

# Ontologies et l'éditeur Protégé

---

## Application à la formalisation des concepts de description d'IMGT-ONTOLOGY

Véronique Giudicelli  
LIGM, IGH, UPR CNRS 1142  
141 rue de la Cardonille  
34296 Montpellier Cedex 5  
tel: 04.99.61.99.28 fax: 04.99.61.99.01  
e-mail: Veronique.Giudicelli@igh.cnrs.fr

5 mai 2010

# ONTOLOGIE

En philosophie, l'**ontologie** est l'étude de l'être en tant qu'être, c'est-à-dire l'étude des propriétés générales de ce qui existe.

Une ontologie est un **ensemble structuré de connaissance** dans un domaine particulier comme l'immunogénétique. Une ontologie cherche à représenter le sens des concepts et des relations qui les lient.

Elle comprend une **partie terminologique**, inventaire du vocabulaire pour les métadonnées et les concepts, la déclaration des instances (valeurs) et les propriétés particulières qui expriment **relations** entre concepts et instances.

# World Wide Web

(Créé par Tim Berners Lee en 1990)

=> Pour partager des connaissances et de relier les informations partagées.

The screenshot shows a Mozilla Firefox window with the title bar "IMGT Index - Mozilla Firefox". The address bar displays the URL "http://www.imgt.org/textes/IMGTindex/ontology.html". The page content is the "IMGT Index" website, specifically the "Ontology (IMGT-ONTOLOGY)" section. The page includes a logo for "IMMunoGeneTics Information system®" and a navigation menu with links like "Introduction" and "IMGT-ONTOLOGY axioms and concepts". A yellow callout box highlights a section titled "Citing IMGT-ONTOLOGY:" containing several bibliographic entries. Below this, the "Introduction" section discusses the complexity of immunogenetics and the need for a formal ontology. The "IMGT-ONTOLOGY axioms and concepts" section explains the seven defined axioms and their applications. The bottom of the page includes a footer with the word "Terminé".

# World Wide Web

- Les ressources sous forme de page HTML (Langage HTML)
- Les relations entre ses pages qui se traduisent par des liens hypertextes

The screenshot shows a Mozilla Firefox browser window with two tabs open. The left tab is titled "IMGT Index - Mozilla Firefox" and displays the IMGT Index website. The right tab is titled "Source of: http://imgt.cines.fr/textes/IMGIndex/ontology.html - Mozilla Firefox" and displays the source code of the ontology page.

**IMGT Index Tab Content:**

- IMGT Index**
- Here you are IMGT web resources > IMGT index
- Ontology (IMGT-ONTOLOGY)**
- An ontology is a concise and non ambiguous description of the IMGT-ONTOLOGY [1], is the first ontology which allows the ma
- Citing IMGT-ONTOLOGY:**
  - Giudicelli, V. and Lefranc, M.-P. "Ontology for Immunogenetics"
  - Lefranc, M.-P. et al. "IMGT-ONTOLOGY for Immunogenetics"
  - Lefranc, M.-P. et al. "IMGT-Choreography for Immunogenetics"
- Introduction**
- The molecular synthesis and genetics of the immunoglobulin formal specification of the terms to be used in the domain of immunogenetics
- IMGT-ONTOLOGY includes a controlled vocabulary and annotates terms with the same meaning. This provides a semantic representation of the vocabulary.
- IMGT-ONTOLOGY axioms and concepts**
- Seven IMGT-ONTOLOGY axioms have been defined [1,6-8]: "IDENTIFICATION", "DESCRIPTION", "CLASSIFICATION", "NUMEROTATION", "LOCALIZATION", "ORIENTATION", and "OBTENTION". They constitute the Formal IMGT-ONTOLOGY or IMGT-Kaleidoscope [9,10].
- The IMGT-ONTOLOGY concepts of identification, description, classification, numerotation, localization, orientation and obtention were generated from these axioms and described in ref. [1,2,6-8].
- Click here for:
  - Figures illustrating some of the IMGT-ONTOLOGY axioms and concepts (IMGT Education).
  - Correspondence between the IMGT-ONTOLOGY concepts and the IMGT Scientific chart rules (IMGT Scientific chart).
  - Scientific chart rules and ontologies report [PDF](#)
- Marie-Paule Lefranc, François Ehrenmann, Patrice Duroux and Véronique Giudicelli  
(D1.2 ImmunoGrid, The European Virtual Human Immune System Project, IST-2004-028069)  
Describes the IMGT-ONTOLOGY concepts of identification, description and classification at the molecular level, generated from the IDENTIFICATION, DESCRIPTION and CLASSIFICATION axioms of the Formal IMGT-ONTOLOGY (IMGT-Kaleidoscope).

