

A semantic query interface for the OGO platform

José Antonio Miñarro-Giménez (jose.minyarro@um.es)

Mikel Egaña Aranguren, Ph.D. (mikel.egana.aranguren@gmail.com)

Francisco García-Sánchez, Ph.D. (frgarcia@um.es)

Jesualdo Tomás Fernández-Breis, Ph.D. (jfernand@um.es)

Faculty of Computer Science
University of Murcia
Spain

ITBAM (DEXA)
Bilbo 2010

Overview

Orthologs

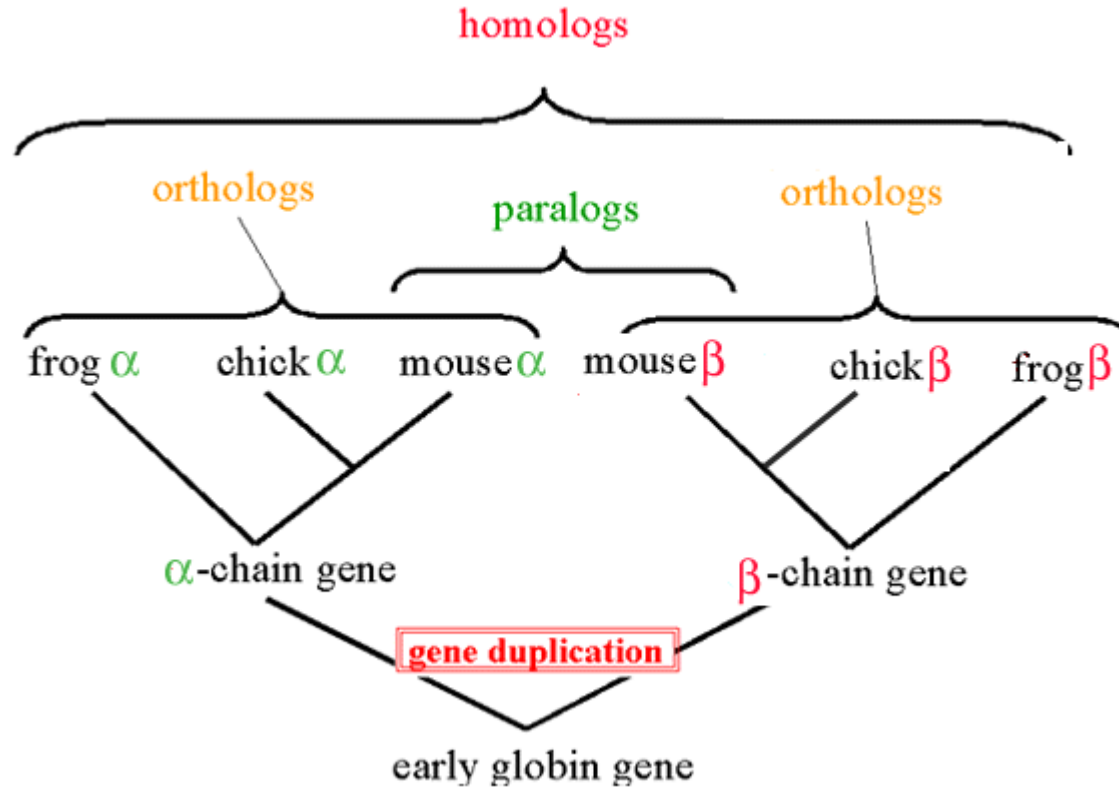
Information about orthologs and diseases

