

# SESIÓN IV: Integration of Phenotypic and Functional Networks

**Junio, 2013**

Armando Reyes-Palomares

Universidad de Málaga  
Unidad 741 CIBERER

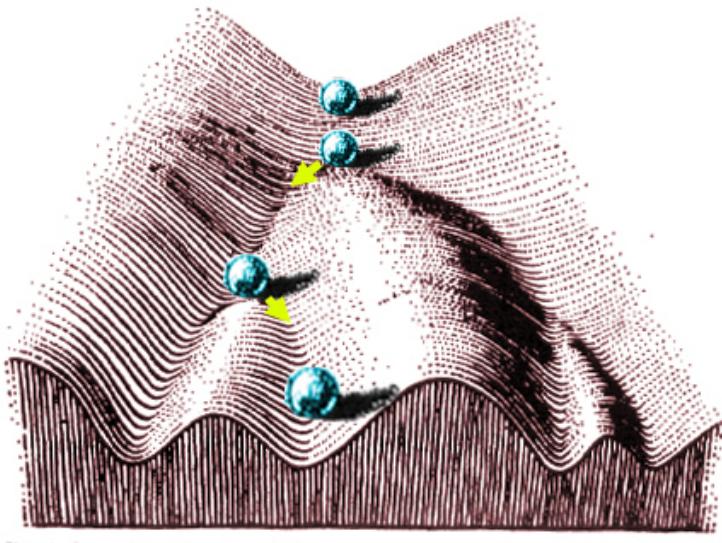
# Biological robustness

“Biological systems, from macromolecules to whole organisms, are **robust** if they continue to function, survive, or reproduce when faced with mutations, environmental change, and internal noise”

**Identify how biological systems lose robustness**

# Landscape Metaphor: Canalization in development biology

## Rolling balls down-hill

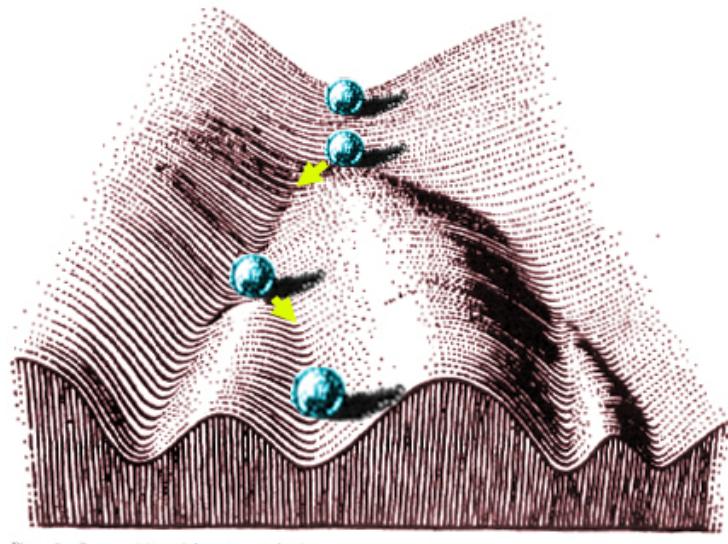


1. Genetic
2. Environmental
3. Stochastic  
(noise or random events)

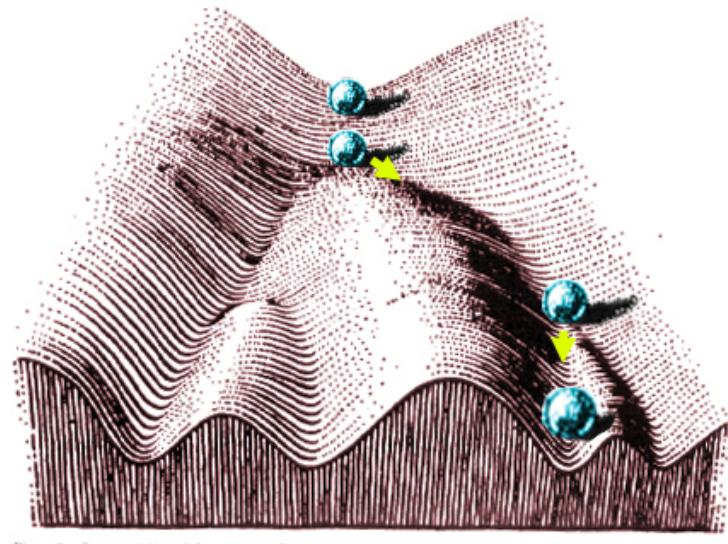
Conrad Hal Waddington

# Landscape Metaphor: Robustness for phenotypic variability

**Wild type**

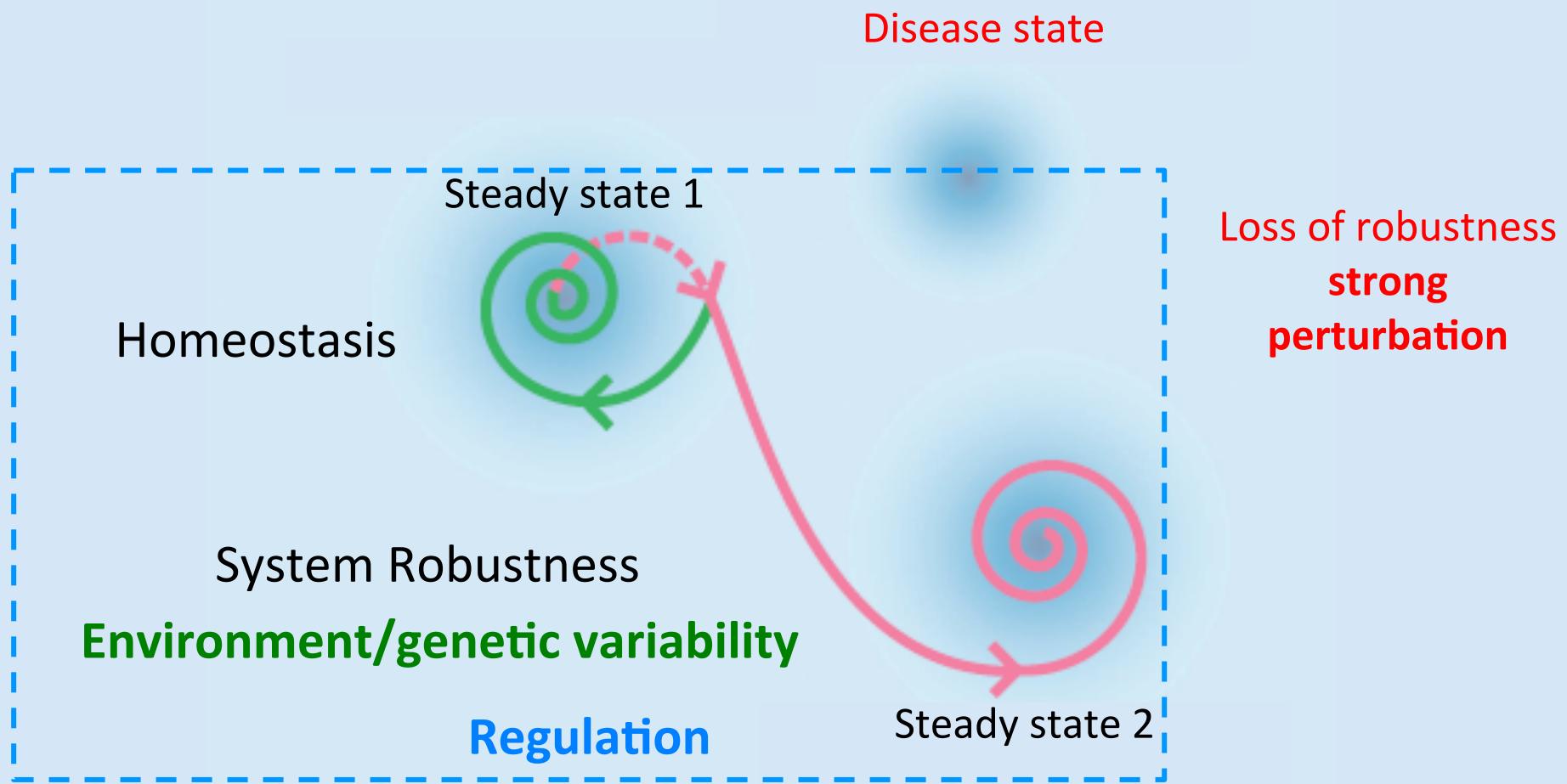


**Disease state**

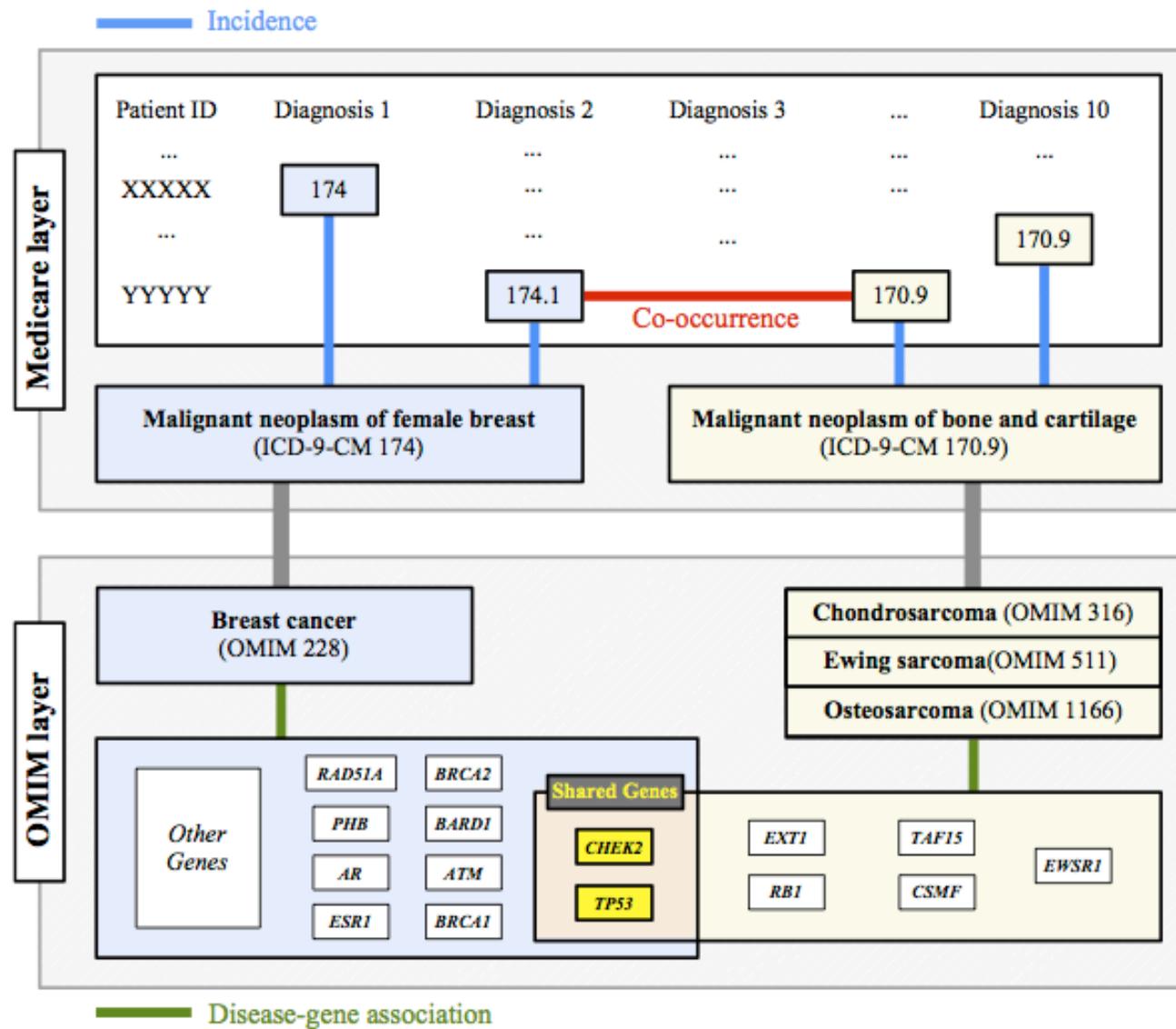


Conrad Hal Waddington

# Loss of robustness induces pathogenesis



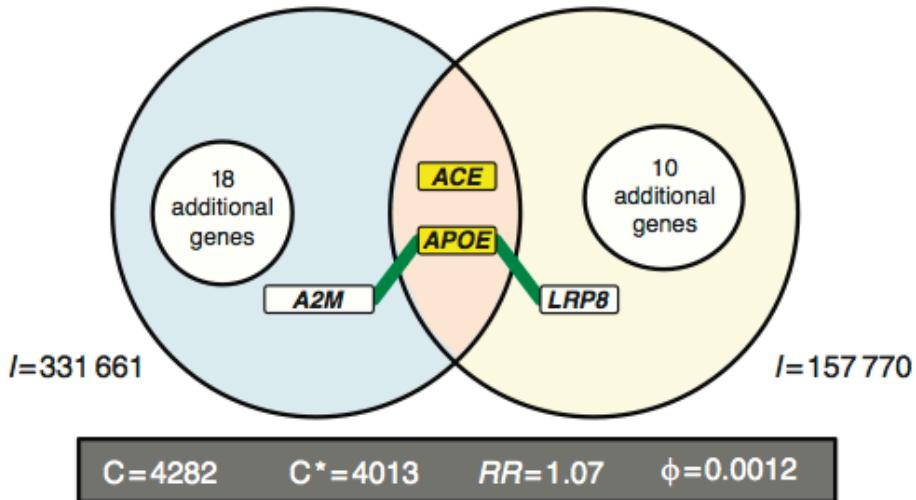
# COMORBIDITY



# COMORBIDITY

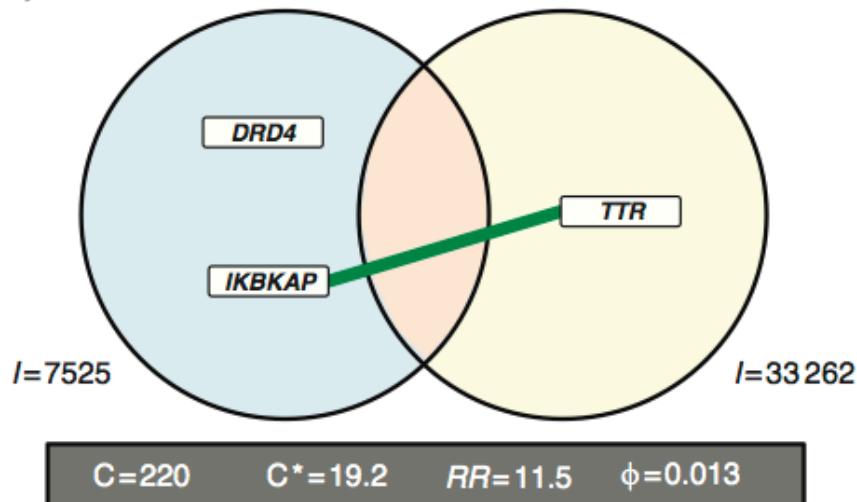
— PPI

Alzheimer's disease      Myocardial infarction



Autonomic nervous  
system disorder

Carpal tunnel syndrome

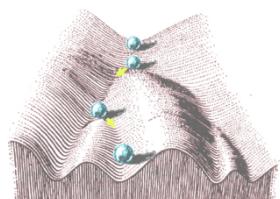


# Starting point

- Most genetic influences on a trait/phenotype are usually unknown
- Phenotypic variation is not just due to genetics.



