Sesión Práctica I: Introducción a

14 Junio de 2013, E.T.S.I Informática, Málaga

UNIVERSIDAD DE MÁLAGA

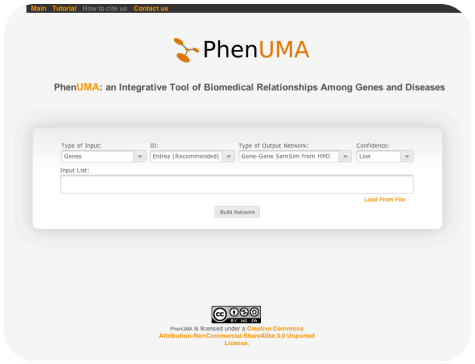
ciberer
CENTRO DE INVESTIGACIÓN BIOMÉDICA EN RED DE ENFERMEDADES RARAS

Sesión Práctica: Esquema

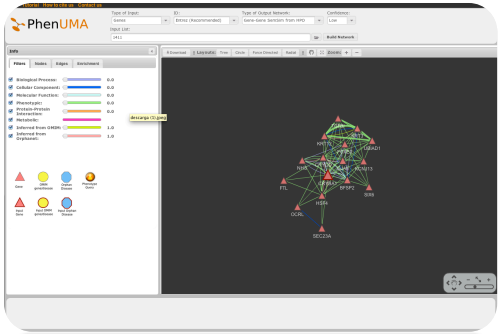
- **Sesión Práctica I**:
  - Entradas de Datos
    1. Opciones de Entrada
    2. Datos de Entrada
    3. Opciones de la Red de Salida
       - Knowledge Base
       - Tipos de Redes
       - Niveles de Confianza
  - Construcción de Redes
  - Funcionalidades de la Aplicación
- **Sesión Práctica II**:
  - Gestión de los Datos de Salida
  - Ejemplos y casos de uso
- **Sesión Práctica III**:
  - Ejemplos y casos de uso

Interfaz de PhenUMA

- Página Principal

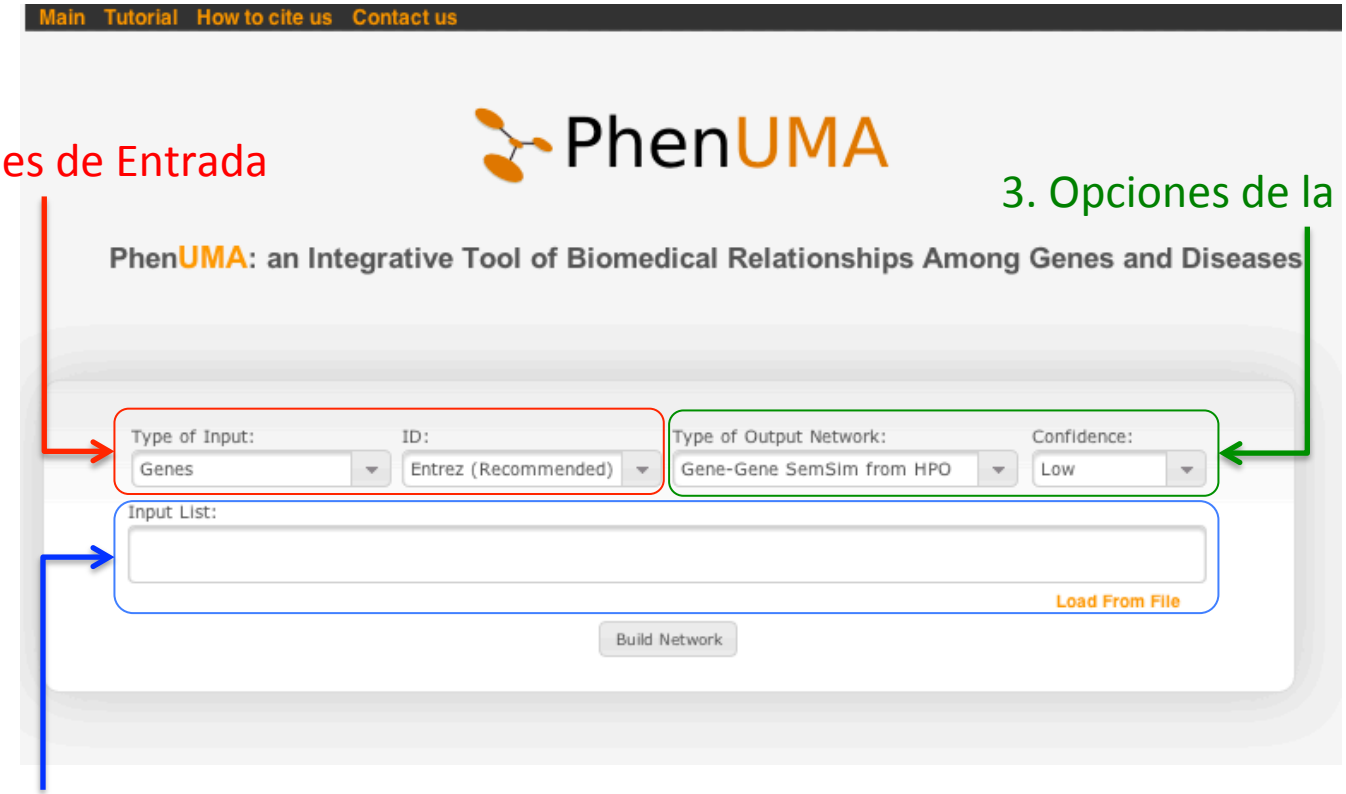


- Página de Visualización



Entrada a PhenUMA

1. Opciones de Entrada



3. Opciones de la Red de Salida

2. Datos de Entrada

1. Opciones de Entrada

- Tipo de Entrada:
  - Genes
  - OMIM gene/diseases
  - Orphan Diseases
  - Phenotypes

Identificadores de Genes

- Para cada tipo de entrada debe usarse el identificador adecuado.
- Identificadores permitidos para genes:
  - Código Entrez (Recomendado)
  - Official Symbol
  - HGNC
  - MIM
  - Ensembl
  - Orphanum

- No es posible usar:
  - Full Name: crystallin, beta A1
  - Sinónimos: CRYB1, CTRCT10
  - Proteína: Beta-crystallin A3

Identificadores de Genes

¿Cómo obtener los identificadores?

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

Ensembl
http://www.ensembl.org/

HGNC
http://www.genenames.org/

MIM
http://www.omim.org/

Orphanum
http://www.orpha.net/

Indetificadores de Genes: NCBI

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: NCBI

Entrez/Gene ID

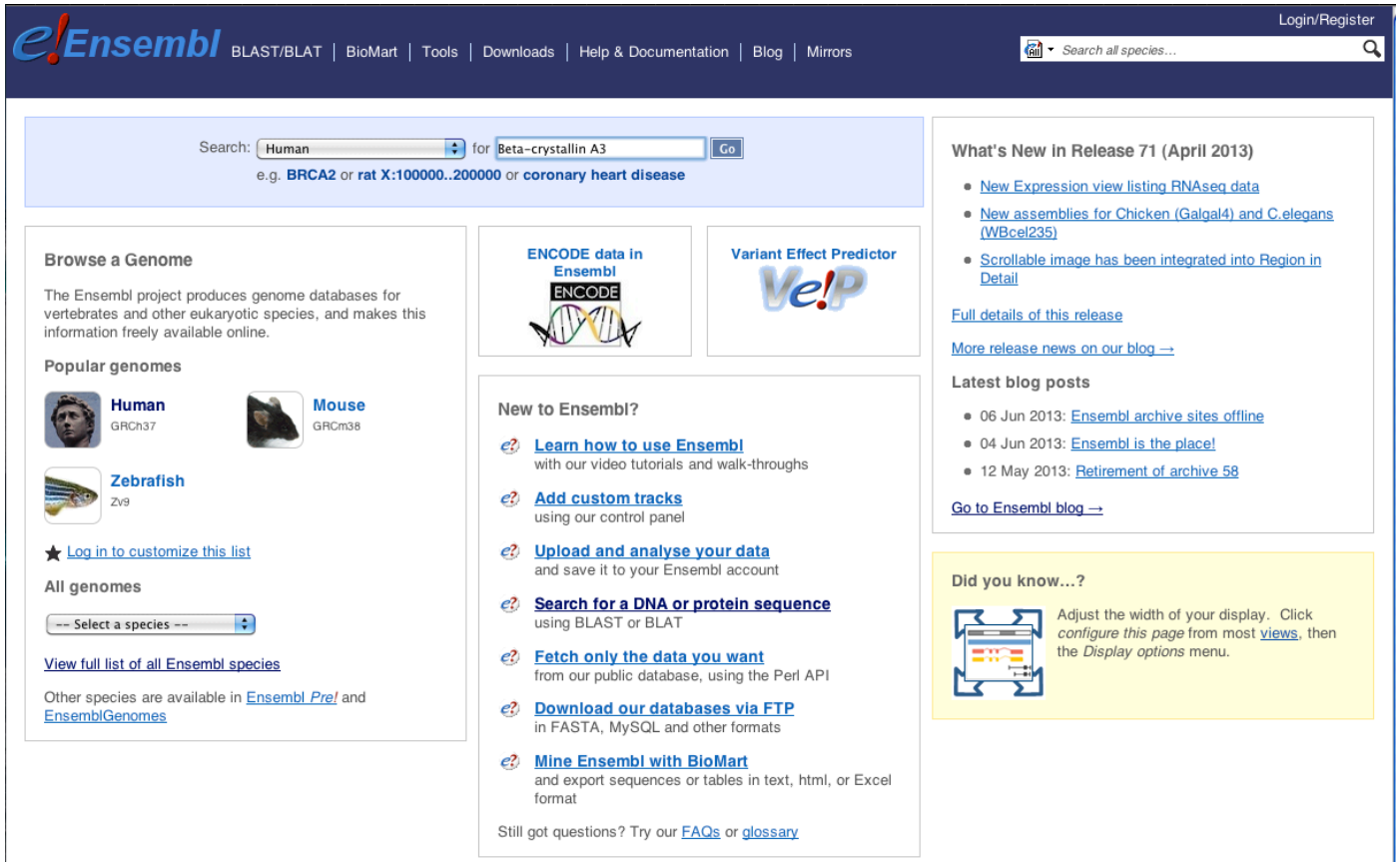
Official Symbol

Más Identificadores

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: Ensembl

http://www.ensembl.org/



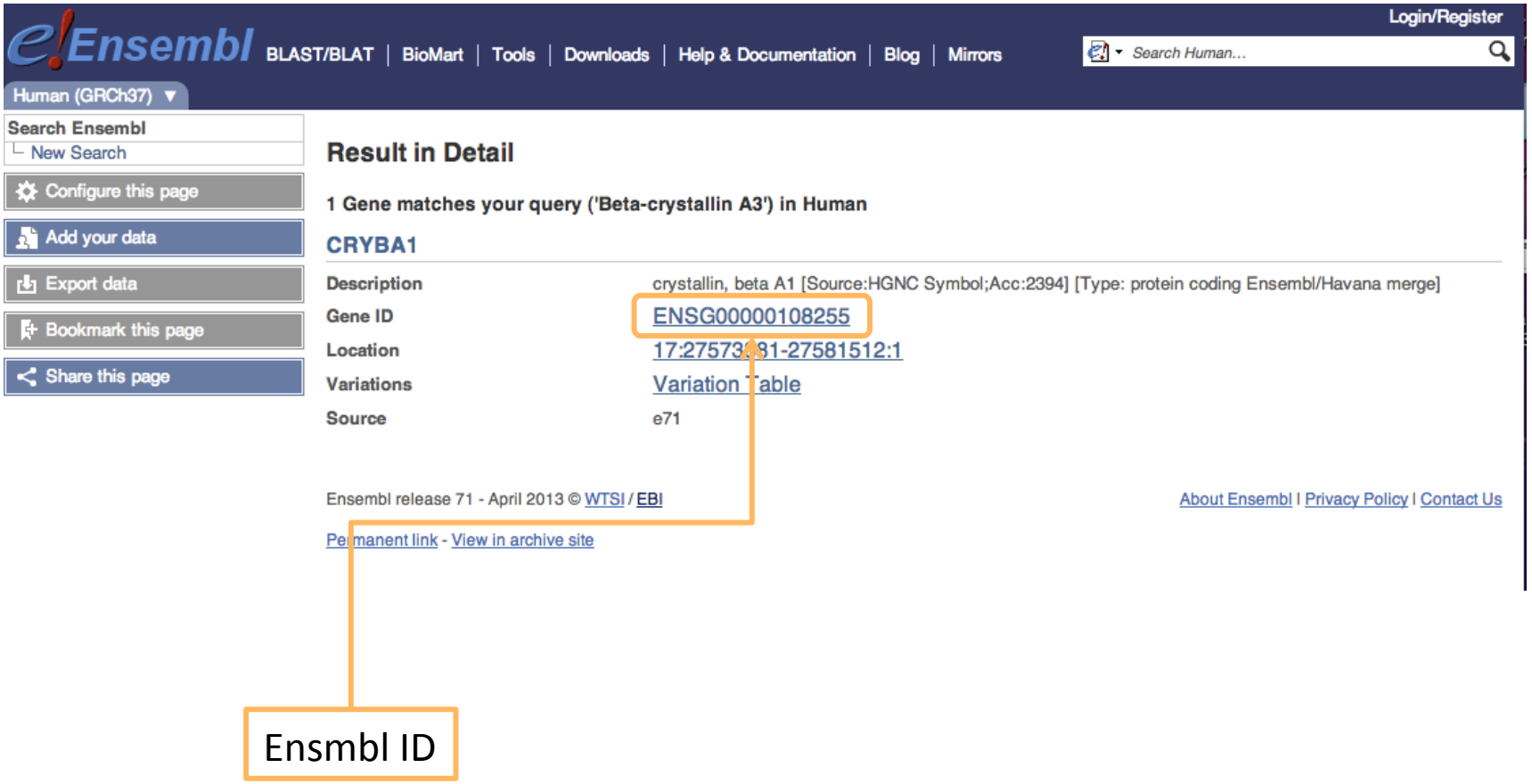
Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: Ensembl

Resultado de la búsqueda de “Beta-crystallin A3”

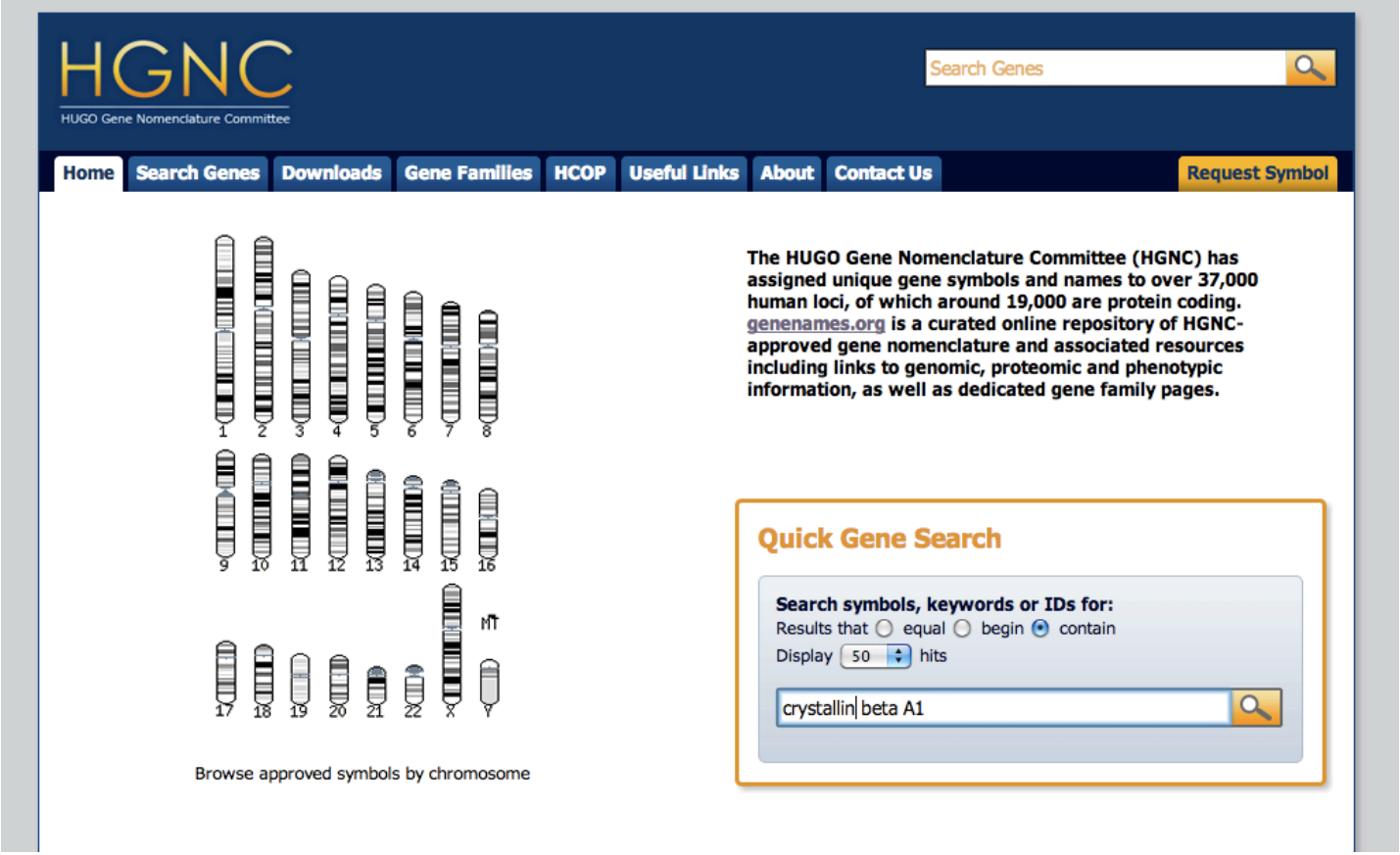
http://www.ensembl.org/Homo_sapiens/Search/Results?species=Homo_sapiens;idx=;q=Beta-crystallin%20A3