**Source Tab Content:**

```
<h3 id="keyConcepts">IMGT-ONTOLOGY axioms and concepts</h3>

<p>
Seven IMGT-ONTOLOGY axioms have been defined <a href="#refs">[1,6-8]</a>:
<strong><a href="Identification.html">IDENTIFICATION</a>,
<a href="Description.html">DESCRIPTION</a>,
<a href="Classification.html">CLASSIFICATION</a>,
<a href="Numerotation.html">NUMEROTATION</a>,
'LOCALIZATION',
<a href="Orientation.html">ORIENTATION</a>,</strong>
and <strong><a href="Obtention.html">OBTENTION</a>'</strong>.
They constitute the Formal IMGT-ONTOLOGY or IMGT-Kaleidoscope <a href="#refs">[9,10]</a>.</p>
<p>
The IMGT-ONTOLOGY concepts of identification, description, classification,
numerotation, localization, orientation and obtention
were generated from these axioms and described in ref. <a href="#refs">[1,2,6-8]</a>.
</p>
```

Line 643, Col 1

<http://imgt.cines.fr/textes/IMGIndex/ontology.html>

# World Wide Web

## Limites

- Balises HTML: la connaissance est encapsulée dans une couche de présentation
- Introduction de balises propriétaires (Microsoft, Netscape)
- Les liens hypertexte ne permettent pas de qualifier les relations entre les informations
- La qualité des pages est très hétérogène

⇒ Ces informations ne sont pas  
exploitables automatiquement

# World Wide Web Consortium (W3C)

(créé en 1994)

- Développement de technologies (spécifications, guides, logiciels et outils)
- Proposition des standards, pour la croissance et l'exploitation du Web

**En particulier, en 1998, le langage de balisage extensible XML (Extensible Markup Language) devient une recommandation du W3C**

- XML devrait pouvoir être utilisé sans difficulté sur Internet
- Les documents XML devraient être lisibles par l'homme et raisonnablement clairs
- Il devrait être facile de créer des documents XML
- Il devrait être facile d'écrire des programmes traitant les documents XML

# Séquence EMBL AM947564

Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

ID AM947564; SV 1; linear; genomic DNA; STD; HUM; 287 BP.

XX

AC AM947564;

XX

DT 01-APR-2008 (Rel. 95, Created)

DT 01-APR-2008 (Rel. 95, Last updated, Version 1)

XX

DE Homo sapiens partial IGHV gene for immunoglobulin heavy chain variable  
region, patient 4

XX

KW .

XX

OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae;

OC Homo.

XX

RN [1]

RP 1-287

RA Capello D.;

RT ;

RL Submitted (13-MAR-2008) to the EMBL/GenBank/DDBJ databases.

RL Capello D., Clinical & Experimental Medicine, University of Eastern  
RL Piedmont, Via Solaroli, 17 Novara, 28100, ITALY.