OGO system

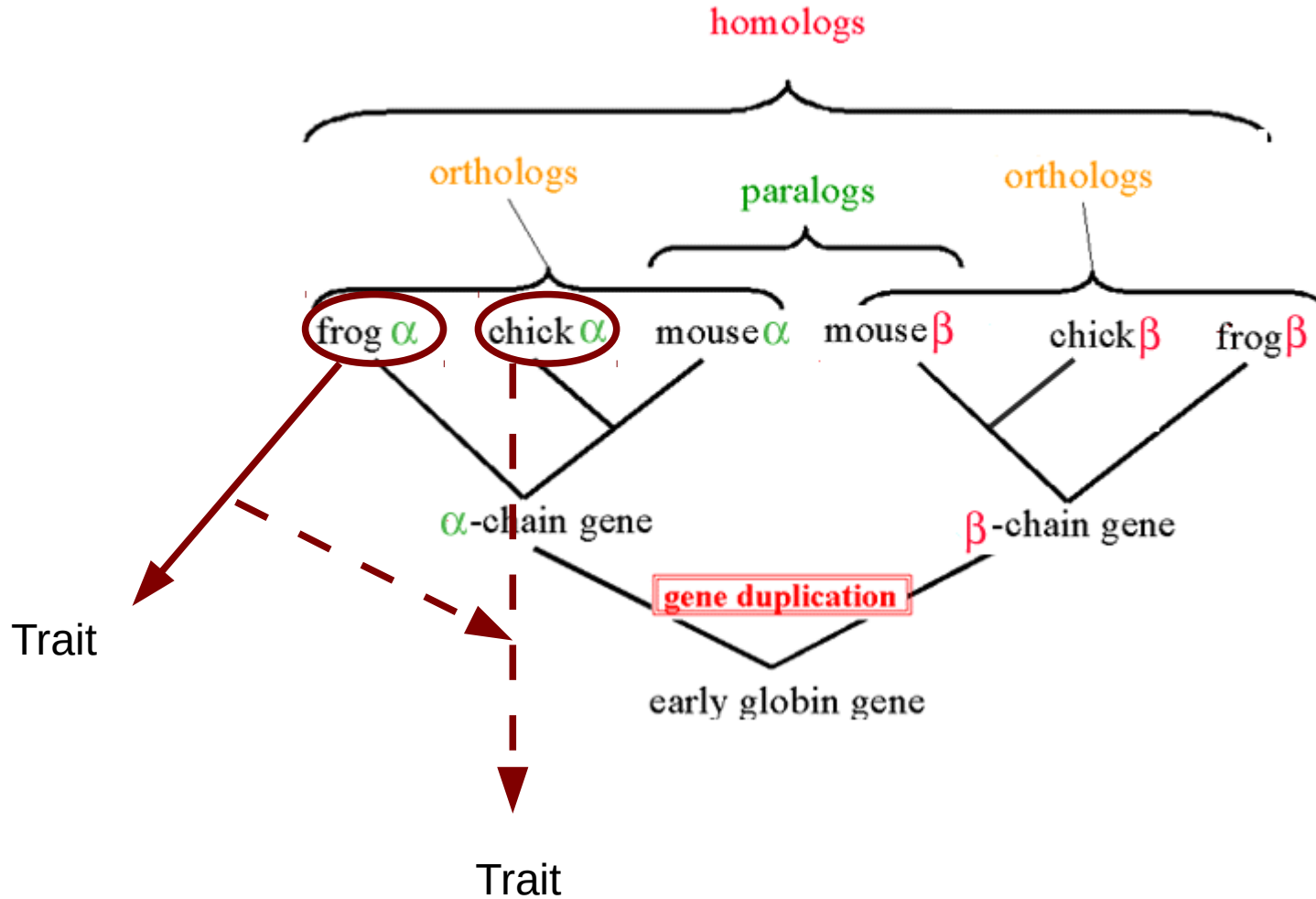
A semantic query interface for the OGO system