1. It is NOT necessary detailed mechanistic models of a biological system in order to predict phenotypic variation, pragmatic statistical approaches may suffice (Genotype combinations studies).
2. Natural (biological) networks evolved to be robust to genetic, environmental and stochastic perturbations. GOOD level of abstraction, less precision



# Sources of information

1. Genetic diseases
2. Pathological Phenotypes variability
3. Functional interactions between genes

**“All of this data can be modeled in networks”**

OPEN  ACCESS Freely available online



## Global Analysis of the Human Pathophenotypic Similarity Gene Network Merges Disease Module Components

Armando Reyes-Palomares<sup>1,2</sup>, Rocío Rodríguez-López<sup>1,2</sup>, Juan A. G. Ranea<sup>1,2</sup>, Francisca Sánchez Jiménez<sup>1,2</sup>, Miguel Angel Medina<sup>1,2\*</sup>

<sup>1</sup> Department of Molecular Biology and Biochemistry, Faculty of Sciences, University of Málaga, Málaga, Spain, <sup>2</sup>CIBER de Enfermedades Raras (CIBERER), Málaga, Spain

# Databases



[www.omim.org](http://www.omim.org)



[www.orpha.net](http://www.orpha.net)

- Dr. Victor A. McKusick
- Repository of studies about clinical features and molecular genetics
- Focused on low prevalence diseases.
- Originally enriched by OMIM diseases
- Actively reviewed by clinical experts

**Genetic disorders:** 3.486

**Genetic disorders:** 2.125

**Mutated genes:** 2.794

**Mutated genes:** 2.331

**All diseases:** 7.263

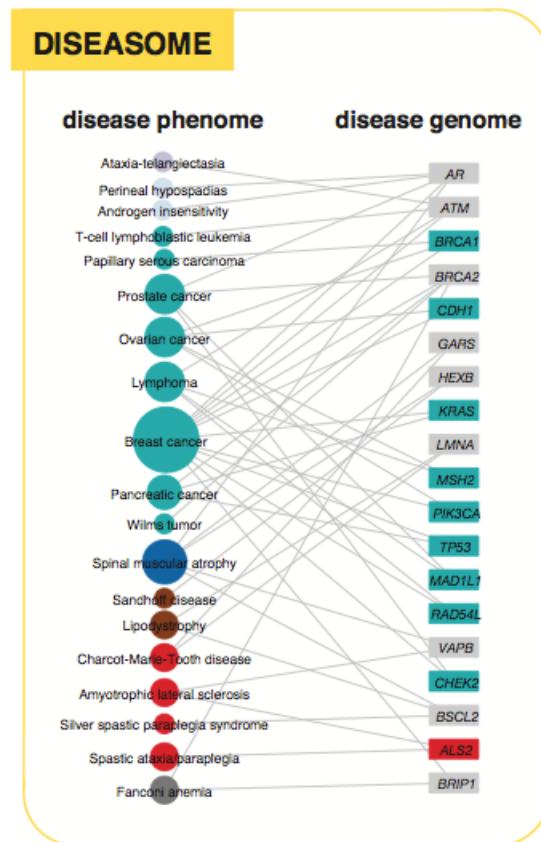
**All diseases:** 5.954

In many cases we know the gene but not the molecular etiology

# Databases 2 Diseases



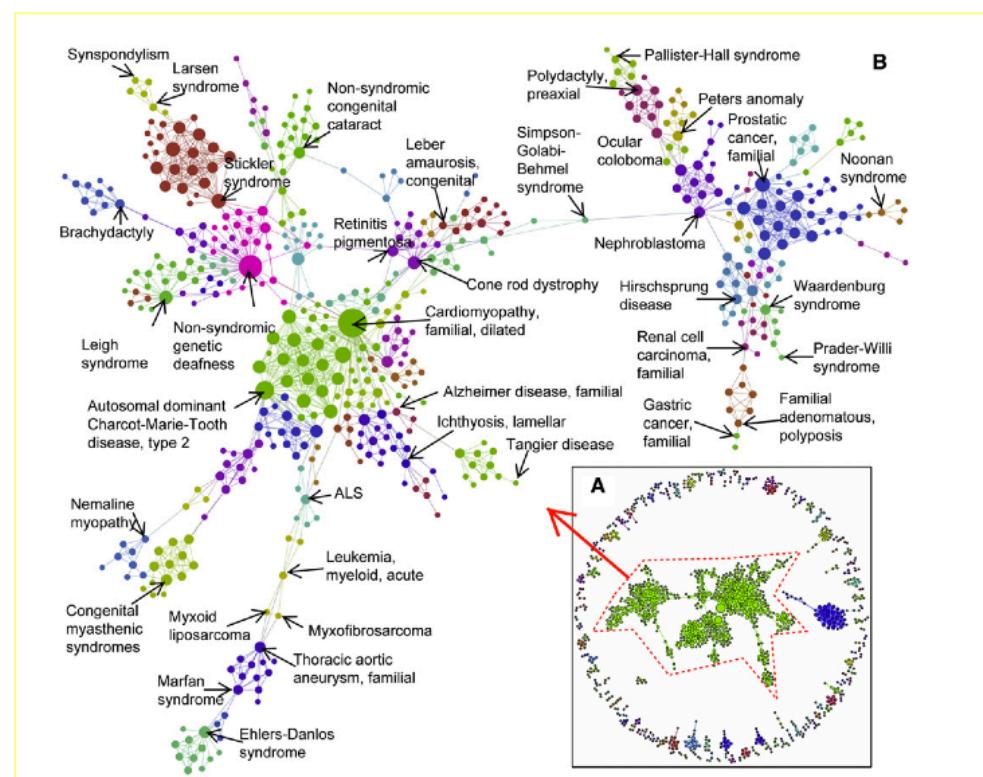
# The Human Diseases Network



Goh et al., PNAS (2007)

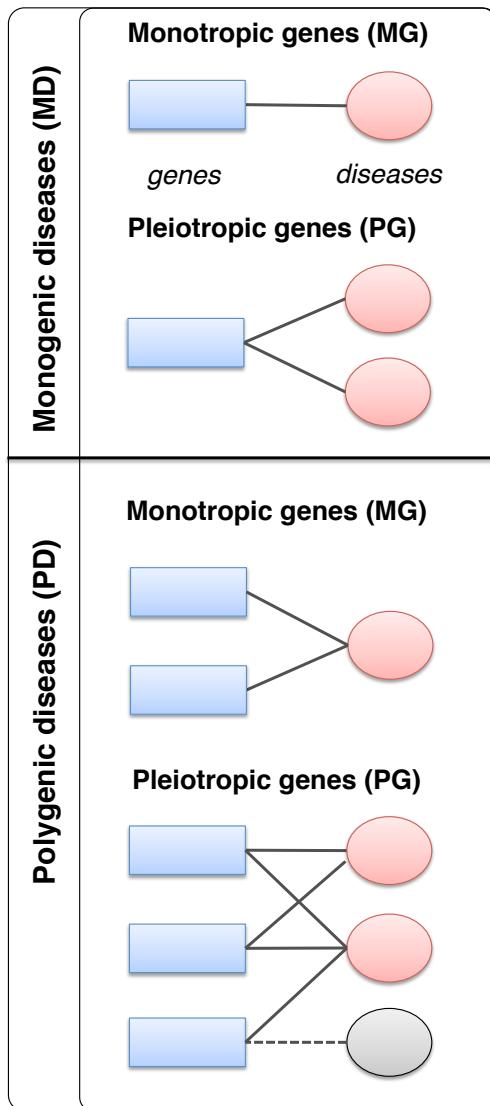


## The Orphan Disease Networks



Zhang et al., Am J Hum Genet (2011)

# Proposed classification (disease-gene)



Subset	Human Diseases Network		Orphan Disease Networks	
	Diseases per gene	Genes (%)	Diseases per gene	Genes (%)
MD-MG	1.00	1431 (56.7)	1.00	717 (30.8)
MD-PG	2.57	639 (25.3)	2.71	435 (18.7)
PD-MG	0.46	379 (15.0)	0.40	908 (39.0)
PD-PG <sup>a</sup>	2.13	371 (14.7)	1.68	584 (25.1)
All genes <sup>b</sup>	1.24	2525 (100)	0.91	2331 (100)

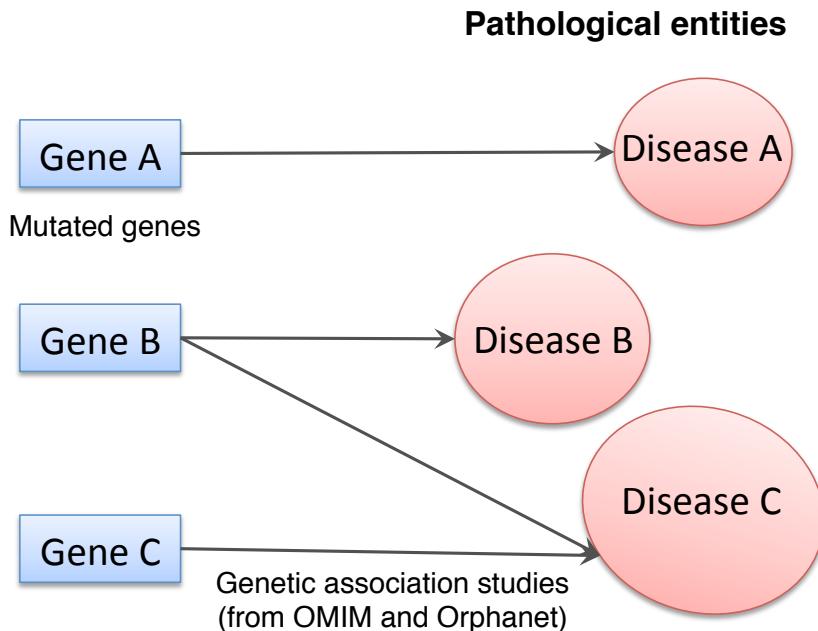
<sup>a</sup>Pleiotropic genes associated with at least one polygenic diseases.

<sup>b</sup>All genes in HDN and ODN respectively.

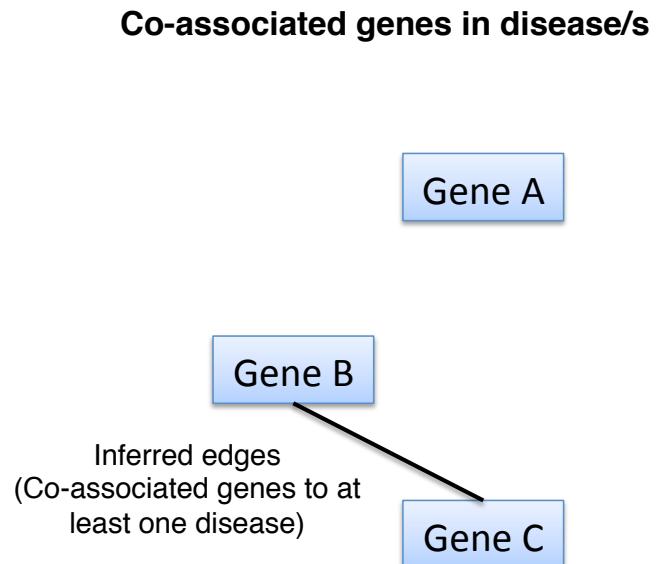
Strong differences between both inherited disease database