XX

RN [2]

RA Capello D., Martini M., Gloghini A., Cerri M., Rasi S., Deambrogi C.,  
RA Rossi D., Spina M., Tirelli U., Larocca L., Carbone A., Gaidano G.;

RT "Molecular analysis of immunoglobulin variable genes in HIV-related non  
RT Hodgkin lymphoma reveals implications for disease pathogenesis and  
RT histogenesis";

RL Unpublished.

XX

DR IMGT/LIGM; AM947564; AM947564.

XX

FH Key Location/Qualifiers

FH

FT source 1..287  
FT /organism="Homo sapiens"  
FT /isolate="patient 4"  
FT /mol\_type="genomic DNA"  
FT /cell\_type="B lymphocyte-derived lymphoma"  
FT /tissue\_type="HIV-related B-cell non-Hodgkin's lymphoma"  
FT /rearranged  
FT /db\_xref="taxon:9606"  
FT <1..>287  
FT /codon\_start=1  
FT /gene="IGHV"  
FT /product="immunoglobulin heavy chain variable region"  
FT /protein\_id="CAQ19576.1"  
FT /translation="LSLTCAVSGGSFSGYSWTWIRQTPGKGLEWIGEIKDGGTTNYNAS  
FT LKSRVISISQDMSTFLFLTLSSVT PADTAIYFCARGRSPNIGFDFWGQQTL"

FT V\_region <1..>287  
FT /gene="IGHV"  
FT /product="immunoglobulin heavy chain variable region"

XX

SQ Sequence 287 BP; 59 A; 80 C; 76 G; 72 T; 0 other;  
ctgtccctca cctgcgcgtgt cagtggaggg tcgttcgtg gctattcgtg gacgtggatc 60