Sample query

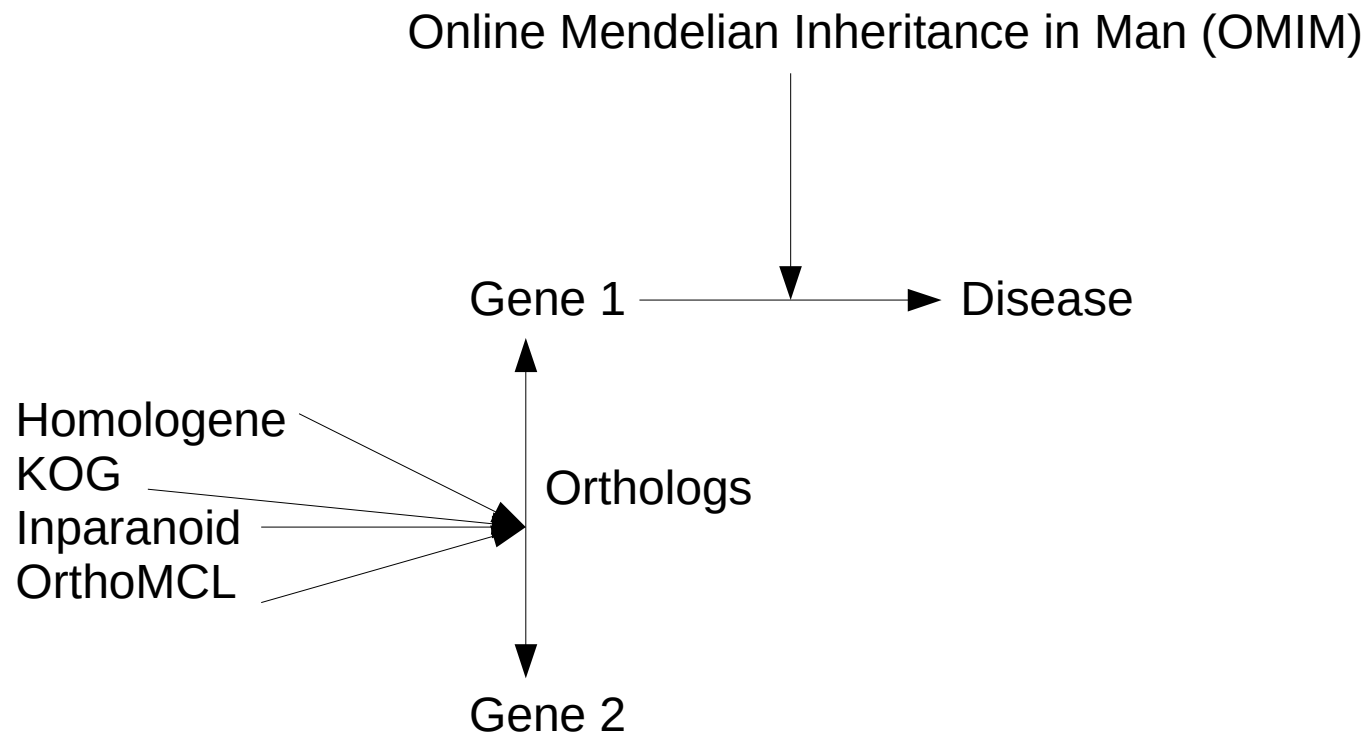
Ortholog sequences



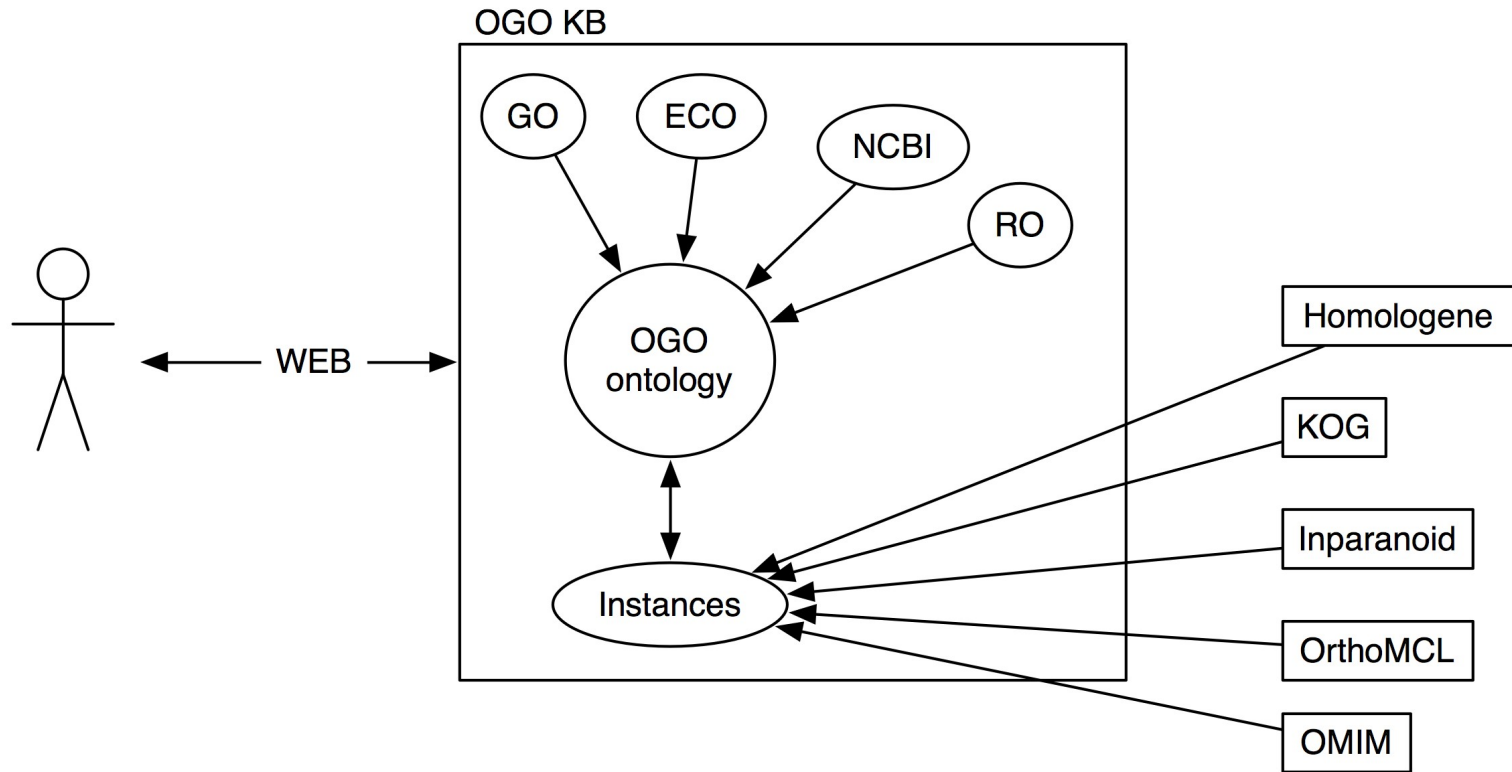
Ortholog sequences



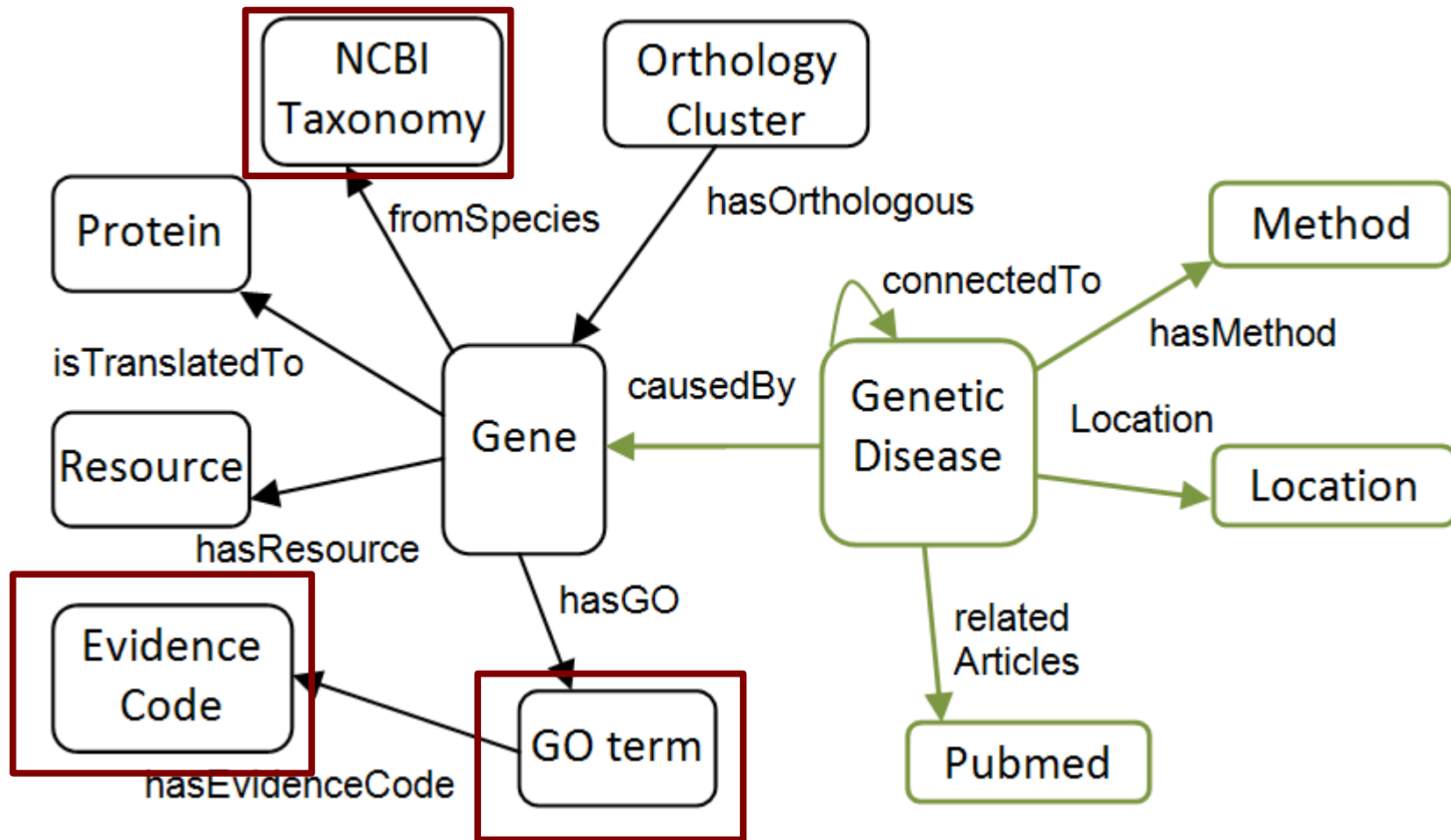
Orthologs and genetic diseases



OGO system



OGO ontology



OGO ontology: imported ontologies

