal defect. A comparative evaluation of tetralogy with perimembranous ventricular septal defect.

Between November 1966 and December 1990, 511 pediatric patients with **tetralogy** of **Fallot** underwent corrective operation at Tenri Hospital. There were 78 patients with **subarterial ventricular septal defect**. Mean age at repair was 5.6 +/- 3.3 years. The method of right ventricular outflow tract reconstruction was simple infundibulectomy in 14 patients, right ventricular ventricular outflow patch in 36, and transannular patch in 28. There were 7 (9.0%) early deaths as a result of low cardiac output syndrome and acute renal failure. The pressure ratio of the right ventricle to the left ventricle was 0.62 +/- 0.18 during the early postoperative catheterization. Follow-up was achieved for 442.6 patient-years and ranged from 0.5 to 27 years, with an average of 8.5 +/- 6.7 years. There were three late deaths (2 cardiac and 1 noncardiac). Actuarial survival was 94.8% +/- 4.0% at 20 years. Catheterization during late follow-up (6.8 +/- 4.7 years after repair) was done in 53 patients and the pressure ratio of the right ventricle to the left ventricle was 0.48 +/- 0.21. Fifteen patients underwent subsequent operation because of residual lesions, including ventricular septal defect in four patients, pulmonary stenosis in nine, combined ventricular septal defect and pulmonary stenosis in one, and pulmonary regurgitation in one, with no mortality. Actuarial rate of freedom from reoperation was 71.1% +/- 8.0% at 10 years and 58.8% +/- 16.8% at 20 years. Patients with tetralogy and **subarterial ventricular septal defect** were more likely to have the development of residual obstruction at the level of the pulmonary valve anulus after repair than were those with tetralogy and **perimembranous ventricular septal defect**.

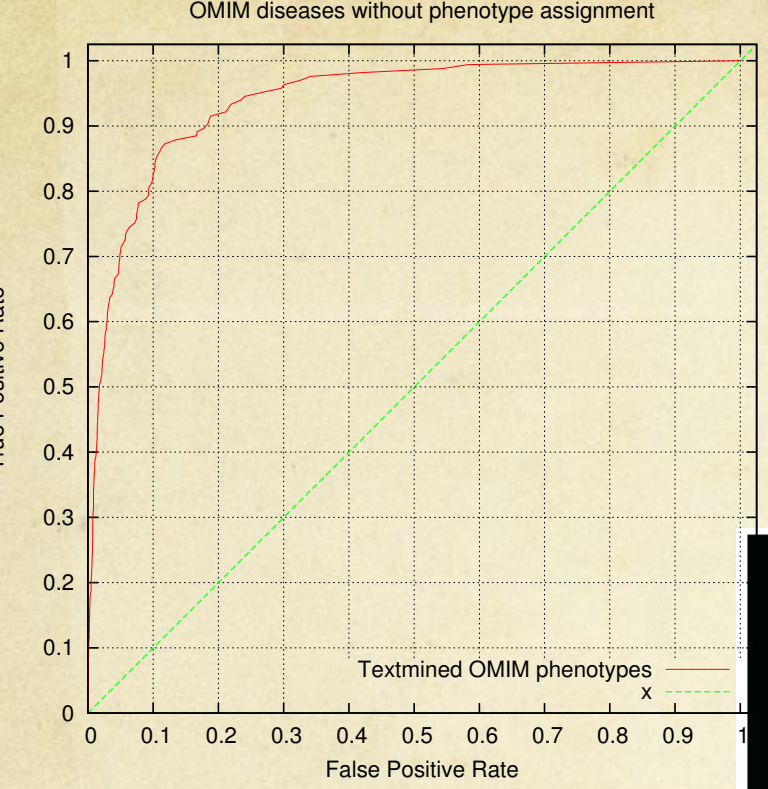
[A case of coronary artery-pulmonary artery fistula in tetralogy of Fallot with pulmonary atresia and major aortopulmonary collateral arteries (MAPCA)].