Done

# Séquence AM947564 en format XML

Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

Query Results file:///C:/D.../Temp/wgetz

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
- <EMBL_Services xsi:noNamespaceSchemaLocation="http://www.ebi.ac.uk/ebi/schema/EMBL_Services_V1.1.xsd">
- <entry accession="AM947564" version="1" dataClass="STD" taxonomicDivision="HUM" created="2008-04-01" lastUpdated="2008-04-01" releaseCreated="95" releaseLastUpdated="95">
- <description>
  Homo sapiens partial IGHV gene for immunoglobulin heavy chain variable region, patient 4
</description>
<reference>
- <citation id="1" type="submission" date="2008-03-13">
  <author>Capello D.</author>
  - <locator>
    Capello D., Clinical & Experimental Medicine, University of Eastern Piedmont, Via Solaroli, 17 Novara, 28100, ITALY.
  </locator>
</citation>
<citationLocation begin="1" end="287"/>
</reference>
- <reference>
- <citation id="2" type="unpublished">
  - <title>
    Molecular analysis of immunoglobulin variable genes in HIV-related non Hodgkin lymphoma reveals implications for disease pathogenesis and histogenesis
  </title>
  <author>Capello D.</author>
  <author>Martini M.</author>
  <author>Gloghini A.</author>
  <author>Cerri M.</author>
  <author>Rasi S.</author>
  <author>Deambrogi C.</author>
  <author>Rossi D.</author>
  <author>Spina M.</author>
  <author>Tirelli U.</author>
  <author>Larocca L.</author>
  <author>Carbone A.</author>
  <author>Gaidano G.</author>
</citation>