Gene Ontology (OBOF): molecular function, biological process and cellular component of gene products

Evidence Codes Ontology (Candidate OBOF):
GO annotations evidence codes

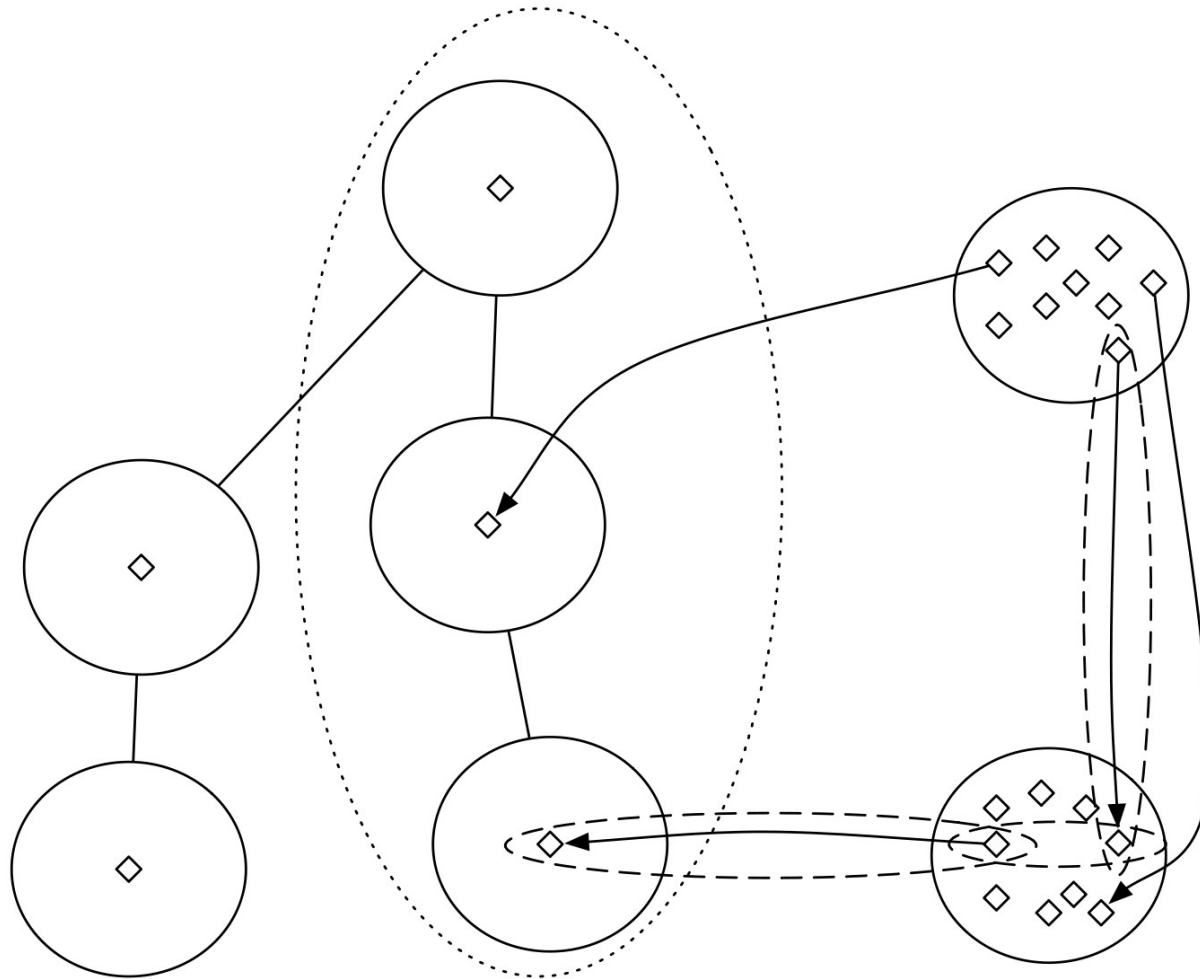
OBO Relationship Types (Candidate OBOF):