# From KNOWN to INFERRRED relationships

## BIPARTITE PROJECTION (disease-to-gene)

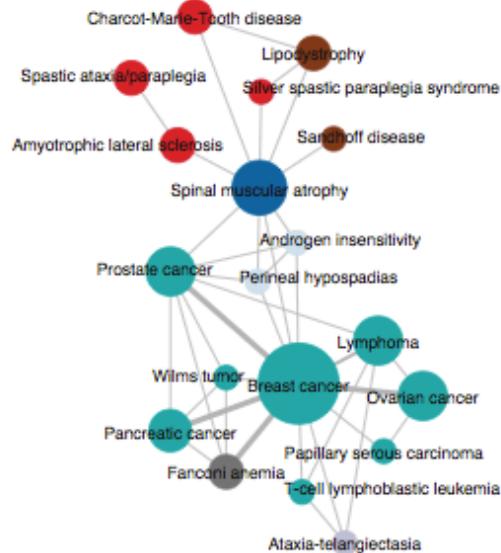


## INFERRRED UNIPARTITE PROJECTION (gene-to-gene)



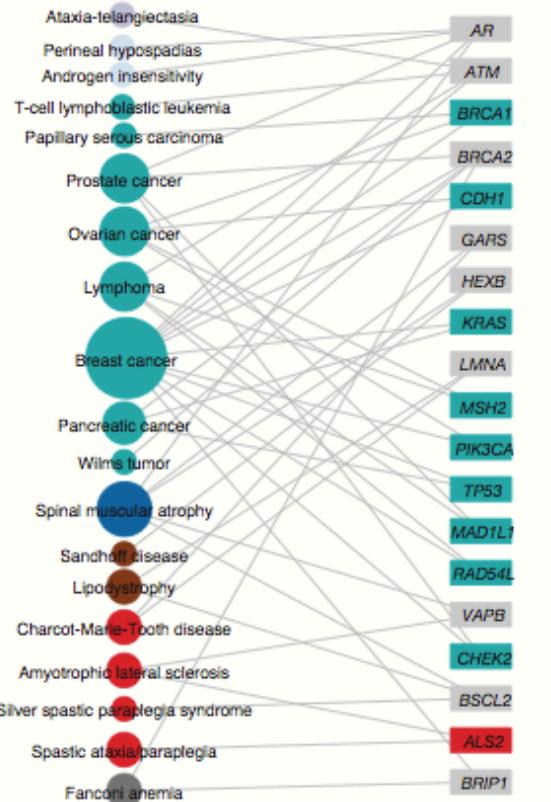
# The Human Diseases Network

*Human Disease Network  
(HDN)*

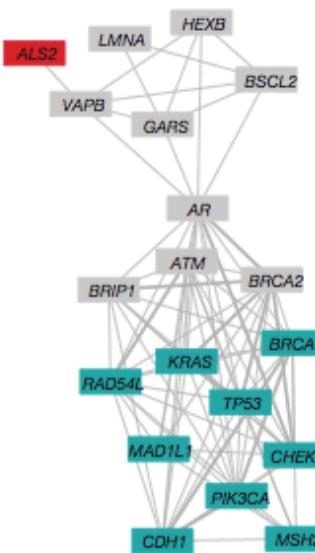


**DISEASOME**

**disease phenotype      disease genome**



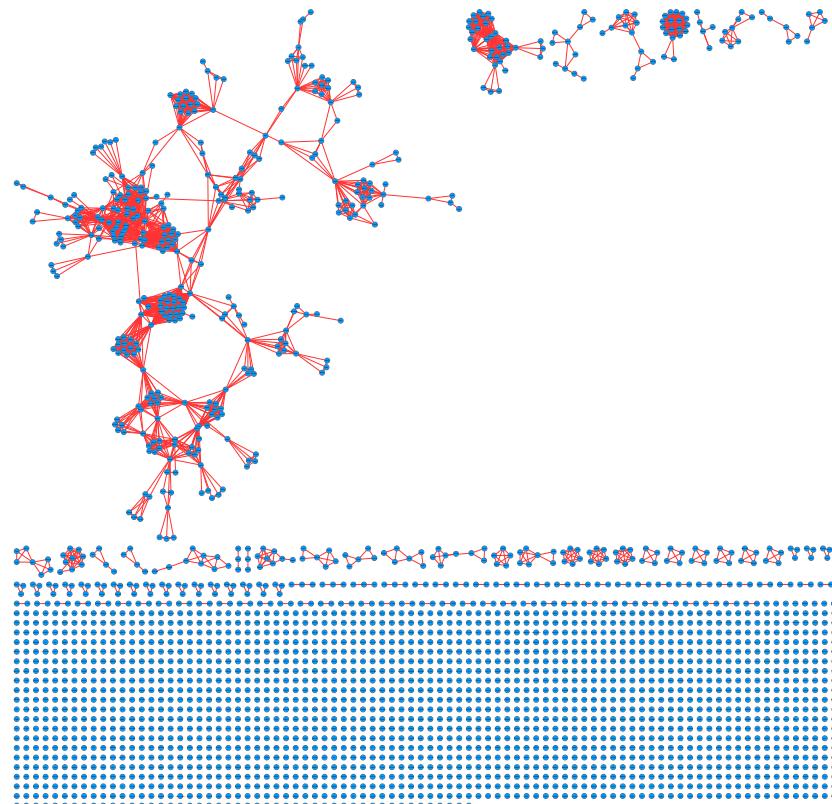
*Disease Gene Network  
(DGN)*



# Inferred gene-gene (unipartite projections)

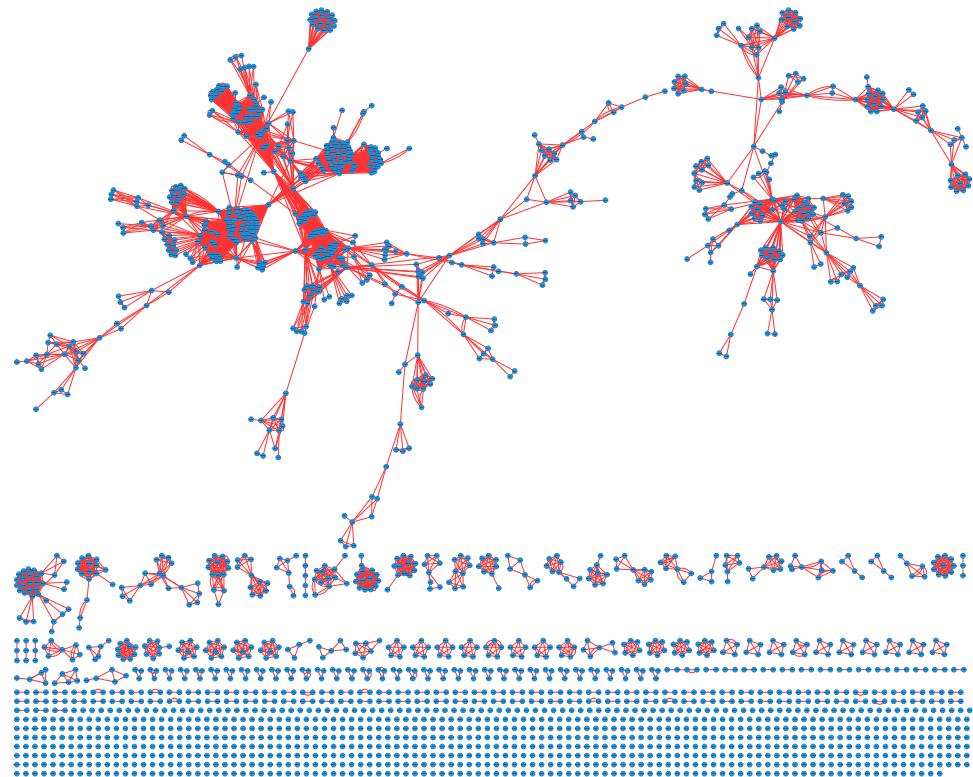
**A) Human Diseases Gene Network  
(gene-to-gene unipartite)**

Connected nodes: 749  
Unconnected nodes: 1776  
Edges: 2654



**B) Orphan Diseases Gene Network  
(gene-to-gene unipartite)**

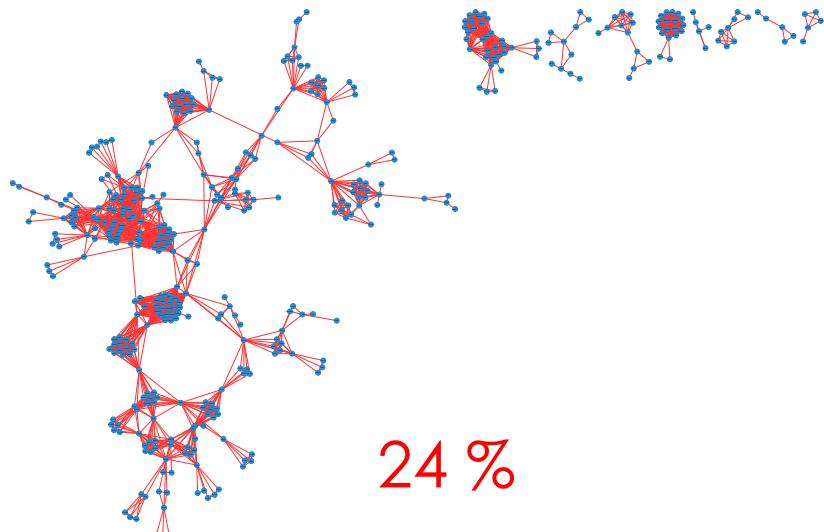
Connected nodes: 1492  
Unconnected nodes: 839  
Edges: 6380



# Network comparison HDGN vs ODGN

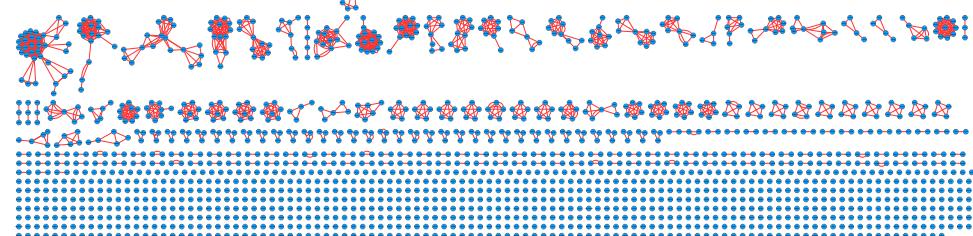
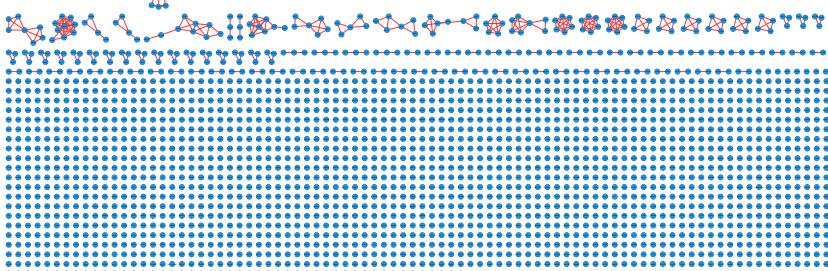
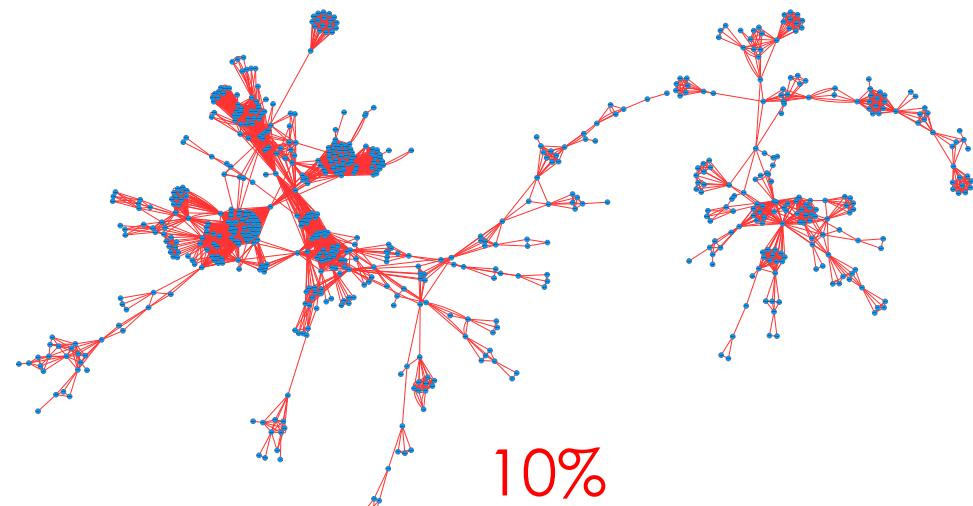
**A) Human Diseases Gene Network  
(gene-to-gene unipartite)**

Connected nodes: 749  
Unconnected nodes: 1776  
Edges: 2654

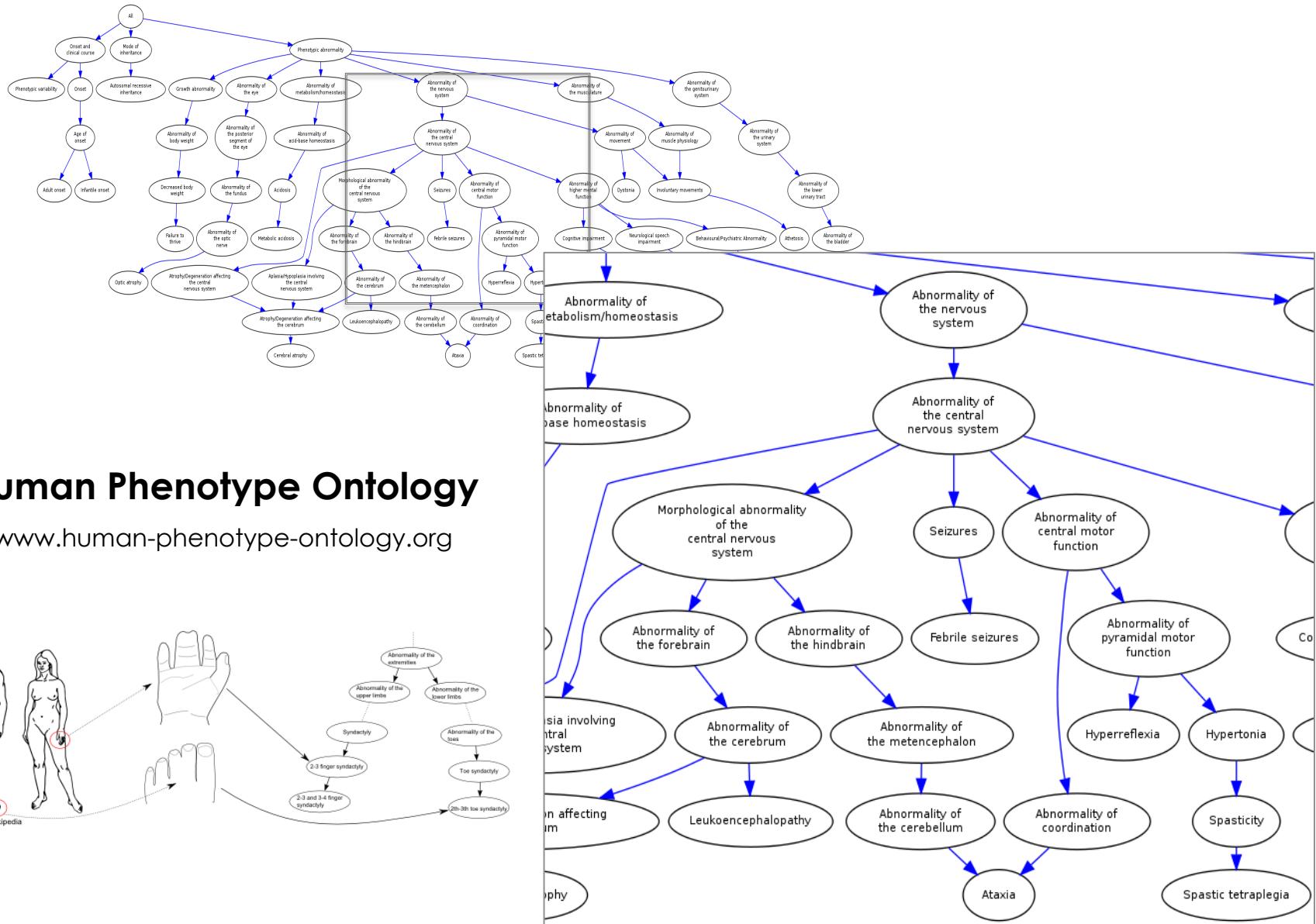


**B) Orphan Diseases Gene Network  
(gene-to-gene unipartite)**

Connected nodes: 1492  
Unconnected nodes: 839  
Edges: 6380

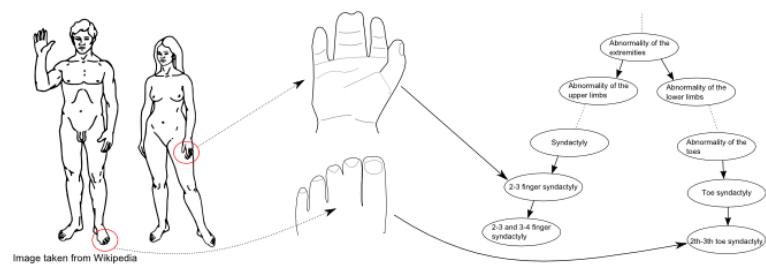


# Understanding diseases as sets of phenotypes



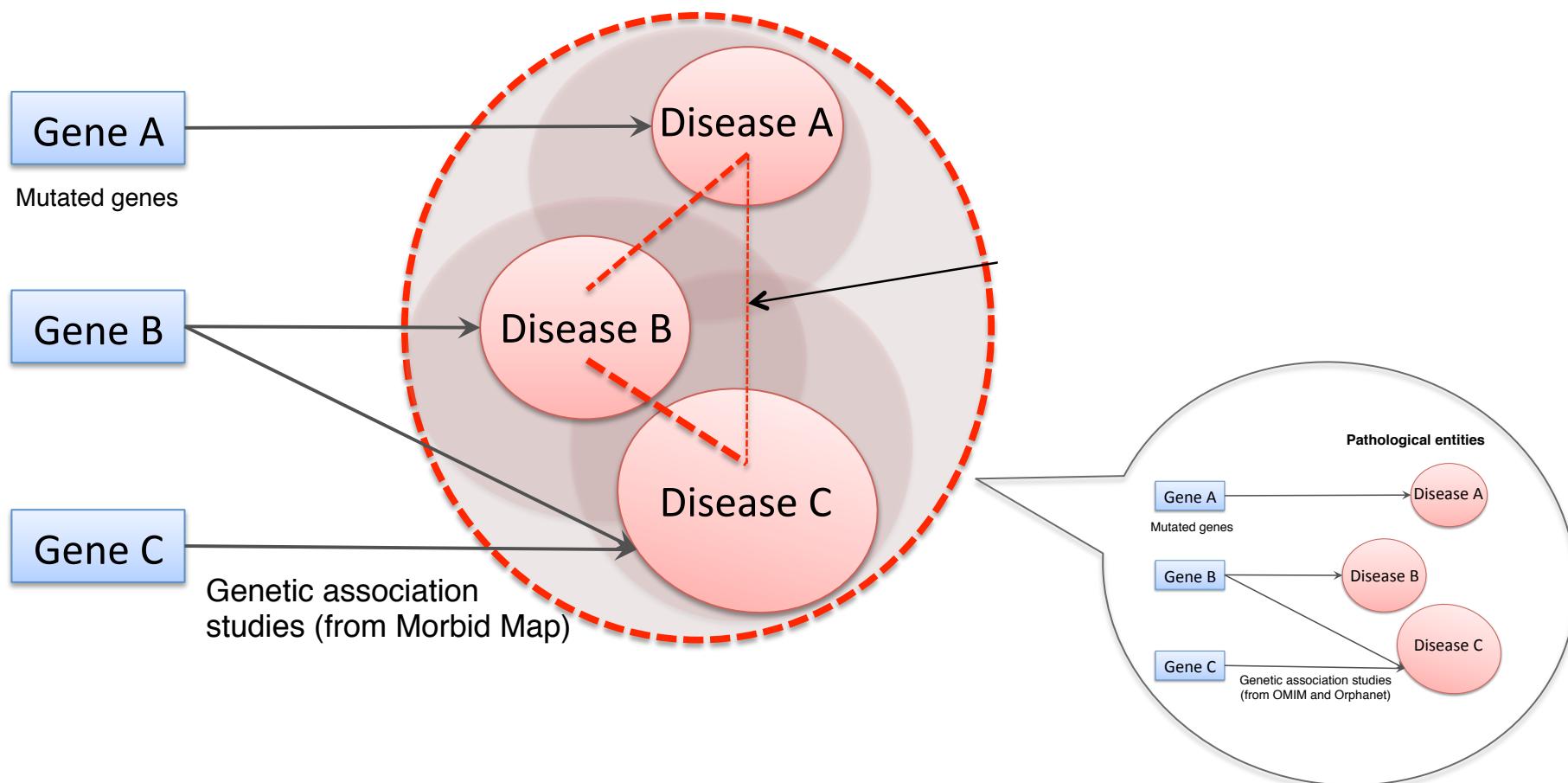
# Human Phenotype Ontology

[www.human-phenotype-ontology.org](http://www.human-phenotype-ontology.org)