</reference>
<dbreference db="IMGT/LIGM" primary="AM947564" secondary="AM947564"/>
+ <feature name="source"></feature>
+ <feature name="CDS"></feature>
+ <feature name="V_region"></feature>
- <sequence type="genomic DNA" length="287" topology="linear" version="1">
  ctgtccctcacctgcgtcgtggagggtcgtagtggctatcggtggacgtggatccggccagaccccaggaaaaggctggatggatcgccaaatcaaagatggggactactaaactacaatgcgtccctcaagactgagttccatttcacaagataitgcgacgttccctcttc
</sequence>
</entry>
</EMBL_Services>
```

Done

# RDF : Resource Description Framework

Créé en 1999 pour décrire les métadonnées dans l'objectif de traiter l'information automatiquement, de favoriser l'interopérabilité des connaissances et de structurer les informations.

**Métadonnées:** information permettant d'en décrire une autre

- Les métadonnées sont habituellement comprises comme des données à propos des données.
- Un catalogue de bibliothèque contient de l'information (métadonnées) à propos de publications (données)
- Un système de fichier informatique définit des droits de lecture, écriture, etc. (métadonnées) à propos des fichiers (données).

# RDF

- Standard du W3C
- Basé sur XML
- Un début vers un Web structuré
- RDF emploie les URIs comme schéma de nommage  
(pour éviter la confusion, entre termes de sémantique différents)
- On peut décrire n'importe quoi... même une personne
- Grande souplesse quand à l'extensibilité

# RDF

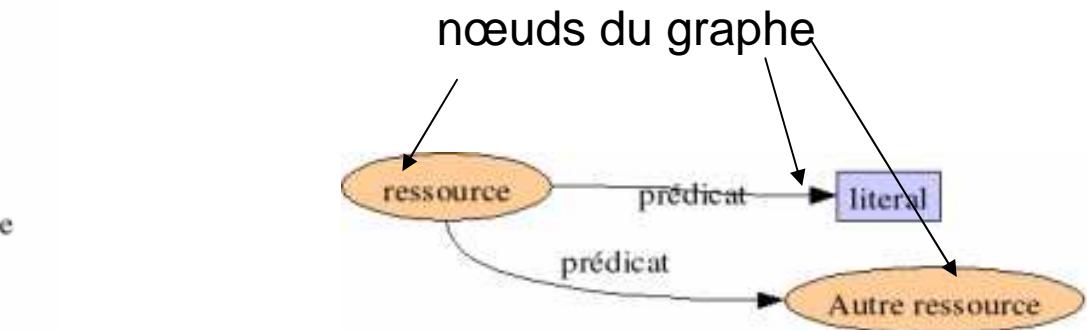
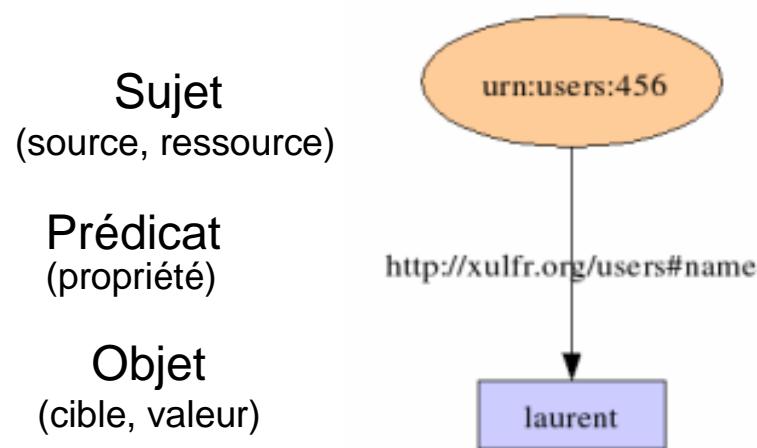