Gene product **participates in** some (molecular function or biological process)

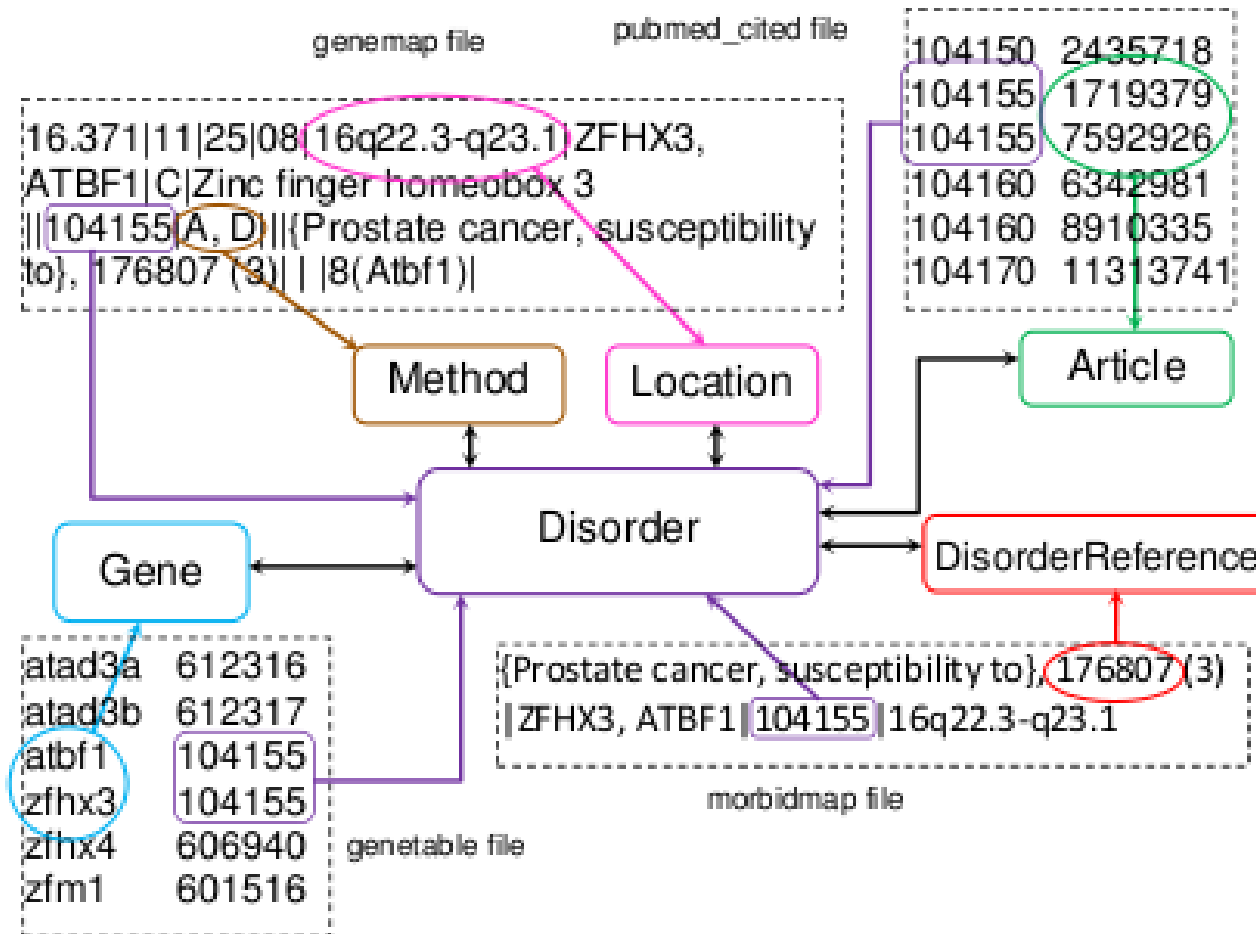
Gene product **located in** some cellular component