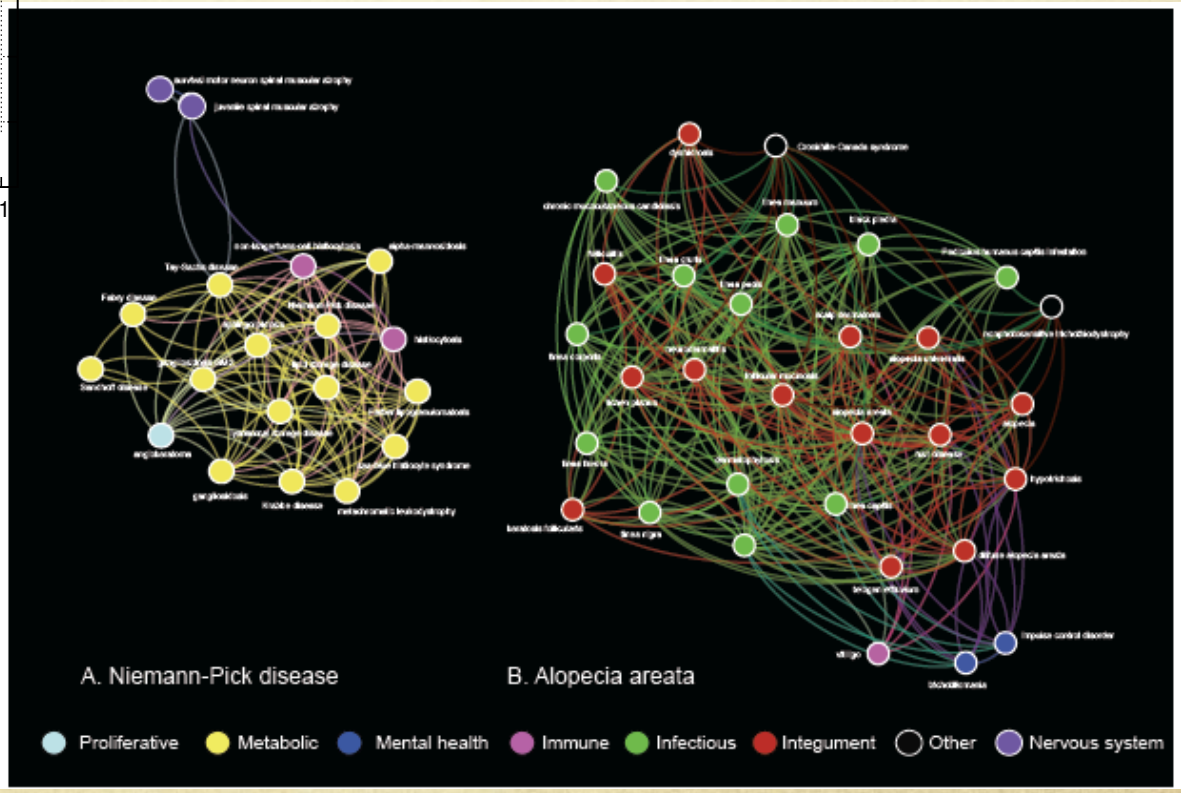
Aber-OWL: Common diseases





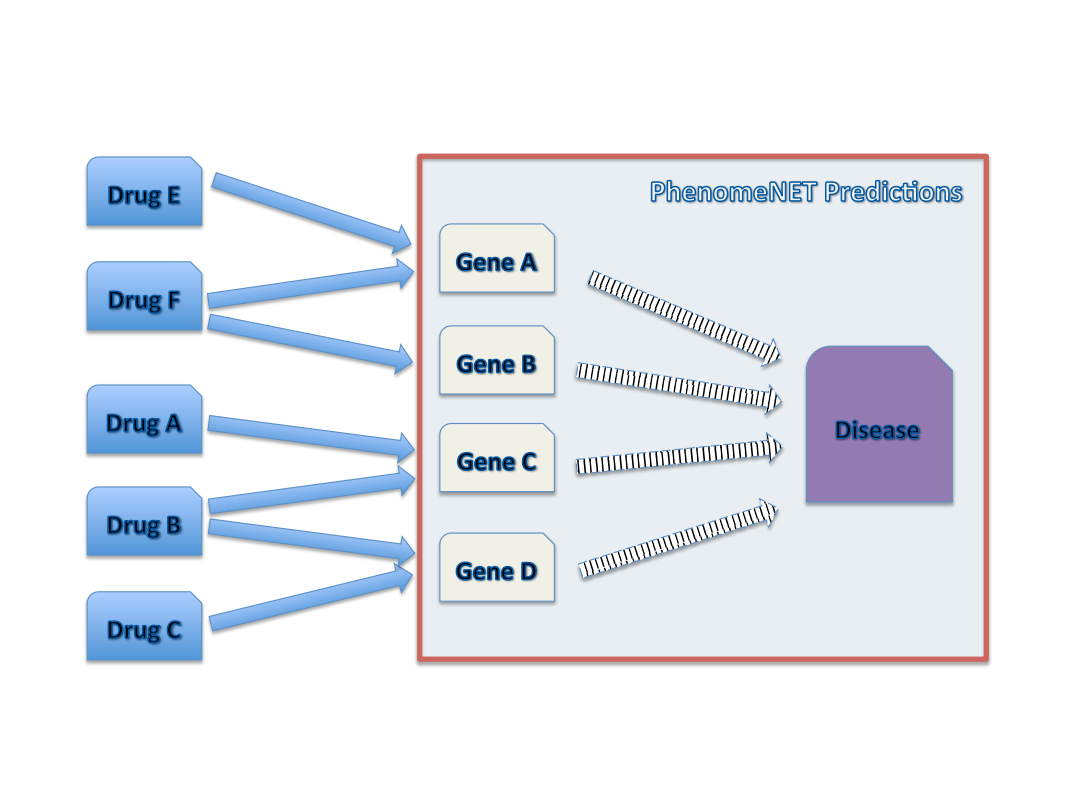
Nat. Scientific Reports. 2015 *in press*.
Analysis of the human diseasome reveals phenotype modules across common, genetic, and infectious diseases
Hoehndorf R, Schofield PN, Gkoutos GV.

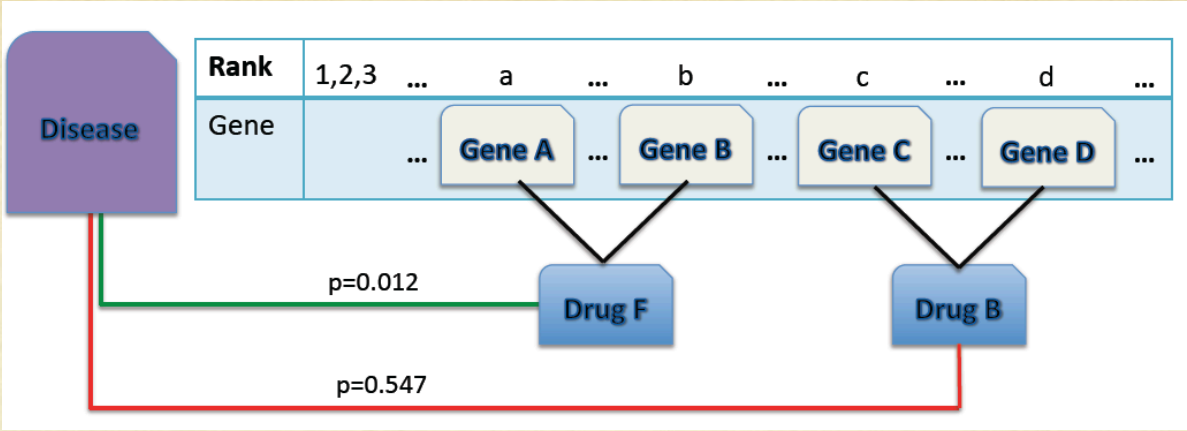

OMIM diseases without phenotype assignment
True Positive Rate
False Positive Rate
Textmined OMIM phenotypes
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

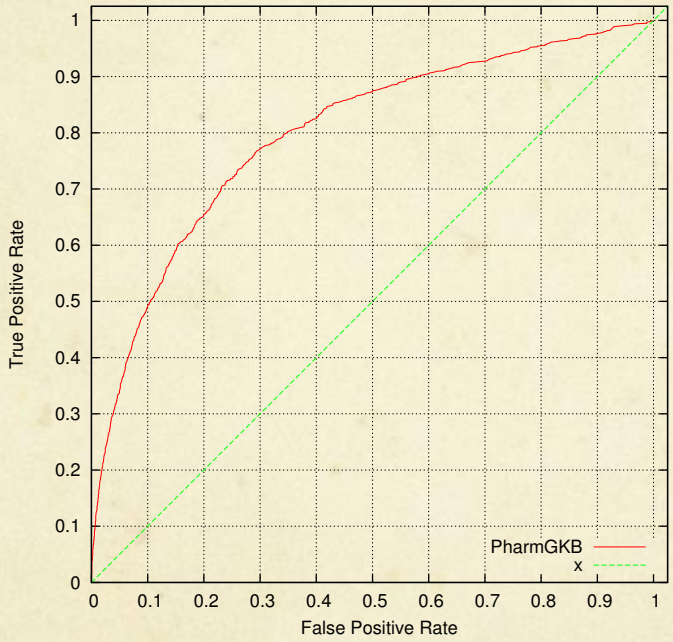

A. Niemann-Pick disease
B. Alopecia areata
Proliferative
Metabolic
Mental health
Immune
Infectious
Integument
Other
Nervous system

Pharmacogenomics

Can a phenotype of gene which the drug interacts be used to predict diseases in which the drug is active?







AUC: 0.76

Extend PhenomeNET to include side effects, drug indications, drug targets

Pac Symp Biocomput. 2012:388-99.
Linking PharmGKB to phenotype studies and animal models of disease for drug repurposing.
Hoehndorf R, Oellrich A, Rebholz-Schuhmann D, Schofield PN, Gkoutos GV.