RDF se base sur une description des connaissances à l'aide de phrases simples :

C'est un moyen d'exprimer des relations.

Ces relations sont décrites sous forme de graphe.

Chaque nœud du graphe est une ressource ou une valeur.

Et chaque nœud est relié à un autre par un arc "nommé"



Shémas issus de <http://xulfr.org/wiki/FormatRdf/Introduction>

=> Ceci correspond à un lien qui comporte un sens :  
c'est du Web sémantique

Les objectifs du Web sémantique sont de partager les connaissances et de pouvoir les manipuler automatiquement.

Pour ceci, il faut qualifier le savoir (à l'aide de métadonnées), le formaliser en utilisant une syntaxe extensible (par exemple avec XML) et le structurer pour éviter les duplications.

# ONTOLOGIE

## Sur le plan informatique:

Les ressources sont définies les unes par rapport aux autres selon un graphe. Cette structure permet une automatisation de la manipulation des données.

Différence entre une thésaurus et une ontologie :  
un thésaurus relie des concepts entre eux selon des relations précises : synonyme, homonyme, hiérarchie, terme associé.

L'ontologie ajoute des règles et des outils de comparaison sur et entre les termes, groupes de termes et relations :  
équivalence, symétrie, contraire, cardinalité, transitivité.

=> Dans cet objectif RDF est trop limité en particulier, il ne permet pas de donner la nature des relations (transitive inverse, ...), possède une logique trop limitée pour faire du raisonnement