Indetificadores de Genes: Ensembl



Indetificadores de Genes: HGNC

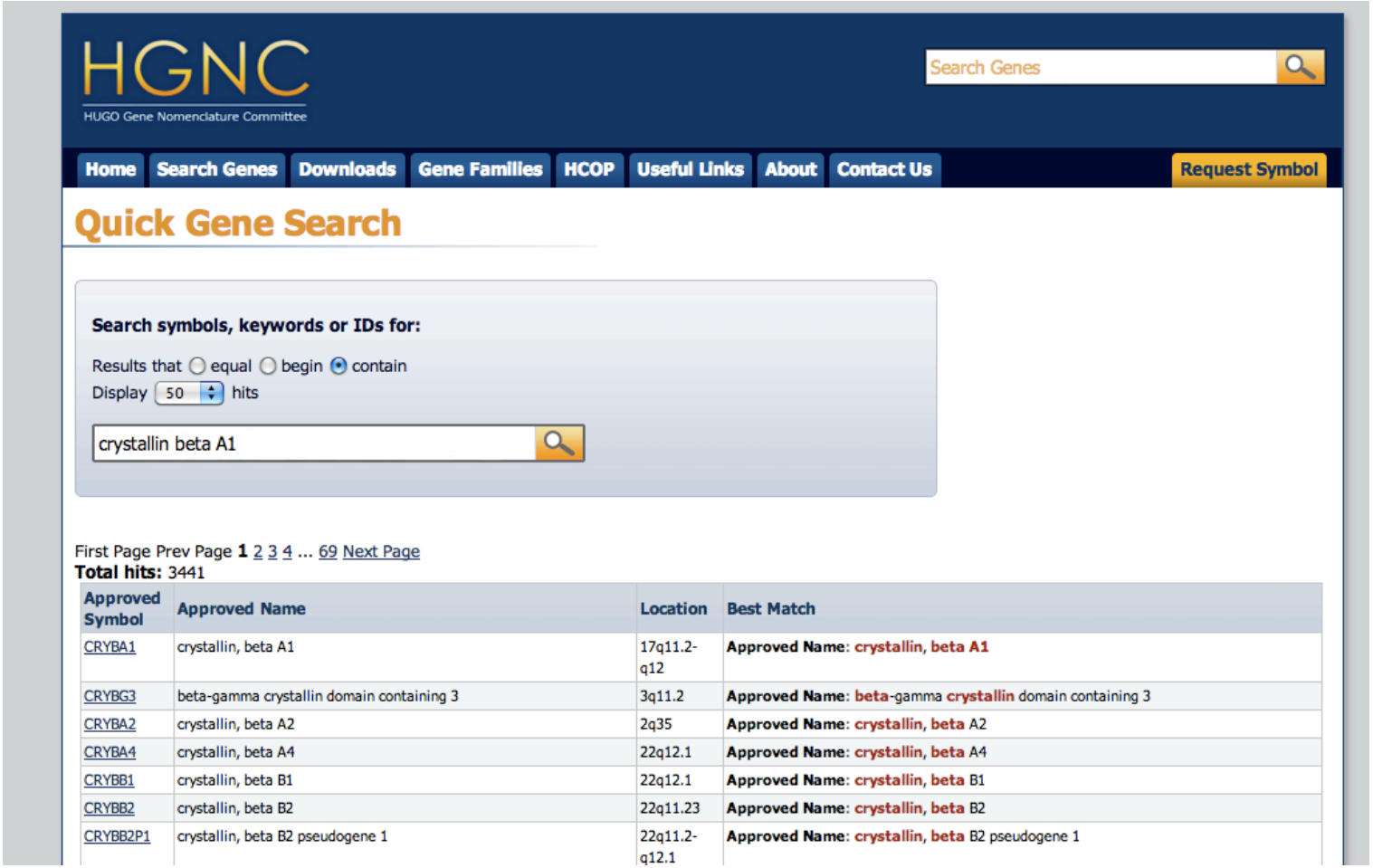
http://www.genenames.org/



Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: HGNC

http://www.genenames.org/



Approved Symbol	Approved Name	Location	Best Match
CRYBA1	crystallin, beta A1	17q11.2-q12	**Approved Name**: **crystallin, beta A1**
CRYBG3	beta-gamma crystallin domain containing 3	3q11.2	**Approved Name**: **beta**-gamma **crystallin** domain containing 3
CRYBA2	crystallin, beta A2	2q35	**Approved Name**: **crystallin, beta** A2
CRYBA4	crystallin, beta A4	22q12.1	**Approved Name**: **crystallin, beta** A4
CRYBB1	crystallin, beta B1	22q12.1	**Approved Name**: **crystallin, beta** B1
CRYBB2	crystallin, beta B2	22q11.23	**Approved Name**: **crystallin, beta** B2
CRYBB2P1	crystallin, beta B2 pseudogene 1	22q11.2-q12.1	**Approved Name**: **crystallin, beta** B2 pseudogene 1

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: HGNC

http://www.genenames.org/

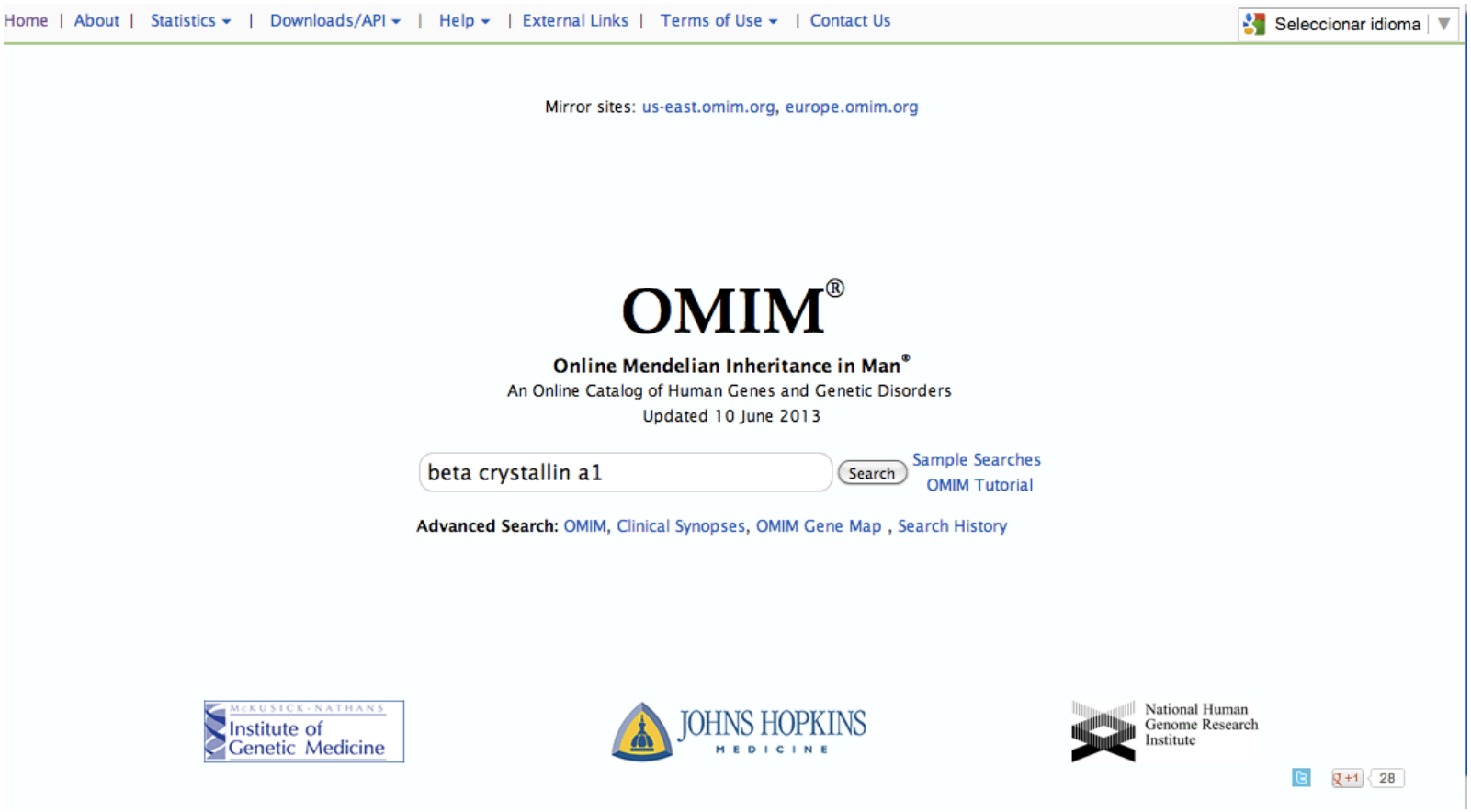
Symbol

HGNC ID

Más Identificadores

Indetificadores de Genes: MIM

http://www.omim.org/



Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: MIM

http://www.omim.org/

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: MIM

http://www.omim.org/

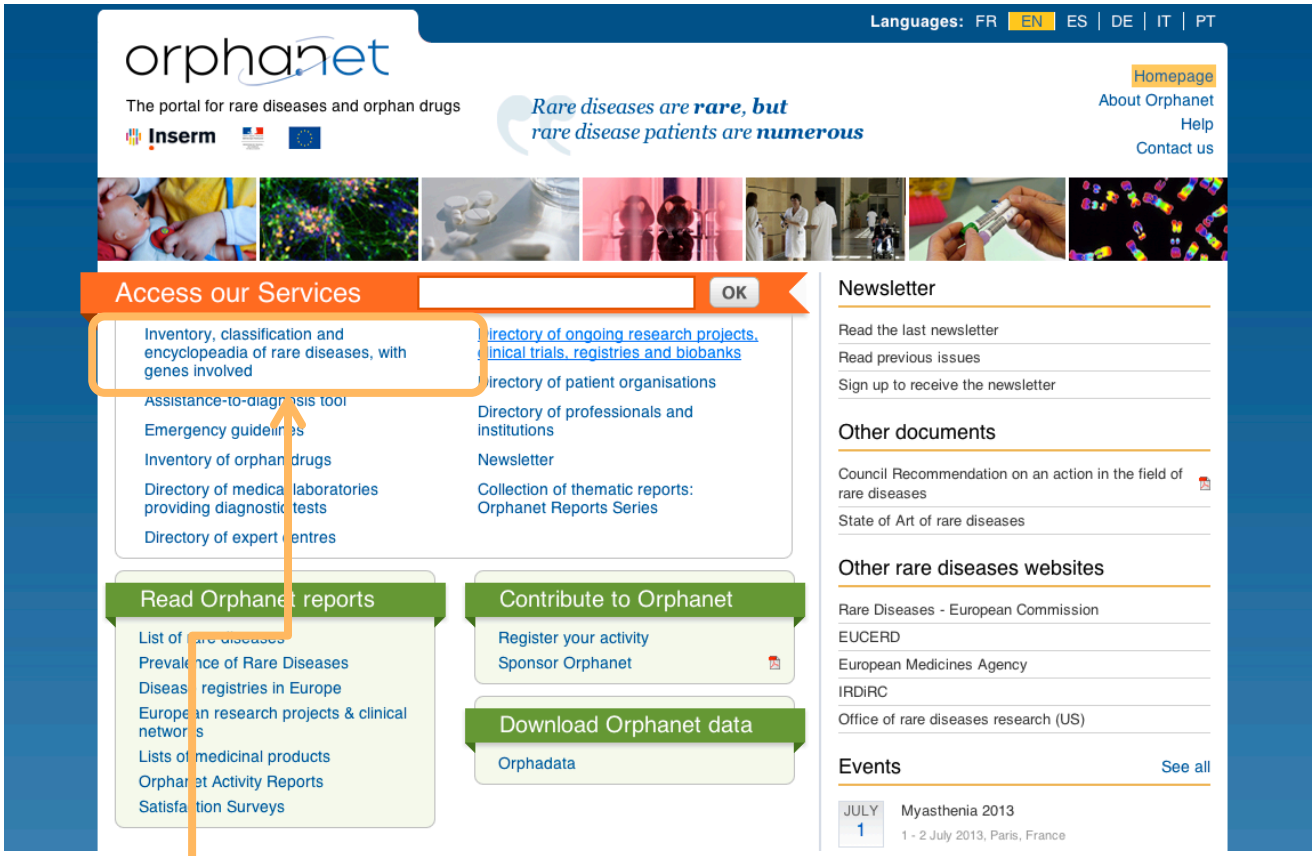
OMIM ID

Symbol

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: Orphanum

http://www.orpha.net/

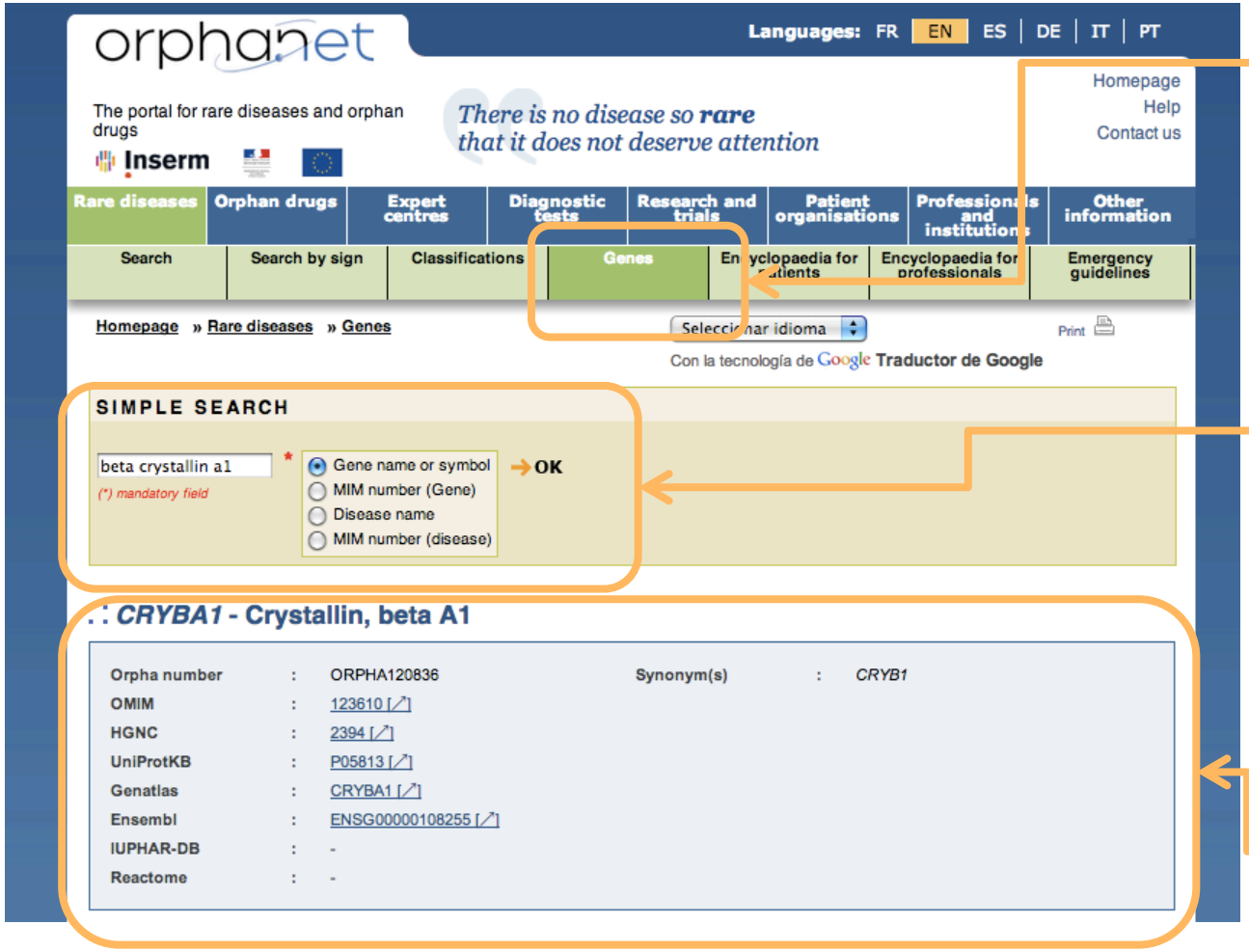


Búsquedas de Genes y Enfermedades

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Indetificadores de Genes: Orphanum

http://www.orpha.net/



Indicar tipo de búsqueda: Gen

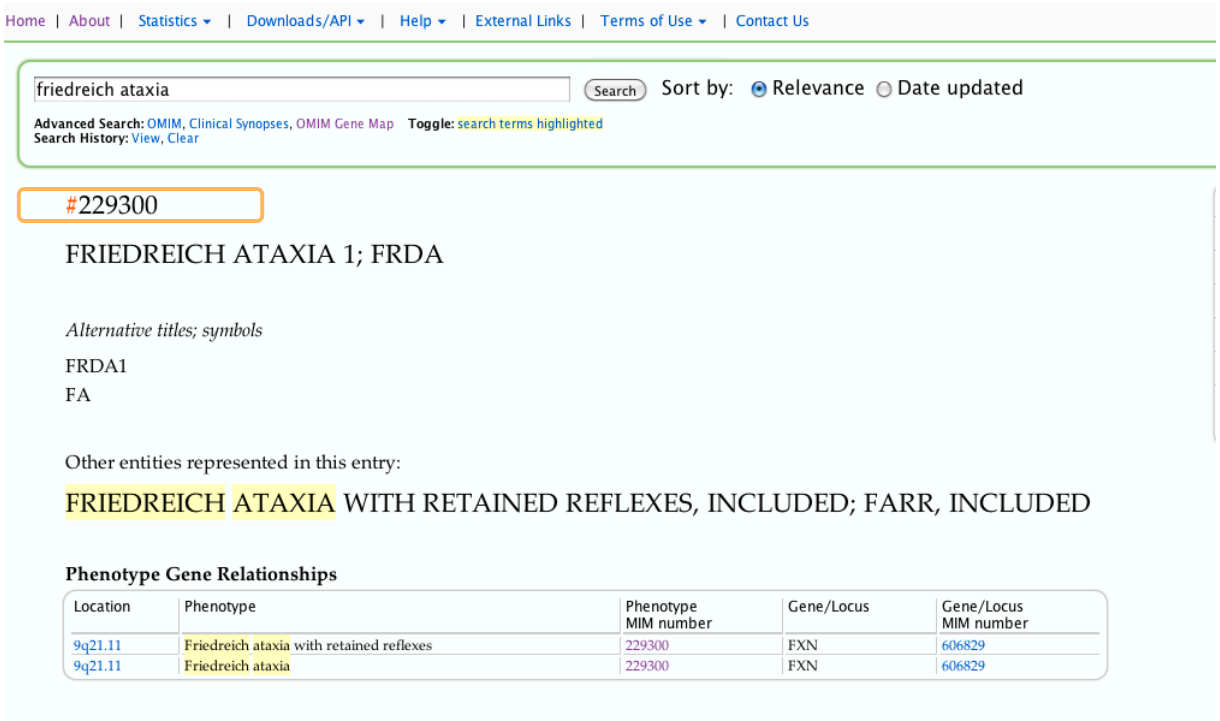
Formulario de búsqueda: Indicar tipo de identificador

Salida: Código ORPHA, OMIM, HGNC, Ensembl….

Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Identificadores de Enfermedades: OMIM

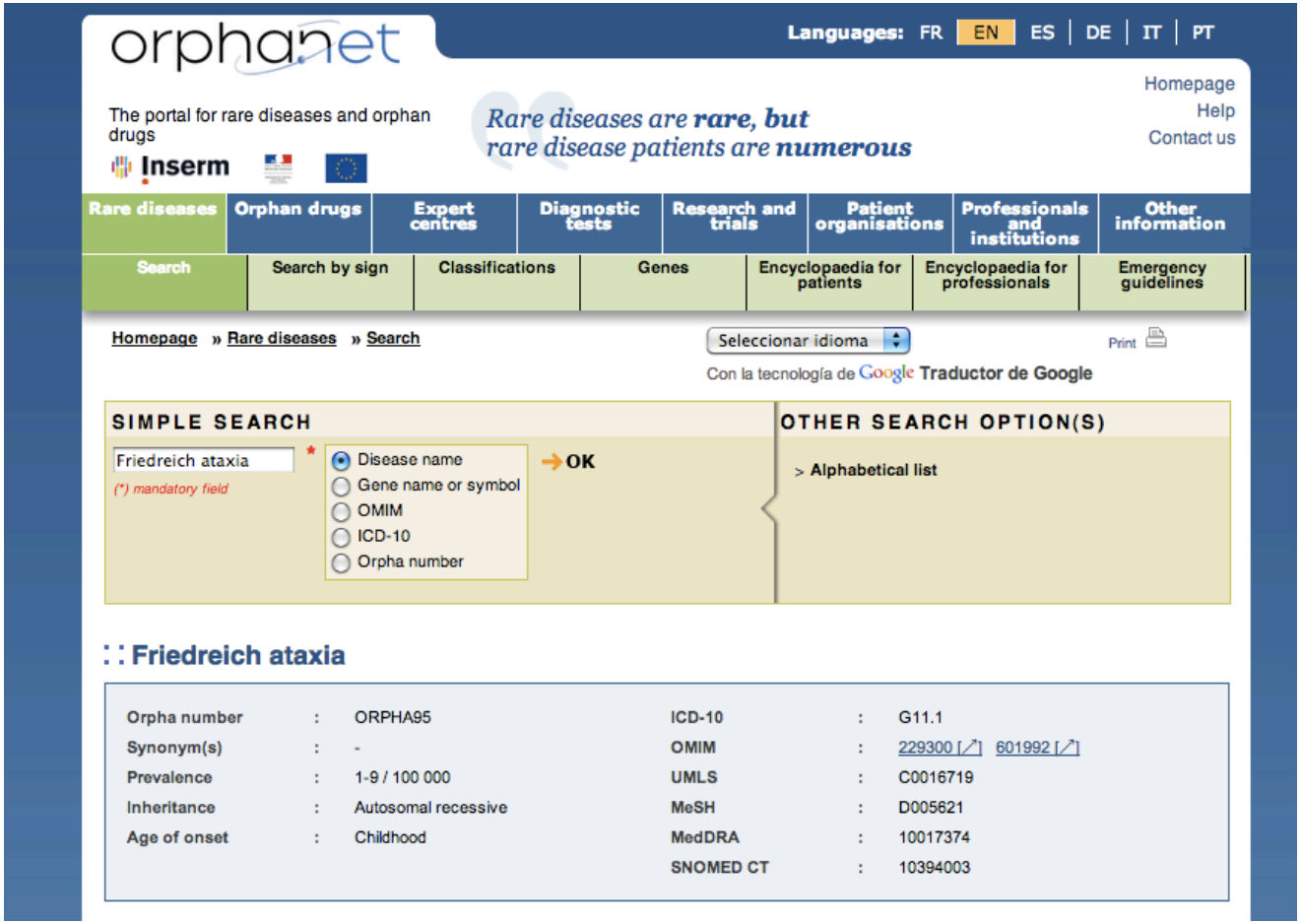
- Enfermedades OMIM
  - http://www.omim.org/
  - Buscar por nombre de enfermedad:
    - Ej: Friedreich ataxia



Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Enfermedades

Identificadores de Enfermedades: Orphanum

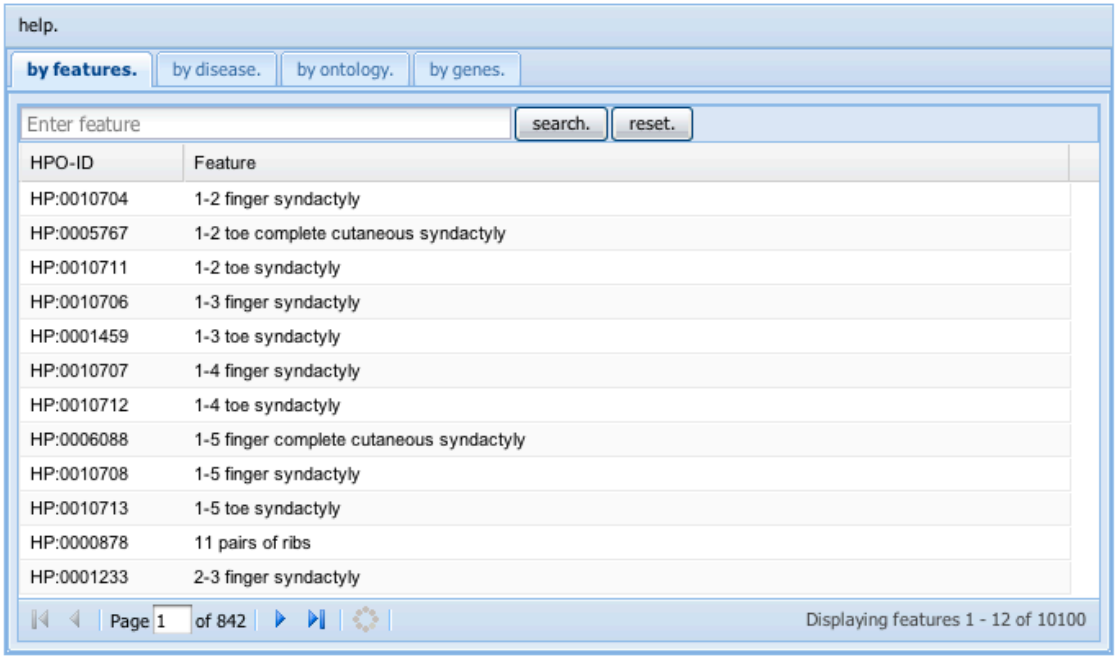
- Enfermedades Raras



Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Genes

Fenotipos

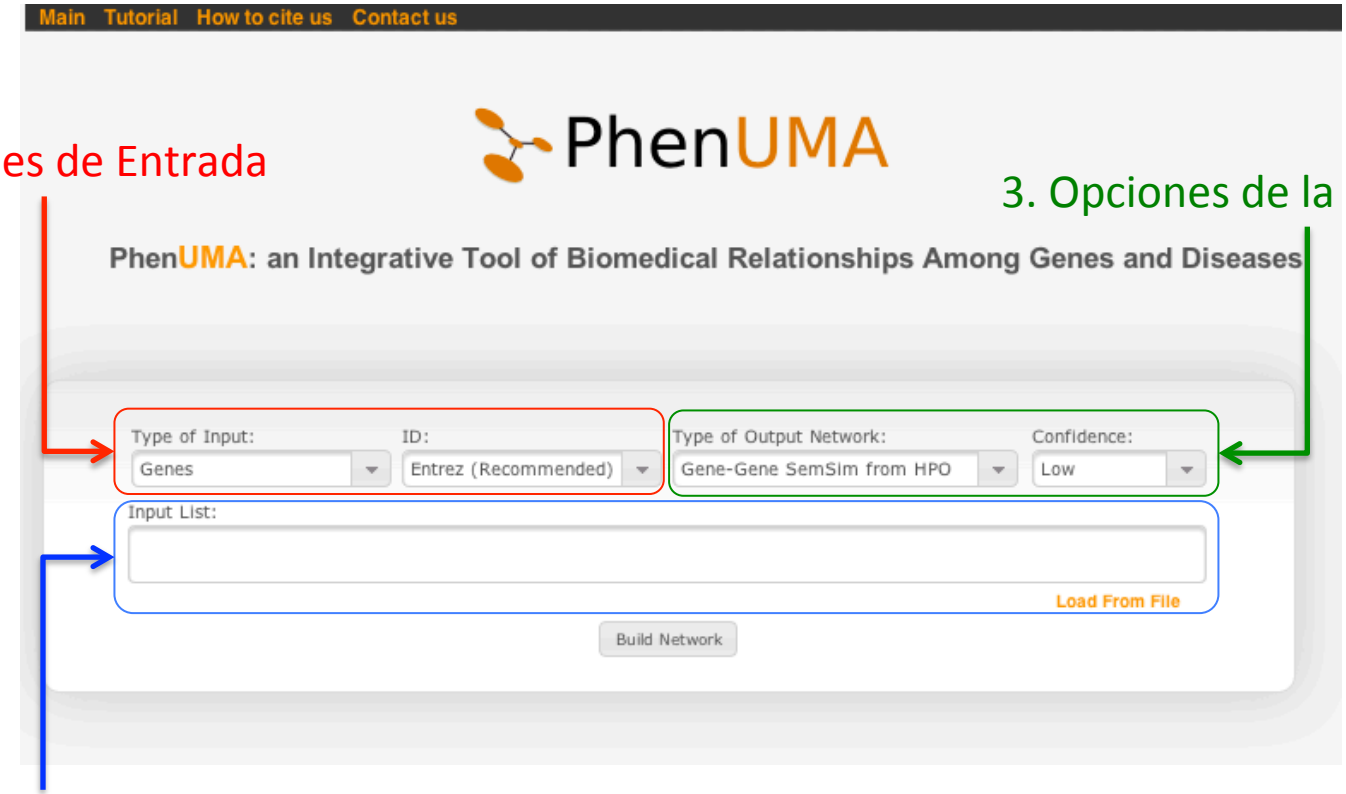
- PhenExplorer
  - http://compbio.charite.de/phenexplorer/



Entrada a PhenUMA → 1. Opciones de la Entrada → Identificadores de Fenotipos

Entrada a PhenUMA

1. Opciones de Entrada

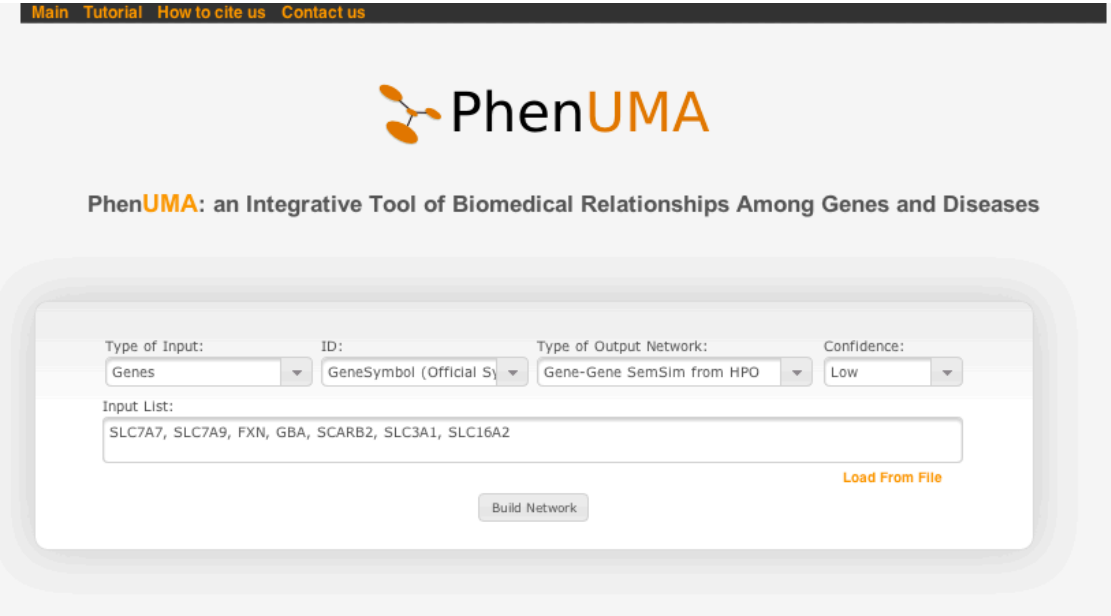


3. Opciones de la Red de Salida

2. Datos de Entrada

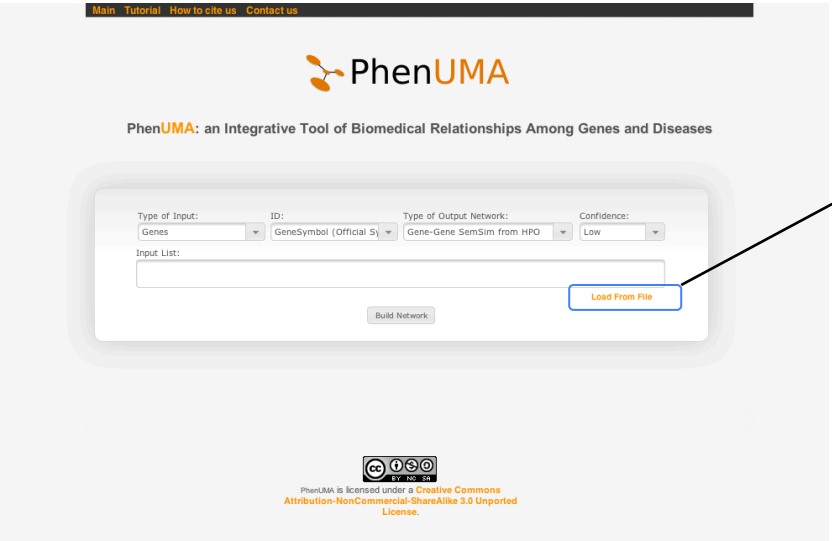
2. Datos de Entrada

- Separada por comas:
  - **Ejemplo1**
    - SLC7A7, SLC7A9, FXN, GBA, SCARB2, SLC3A1, SLC16A2

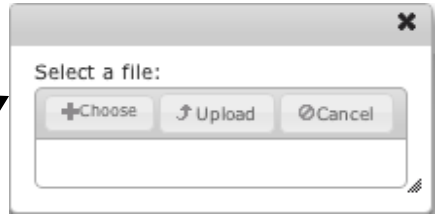


2. Datos de Entrada

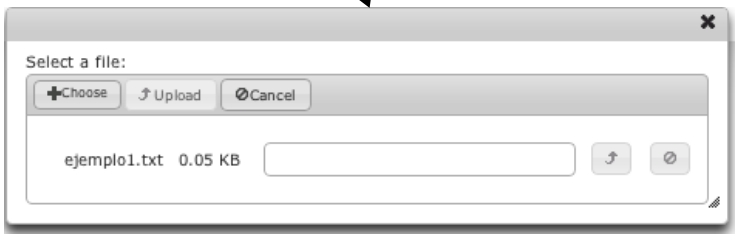
- Cargar entrada desde fichero de texto



1. **Presionar “Load From File”**



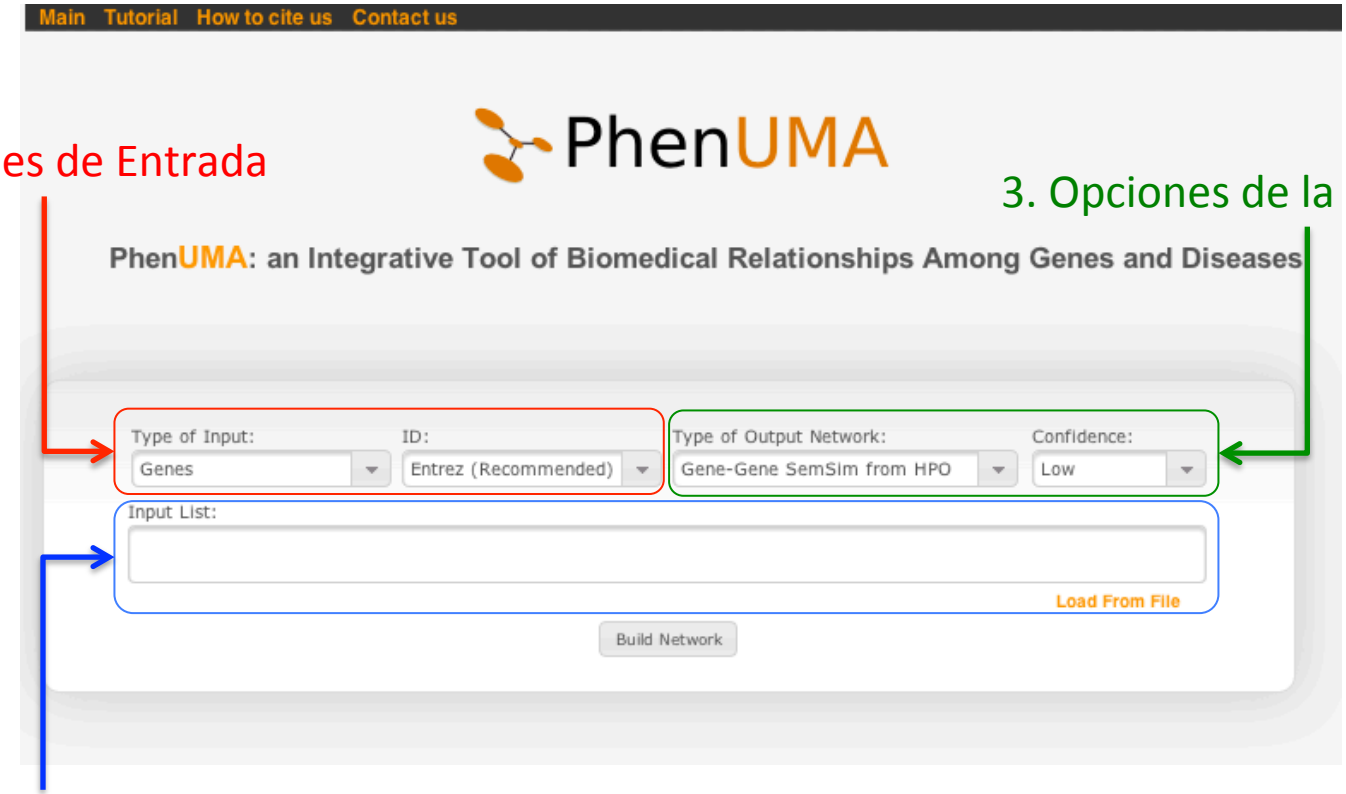
2. **En la ventana emergente pulsar “Choose” para seleccionar el archivo**



3. **“Upload” para cargar el archivo seleccionado y “Cancel” para eliminar el archivo seleccionado**