Disease and drug pathways

Integrate:

- chemical ontologies and disease ontologies
  - e.g. ATC, ChEBI, MeSH (chemicals, diseases) HumanDO

- databases containing drug, gene and disease information
  - DrugBank, PharmGKB, CTD, Pathway Interaction Database
  - associations: drug-gene, gene-disease, drug-disease

Identify:

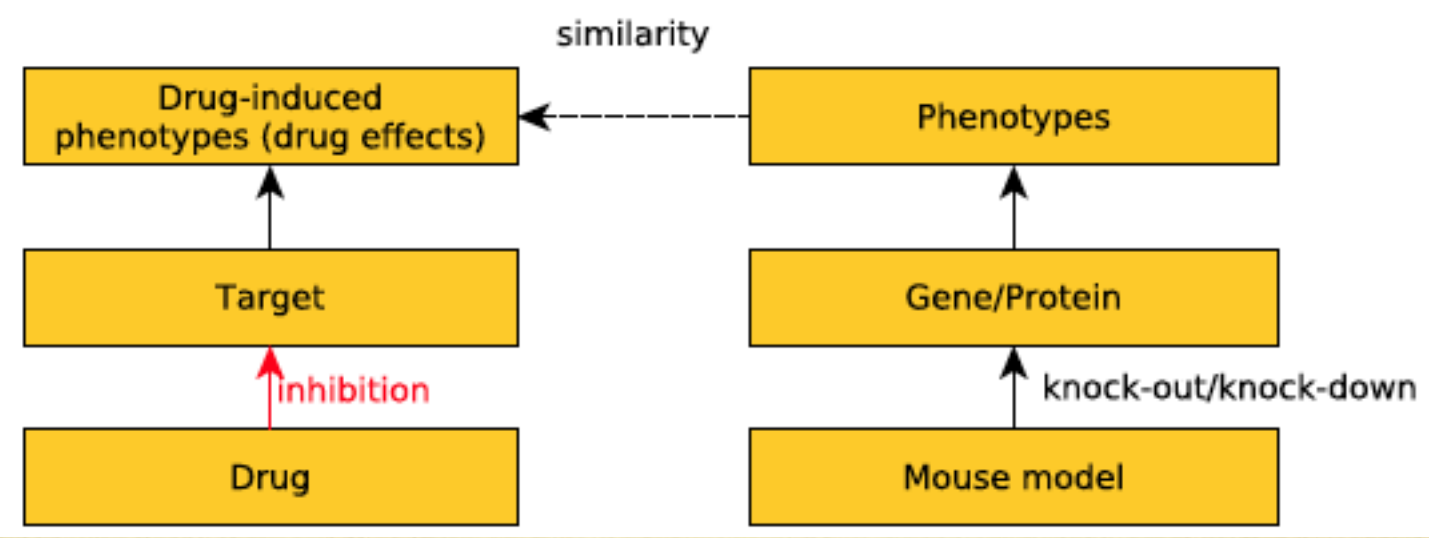
- 22,653 disease-pathway associations (disease pathways)
  - e.g. Mood disorder and Zidovudine Pathway ($p < 10^{-10}$)

- 13,826 pathway-chemical associations (drug pathways)
  - e.g. Clopidogrel and Endothelin signaling pathway ($p < 10^{-3}$)

Bioinformatics. 2012 Aug 15;28(16):2169-75.
Identifying aberrant pathways through integrated analysis of knowledge in pharmacogenomics.
Hoehndorf R, Dumontier M, Gkoutos GV.

Drug targets and indications identification

A similarity between drug D's effects and phenotypes resulting from *knock-out/knock-down* of a gene/protein in an animal model may indicate that D *inhibits* the gene/protein (or its human ortholog).



PhenomeDrug

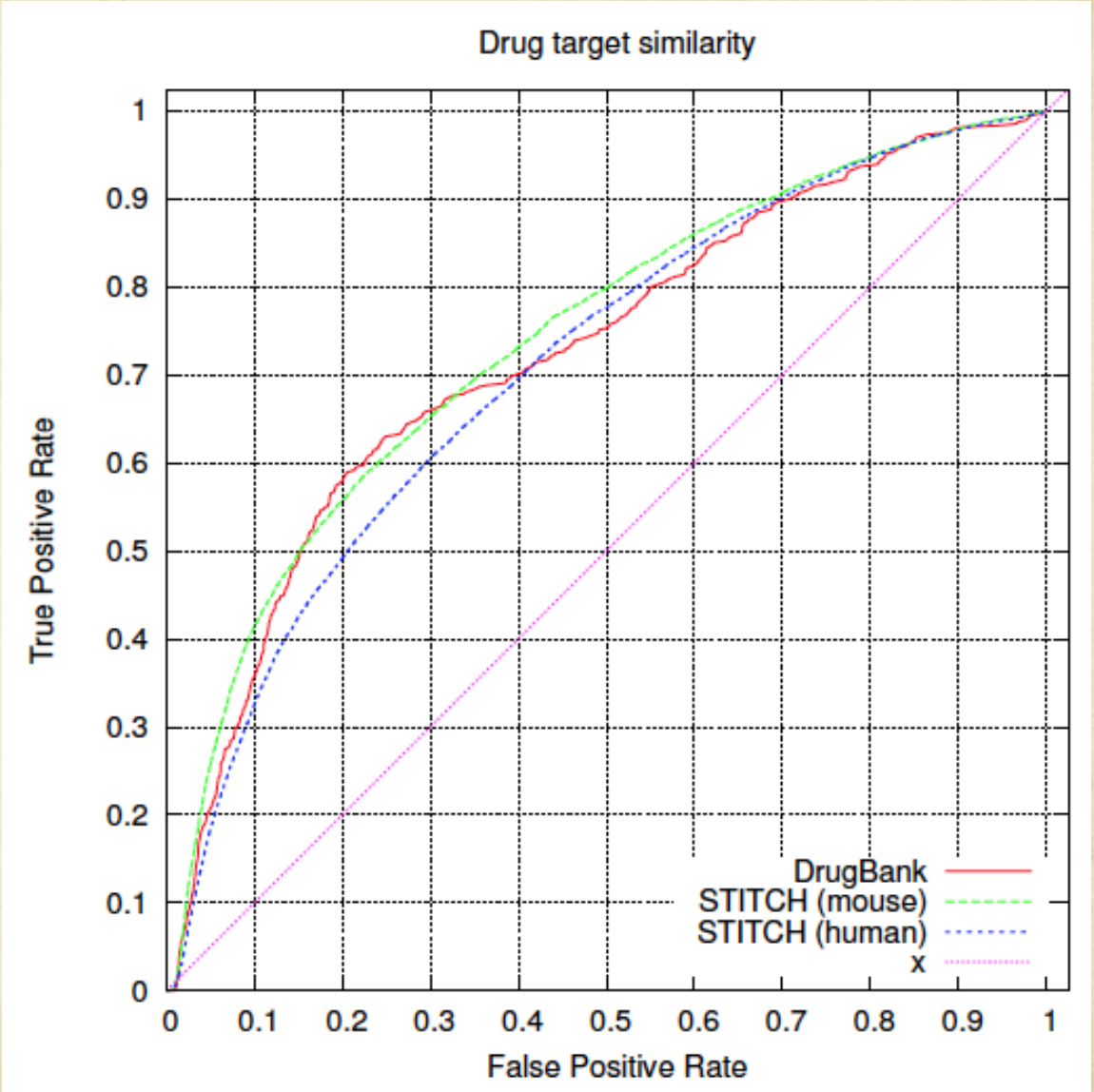
Evaluation using human and mouse drug targets