# ONTOLOGIE

Une **ontologie** définit formellement les **termes** employés pour décrire et représenter un **domaine de connaissance**.

Les ontologies sont destinées à être utilisées par:

- des personnes
- des bases de données
- des applications

ayant besoin de partager des informations

# ONTOLOGIE

Au sein d'une ontologie, les termes sont regroupés sous forme de **concepts** (ou classes) sémantiques: ils définissent un groupe d'individus (instances) possédant des propriétés similaires.

Les ontologies incluent les **définitions**, informatiquement exploitables, des **concepts** élémentaires et de leurs **relations**.

Les ontologies doivent permettre le **partage** et la **réutilisation des connaissances**.

## Bases de connaissances:

Une ontologie ainsi que l'ensemble des **instances** individuelles des **concepts** constituent une base de connaissances. Une frontière subtile marque la fin d'une ontologie et le début d'une base de connaissances.

# OWL (Web Ontology Language)

En 2004, OWL devient une recommandation du W3C

OWL découle de RDF et RDFS, possède des connecteurs logiques, d'exprimer des cardinalités sur les propriétés et d'en spécifier la nature.

Une ontologie formalisée en OWL comprend

Un espace de nom, :

L'entête `<owl :Ontology>` pour décrire l'ontologie

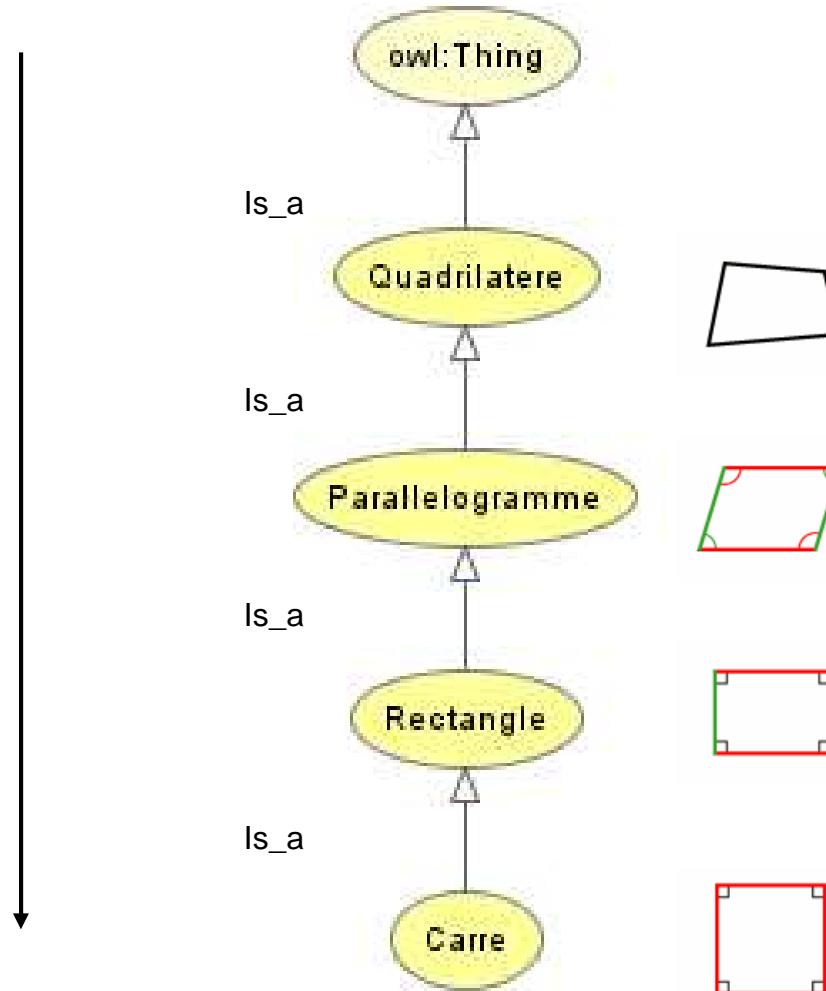
La définition des classes, des propriétés et des instances

# Classes, sous-classes, héritage et spécialisation

Notion de Classe :

Définit un groupe d'individus possédant des propriétés similaires.  