NCBI taxonomy: organisms classification

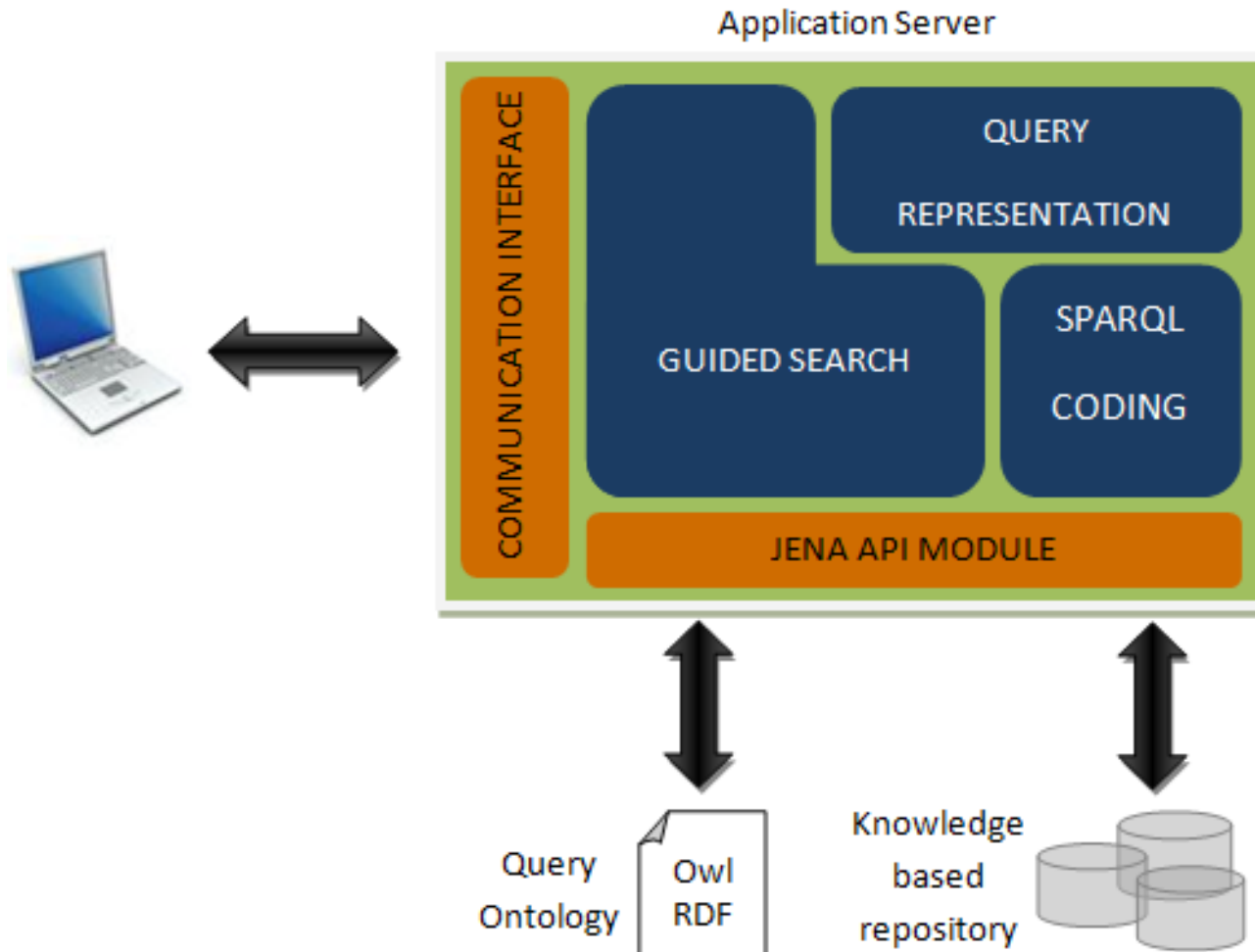
Imported ontologies: OWL punning



OGO ontology: mappings to OMIM



Implementation of the OGO system



Interfaces of the OGO system

The image shows a screenshot of the OGO Portal interface. The page has a yellow header with 'UM | Home' on the right. Below the header is a grey bar with the 'OGO' logo on the left and 'OGO Portal' on the right. A green sidebar on the left contains navigation links: 'Home', 'Contact Us', 'Search', 'Disease Search', and 'Advanced Search'. The 'Search' and 'Advanced Search' links are highlighted with blue and red boxes, respectively. A blue arrow points from the 'Search' link to the 'Keyword based querying' text. A red arrow points from the 'Advanced Search' link to the 'Semantic querying' text. The main content area is titled 'SEARCH FORM' and includes a dropdown menu set to 'Orthology' with the text 'Search for orthologous genes by ID or name:'. Below this is a search input field and a 'Search' button. Underneath are two dropdown menus: 'Select Resources' (with options: ALL_RESOURCES, Homologene, Inparanoid, KOG, OrthoMCL) and 'Select Organisms' (with options: ALL_ORGANISMS, Aedes aegypti, Agrobacterium tumefaciens, Anopheles gambiae, Apis mellifera). At the bottom of the search form are five checkboxes: 'Show synonyms', 'Show Gene Ontology terms', 'Show proteins', 'Show resources', and 'Show Disorders'. The footer of the page reads 'University of Murcia'.