Entrada a PhenUMA

1. Opciones de Entrada

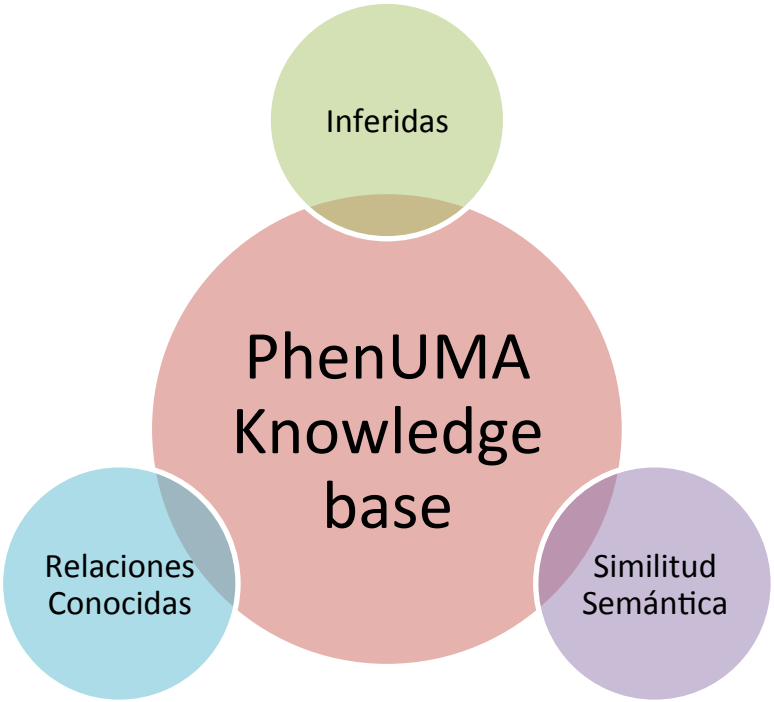


3. Opciones de la Red de Salida

2. Datos de Entrada

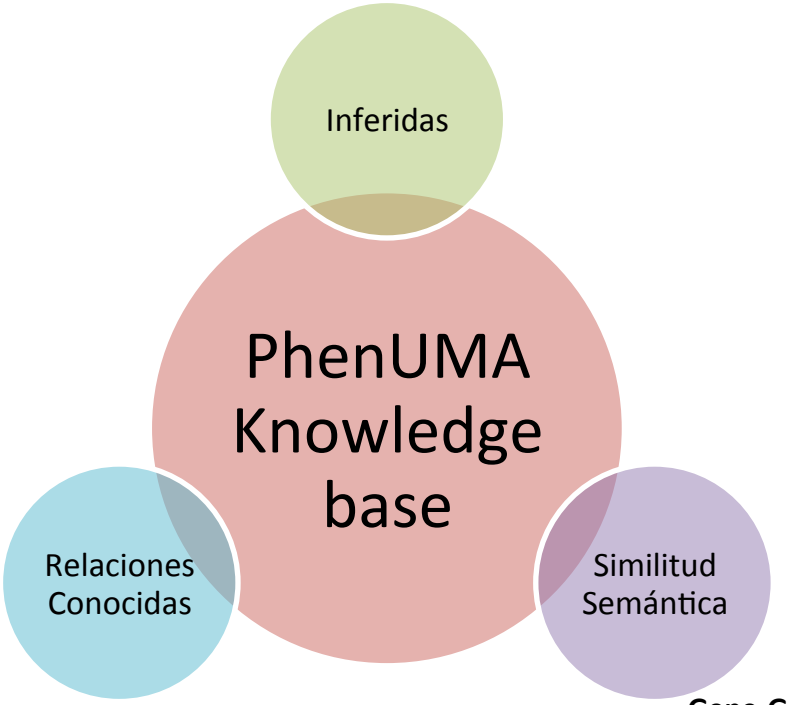
3. Opciones de Salida: Knowledge Base

- Las relaciones entre genes y enfermedades se clasifican en tres grupos:



Knowledge Base

Gene-Gene Inferred from OMIM
Gene-Gene Inferred from Orphanet



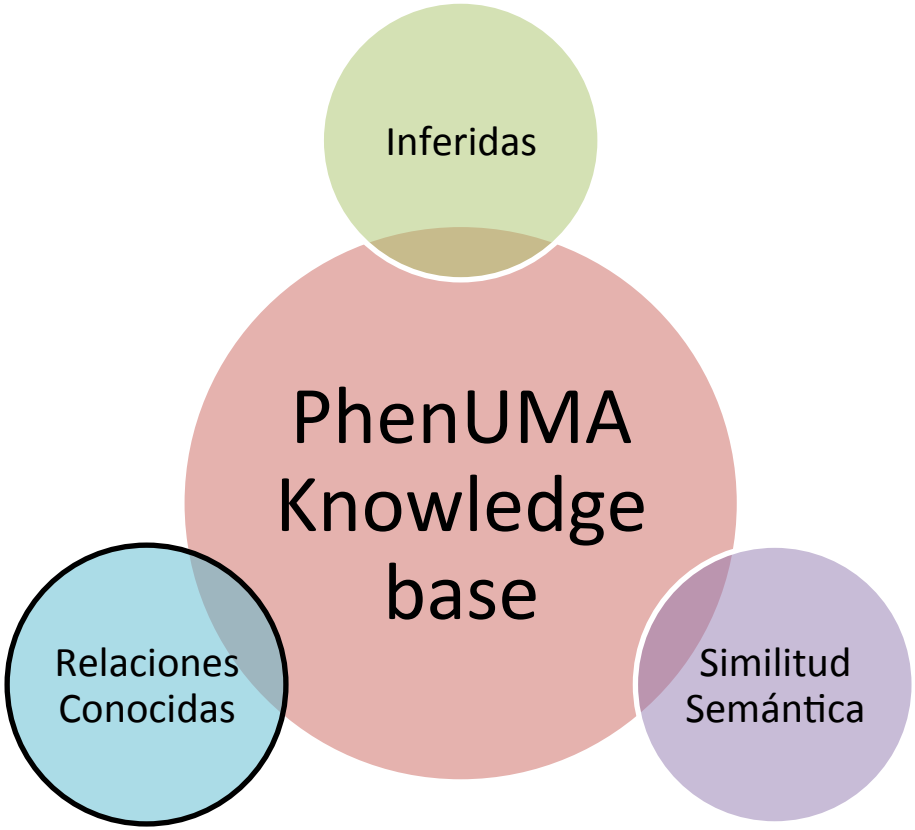
Gene-Gene Known Protein-Protein Interaction from STRING
Gene-Gene Known Metabolic Interactions (Veeramani et al.)

Gene-OMIM Known OMIM Relatioships
Gene-ORPHA Known from Orphanet

Gene-Gene SimSem from HPO
Gene-Gene SimSem from GO (Biological Process)
Gene-Gene SimSem from GO (Cellular Component)
Gene-Gene SimSem from GO (Molecular Function)

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

3. Opciones de Salida: Knowledge Base



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Relaciones Conocidas

- STRING:
  - Base de datos que incluye información de proteínas que interaccionan físicamente.
  - Humanos (Homo spaiens)
  - 96000 interacciones aproximadamente

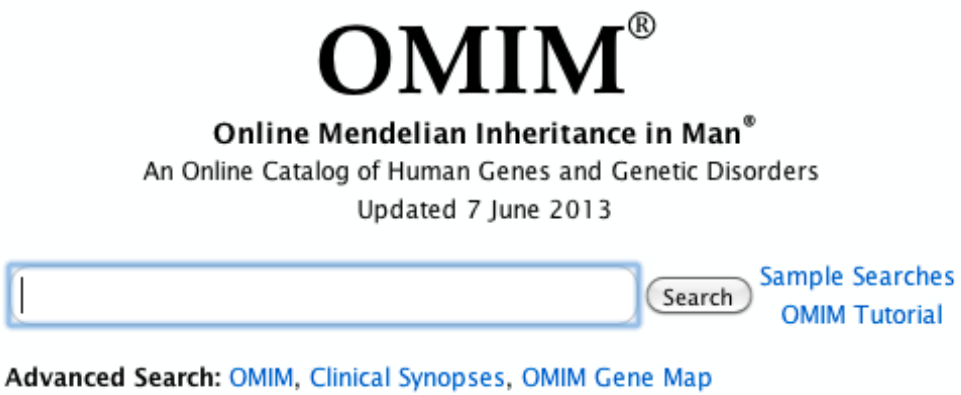
- Metabólicas (Veeramani et al.)
  - Recon 1
  - Futuro : Recon2

http://string-db.org/

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Relaciones Conocidas

- Relaciónes gen-OMIM y gen-Orphan.
- Relaciones moleculares entre variaciones genéticas y expresión de fenotipos.



orphanet

http://www.omim.org/

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Relaciones Conocidas

#271980

ICD+

SUCCINIC SEMIALDEHYDE DEHYDROGENASE DEFICIENCY; SSADHD

Alternative titles; symbols

SSADH DEFICIENCY
4-HYDROXYBUTYRIC ACIDURIA
GABA METABOLIC DEFECT
GAMMA-HYDROXYBUTYRIC ACIDURIA

Phenotype Gene Relationships

Location	Phenotype	Phenotype MIM number	Gene/Locus	Gene/Locus MIM number
6p22.3	Succinic semialdehyde dehydrogenase deficiency	271980	ALDH5A1	610045

Clinical Synopsis

TEXT

A number sign (#) is used with this entry because succinic semialdehyde dehydrogenase (SSADH) deficiency can be caused by homozygous mutation in the ALDH5A1 gene (610045) on chromosome 6p22.

Description

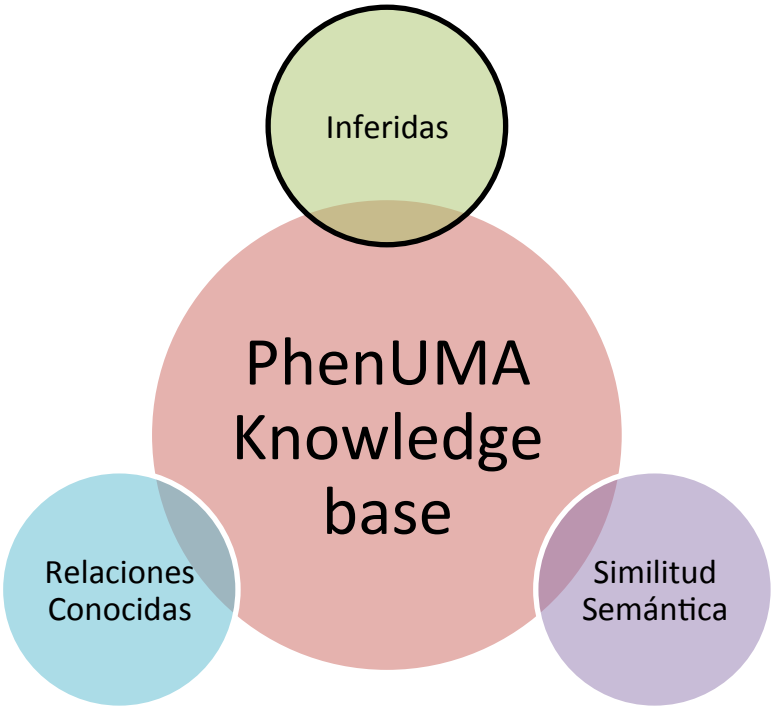
Succinic semialdehyde dehydrogenase deficiency (SSADHD) is a rare autosomal recessive neurologic disorder in which an enzyme defect in the GABA degradation pathway causes a consecutive elevation of gamma-hydroxybutyric acid (GHB) and GABA. The clinical features include developmental delay, hypotonia,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

http://www.omim.org/entry/271980?search=271980&highlight=271980

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

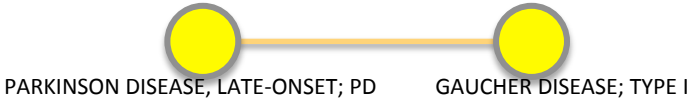
3. Opciones de Salida: Knowledge Base

- Relaciones Inferidas: a partir de las conocidas



Knowledge Base : Relaciones Inferidas

- A partir de las relaciones conocidas de OMIM y Orphanet



Gene-Gene Inferred from OMIM

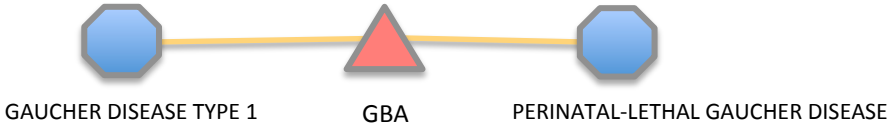
OMIM-OMIM Inferred from Gene

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Relaciones Inferidas

- A partir de las relaciones conocidas de OMIM y Orphanet

Gene-Gene Inferred from Orphanet

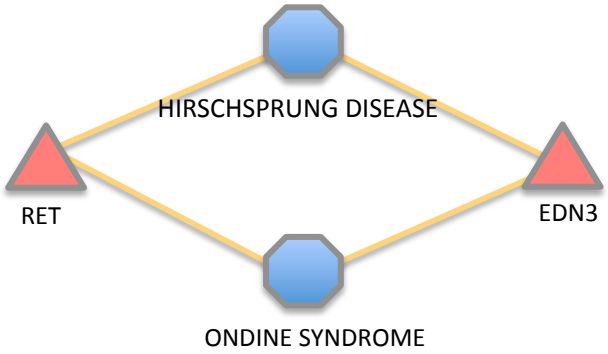


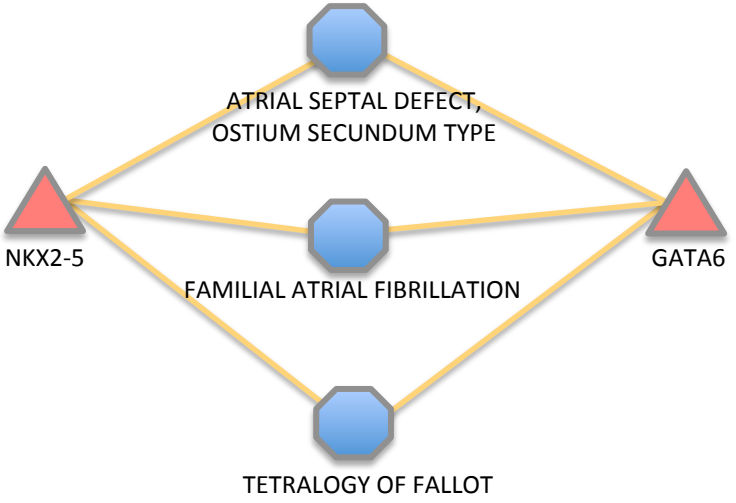
Orpha-Orpha Inferred from Gene

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Relaciones Inferidas

- Relaciones múltiples.
- Valor de la relación determinado por el número de enfermedades compartidas



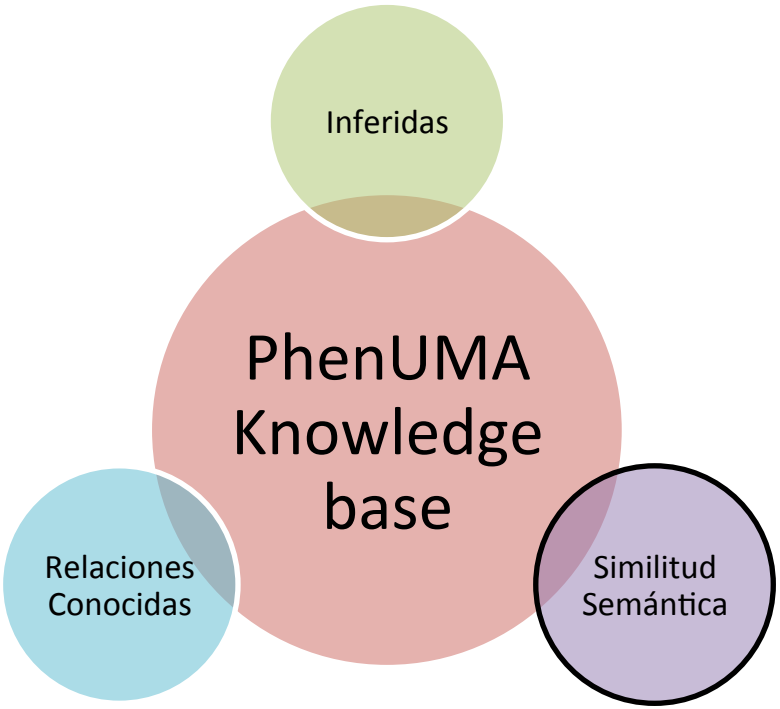


Gene-Gene Inferred from Orphanet

Orpha-Orpha Inferred from Gene

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

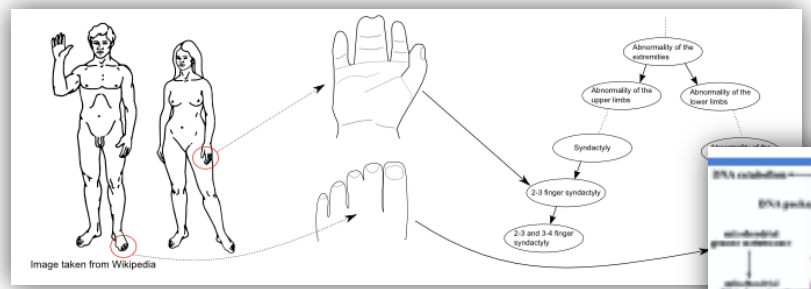
Knowledge Base



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

- Similitud calculada con el uso de dos ontologías:
  - Similitud Fenotípica: Human Phenotype Ontology (HPO)
    - http://www.human-phenotype-ontology.org/
  - Similitud Funcional : Gene Ontology (GO)
    - http://www.geneontology.org/
- **Vocabularios organizados:**
  - Procesos Biológicos, Componentes Celulares, Funciones Moleculares → GO
  - Fenotipos relacionados con patologías → HPO
- **Estructura jerárquica**