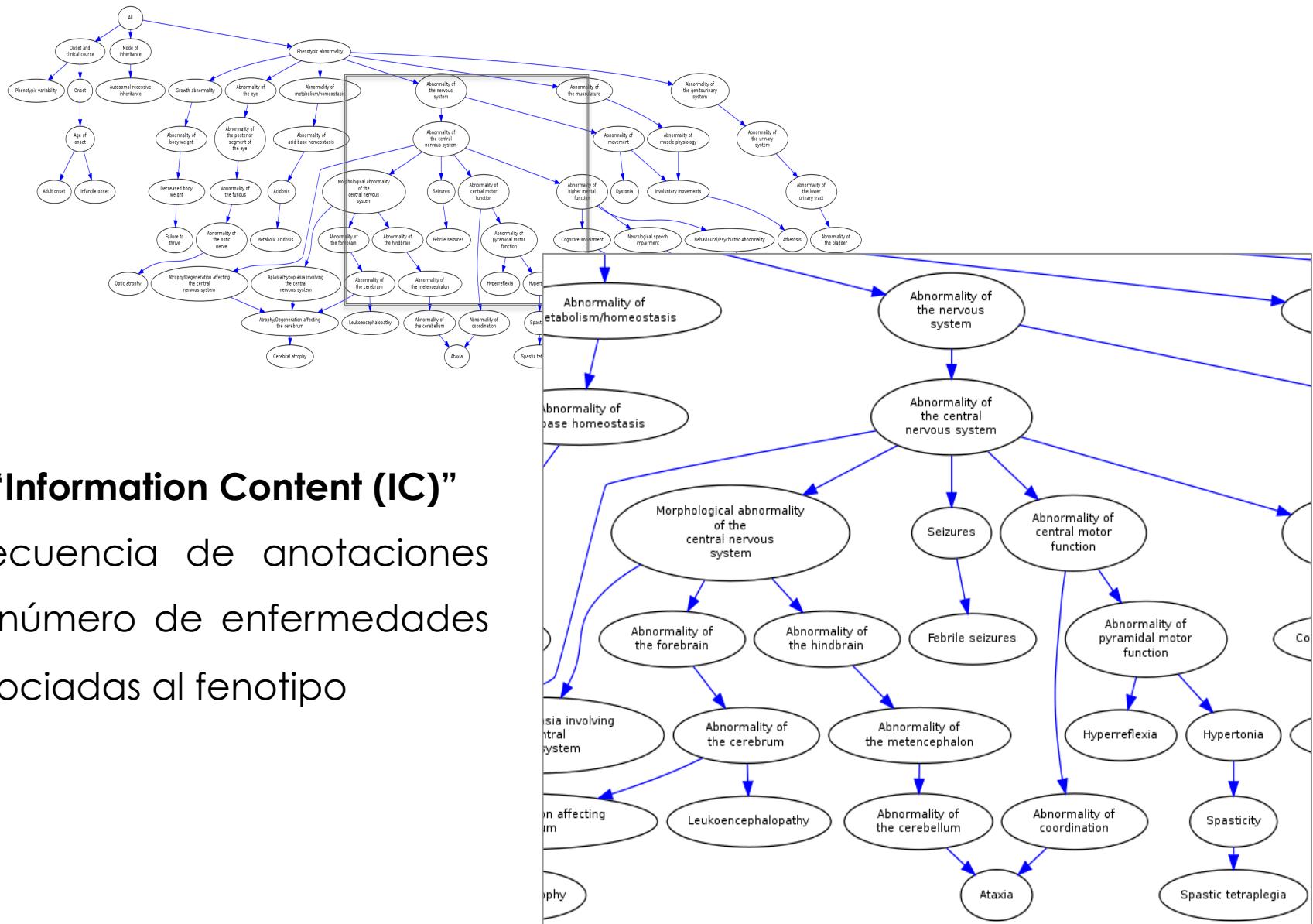


# Pathophenotypic space instead of diseases

## Pathophenotypic space of diseases



# Understanding diseases as sets of phenotypes

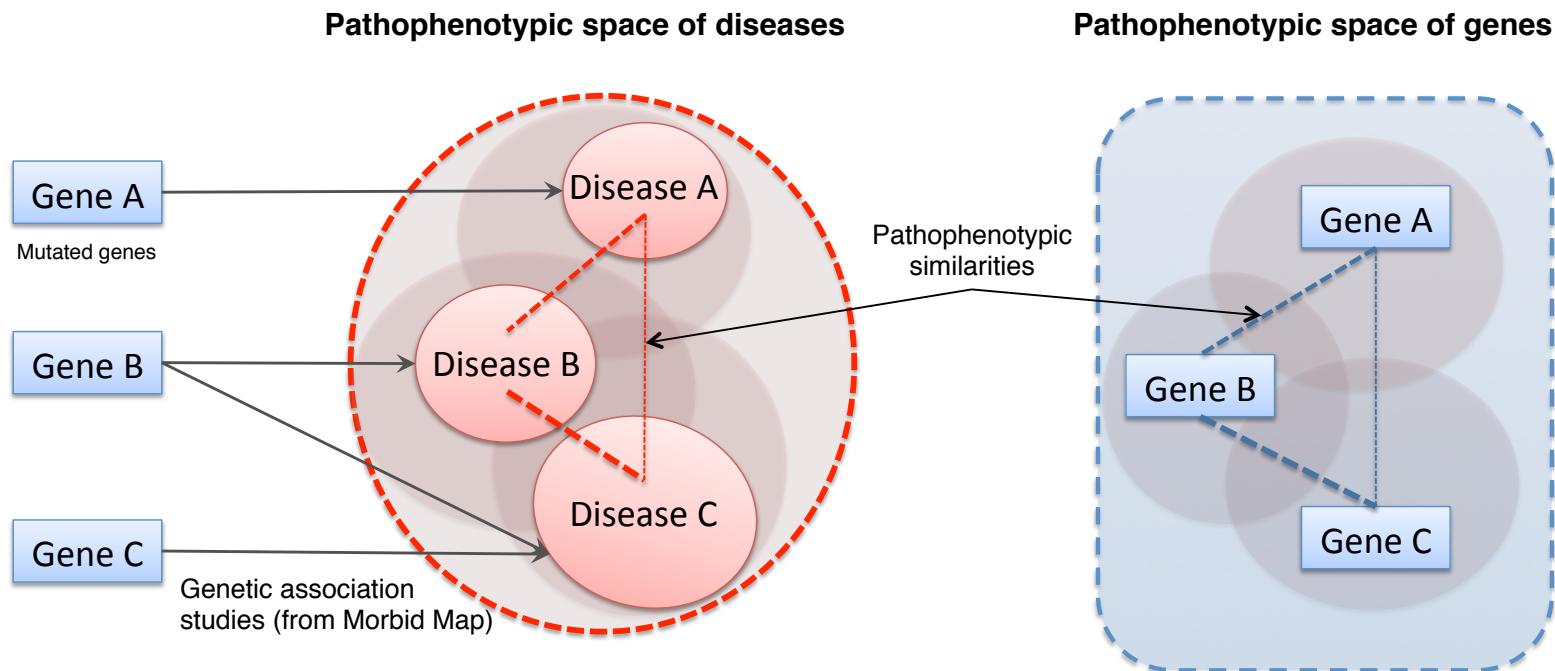


## **“Information Content (IC)”**

Frecuencia de anotaciones  
o número de enfermedades  
asociadas al fenotipo

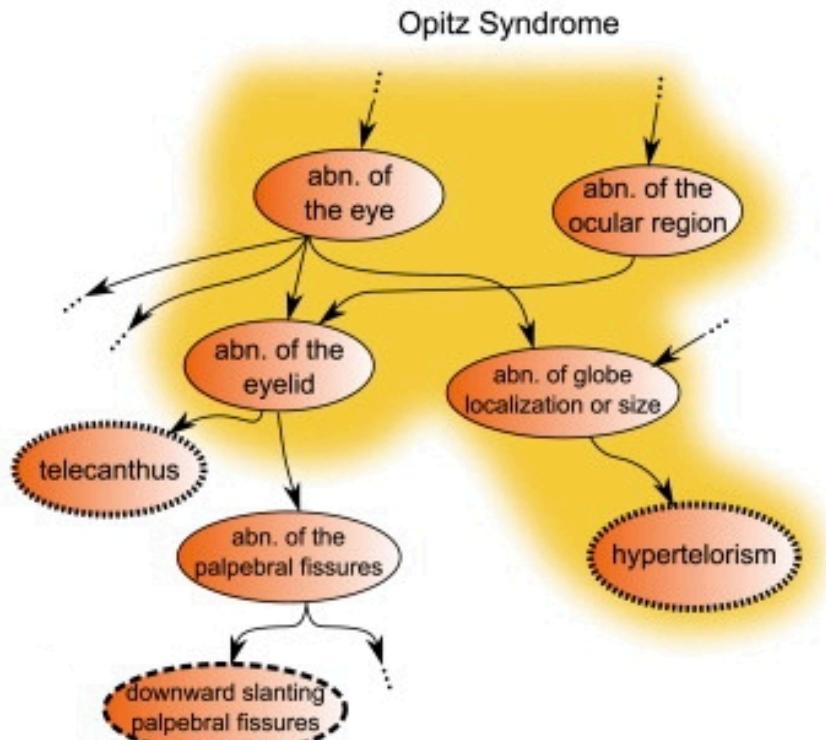
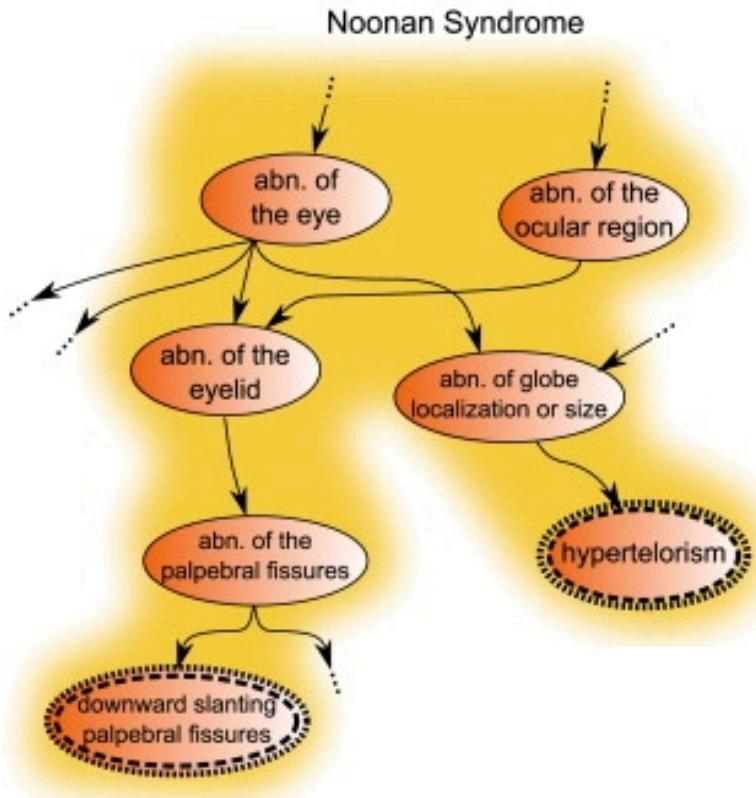
# Pathophenotypic space for genes

## PATHOPHENOTYPIC SPACES BASED ON HPO TERMS ANNOTATIONS



**Phenotypes show different Information Content (IC) for gene and disease annotations**

# Phenotypic similarity



**Renisk** uses the most specific phenotypes

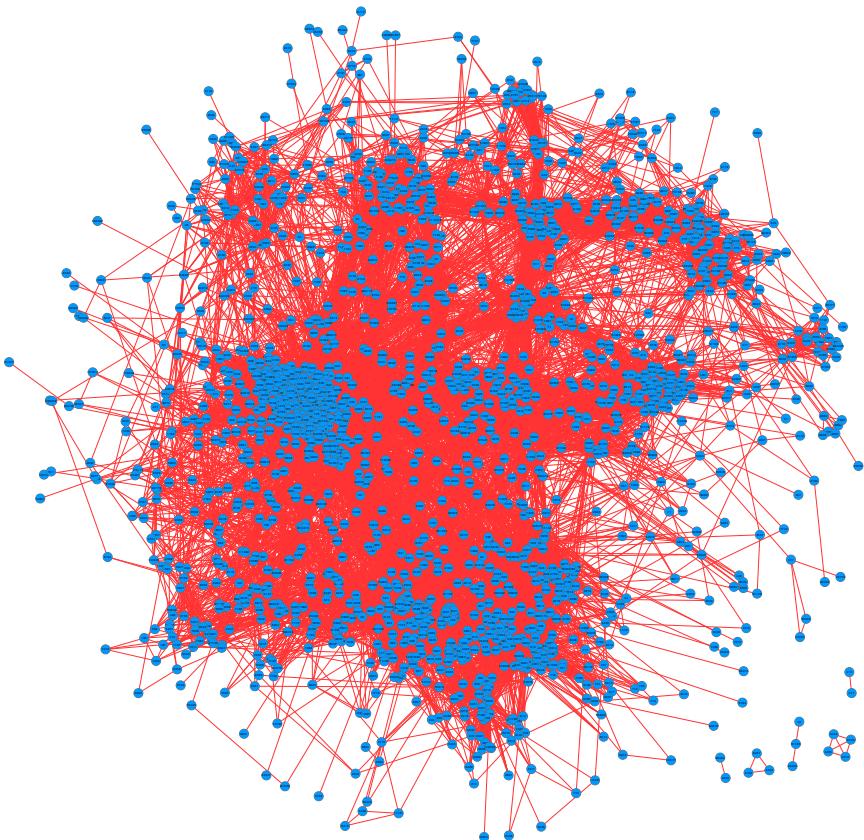
**Robinson** uses the media of all phenotypes