- DrugBank
- STITCH

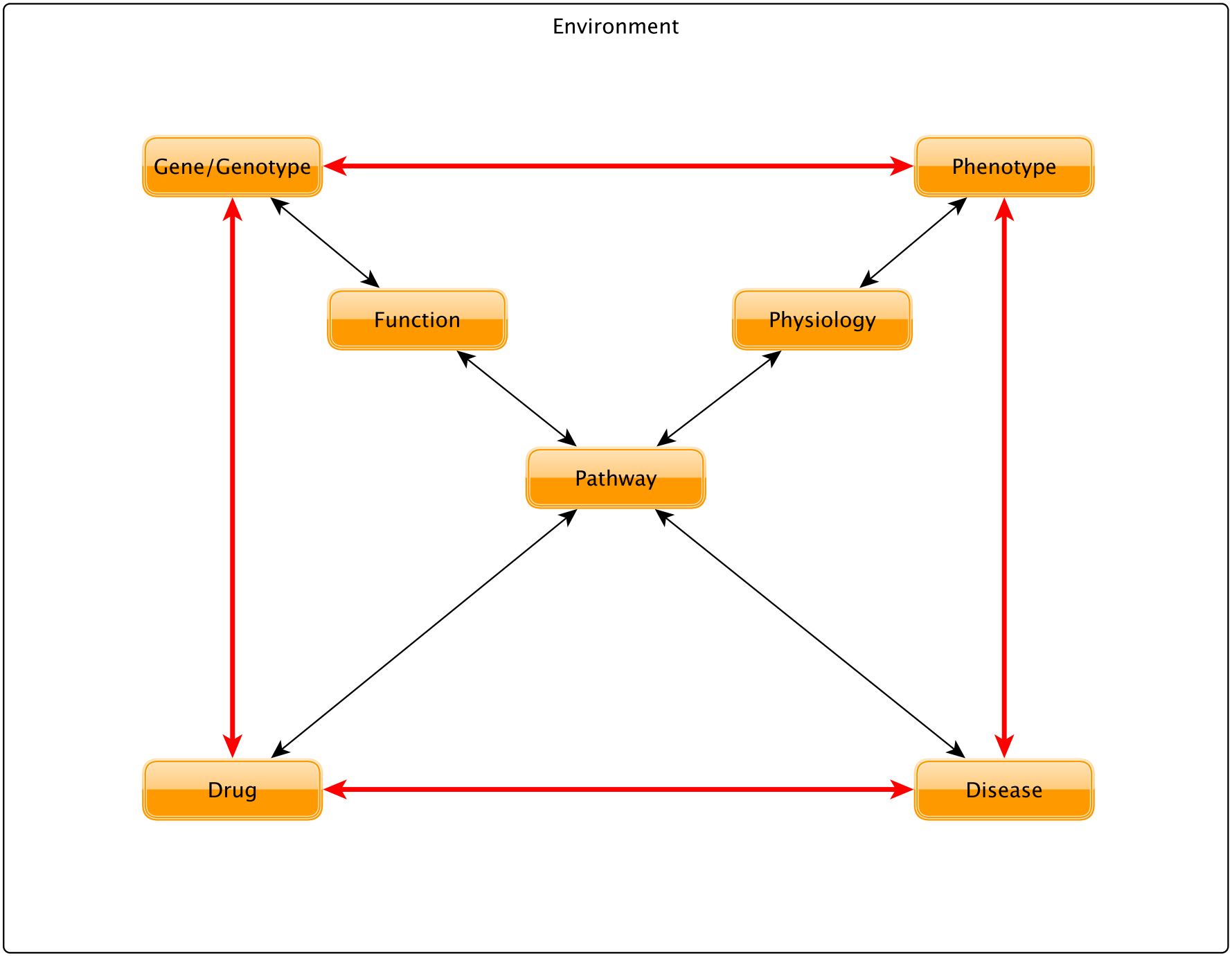
Drug side effects
- SIDER

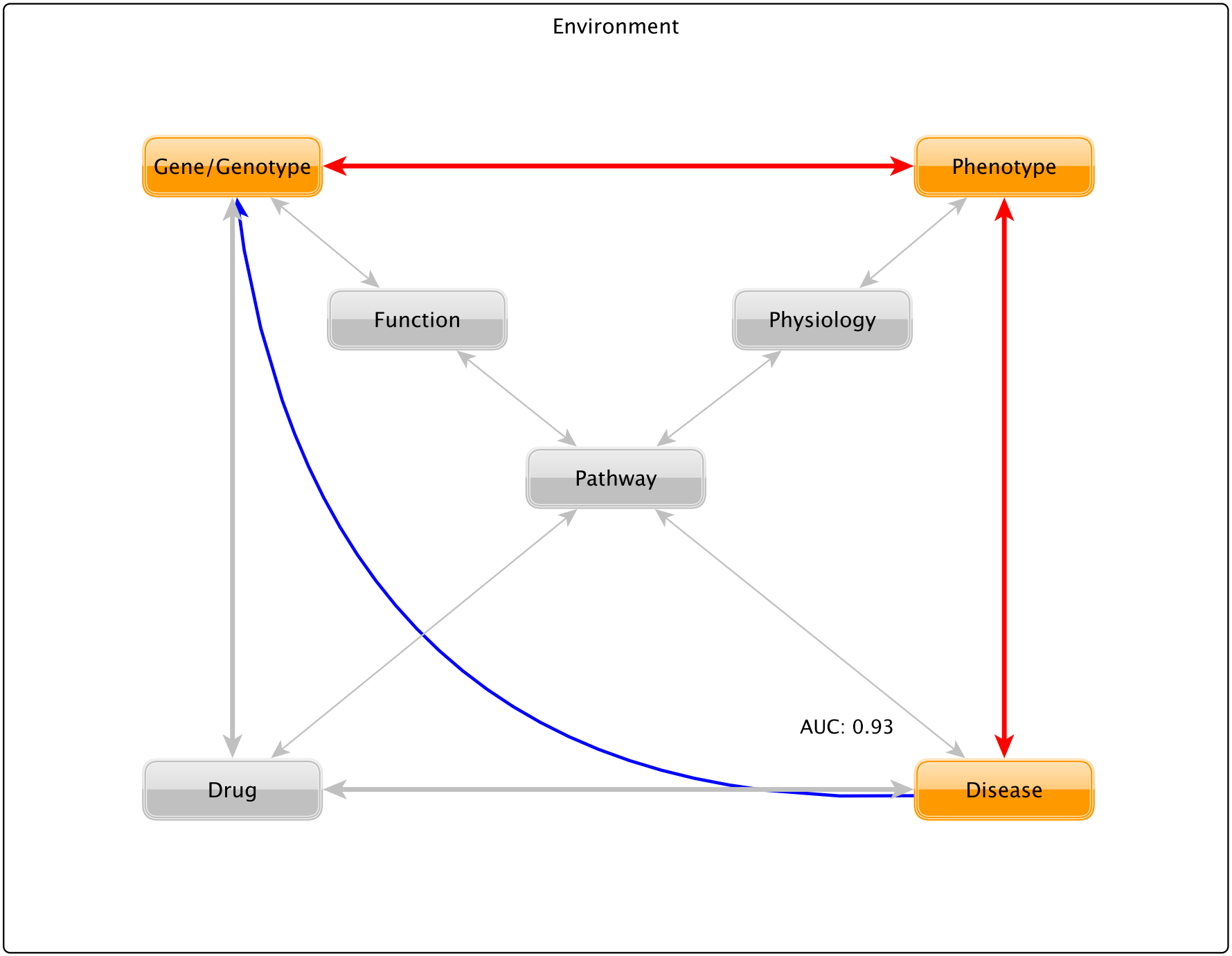
Overall AUC: 0.82

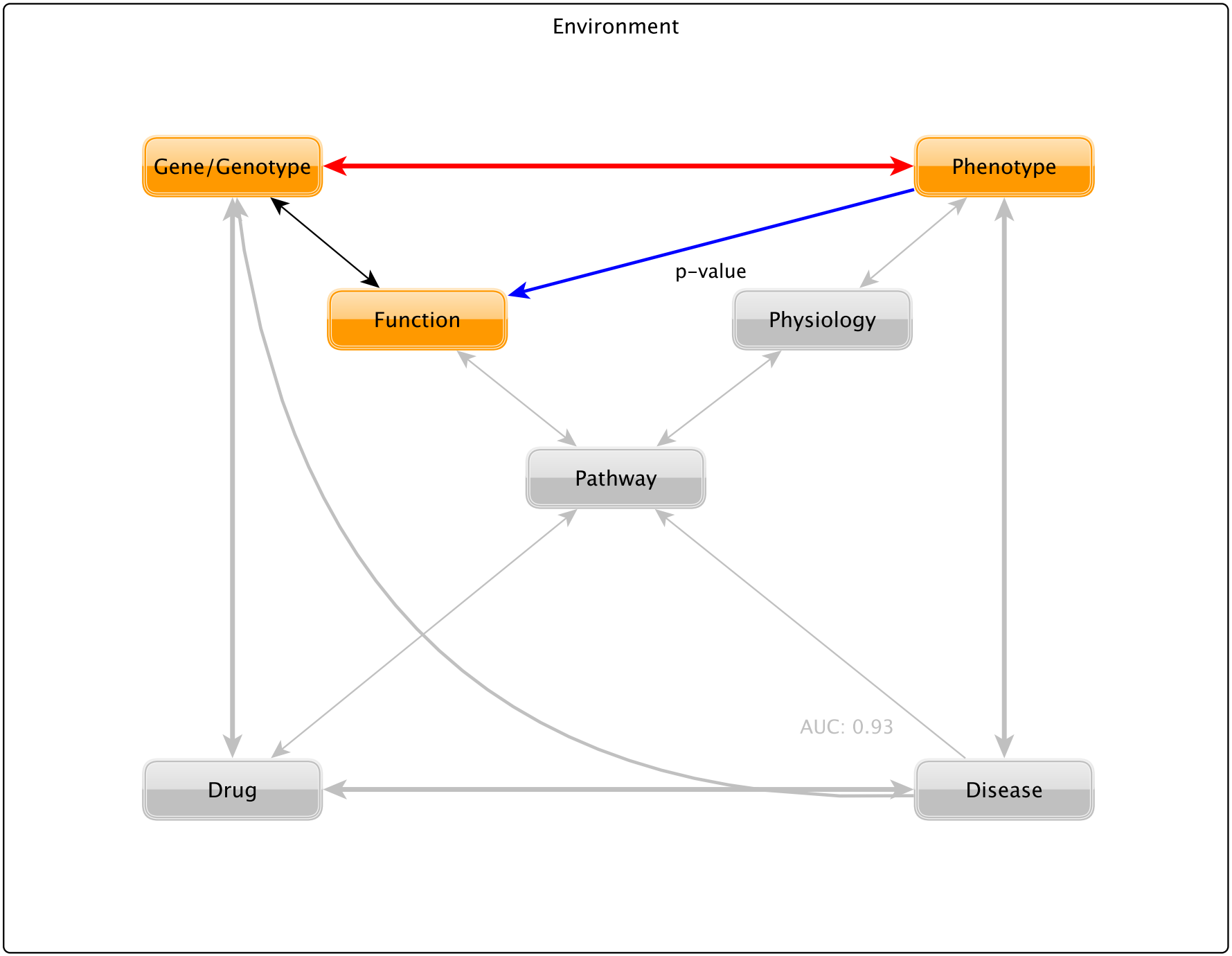
InterPro family	ROCAUC (STITCH)
G protein-coupled receptor, rhodopsinlike	0:800
Peptidase S1A, chymotrypsin-type	0:892
Steroid hormone receptor	0:916
Neurotransmitter-gated ion-channel	0:581