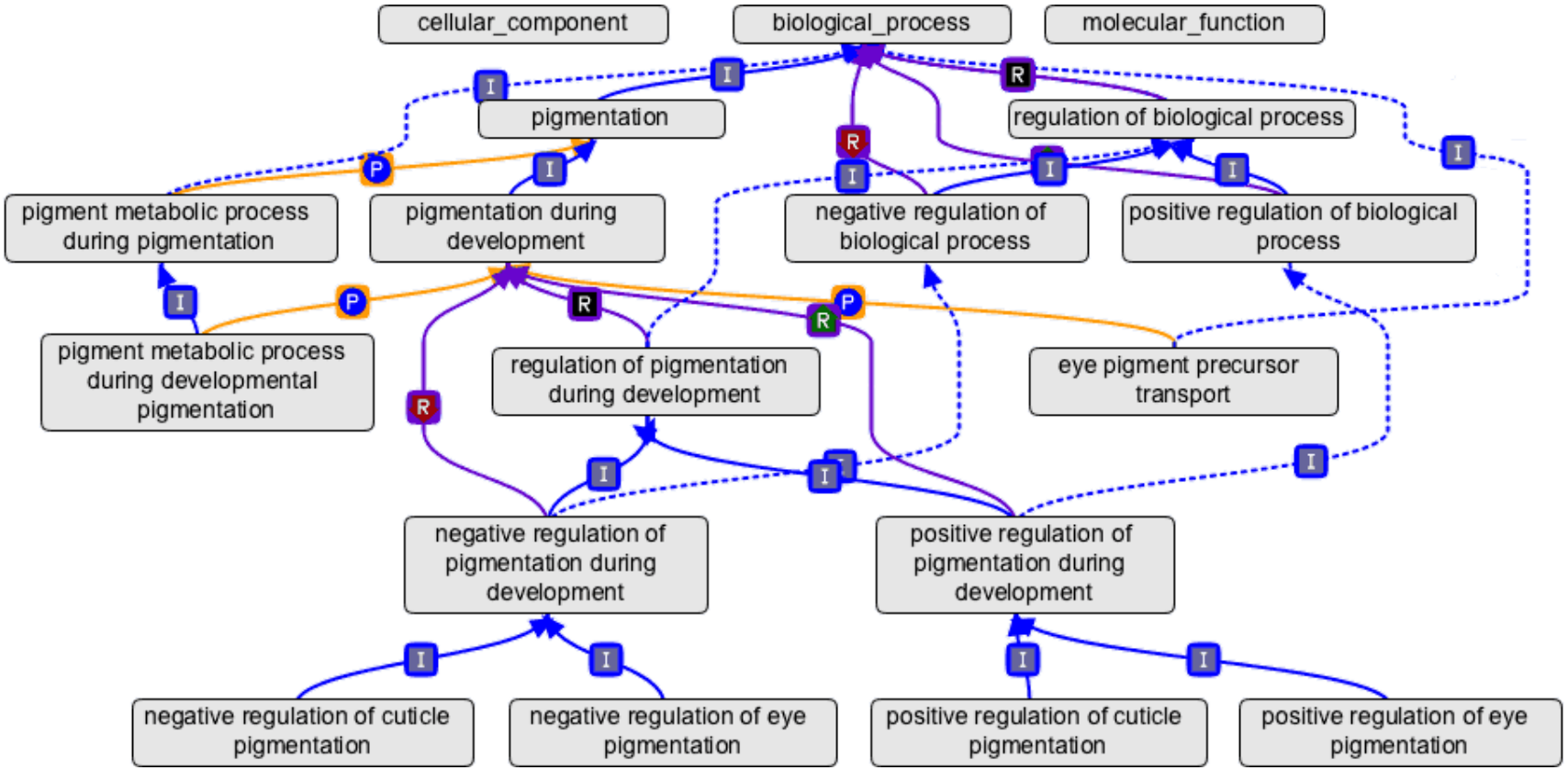


Human Phenotype Ontology

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

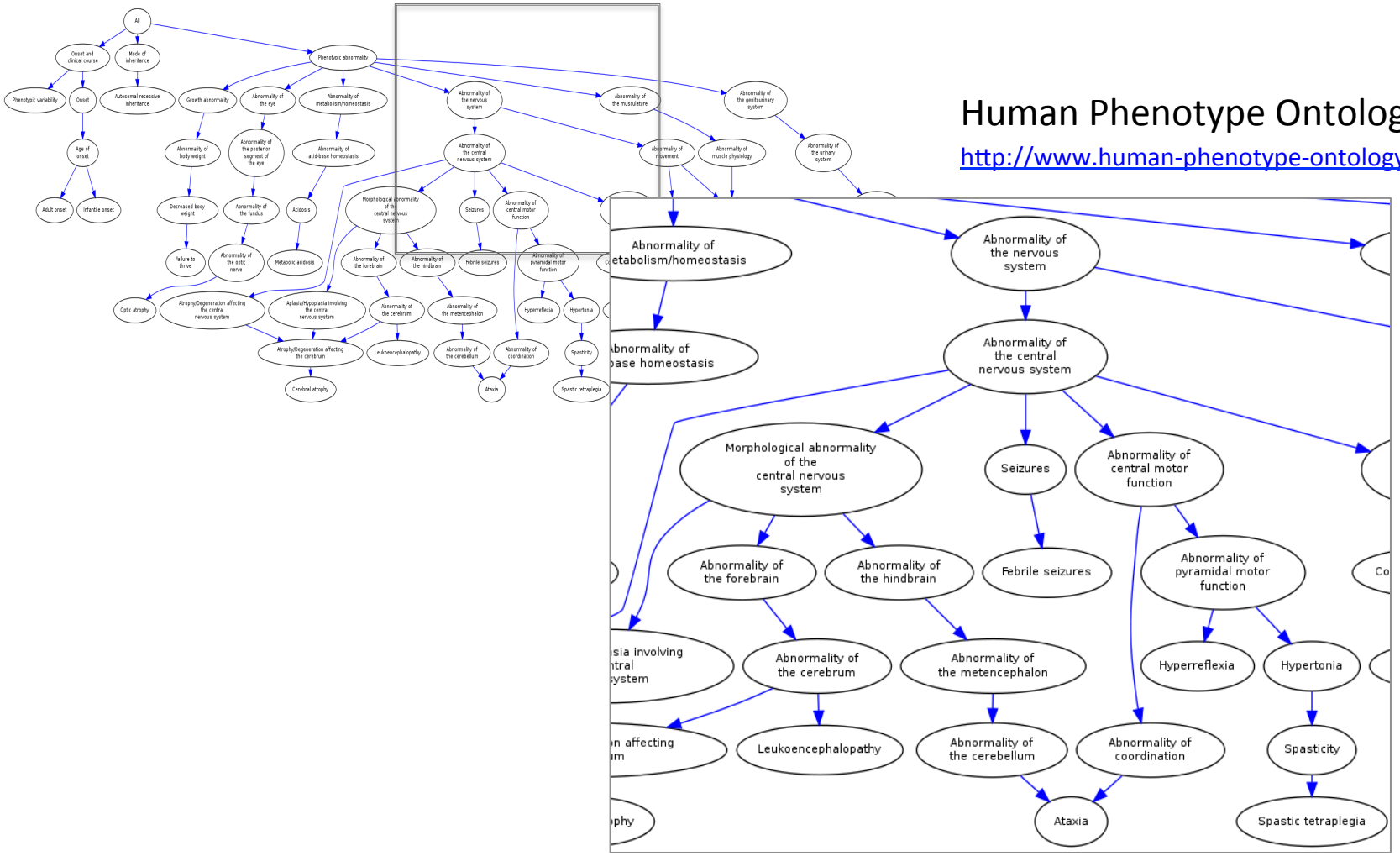
Knowledge Base : Similitud Semántica

- Gene Ontology
  - http://www.geneontology.org/



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica



Human Phenotype Ontology

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

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

- **Espacio fenotípico:** conjunto de fenotipos de una enfermedad a partir de la descripción de su sintomatología.

… Many showed parkinsonian features and-or other neurologic abnormalities, as well as **mental retardation** and **transient psychotic episodes**. In children, **head nodding** was a frequent precursor of later grand mal **seizures**…

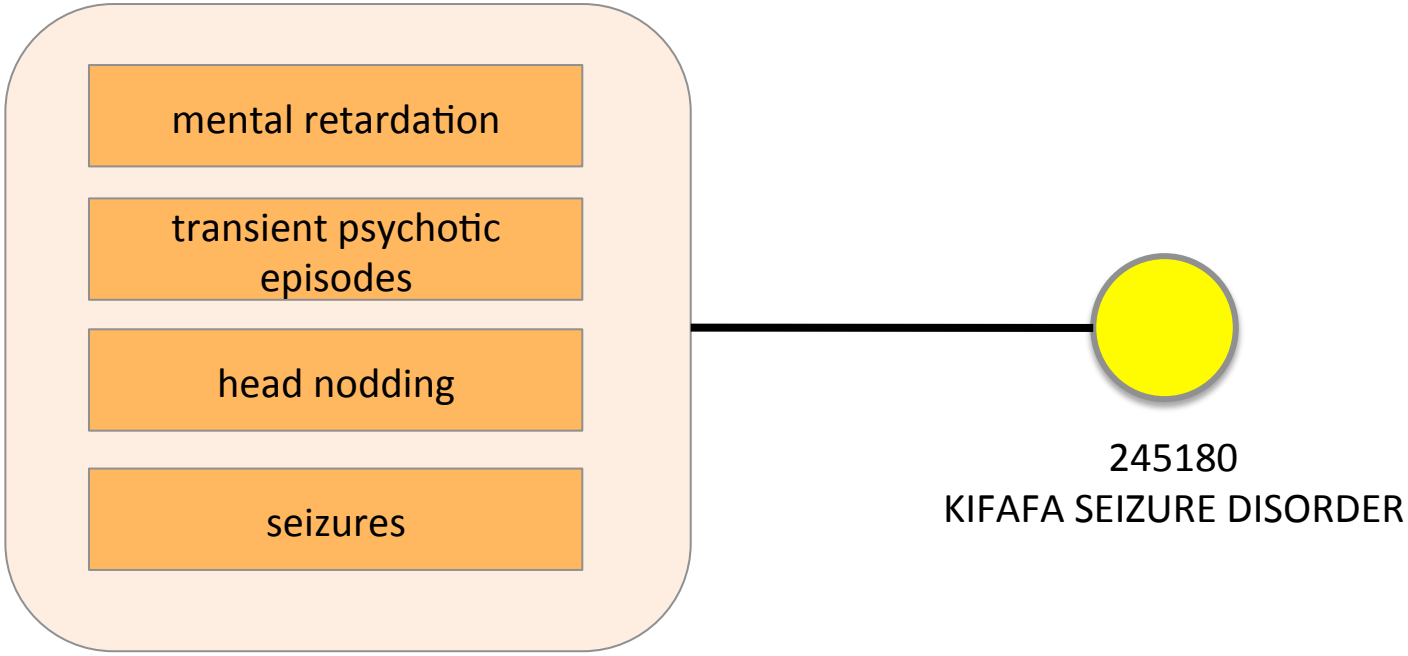
245180
KIFAFA SEIZURE DISORDER

http://omim.org/entry/245180

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

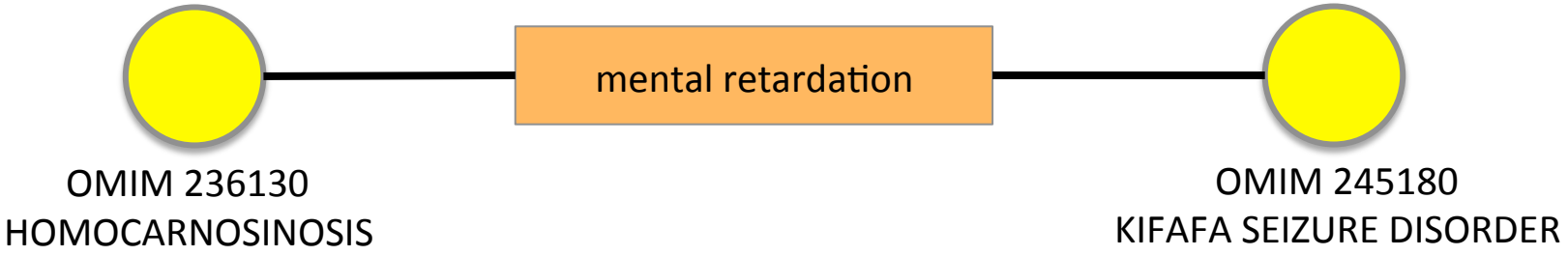
- El vocabulario organizado permite seleccionar el conjunto de términos con los que definir la enfermedad.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

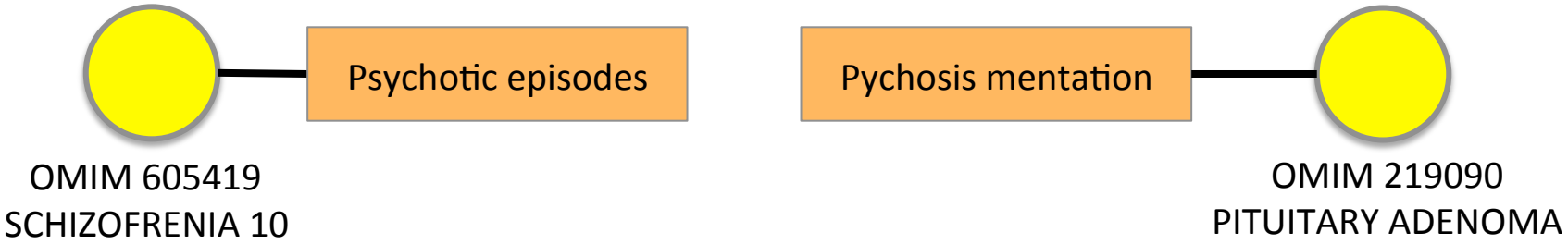
- Relaciones entre enfermedades a partir de los fenotipos.
- Similar a la relaciones inferidas.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

- La estructura jerárquica de la ontología permite ampliar las relaciones que pueden establecerse a partir de inferencias.

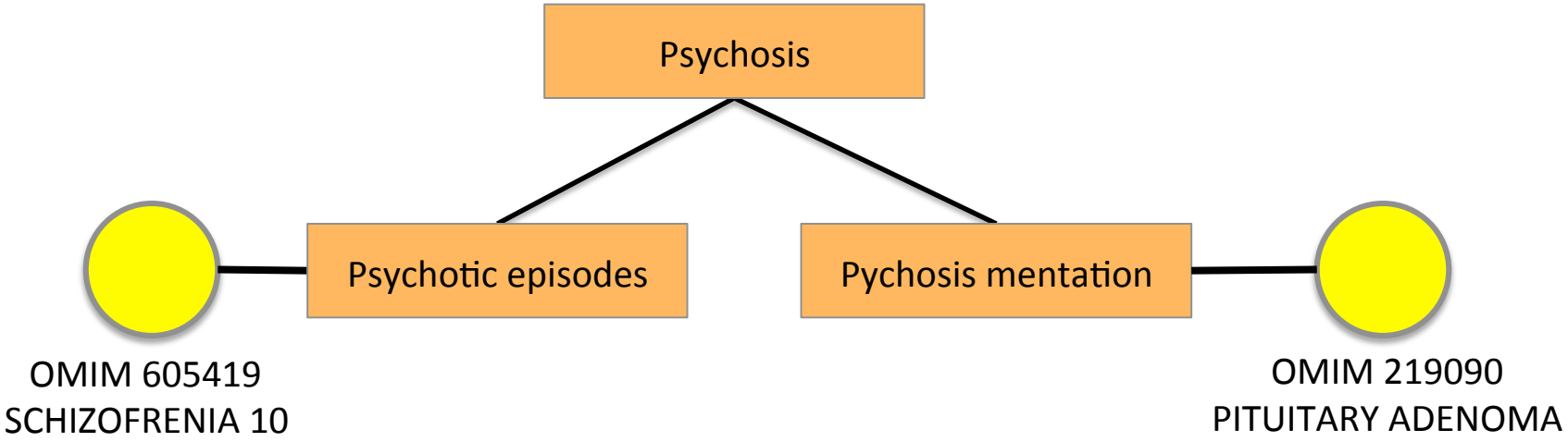


¿Deberían estar relacionadas?

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

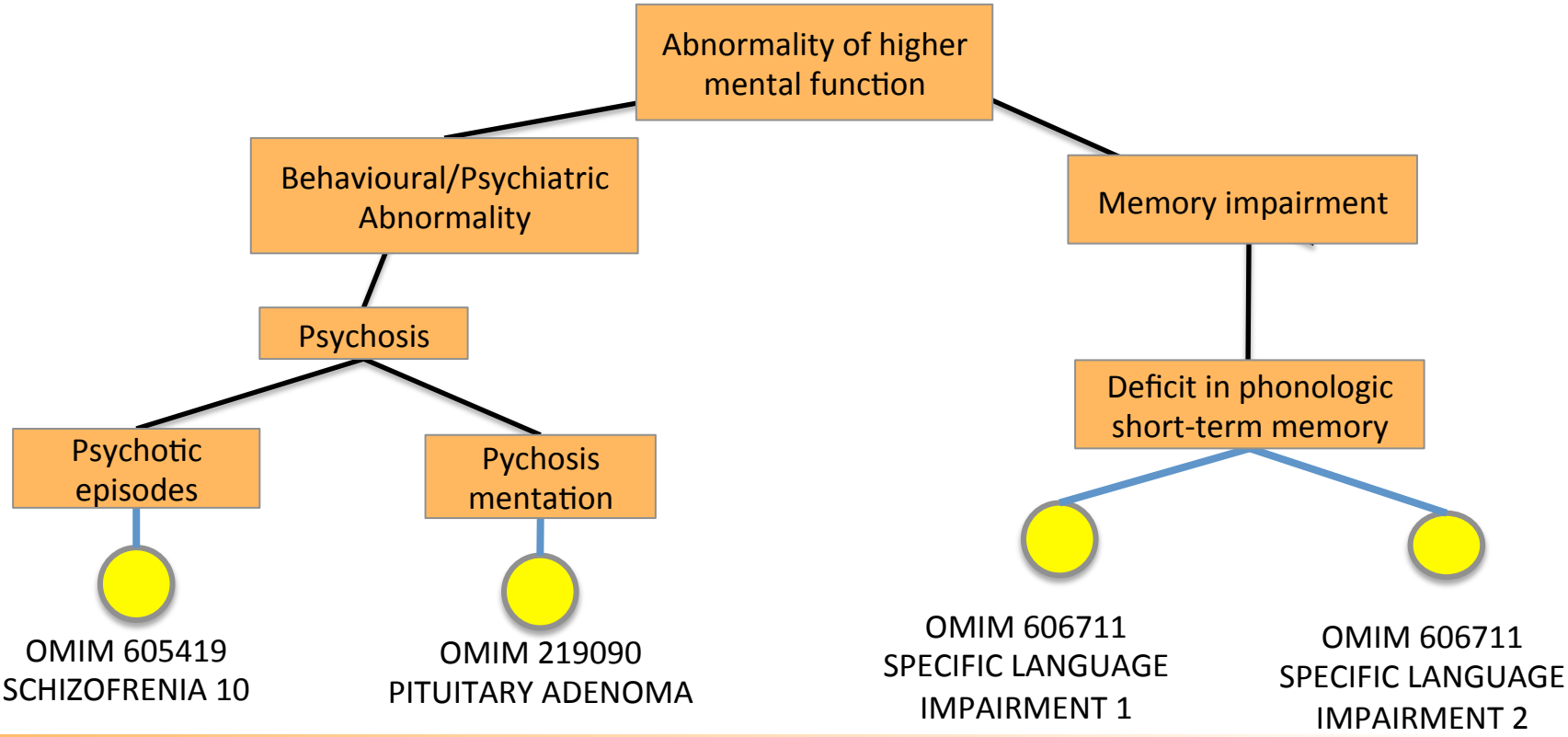
- El fenotipo “pychosis” permite establecer una relación entre las dos enfermedades.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