# Human Pathophenotypic Similarity Gene Network (PSGN)

Connected nodes: 1705

Unconnected nodes: 0

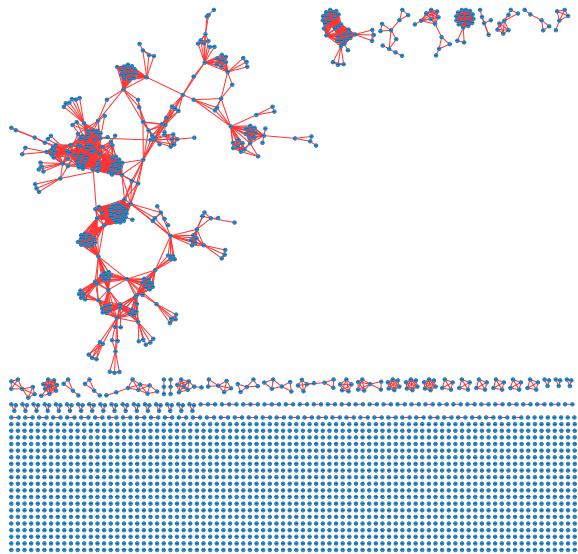
Edges: 26197



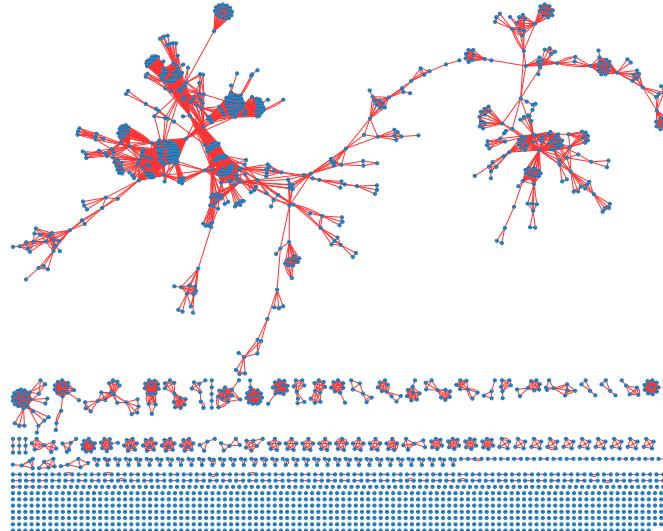
Expanding phenotypic relationships  
to decrease the specificity

# Network comparison HDGN, ODGN and PSGN

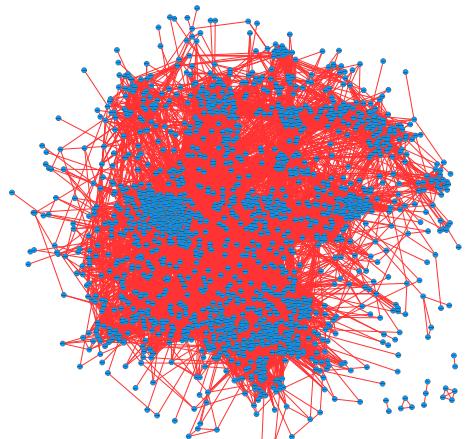
**A) Human Diseases Gene Network  
(gene-to-gene unipartite)**  
Connected nodes: 749  
Unconnected nodes: 1776  
Edges: 2654



**B) Orphan Diseases Gene Network  
(gene-to-gene unipartite)**  
Connected nodes: 1492  
Unconnected nodes: 839  
Edges: 6380



**C) Pathophenotypic Similarity Gene Net  
(gene-to-gene semantic similarity)**  
Connected nodes: 1705  
Unconnected nodes: 0  
Edges: 26197

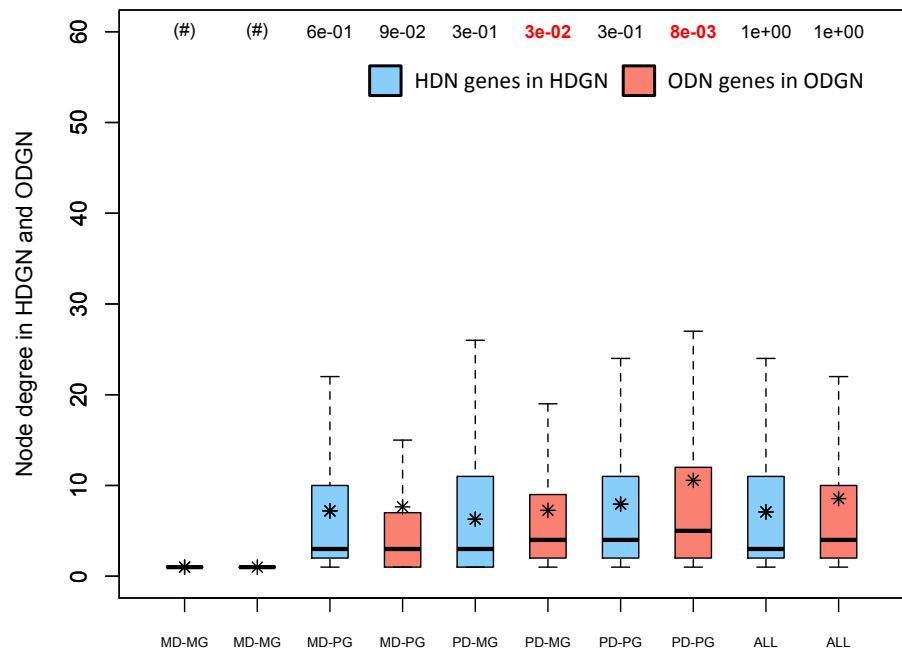


Genes taking part in bi-univocal relationships (MD-MG) are unconnected in unipartite projections of diseasesomes

# Degree of genes in HDGN, ODGN and PSGN

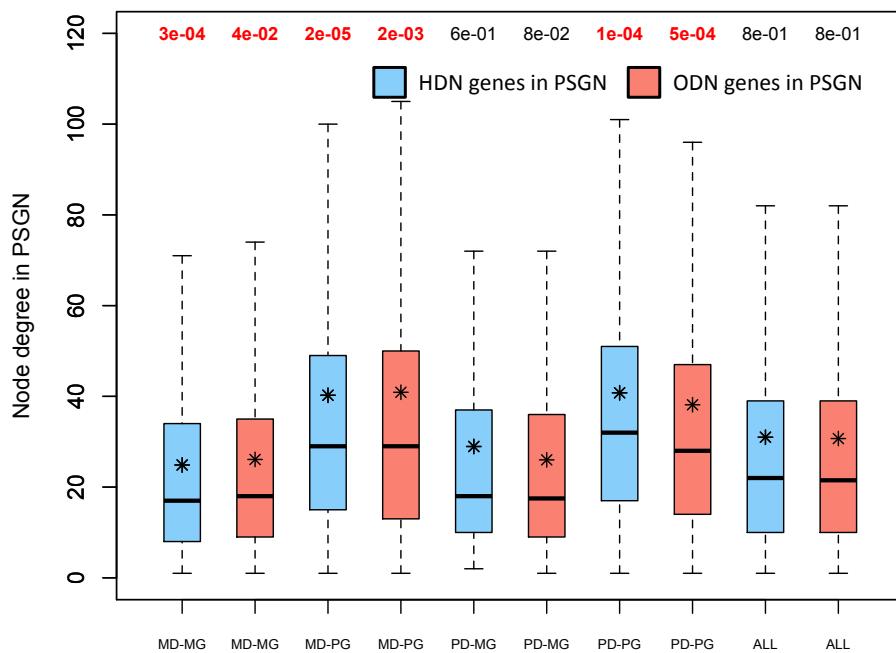
A

**Degree distribution of gene subsets in HDGN or ODGN**



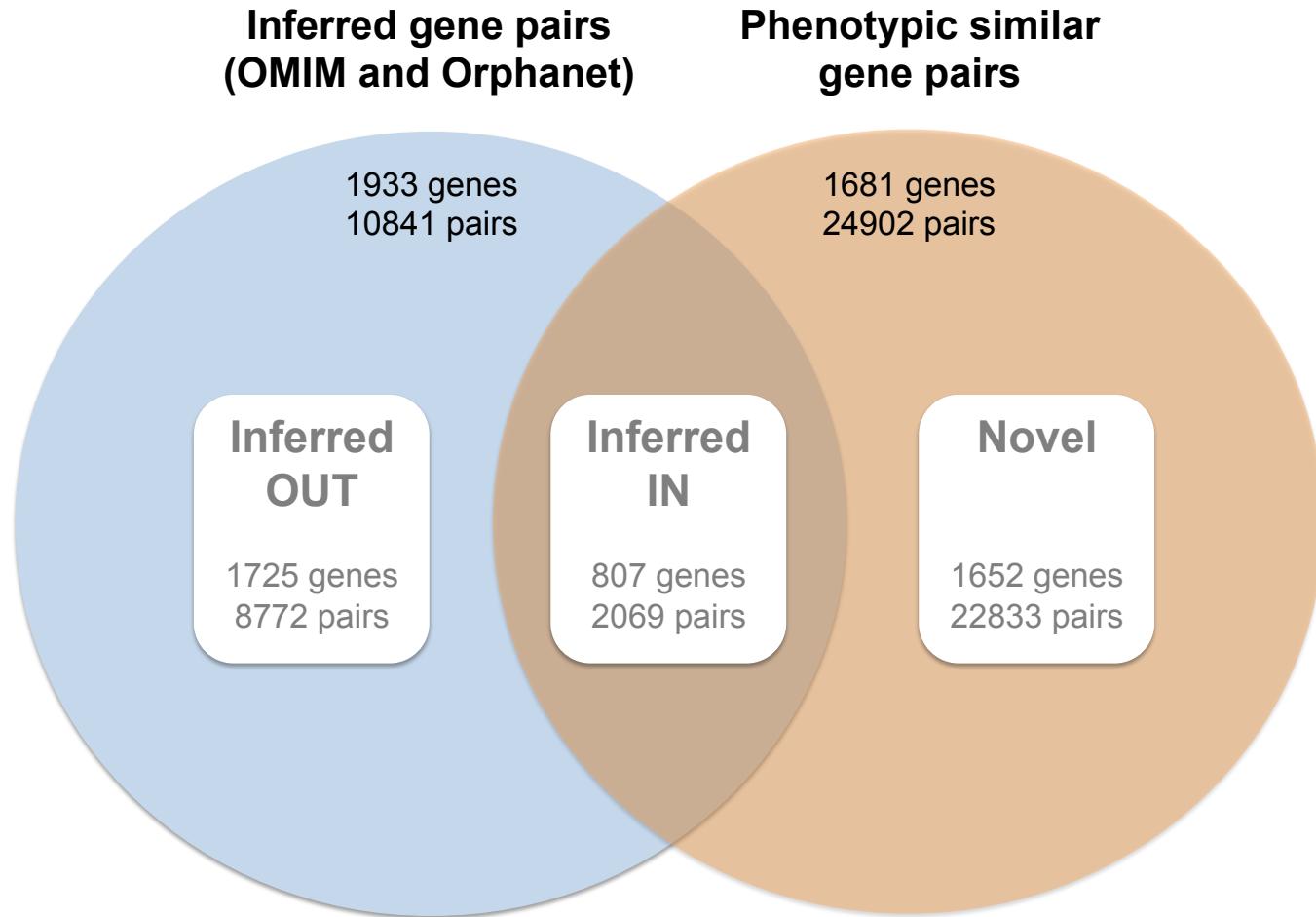
B

**Degree distribution of gene subsets in PSGN**



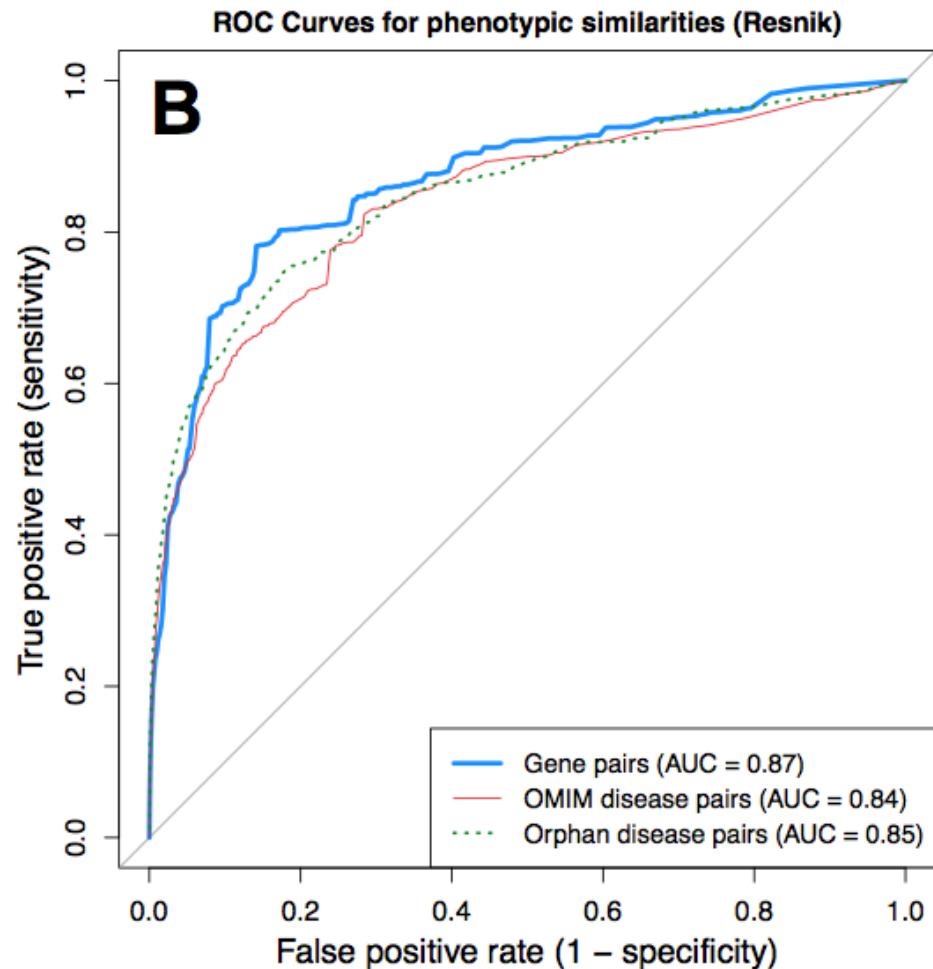
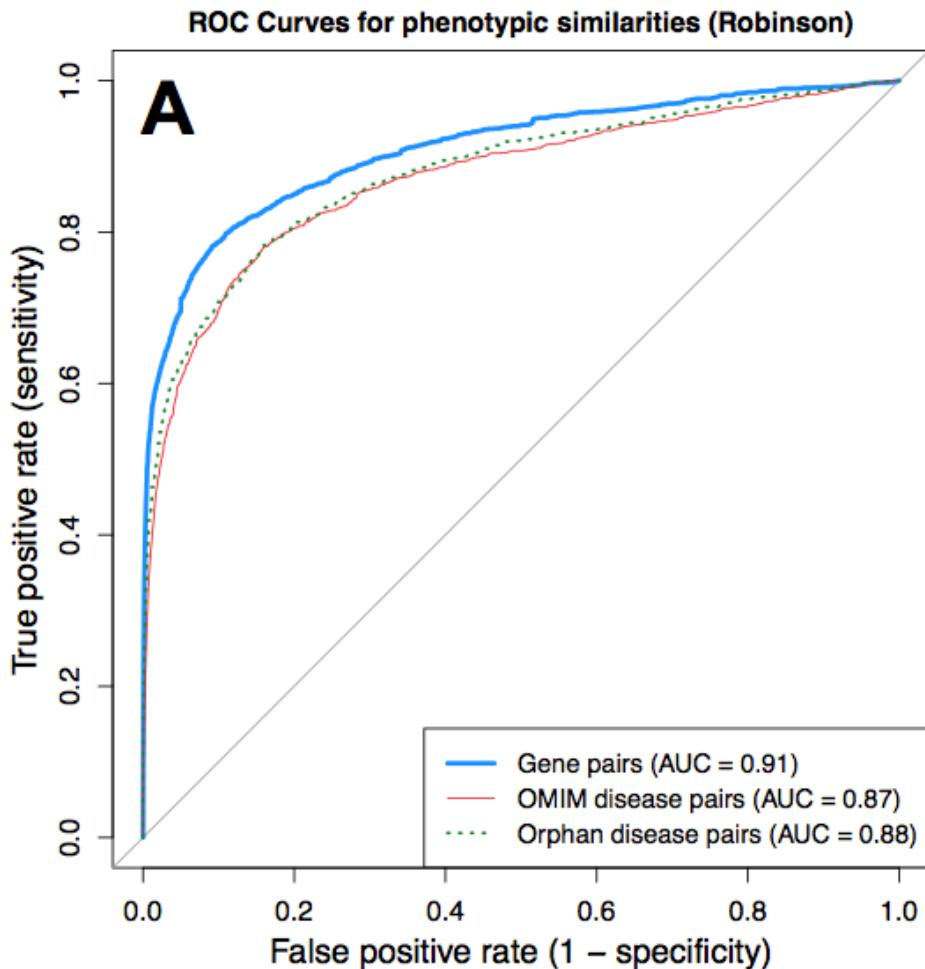
- Bi-univocal relationships (one gene one disease) an average degree of 25
- Uniform connectivity for subset of disease causing genes in PSGN respect to HDGN and ODGN