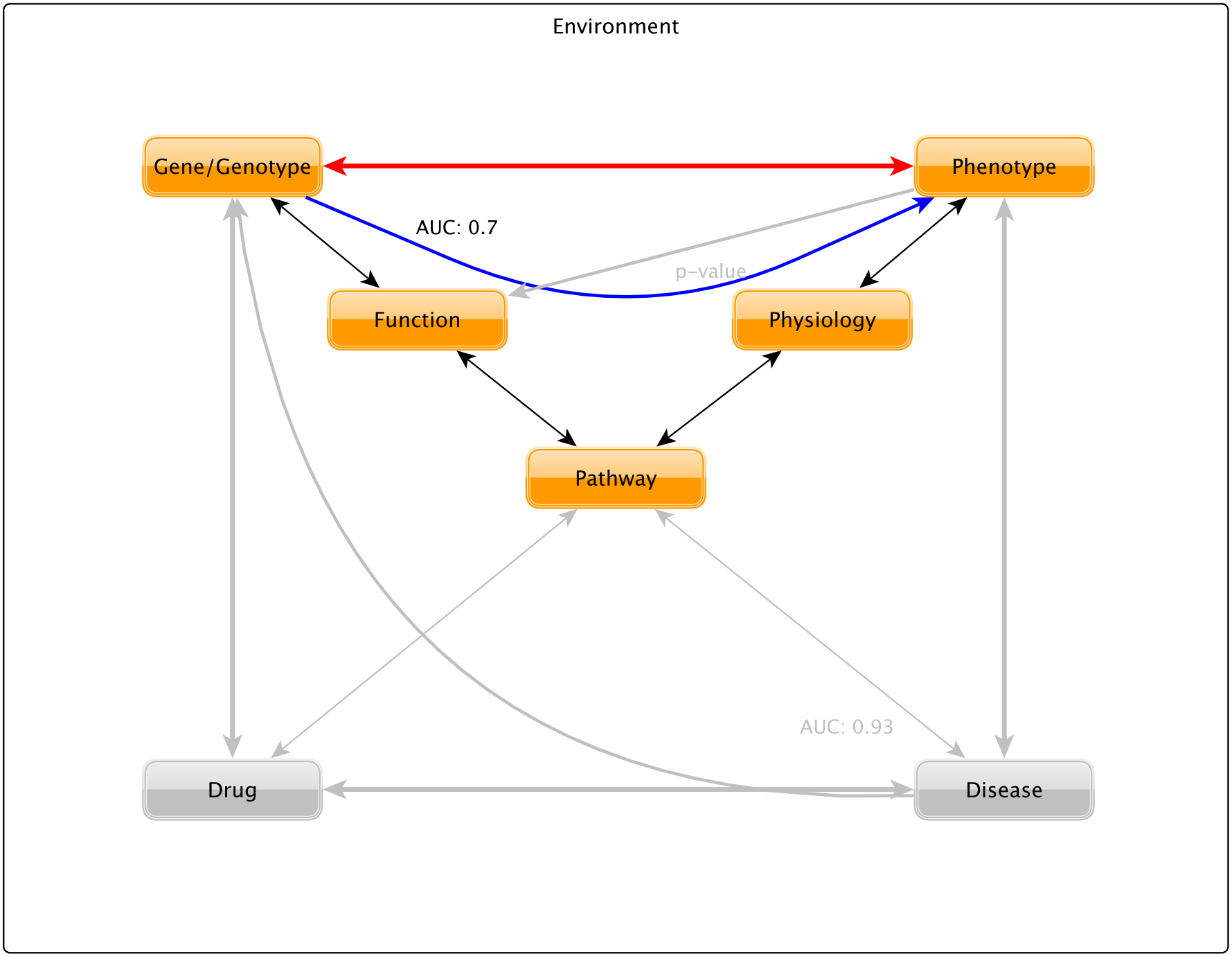


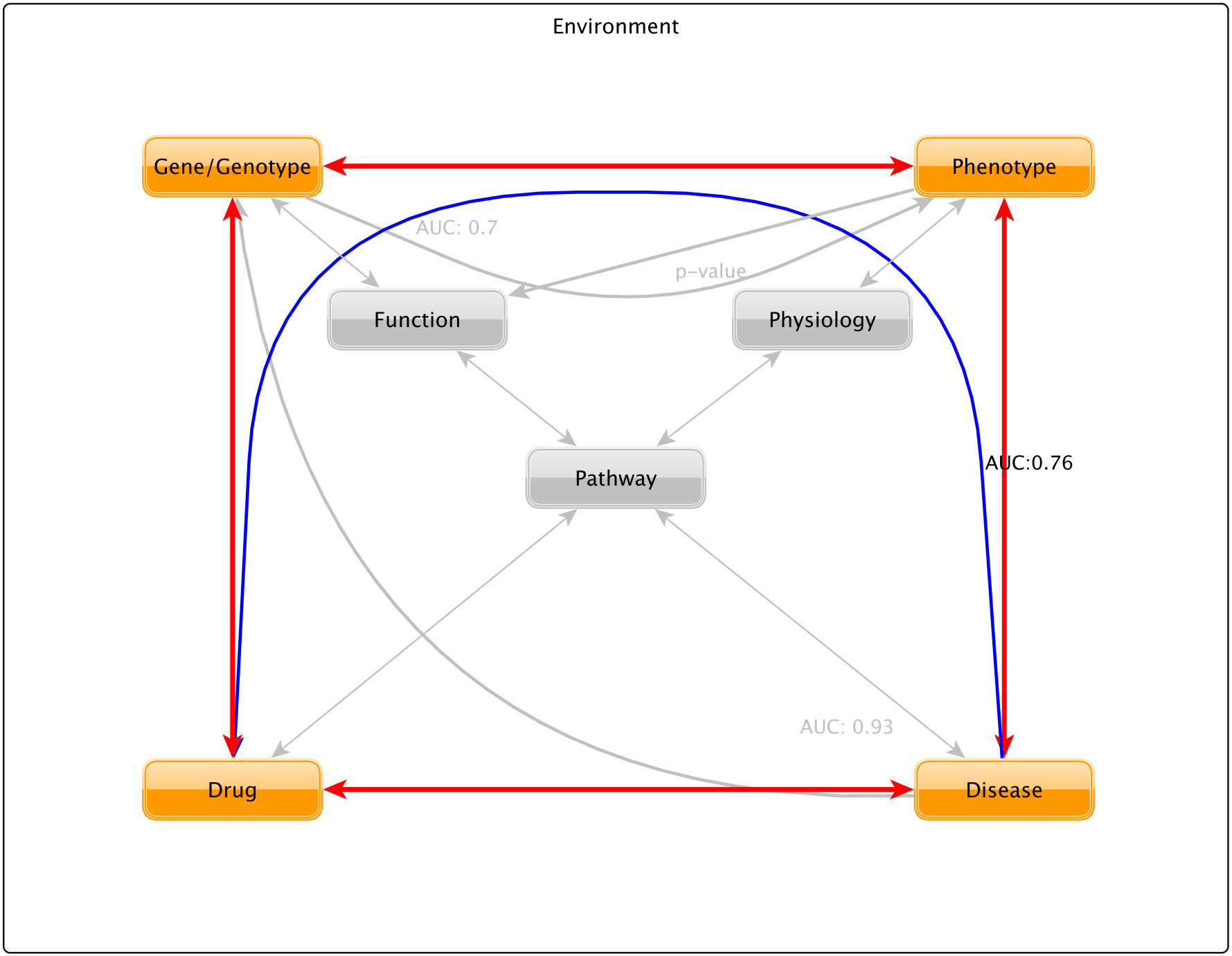
Bioinformatics. 2013 Oct 24, PMID: 24158600
Mouse model phenotypes provide information about human drug targets.
Hoehndorf R, Hiebert T, Hardy NW, Schofield PN, Gkoutos GV