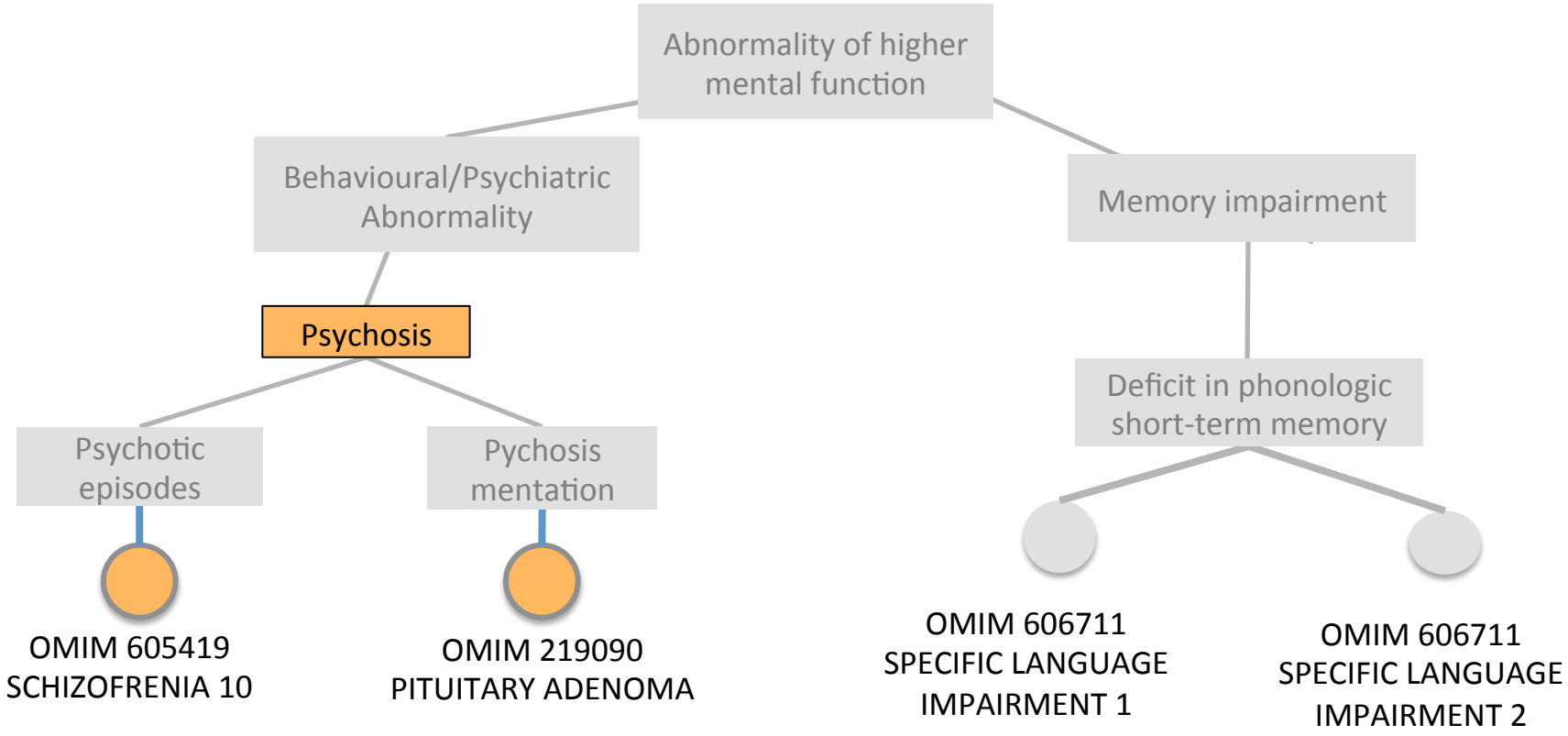
- La lejanía del término común entre los síntomas de dos enfermedades van a determinar el valor de la similitud semántica.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

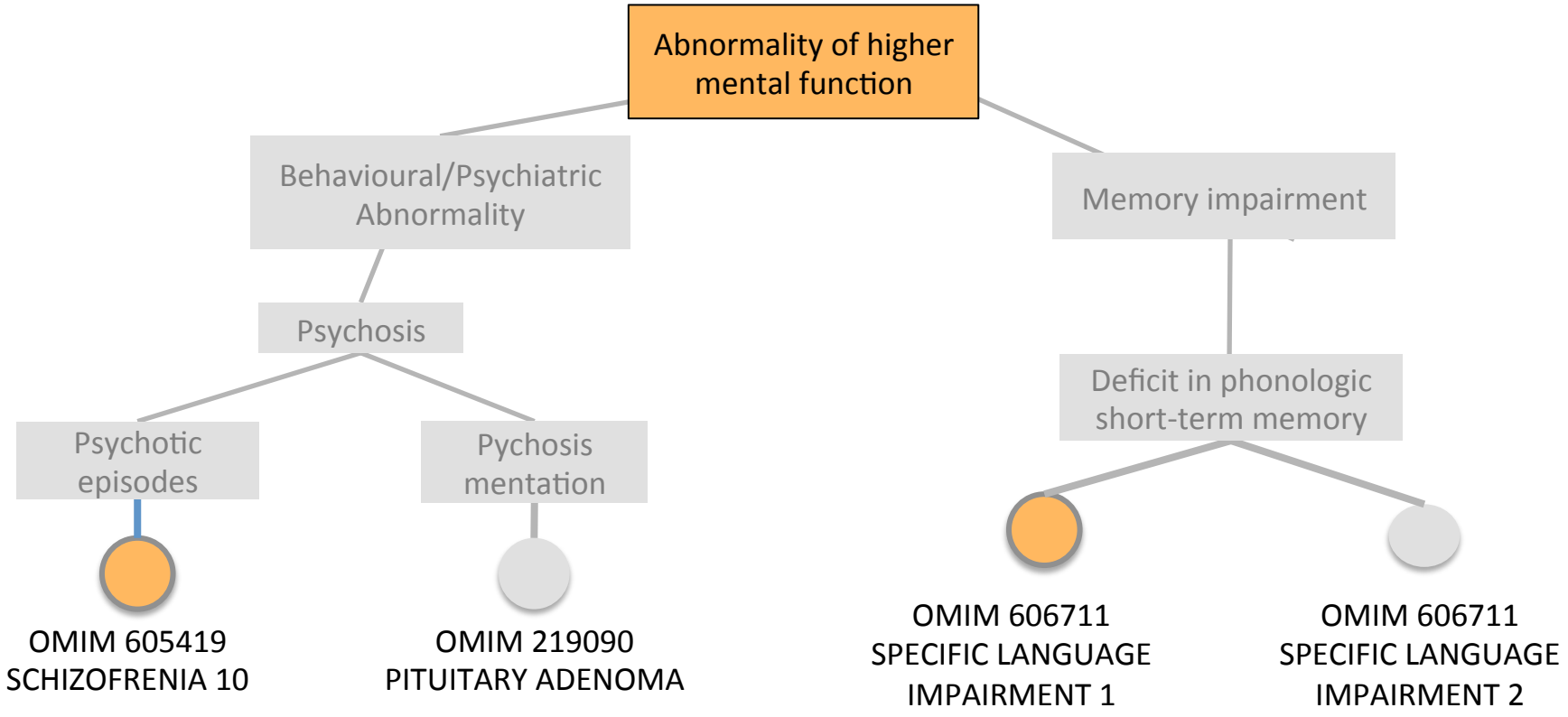
- La similitud entre **Schizofrenia** y **Pituitary** adenoma será mayor que la similitud entre **Schizofrenia** y **Specific Language Impairmet 1**.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

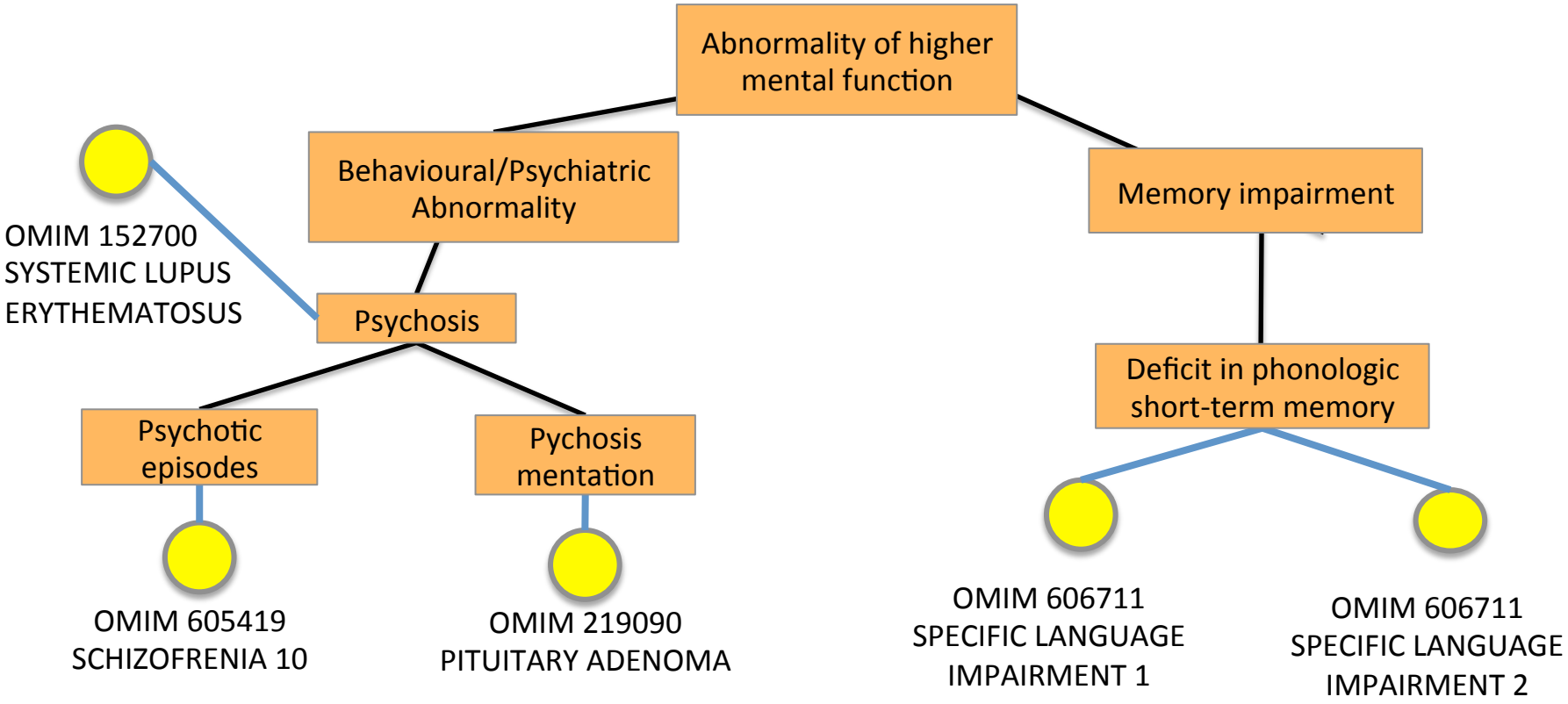
- La similitud entre **Schizofrenia** y **Pituitary adenoma** será mayor que la similitud entre **Schizofrenia** y **Specific Language Impairmet 1**.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

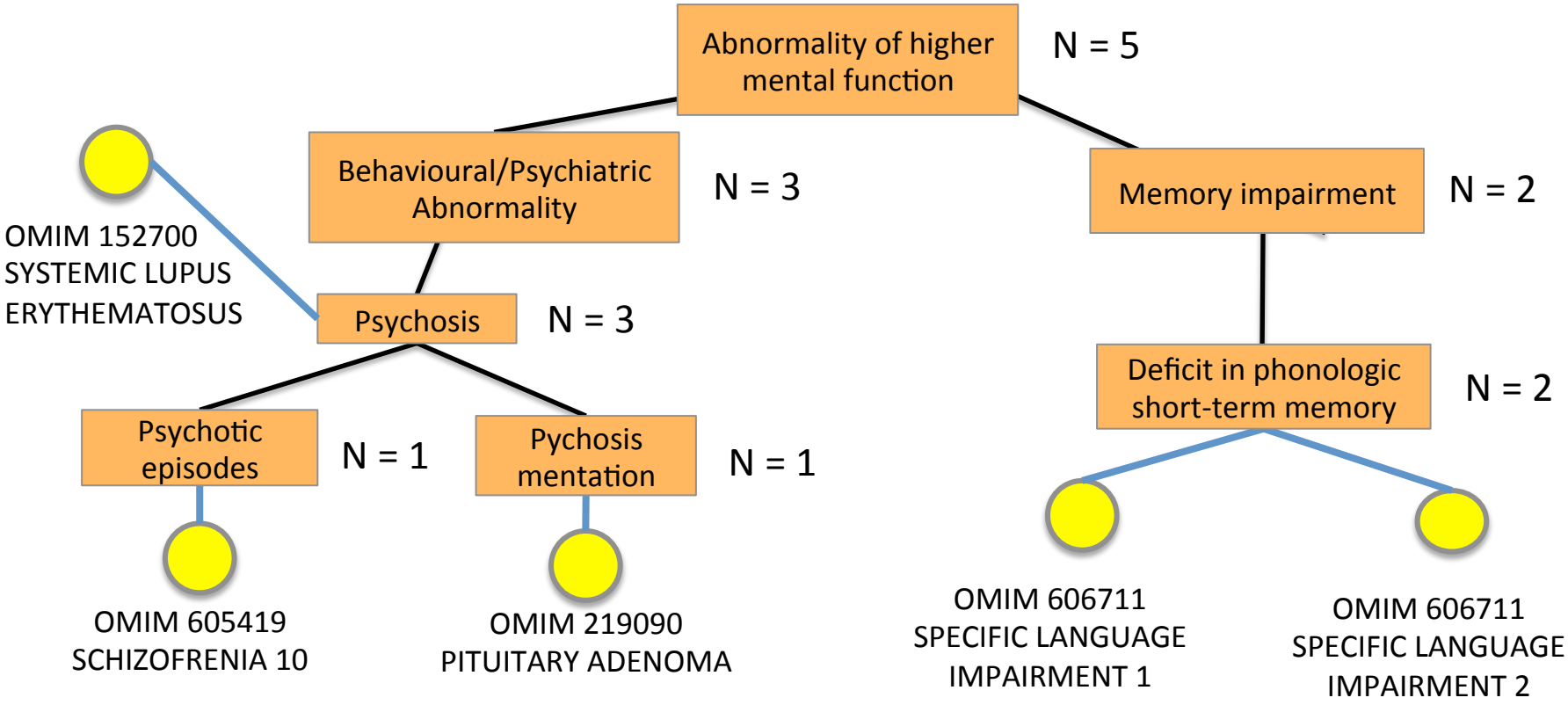
- A cada fenotipo del árbol se le asigna un valor numérico (IC) que será mayor a medida que nos alejamos de la raíz.
- Relaciones fenotipo – enfermedad se propagan hacía la raíz



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

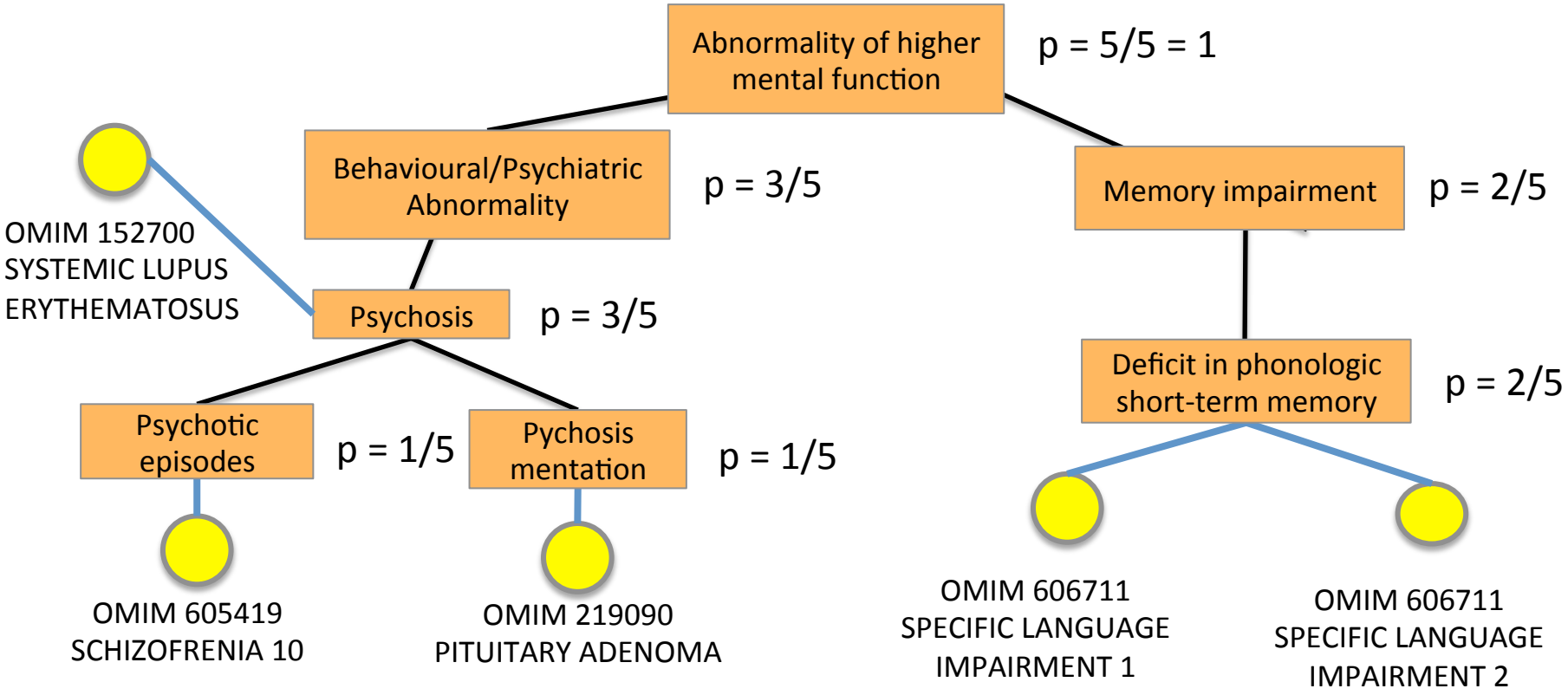
- N = número de enfermedades asociadas a un fenotipo