# Venn diagram $[(HDGN \cup ODGN) \cap PSGN]$



All disease causing genes at least one phentypic relationship

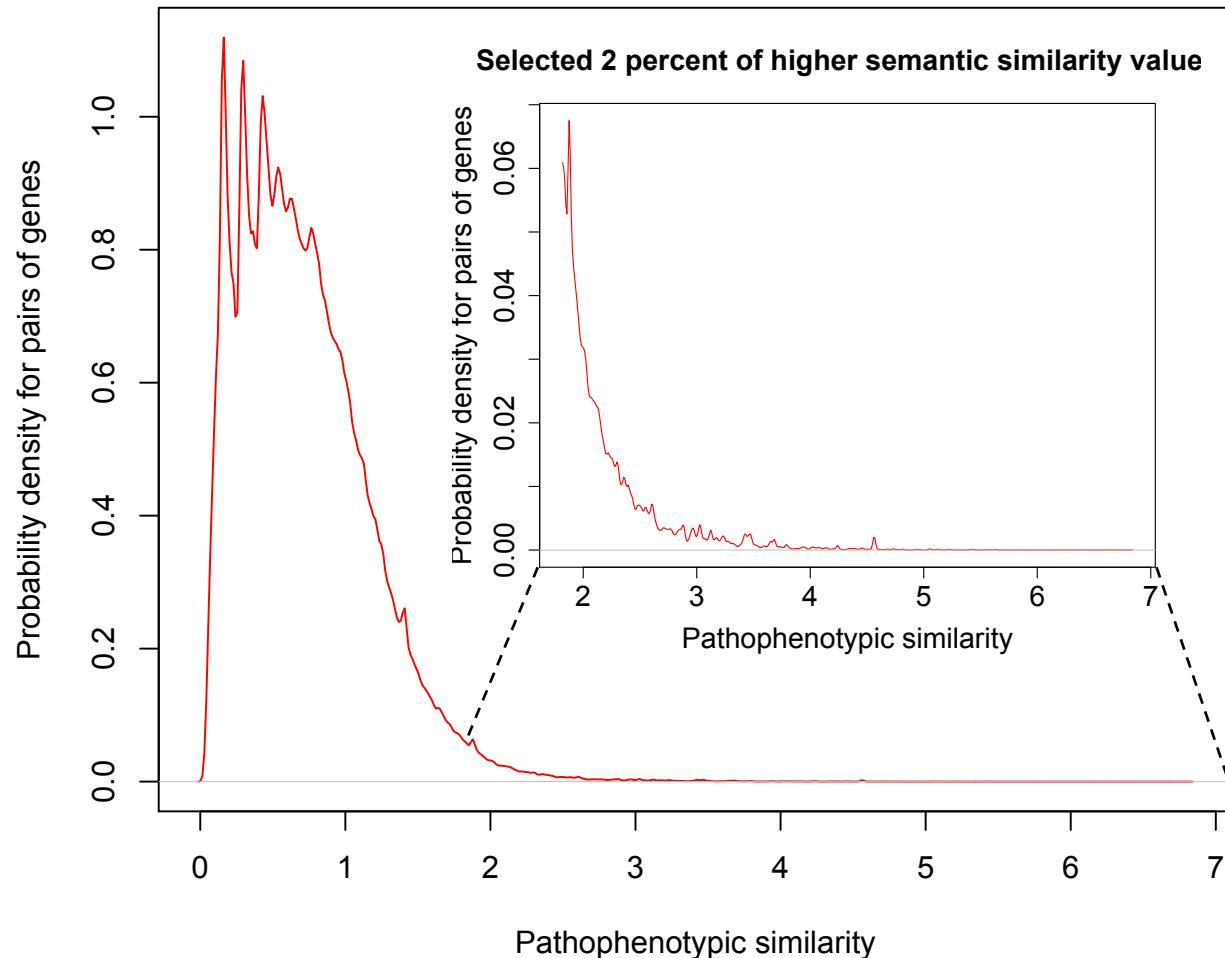
# Performance validation (Phenotypic similarity vs. inference)



Robinson approach is a good measurement for phenotypic similarity

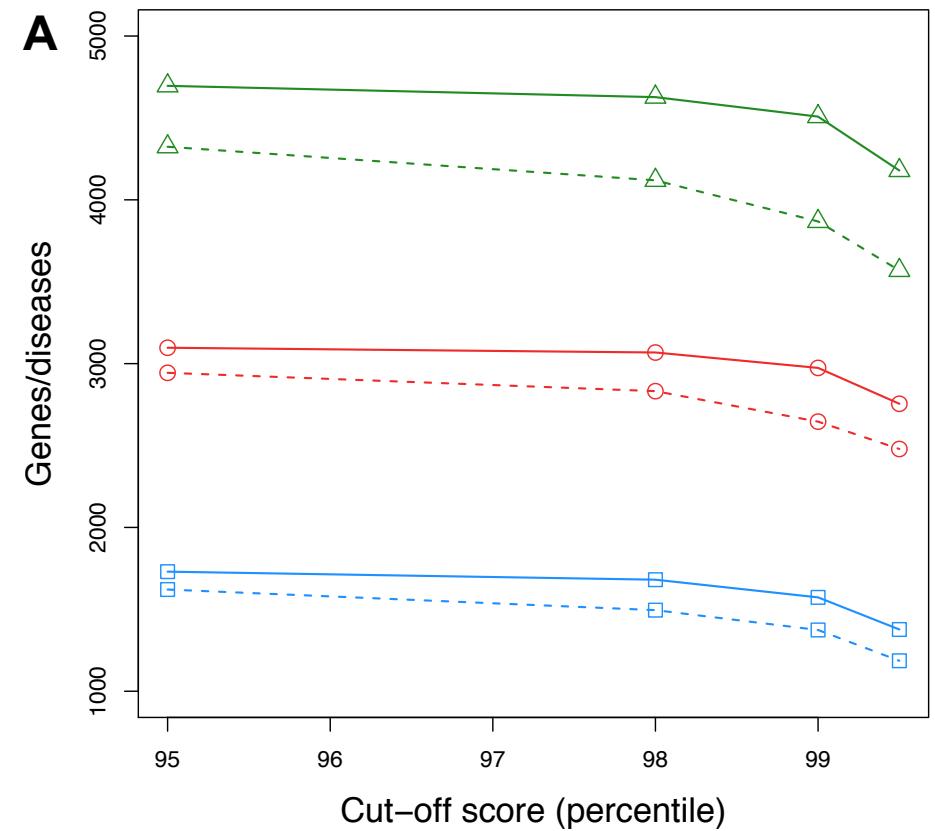
# Density function probability of phenotypic similarities

All semantic similarity values calculated in HPO

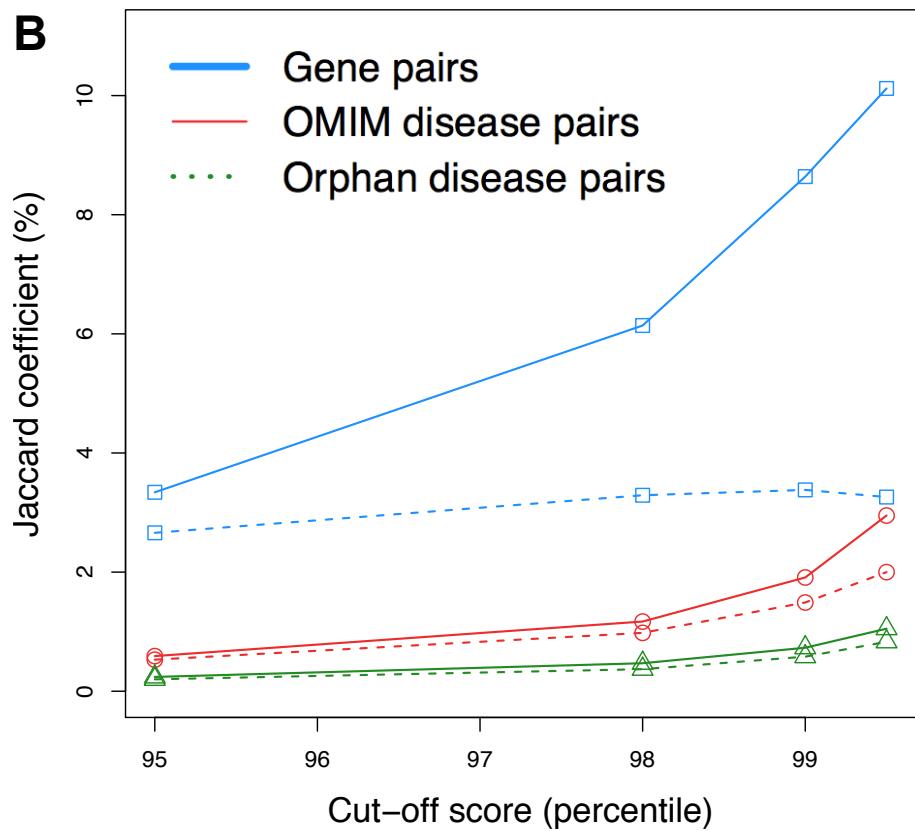


Robinson approach is a good measurement for phenotypic similarity

# Optimal Statistical Cut-Off (pragmatic)

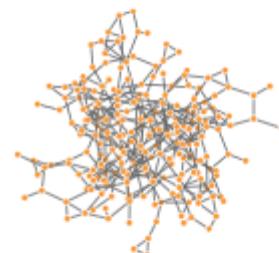
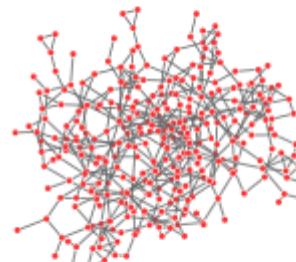
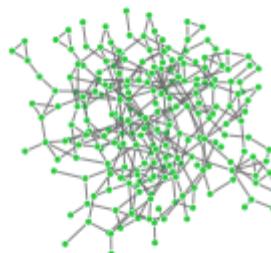
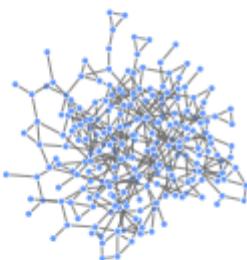
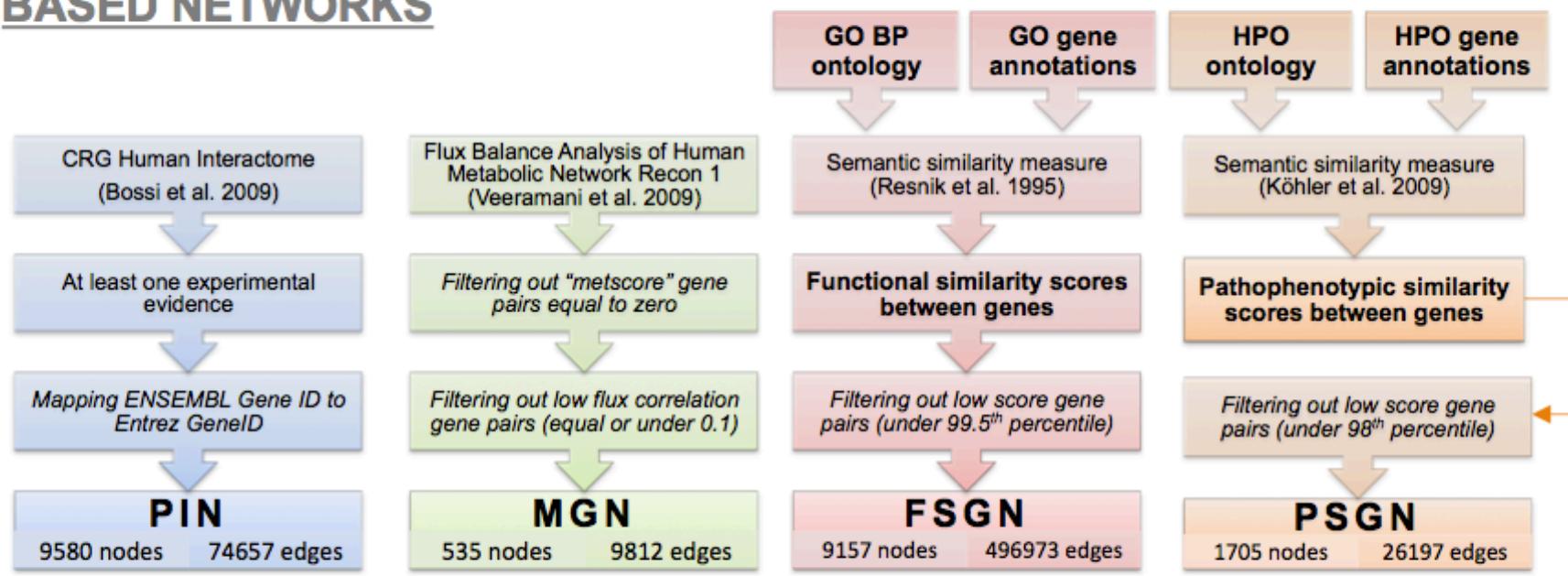