Keyword based querying

Semantic querying

Semantic interface



The screenshot displays the OGO Semantic Interface. At the top left is the OGO logo, consisting of three orange rings forming the letters 'OGO'. Below the logo, a breadcrumb trail reads 'Estás en: Proyecto OGO | OGO'. The main content area is titled 'SEARCH GUIDED' in red, underlined text. Below this title, there are two main sections: 'Search for' and 'Query Requirements'. The 'Search for' section features a large, empty text input field. To its right are two buttons: 'Select Concept' and 'Delete Concept'. The 'Query Requirements' section features a large, empty text input field. To its right are two buttons: 'Add new requirements' and 'Delete requirement'. At the bottom of the interface, there are two buttons: 'Execute Query' and 'Clear Query'.

OGO

Estás en: [Proyecto OGO](#) | OGO

SEARCH GUIDED

Search for

Select Concept

Delete Concept

Query Requirements

Add new requirements

Delete requirement

Execute Query

Clear Query

Semantic interface

SEARCH

Search for

Select Concept
Delete Concept

Query Requirements

Add new requirements
Delete requirement

Execute Query Clear Query

Select a concept

- Protein
- Pubmed
- Gene
- + ● Method
- + ● Functional Category
- + ● organisms
- + ● Evidence code
- Genetic disease
- Resource
- + ● GO_term
- Cluster of Orthologous genes

Select a Requirement

- Gene[0]
 - Identifier **RANGE** string
 - Name **RANGE** string
 - hasResource **RANGE** Resource
 - hasFunctionalCategory **RANGE** Functional Category
 - fromSpecies **RANGE** organisms
 - hasGO **RANGE** GO_term
 - isTranslatedTo **RANGE** Protein
 - hasOrthologous **DOMAIN** Cluster of Orthologous genes
 - causedBy **DOMAIN** Genetic disease
 - label literal[0]

Edit Requirement

Subject

Property

Object

Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

SEARCH GUIDED

Search for

Gene[0]
Genetic disease[1]

Select Concept
Delete Concept

Query Requirements

Add new requirements
Delete requirement

Execute Query
Clear Query

Select a concept

- Protein
- Pubmed
- Gene
- + ● Method
- + ● Functional Category
- + ● organisms
- + ● Evidence code
- Genetic disease
- Resource
- + ● GO_term
- Cluster of Orthologous genes

The image shows a web-based search interface. At the top, the text 'SEARCH GUIDED' is underlined. Below it, the 'Search for' section contains a text input field with the query 'Gene[0] Genetic disease[1]'. To the right of this field are two buttons: 'Select Concept' and 'Delete Concept'. Below the search field is the 'Query Requirements' section, which is currently empty. To its right are two buttons: 'Add new requirements' and 'Delete requirement'. At the bottom left are two buttons: 'Execute Query' and 'Clear Query'. On the right side of the interface is a 'Select a concept' menu, which is a scrollable list of concept types, each preceded by a yellow circle. The concepts listed are: Protein, Pubmed, Gene, Method, Functional Category, organisms, Evidence code, Genetic disease, Resource, GO_term, and Cluster of Orthologous genes. Red arrows point from the 'Select Concept' button to the 'Select a concept' menu, and from the 'Delete Concept' button to the 'Select a concept' menu. Another red arrow points from the 'Select a concept' menu back to the search field, indicating that a selected concept is added to the query.

Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

SEARCH GUIDED

Search for

Gene[0]
Genetic disease[1]

Select Concept
Delete Concept

Query Requirements

Add new requirements
Delete requirement

Execute Query
Clear Query

Select a Requirement

- Gene[0]
 - Genetic disease[1]->causedBy Gene[0]
 - Identifier RANGE string
 - Name RANGE string
 - hasResource RANGE Resource
 - hasFunctionalCategory RANGE Functional Category
 - fromSpecies RANGE organisms
 - hasGO RANGE GO_term
 - isTranslatedTo RANGE Protein
 - hasOrthologous DOMAIN Cluster of Orthologous genes
 - causedBy DOMAIN Genetic disease
 - label literal[0]
- Genetic disease[1]
 - causedBy Gene[0]
 - Location RANGE string
 - Name RANGE string
 - OimReference RANGE string
 - GeneReference RANGE string
 - relatedArticles RANGE Pubmed
 - connectedTo RANGE Genetic disease
 - hasMethod RANGE Method
 - hasDisorderReference RANGE Genetic disease
 - causedBy RANGE Gene
 - hasDisorderReference DOMAIN Genetic disease
 - connectedTo DOMAIN Genetic disease
 - label literal[1]

Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

SEARCH GUIDED

Search for

Gene[0]
Genetic disease[1]