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

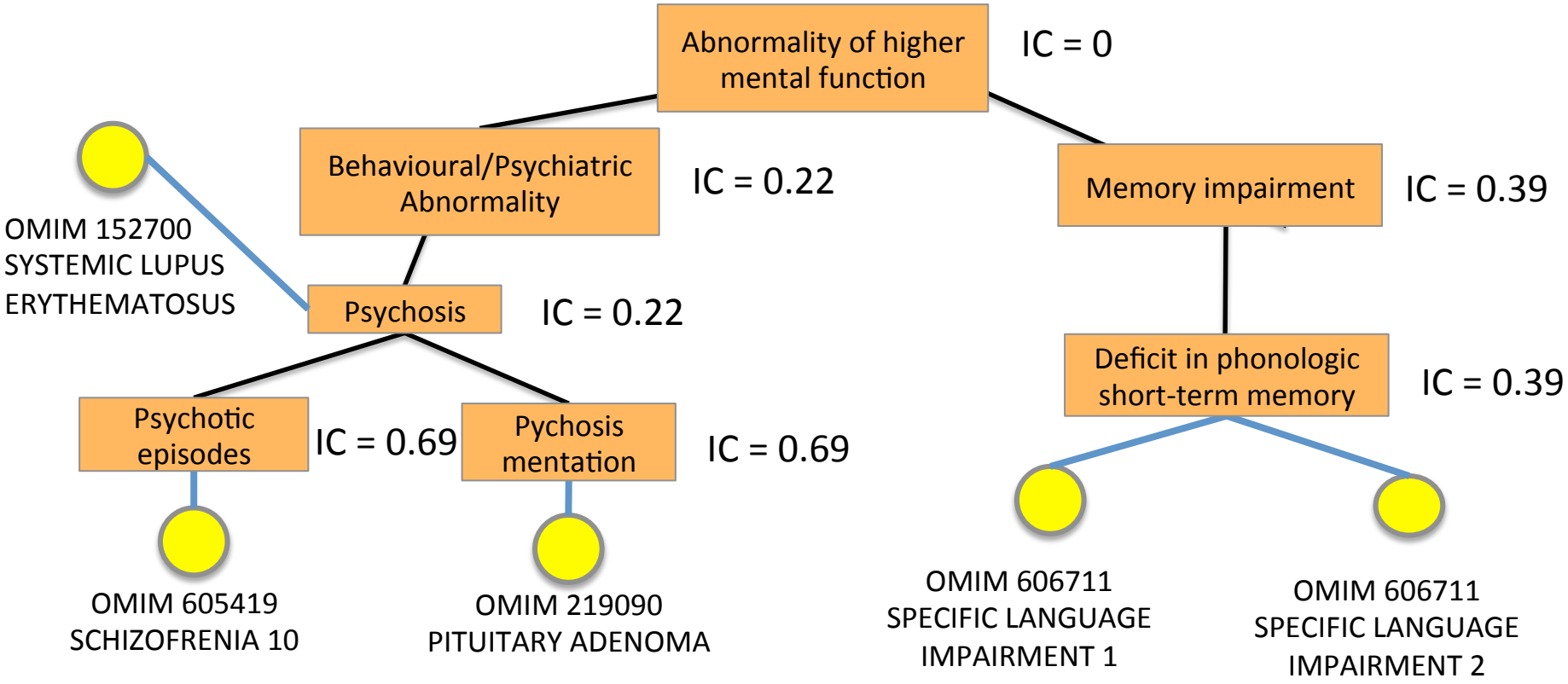
- Probabilidad de cada término p(t)



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base: Similitud Semántica

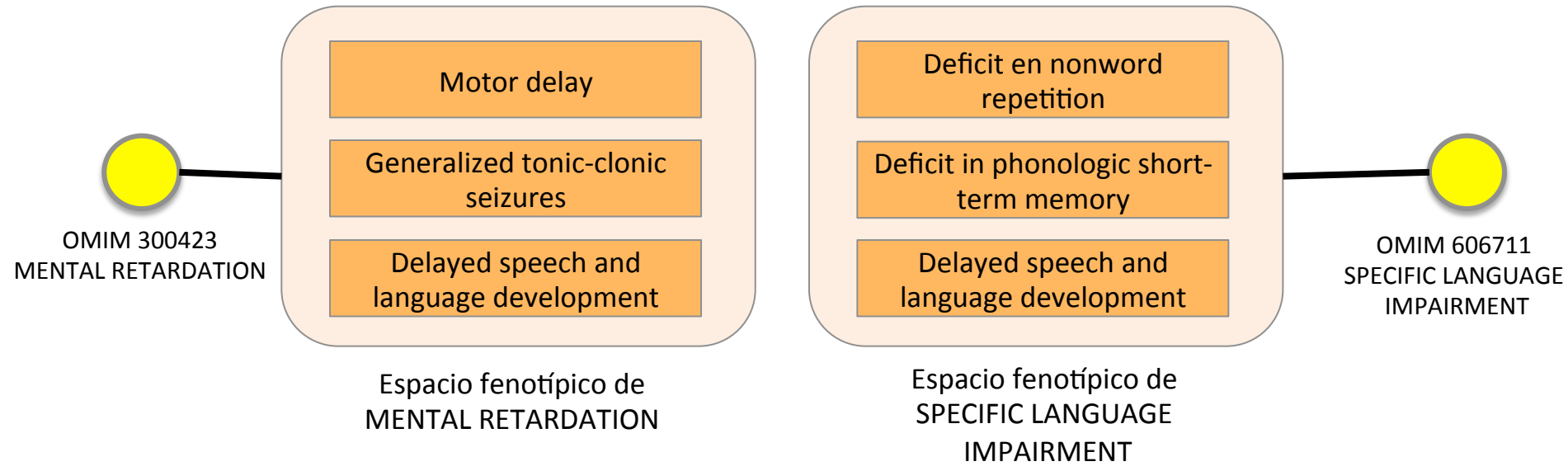
- IC = - log(p(t))



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

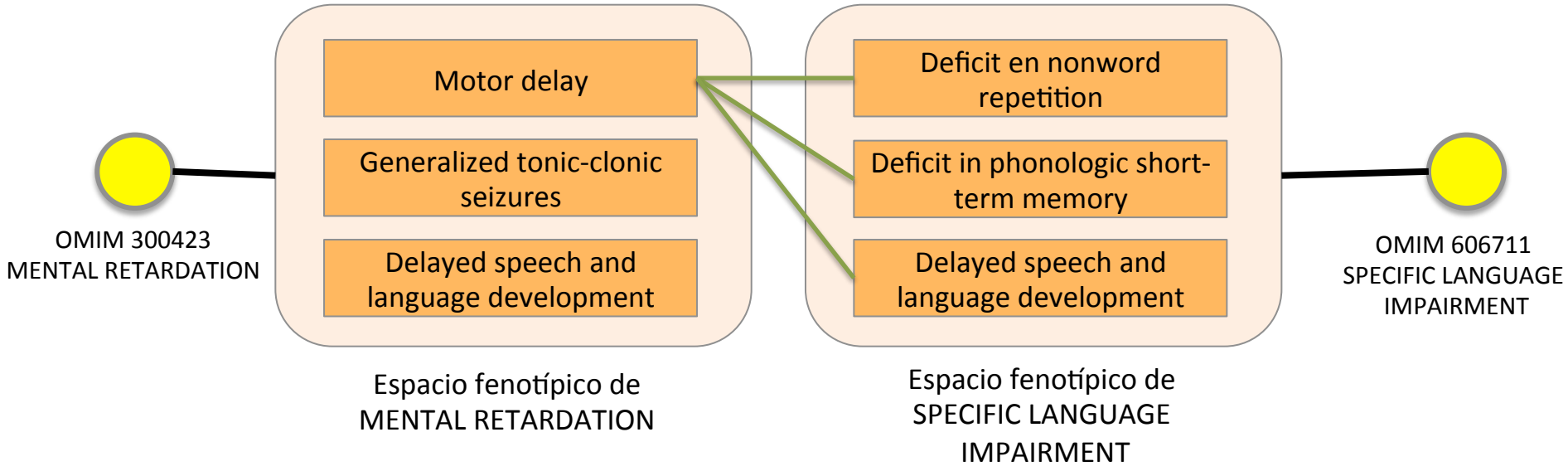
- Una enfermedad no se define por un solo fenotipo sino por un conjunto de fenotipos.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

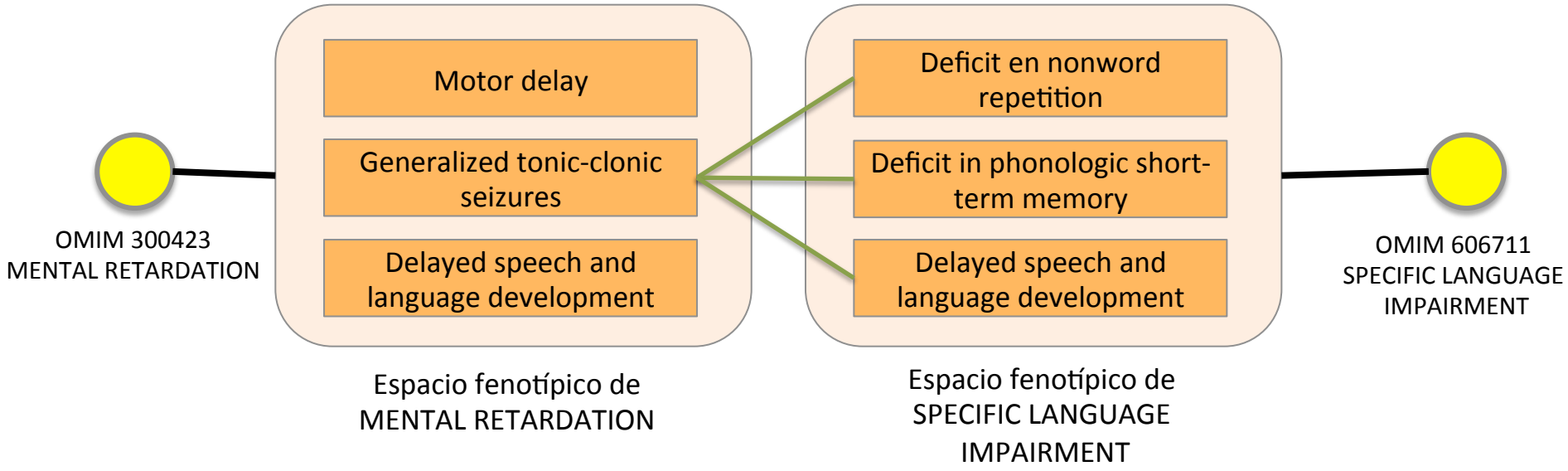
- Para comparar dos grupos de fenotipos se comparan uno a uno los fenotipos de las dos enfermedades.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

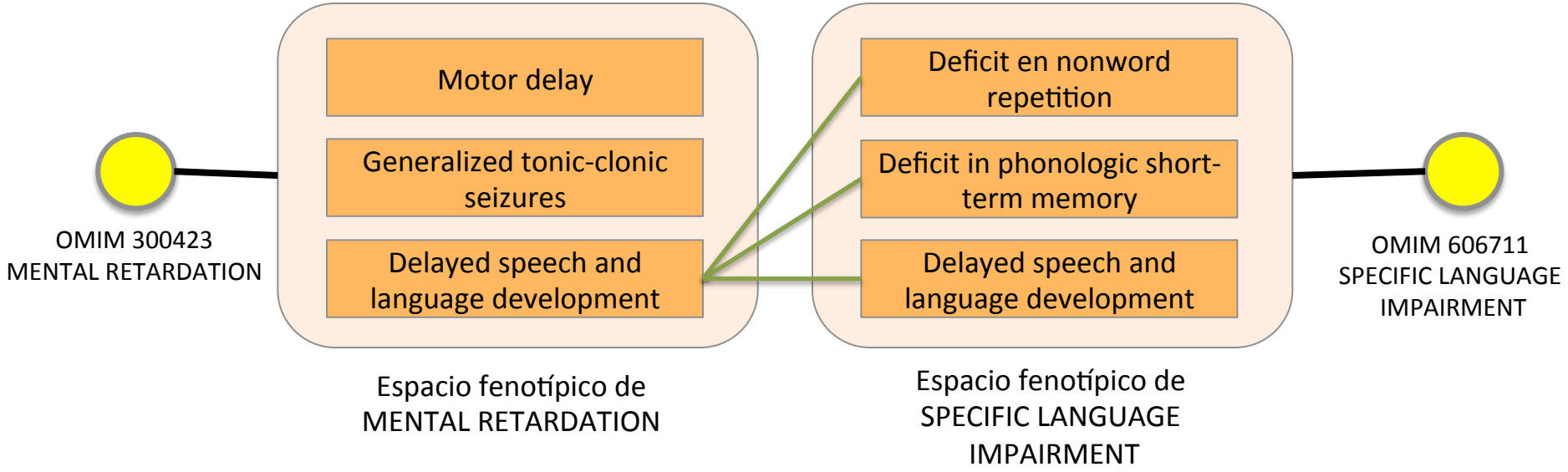
- Para comparar dos grupos de fenotipos se comparan uno a uno los fenotipos de las dos enfermedades.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

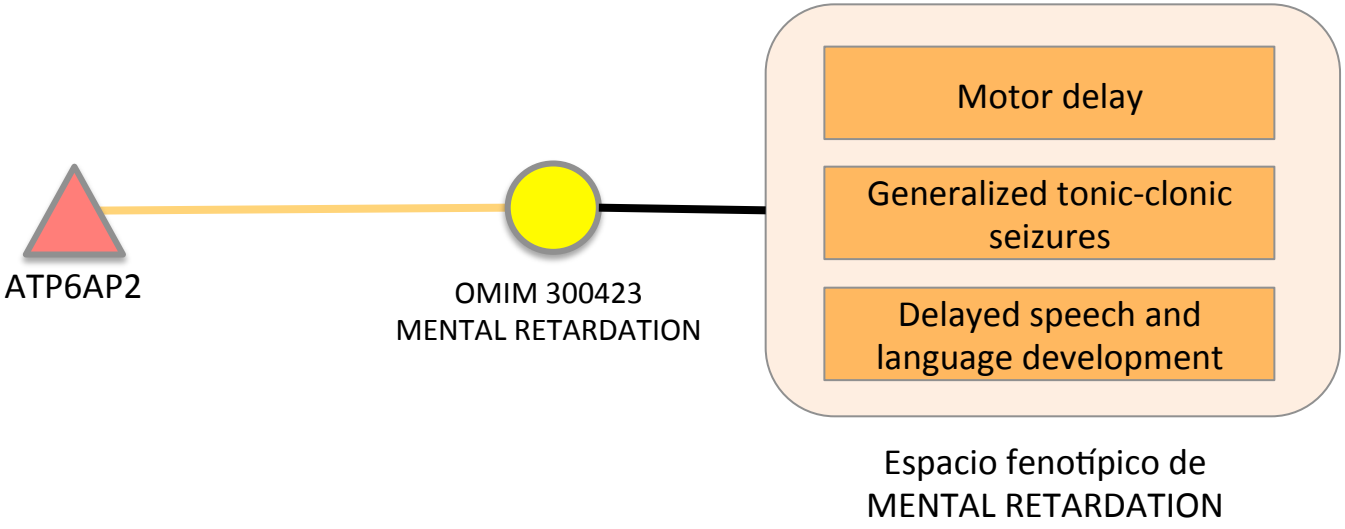
- Para comparar dos grupos de fenotipos se comparan uno a uno los fenotipos de las dos enfermedades.
- El valor de similitud fenotípica entre las dos enfermedades es el **promedio de todas estas comparaciones**



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

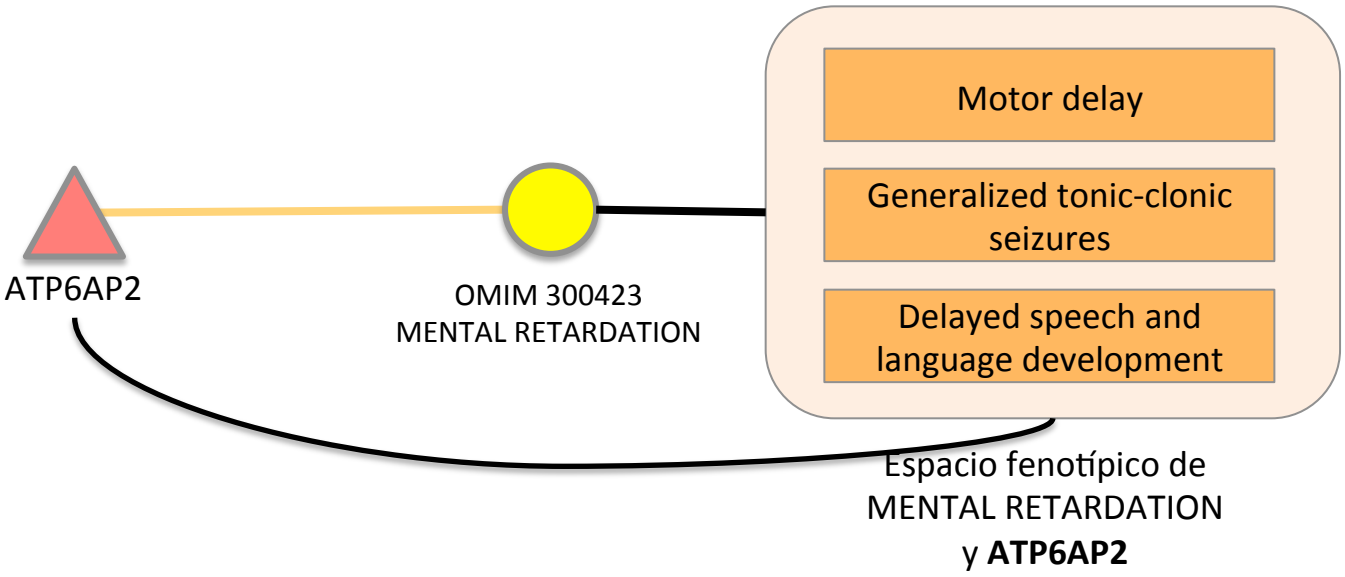
- A partir Orphanet también es posible definir un conjunto de fenotipos a enfermedades raras.
- Las relaciones entre un conjunto de fenotipos y los genes se establecen a partir de las relaciones conocidas: gen-omim.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