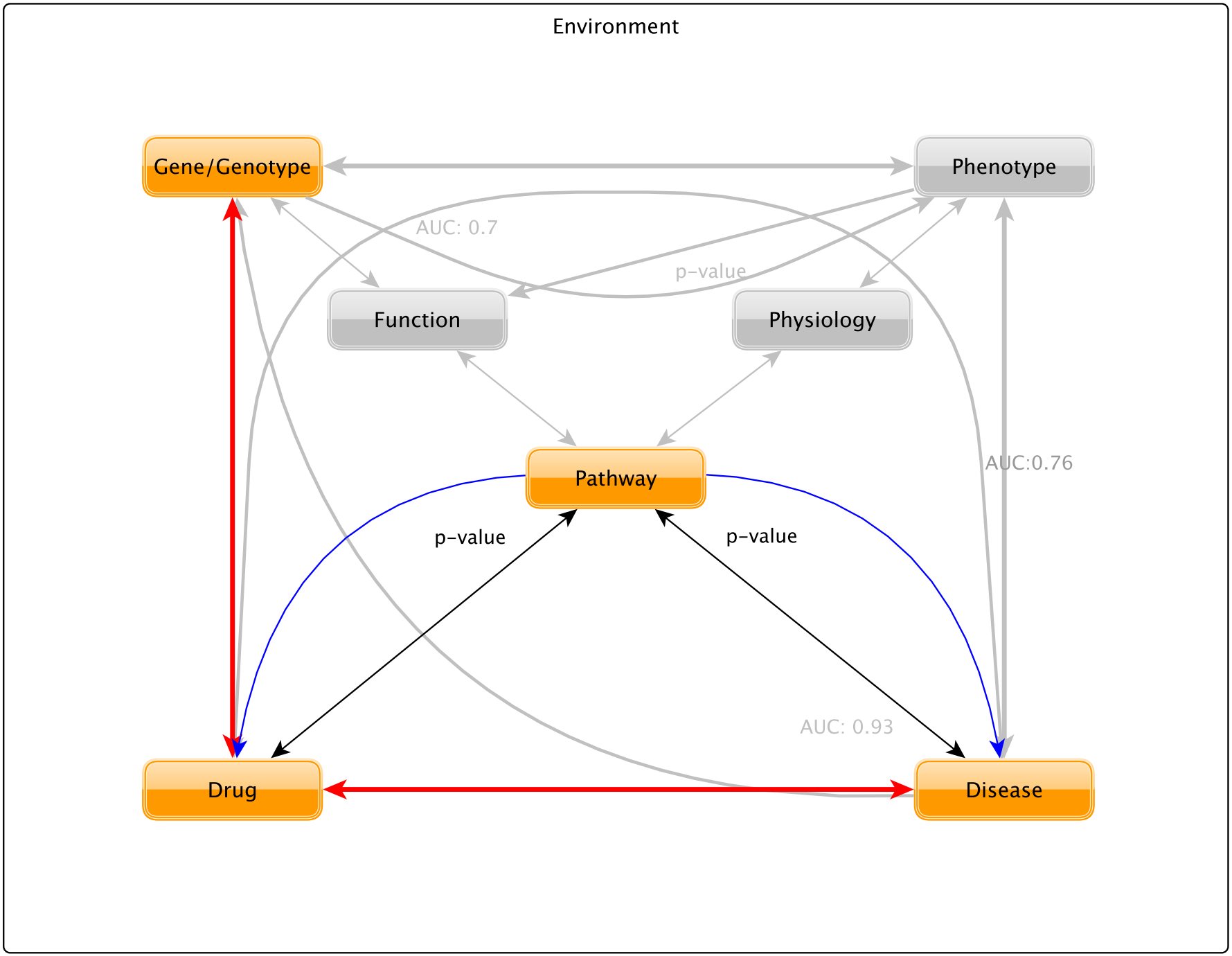

Environment
Gene/Genotype
Phenotype
Function
Physiology
Pathway
Drug
Disease

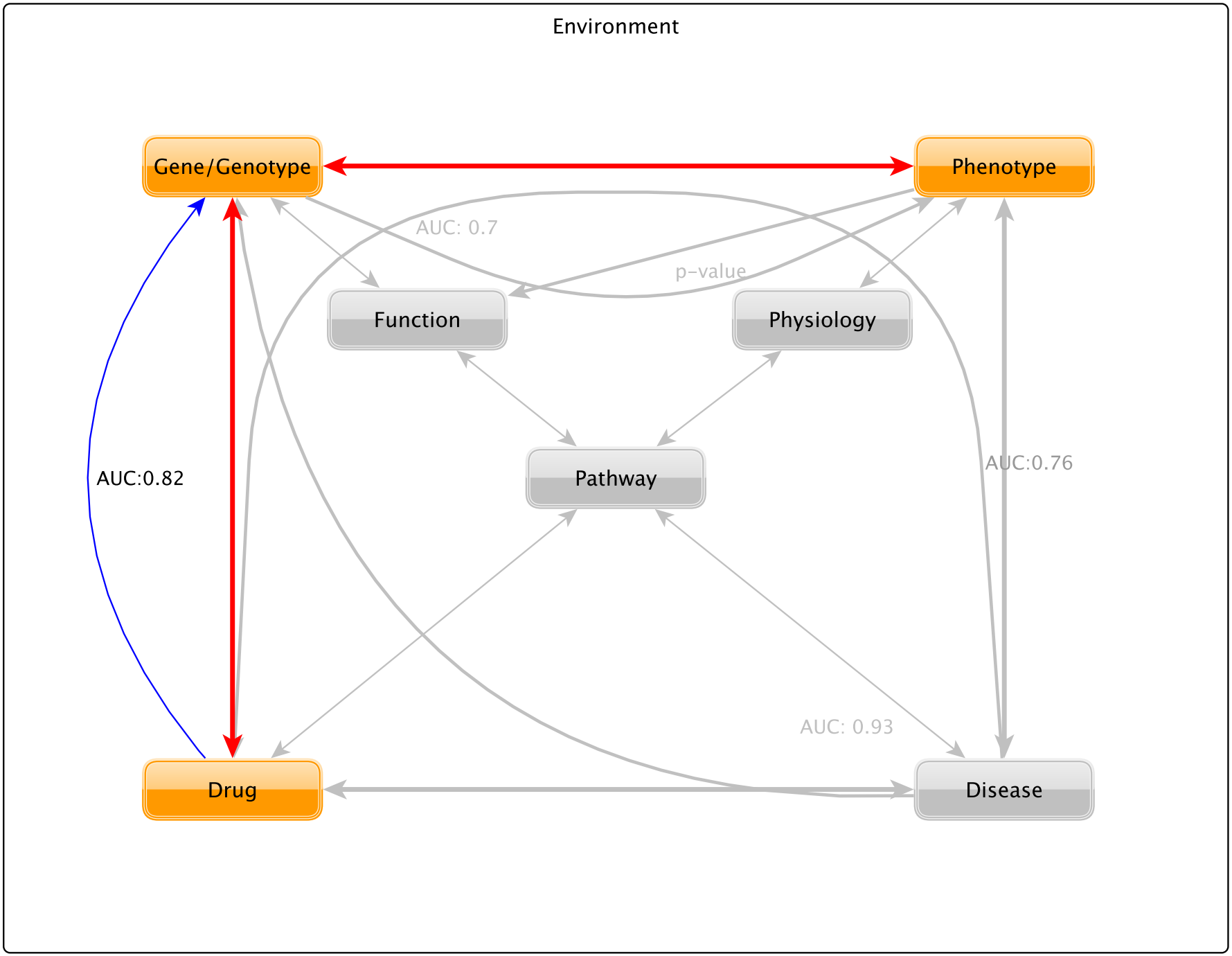

Environment
Gene/Genotype
Phenotype
Function
Physiology
Pathway
AUC: 0.93
Drug
Disease