From 98<sup>th</sup> we loss information



From 98<sup>th</sup> we increase similarity with inferred

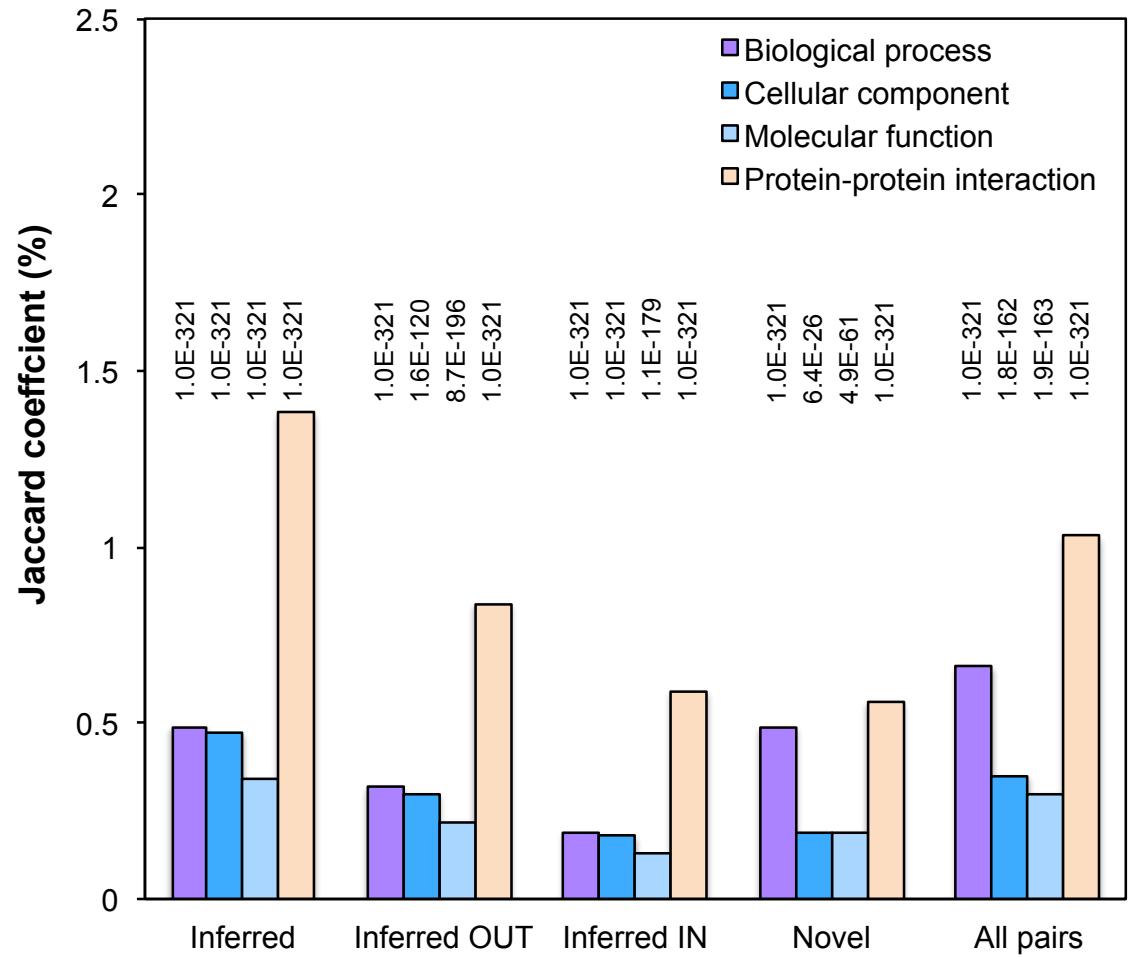
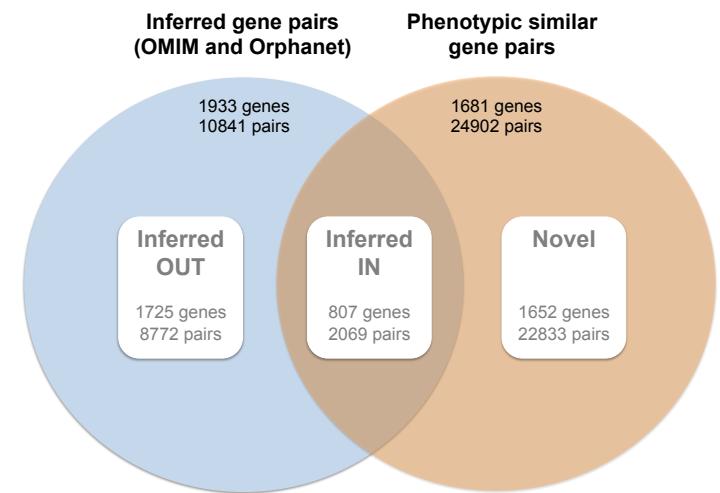
# 1. CONSTRUCTION OF INTERACTOMES AND SEMANTIC SIMILARITY BASED NETWORKS



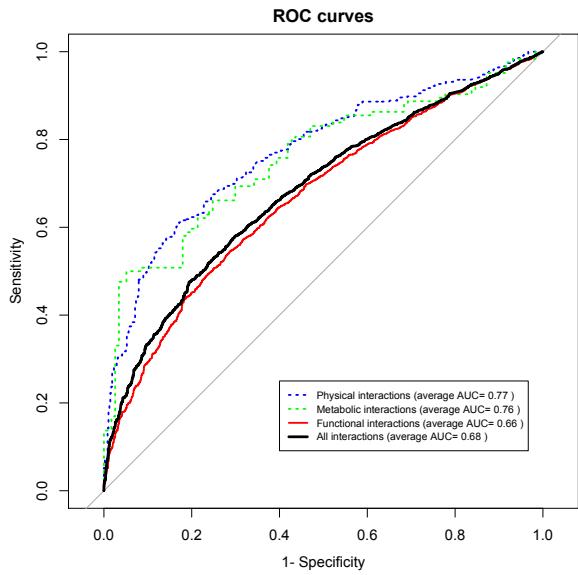
**BIOMOLECULAR INTERACTOMES**

**PSGN**

# Network Comparisons Between Interactomes and PSGN Subsets

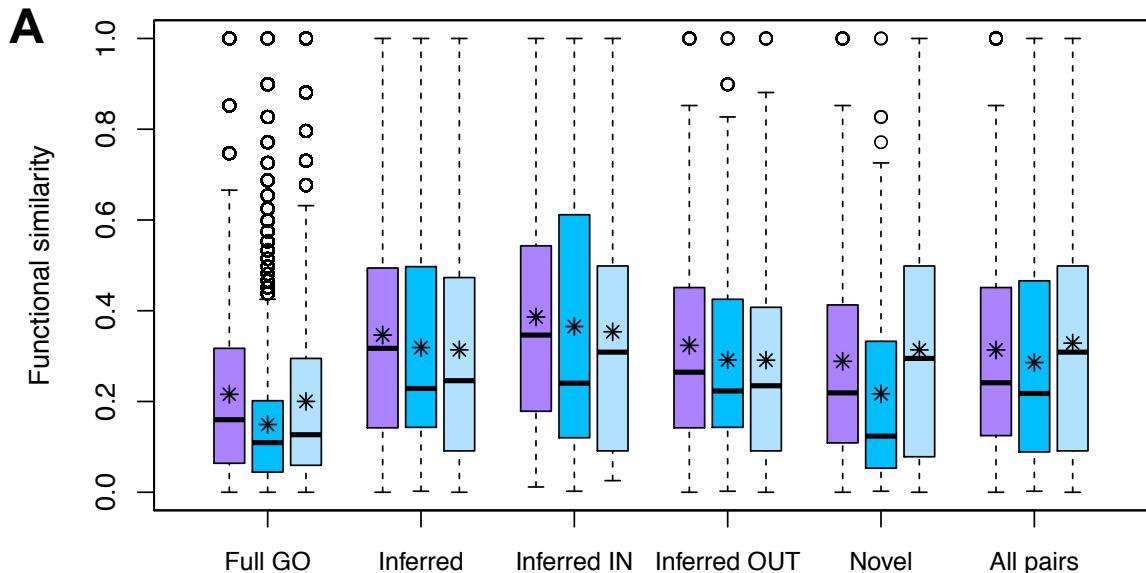
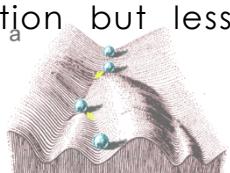


# Intersection between Phenotypic and Functional interactions has higher similarity scores

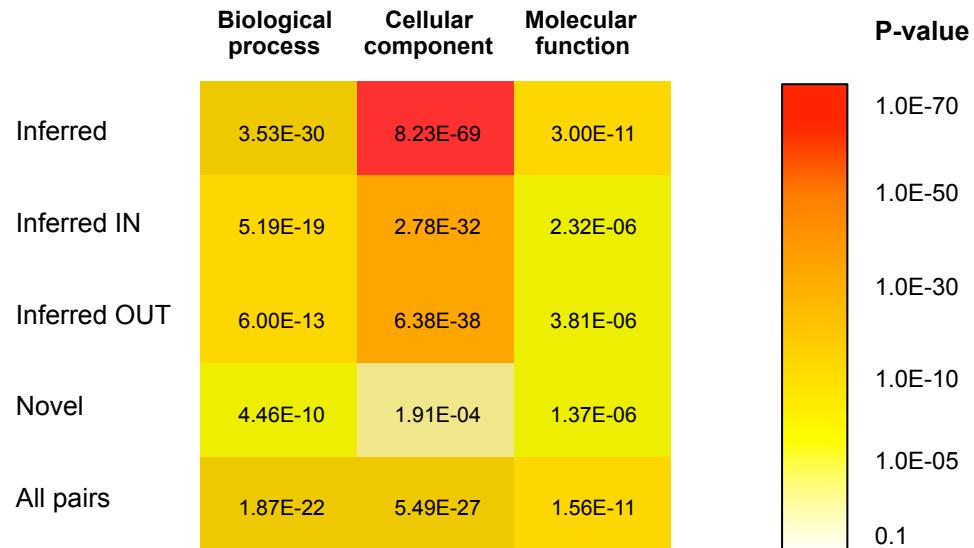


**Natural (biological) networks evolved to be robust** to genetic, environmental and stochastic perturbations.

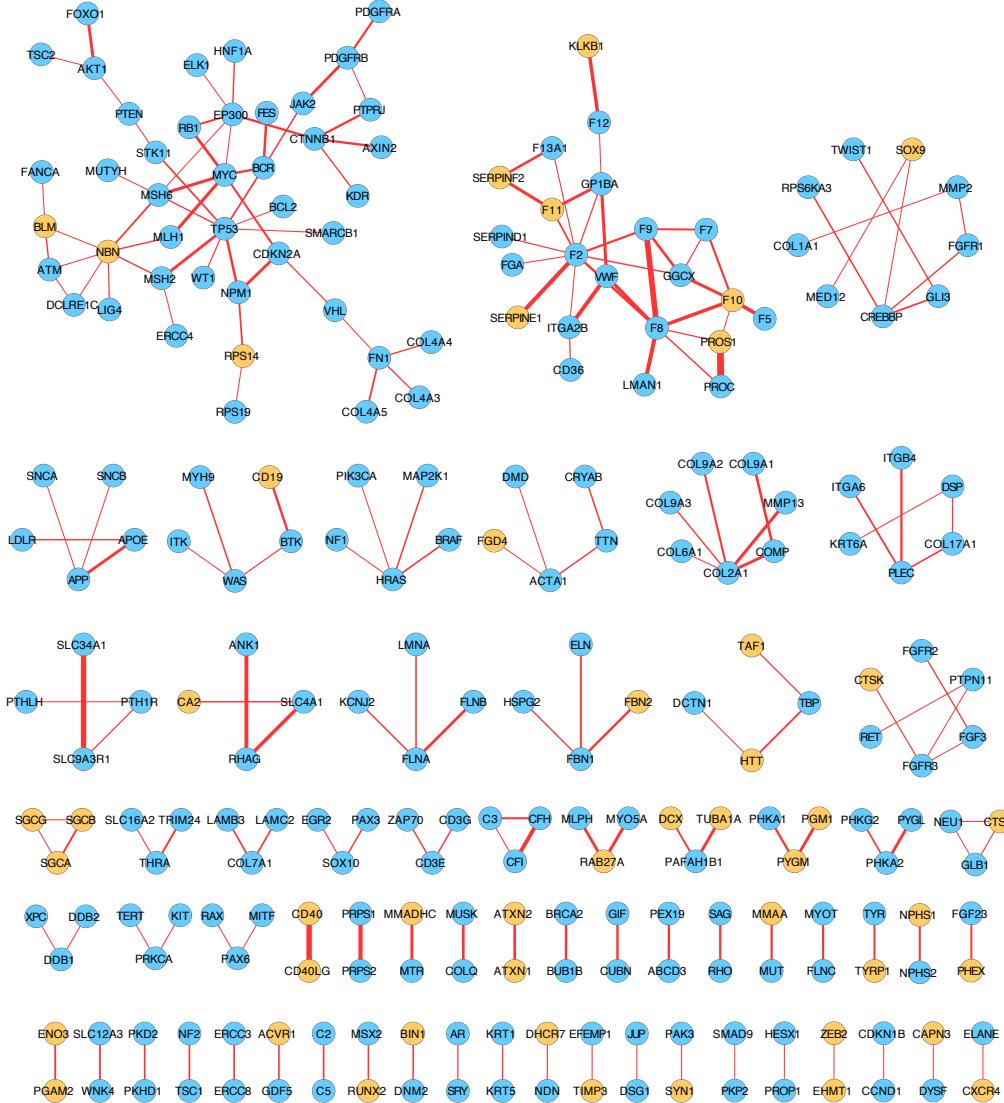
GOOD level of abstraction but less precision



**B Heat map of Mann-Whitney P-values from the distribution comparisons**



# Novel phenotypic similarities ∩ protein interactions



# Conclusion I

Inherited disease database information can be migrated to networks to manage and analyzed this information as a whole. **Network Medicine.**

Advanced Search: OMIM, Clinical Synopses, OMIM Gene Map    Sort by: #271980    Toggle: search terms highlighted  
Search History: View, Clear

**#271980**  
**SUCCINIC SEMIALDEHYDE DEHYDROGENASE DEFICIENCY; SSADHD**

Alternative titles/symbols:  
SSADH DEFICIENCY  
4-HYDROXYBUTYRIC ACIDURIA  
GAMMA-HYDROXYBUTYRIC ACIDURIA  
GAMMA-HYDROXYBUTYRIC ACIDURIA

Phenotype Gene Relationships

Location	Phenotype	Phenotype MIM number	Gene/Locus	Gene/Locus MIM number
6p22.3	Succinic semialdehyde dehydrogenase deficiency	271980	ALDHBA1	610003

Clinical Synopsis

TEXT

A number sign (#) is used with this entry because succinic semialdehyde dehydrogenase (SSADH) deficiency can be caused by hemizygous mutation in the ALDHBA1 gene (610003) on chromosome 6p22.

Description

Succinic semialdehyde dehydrogenase deficiency (SSADHD) is a rare autosomal recessive neurologic disorder in which an enzyme defect in the GABA degradation pathway causes a consecutive elevation of gamma-hydroxybutyric acid (GHB) and GABA. The clinical features include developmental delay, hypotonia, mental retardation, ataxia, seizures, hyperactive behavior, aggressiveness, and sleep disturbances (summary by Jaksic et al., 2012).