Select Concept
Delete Concept

Query Requirements

Add new requirements
Delete requirement

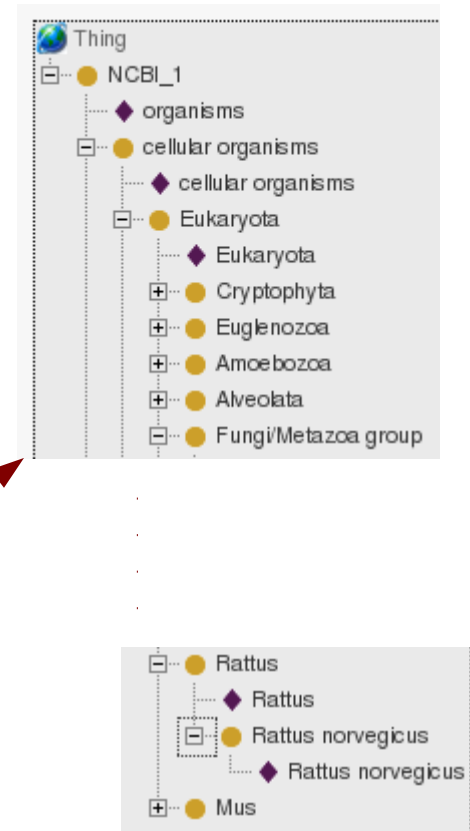
Execute Query
Clear Query
Add Requirement
Clear Requirement

Edit Requirement

Subject: Gene[0]

Property: fromSpecies

Object: organisms Edit Object



Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

SEARCH GUIDED

Search for

Gene[0]
Genetic disease[1]

Select Concept

Delete Concept

Query Requirements

Gene[0] ->fromSpecies->Rattus norvegicus

Add new requirements

Delete requirement

Execute Query

Clear Query

Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

SEARCH GUIDED

Search for

Gene[0]
Genetic disease[1]

Select Concept

Delete Concept

Query Requirements

Genetic disease[1]->Name->prostate cancer,
susceptibility to
Genetic disease[1]->causedBy->Gene[0]
Gene[0]->fromSpecies->Rattus norvegicus
Cluster of Orthologous
genes[5]->hasOrthologous->Gene[0]
Cluster of Orthologous
genes[5]->hasOrthologous->Gene[6]

Add new requirements

Delete requirement

Execute Query

Clear Query

Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

```
@prefix ncbi: <http://um.es/ncbi.owl>.
@prefix ogo: <http://miuras.inf.um.es/ontologies/OGO.owl>.
SELECT
  ?Gene_0
  ?Genetic_disease_1
WHERE {
  ?Gene_0 ogo:fromSpecies ncbi:NCBI_10116
  ?Genetic_disease_1 ogo:Name ?literal_4 .
  FILTER (regex(?literal_4,"Prostate cancer, susceptibility to")) .
  ?Genetic_disease_1 ogo:causedBy ?Gene_2 .
  ?Cluster_of_Orthologous_genes_3 ogo:hasOrthologous ?Gene_2 .
  ?Cluster_of_Orthologous_genes_3 ogo:hasOrthologous ?Gene_0 .
}
```

Sample query

Ortholog genes of the gene that causes prostate cancer on *Rattus norvegicus*?

Genetic disease[0]	Gene[1]
104155, prostate cancer, susceptibility to, zinc finger homeobox 3	pex12, 116718
	rgd1560268, zfhx3.predicted, 307829
	rgd1563022, zfhx4.predicted, 310250
600020, neurofibrosarcoma, prostate cancer, susceptibility to, max-interacting protein 1	clec5a, 679787
	ensrnog00000026306, loc684510
	loc689617, 689617
	loc689629, 689629
	max, mgc124611, 60661
	mxi.wr, mxi1, 25701
	tgap1, 294892

Query grammar

```
Query ::= "SELECT" ListVar (WhereClause)?  
ListVar ::= Var (Var)*  
WhereClause ::= "WHERE {" ConditionClause (ConditionClause)* "}"  
ConditionClause ::= [VarCondition | LiteralCondition] "."  
VarCondition ::= [Var | Individual] Property [Var | Individual]  
LiteralCondition ::= [Var | Individual] Property [Var | Individual] "."  
"FILTER (regex (" Var ", " Literal "))"
```

Var -> This term represents a variable in the query which can be matched to any concept or individual in the ontology.

Individual -> This term represents a concept or individual identified by an URI in the ontology.

Property -> This term represents a relationship or property identified by an URI in the ontology.

Literal -> This term represents any data value dened by the user.

Future plans

OWL reasoning for querying (OWL 2 QL?)

Pellet Integrity Constraint Validator (Pellet ICV):

- OWL as schema language for RDF (CWA)

- Check the gathered information

More bio-ontologies

Clinical archetypes for querying (ISO 13606):

- exchange of ortholog/disease information in a standard biomedical research setting

Conclusions

Orthologs and diseases: new hypotheses

OGO provides a resource for exploiting such combined information

Semantic query interface: “Complex” queries easily (No SPARQL syntax)

<http://miuras.inf.um.es/~ogo/>

Acknowledgements

Spanish Ministry for Science and Education
(grant TSI2007-66575-C02-02)

Comunidad Autónoma de la Región de Murcia
(grant BIO-TEC 06/01-0005)

Fundación Séneca, Servicio de Empleo y
Formación (grant 07836/BPS/07)