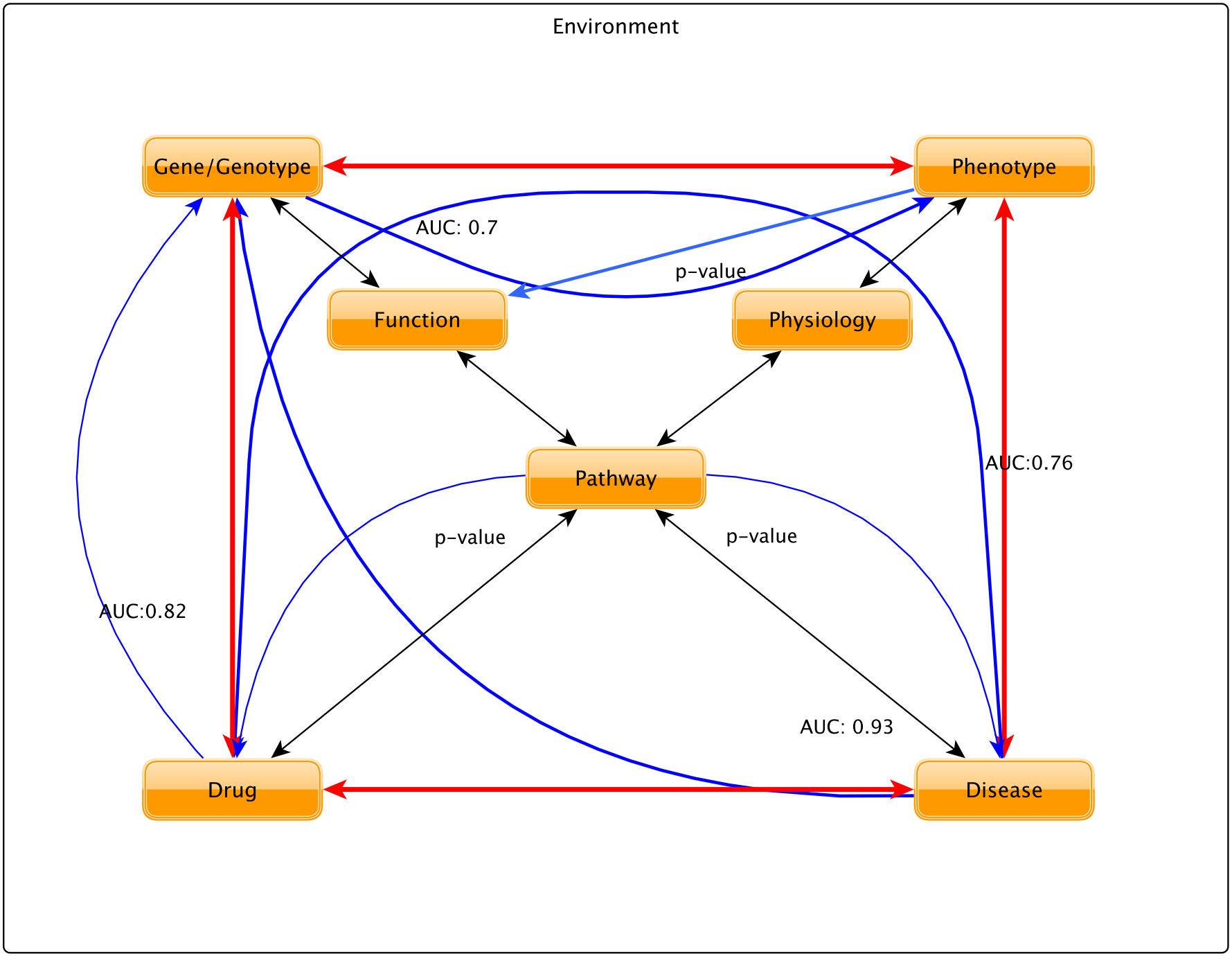

Environment
Gene/Genotype
Phenotype
p-value
Function
Physiology
Pathway
AUC: 0.93
Drug
Disease

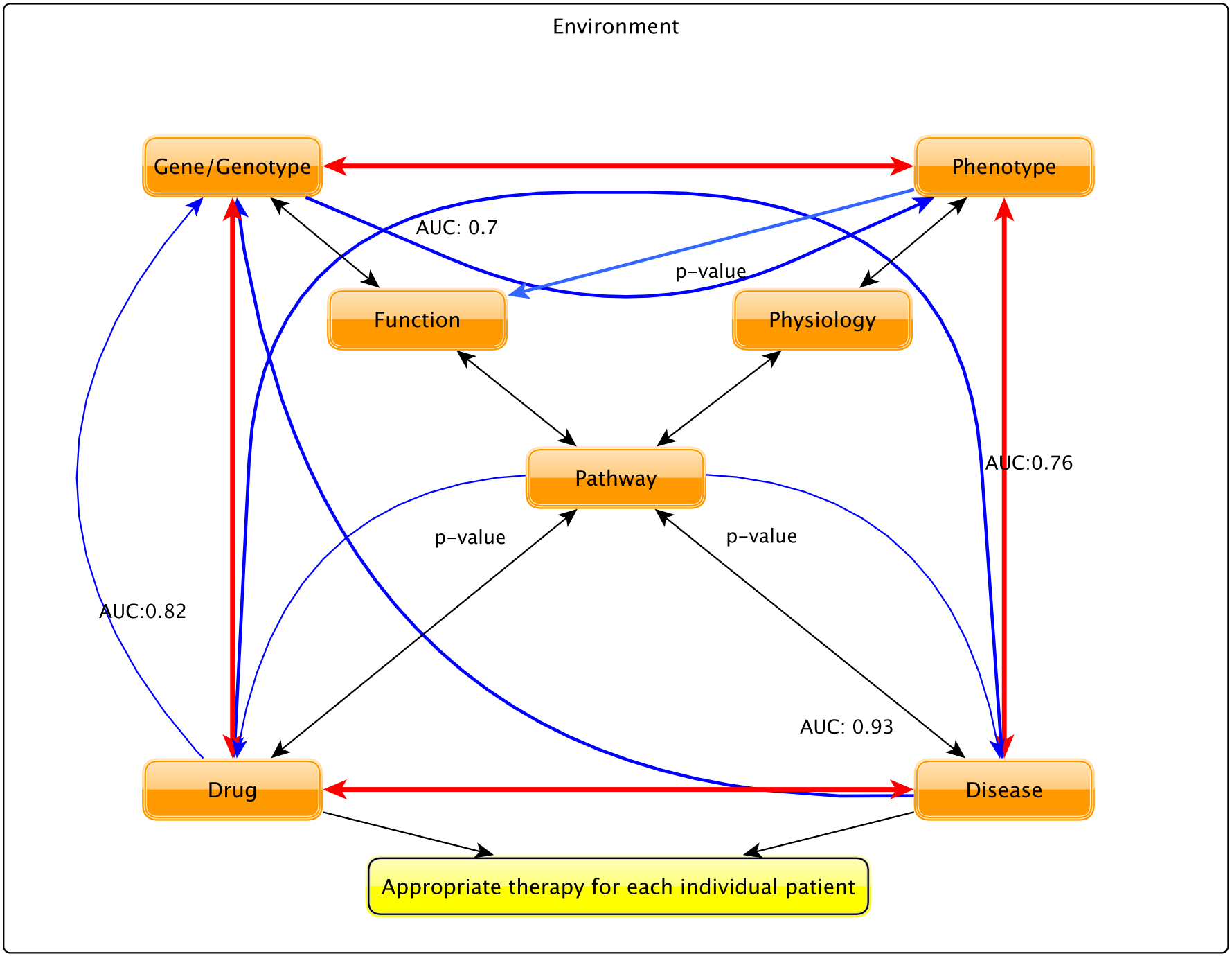

Environment
Gene/Genotype
Phenotype
AUC: 0.7
p-value
Function
Physiology
Pathway
AUC: 0.93
Drug
Disease

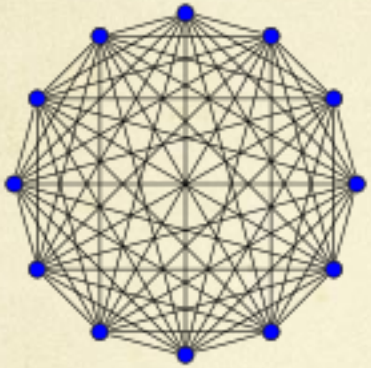

Environment
Gene/Genotype
Phenotype
AUC: 0.7
p-value
Function
Physiology
Pathway
AUC:0.76
AUC: 0.93
Drug
Disease


Environment
Gene/Genotype
Phenotype
AUC: 0.7
p-value
Function
Physiology
AUC:0.76
Pathway
p-value
p-value
AUC: 0.93
Drug
Disease


Environment
Gene/Genotype
Phenotype
AUC: 0.7
p-value
Function
Physiology
AUC:0.82
Pathway
AUC:0.76
AUC: 0.93
Drug
Disease


Environment
Gene/Genotype
Phenotype
AUC: 0.7
p-value
Function
Physiology
Pathway
AUC:0.76
p-value
p-value
AUC:0.82
AUC: 0.93
Drug
Disease


Environment
Gene/Genotype
Phenotype
AUC: 0.7
p-value
Function
Physiology
Pathway
AUC:0.76
p-value
p-value
AUC:0.82
AUC: 0.93
Drug
Disease
Appropriate therapy for each individual patient



Translational Medicine

Personalized Medicine

Therapeutic strategies

Diagnostic Strategies

Systems Genetics

Mathematical modelling

Human Variation

Disease Gene Prioritsation

CNVs

Orphan Diseases

Functional Analysis

etc

translation

Cross Species Integration

Semantic Integration Framework

Systematic genome-wide phenotyping

Model Organism Databases

Human Disease

Drugs

Literature

Experimental Data