Clinical Features

Jaksic et al. (1993) reported a patient with neurologic abnormalities and urinary excretion of gamma-hydroxybutyric acid.

**4-hydroxybutyric aciduria**

Orpha number : ORPHA22	ICD-10 : E72.8
Synonym(s) : Succinic semialdehyde dehydrogenase deficiency	OMIM : 271980 [?]
Prevalence : <1 / 1 000 000	UMLS : -
Inheritance : Autosomal recessive	MeSH : -
Age of onset : Childhood	MedDRA : -
	SNOMED CT : -

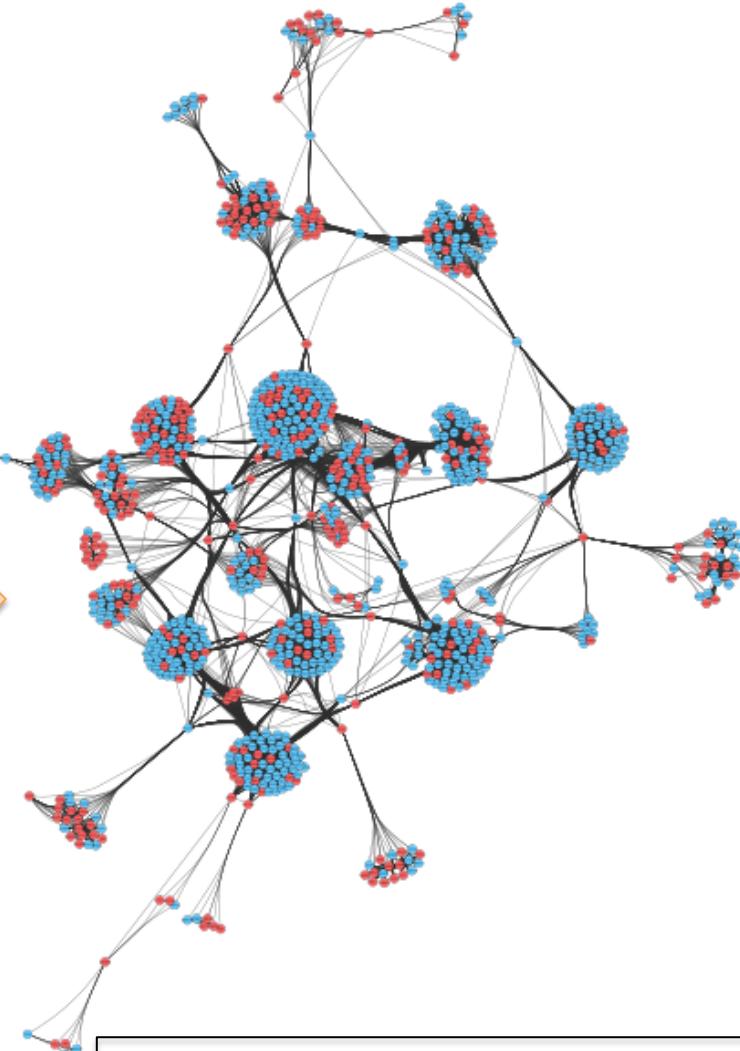
**SUMMARY**

The 4-hydroxybutyricaciduria deficiency is a metabolic disorder with a neurological presentation ranging from mild to severe. It is a rare disease with around 350 cases reported. The most frequent symptom is progressive retardation, often associated with developmental delay and ataxia. Treatment is autosomal recessive and mutations in the SSADH (Succinic Semialdehyde Dehydrogenase NAD(+)-Dependent) gene, located on chromosome 6p22, have been reported. The key biochemical feature is an accumulation of gamma-hydroxybutyrate in urine, plasma and cerebro-spinal fluid. There is no efficient treatment available.

Expert reviewer(s)

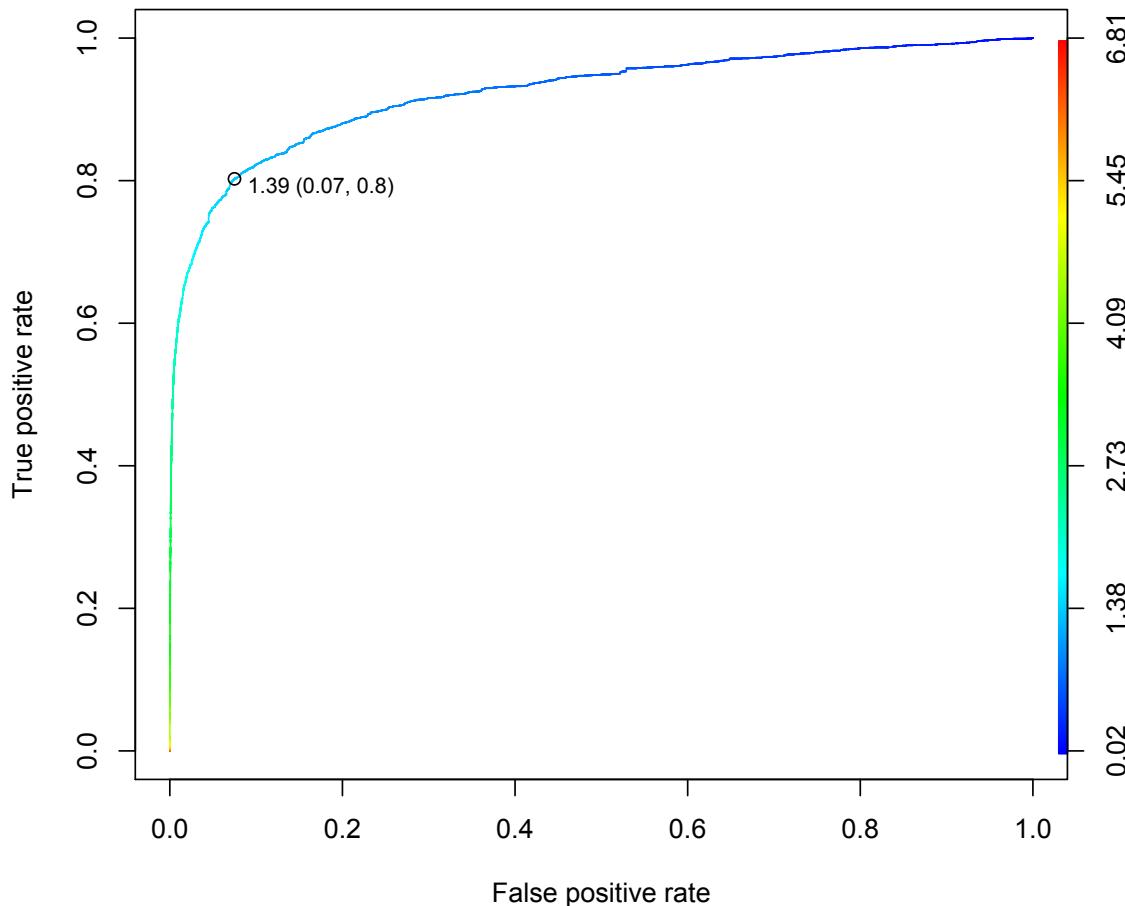
Pr Jaak JAEKEN

Last update: July 2006



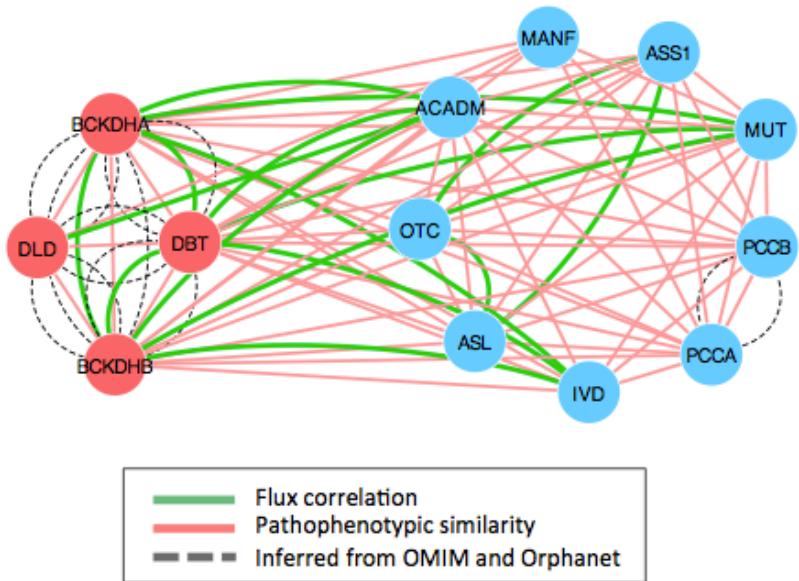
# Conclusion II

ROC curve of similarity scores

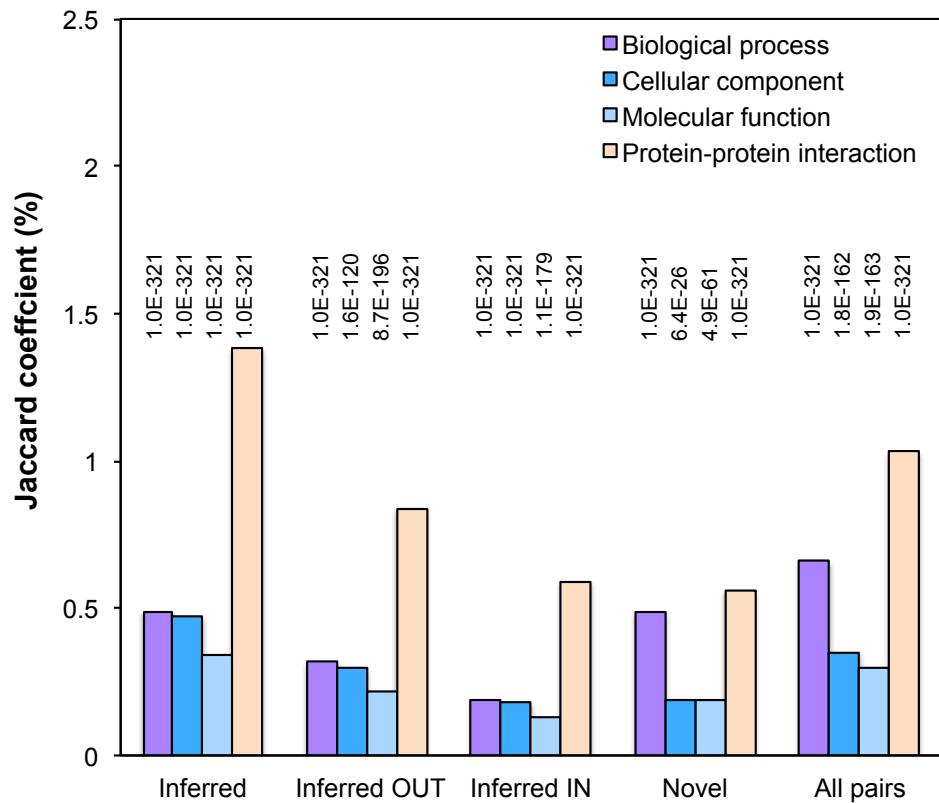


Semantic similarity represent a pragmatic approach to study and model phenotypic relationships

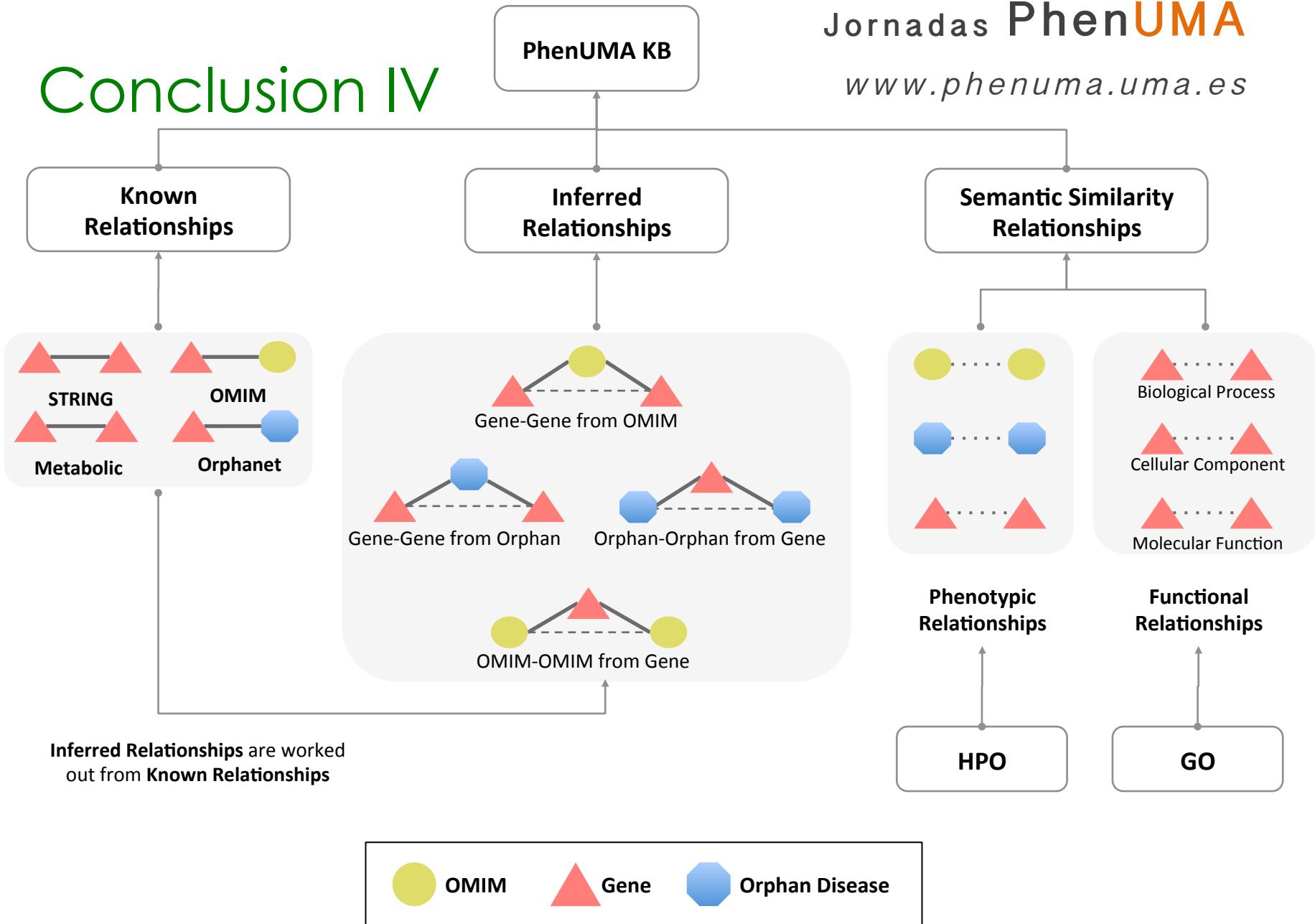
# Conclusion III



Biological Coherence



## Conclusion IV



# ProCel team

- METABOLIC SYSTEMS, Group leader. Dr. F. Sánchez-Jiménez.
- Thesis Director. Dr. Miguel Ángel Medina Torres.
- ProFunc: PROTEIN FUNCTION GROUP, PhD. JA Ranea

**¡¡Muchas gracias y bienvenidos!!**



Kika



Miguel Angel



Rocio Rodríguez-López