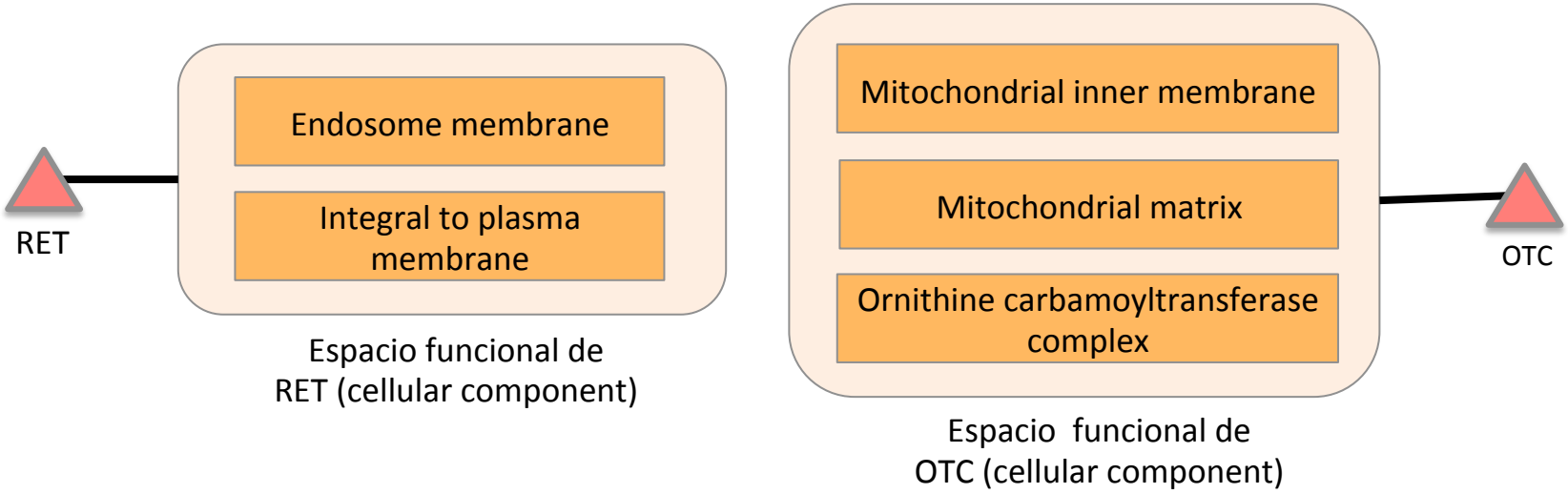
- A partir dela información de **Orphanet** también es posible definir un conjunto de fenotipos a **enfermedades raras**.
- Las relaciones entre un conjunto de fenotipos y un **gen** se establecen a partir de las **relaciones conocidas**: gen-omim.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

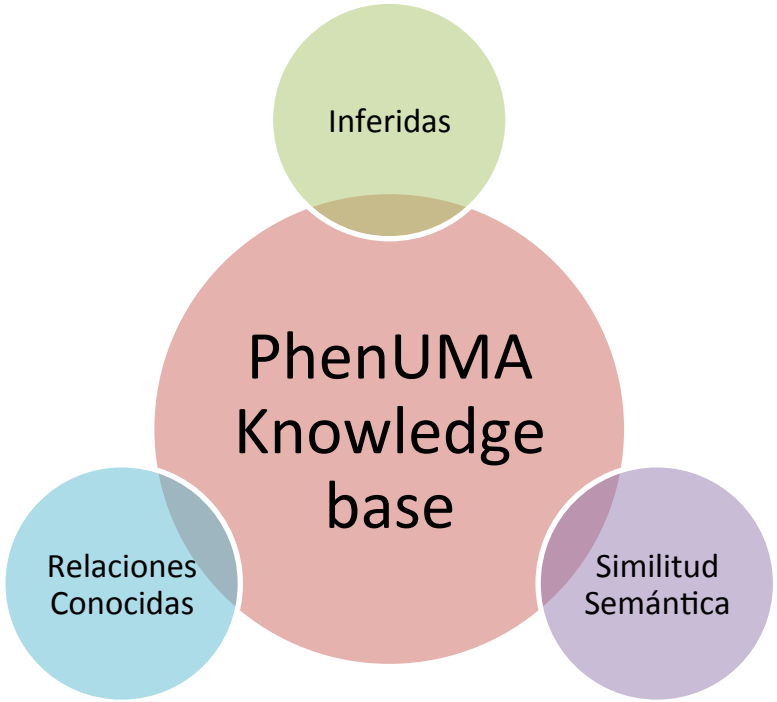
- El cálculo del a similitud semántica funcional entre **genes** es similar en el caso de la Gene Ontology (GO).
- El lugar de fenotipos patológicos los genes se relacionan a partir de **procesos biológicos**, **componentes celulares** y **funciones moleculares**.



Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Knowledge Base : Similitud Semántica

- A partir de los espacios fenotípicos y funcionales se han creado estos tipos de redes de similitud semántica:



OMIM-OMIM SimSem from HPO
ORPHA-ORPHA SimSem from HPO
Gene-Gene SimSem from HPO

Gene-Gene SimSem from GO (Biological Process)
Gene-Gene SimSem from GO (Cellular Component)
Gene-Gene SimSem from GO (Molecular Function)

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

Tipos de Redes

Entrada	Total Relaciones	
Genes	Gene-Gene Known Protein-Protein Interaction from STRING	Conocidas
	Gene-Gene Known Metabolic Interaction (Veeramani et al.)	
	Gene-OMIM Known from OMIM	
	Gene-ORPHA Known from Orphanet	
	Gene-Gene Inferred from OMIM	Inferidas
	Gene-Gene Inferred from Orphanet	
	Gene-Gene SimSem from HPO	Similitud Semántica
	Gene-Gene SimSem from GO (Biological Process)	
	Gene-Gene SimSem from GO (Cellular Component)	
	Gene-Gene SimSem from GO (Molecular Function)	
	ORPHA-ORPHA SimSem from HPO	
	OMIM-OMIM SimSem from HPO	

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Tipos de Redes

Tipos de Redes

Entrada	Tipos de Redes	
Enfermedades OMIM	Gene-OMIM Known from OMIM	Conocidas
	OMIM-OMIM Inferred from Genes	Inferidas
	Gene-Gene SimSem from HPO	Similitud Semántica
	Gene-Gene SimSem from GO (Biological Process)	
	Gene-Gene SimSem from GO (Cellular Component)	
	Gene-Gene SimSem from GO (Molecular Function)	
	OMIM-OMIM SimSem from HPO	

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Tipos de Redes

Tipos de Redes

Entrada	Tipos de Redes	
Enfermedades Orphanet	Gene-ORPHAKnown from OMIM	Conocidas
	ORPHA-ORPHAInferred from Genes	Inferidas
	Gene-Gene SimSem from HPO	Similitud Semántica
	Gene-Gene SimSem from GO (Biological Process)	
	Gene-Gene SimSem from GO (Cellular Component)	
	Gene-Gene SimSem from GO (Molecular Function)	
	ORPHA-ORPHA SimSem from HPO	

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Tipos de Redes

Tipos de Redes

- En consultas de fenotipos las redes de salida consisten en el conjunto de genes, enfermedades OMIM o enfermedades raras (Orphanet) que más se parece al conjunto de fenotipos de entrada.

Entrada	Tipos de Redes
Fenotipos	Genes
	Enfermedades (OMIM)
	Enfermedades raras (Orphanet)

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Tipos de Redes

Nivel de Confianza

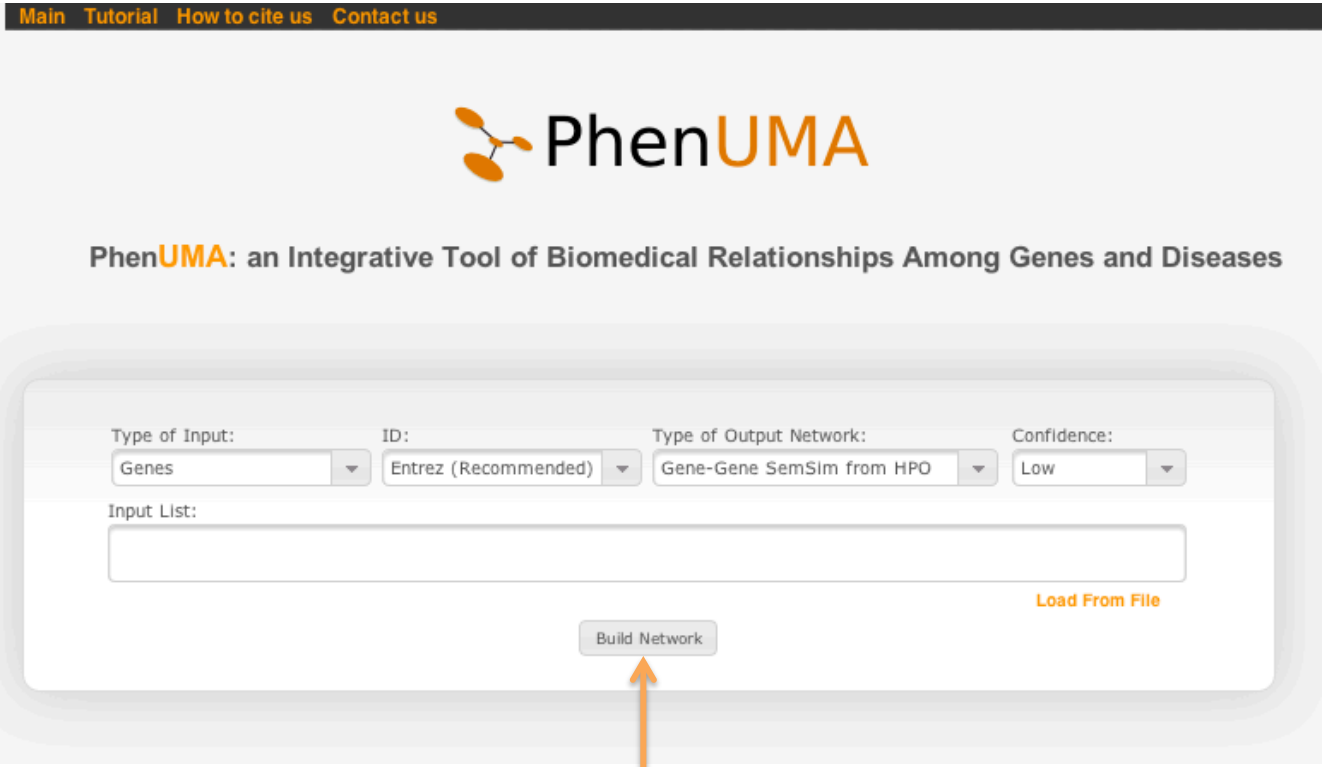
- Cortes y Niveles de Confianza

Red	Total Relaciones	Relaciones TOP 2% (Low)	Relaciones TOP 1%(Medium)	Relaciones TOP 0.5% (High)
ORPHA-ORPHA SimSem from HPO	3804649	75924	37702	18875
OMIM-OMIM SimSem from HPO	7500235	149689	74935	37483
Gene-Gene SimSem from HPO	1245519	24902	12437	6216

Red	Total Relaciones	Relaciones TOP 0.5% (Low)	Relaciones TOP 0.2% (Medium)	Relaciones TOP 0.1%(High)
Gene-Gene SimSem from GO (Biological Process)	99073951	486982	198621	95023
Gene-Gene SimSem from GO (Cellular Component)	113426652	565739	226854	112115
Gene-Gene SimSem from GO (Molecular Function)	80118330	397683	169626	80391

Entrada a PhenUMA → 3.Opciones de la Red de Salida → Knowledge Base

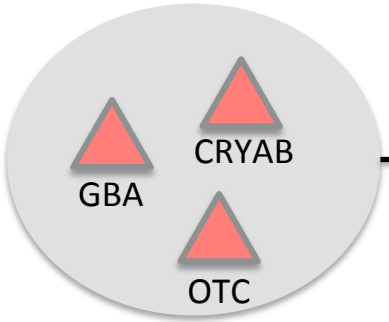
Construcción de Redes



Construcción de la Red

Construcción de Redes

1. Genes de Interés



2. Tipo de Red de Salida:

Ej: Similitud Fenotípica con otros genes (Gene-Gene SimSim from HPO)

3. Nivel de Confianza

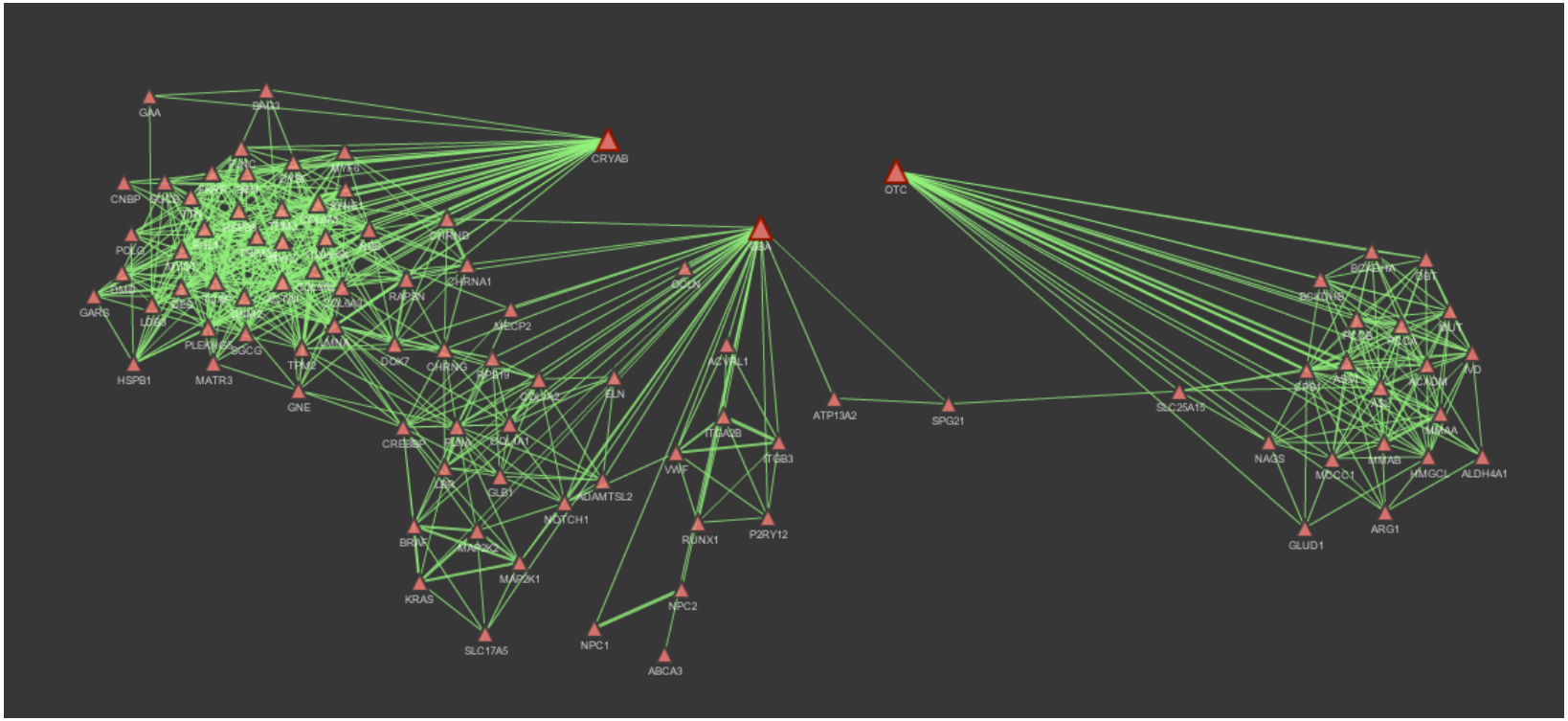
Ej: Low

¿Con que otros genes se relacionan mis genes de interés fenotípicamente?

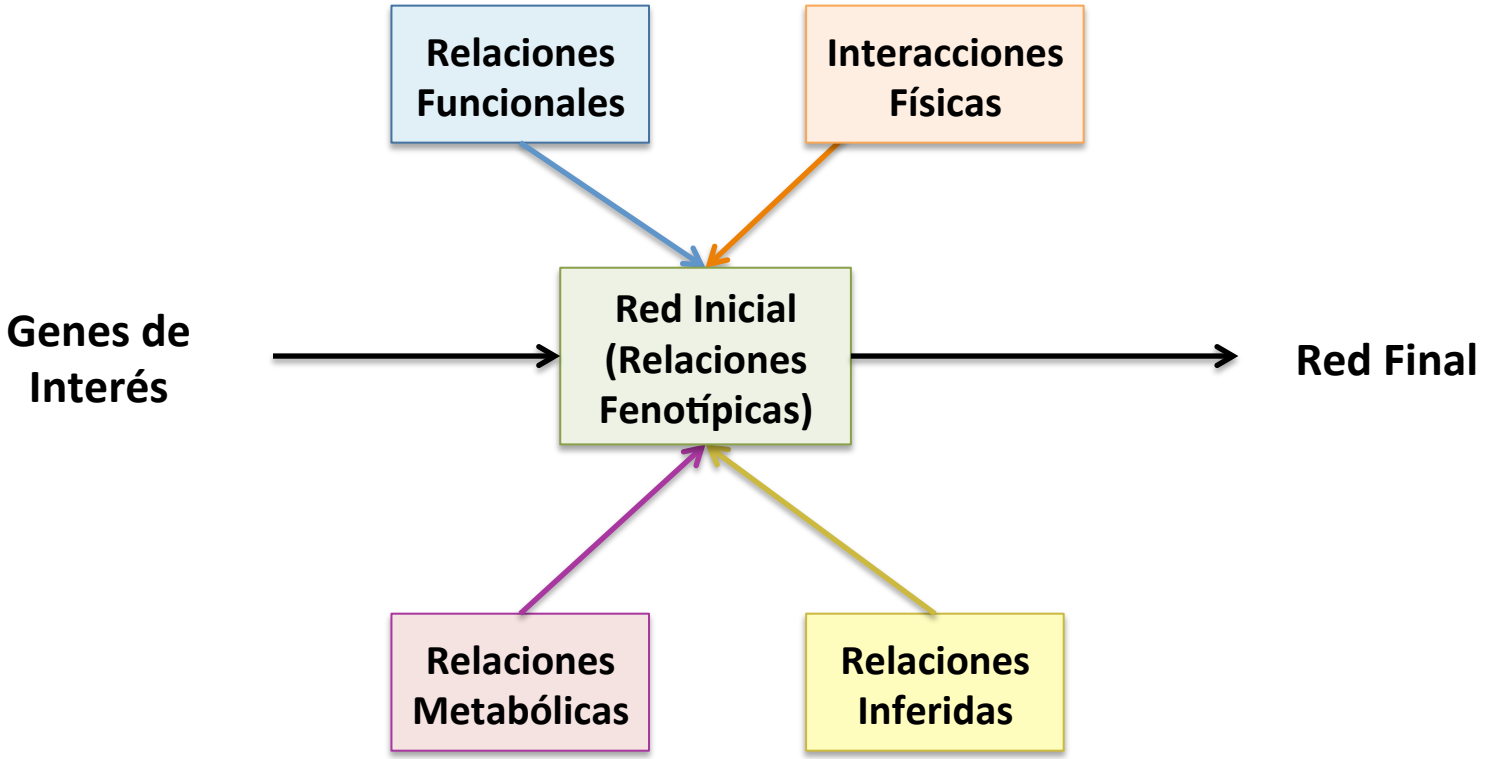
Red Inicial (Relaciones Fenotípicas)

Construcción de Redes

- **Red Inicial**: Tipo de relación seleccionada (relaciones fenotípicas)



Knowledge Base : Construcción de Redes



Construcción de Redes → Enriquecimiento

Construcción de Redes

- **Red Final**: Enriquecida con el resto de relaciones entre los genes de la red resultante