Thing est la classe mère.

spécialisation



Def : polygone à 4 côtés

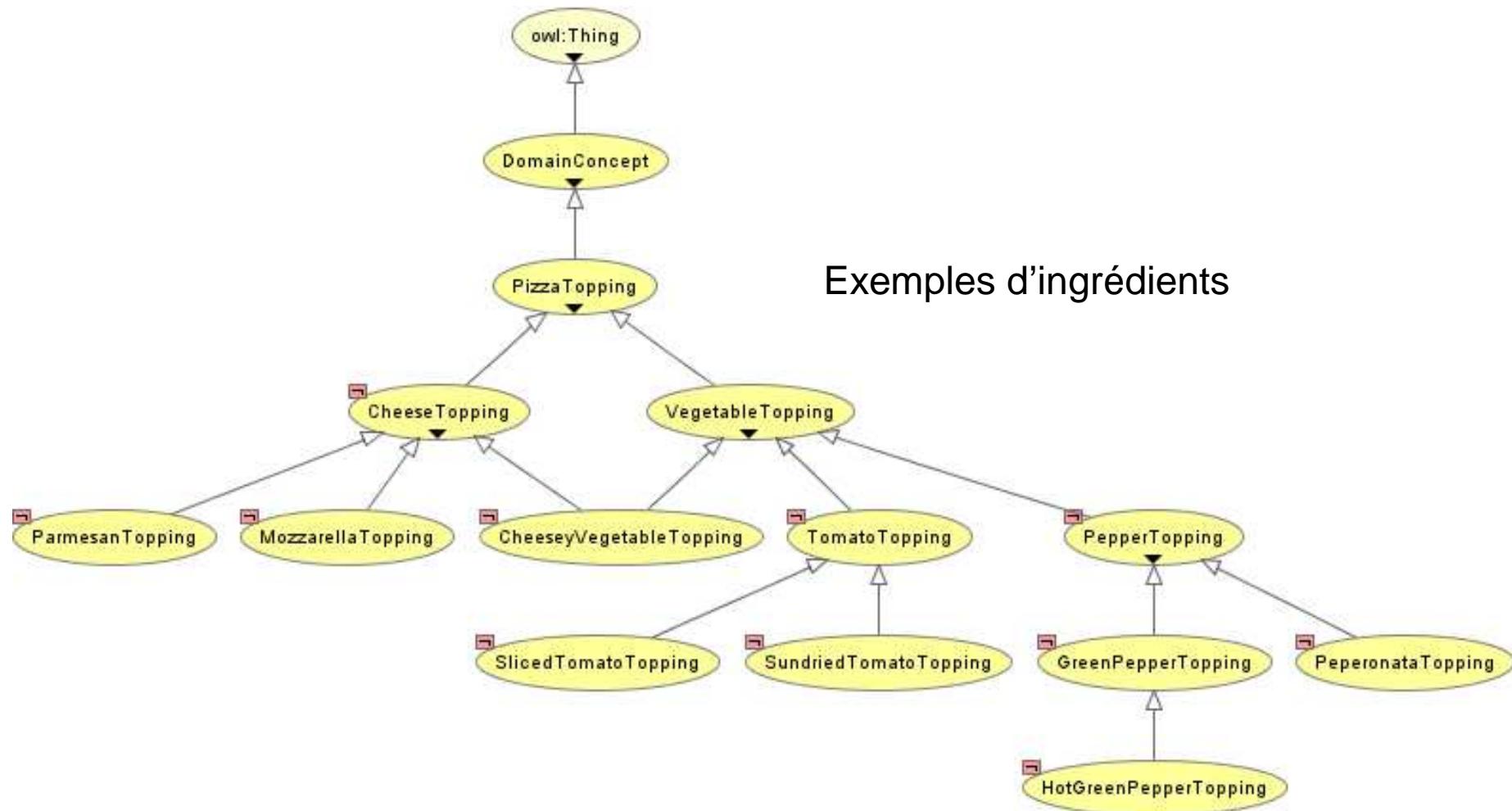
Def: quadrilatère  
(convexe)  
dont les côtés sont  
parallèles deux à deux

Def: Paralléogramme  
à angles droits

Def: rectangle dont les  
côtés sont égaux

# Exemple de hiérarchie : ontologie des pizzas

(exemple fournit avec Protégé)

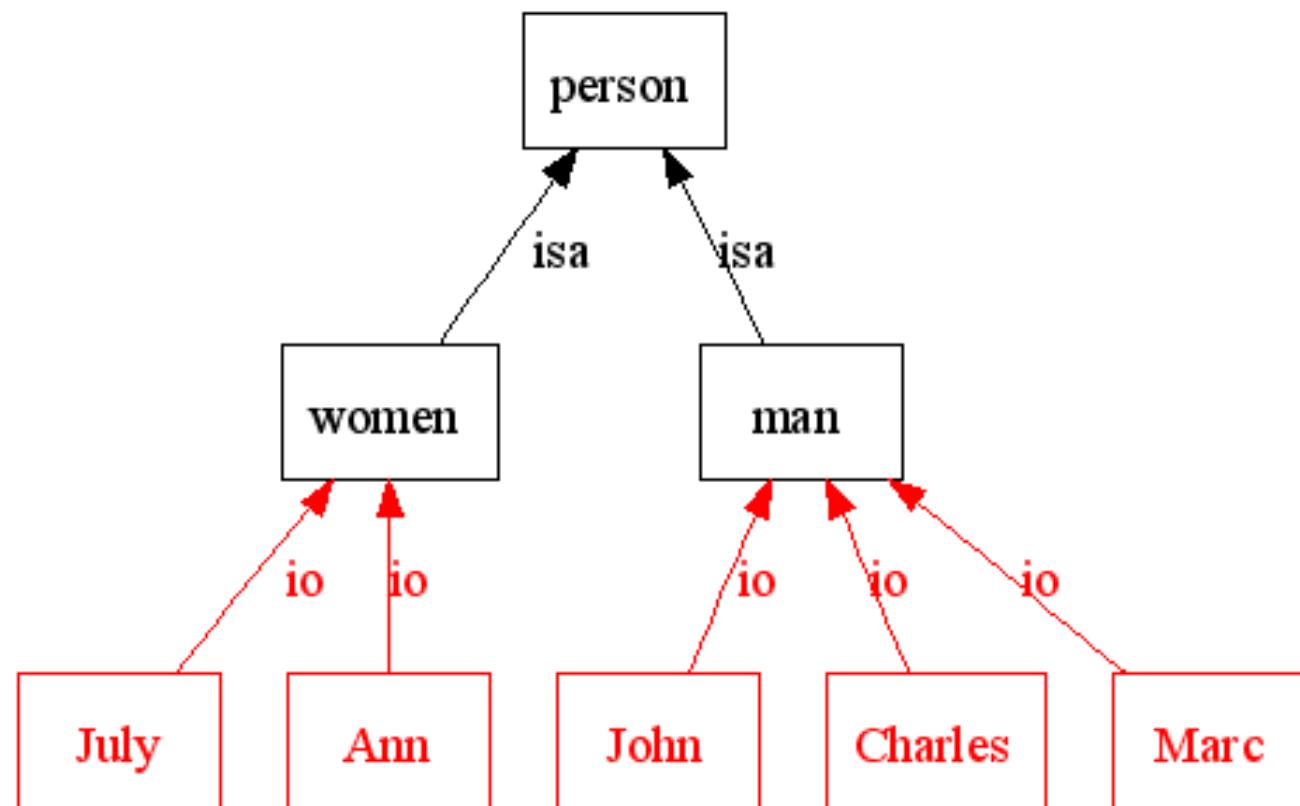


# Exemples de hiérarchie : les concepts d'identification de IMGT-ONTOLOGY



# Instances

Ce sont les « individus » qui peuplent les classes



## Les 3 sous-langages OWL

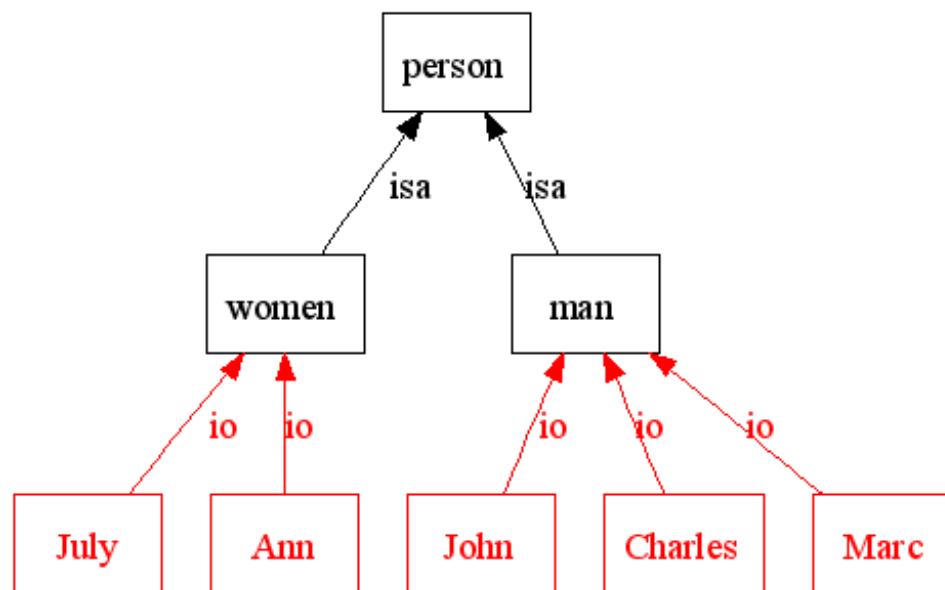
**OWL LITE**: permet d'établir une hiérarchie de concepts simples, contraintes simples.

**OWL DL** (DL pour description logic): comprend toutes les structures de OWL, possède une expressivité plus importante, avec complétude de calcul.

**OWL FULL** expressivité maximale, liberté syntaxique sans garantie de calcul, une classe peut aussi correspondre à l'instance d'une autre classe.

# Classes équivalentes et classes disjointes

- OWL permet de déclarer que 2 classes sont équivalentes: «**equivalentClass**»  
2 classes sont équivalentes lorsqu'elles ont les mêmes instances  
exemple: classe « Voiture » et « Automobile »
- Inversement on peut déclarer que 2 classes sont disjointes: «**disjointWith**»  
2 classes disjointes lorsqu'elles ne peuvent avoir des instances communes  
exemple: les cercles ne fait pas partie des parallogrammes.  
exemple: classe « Women » et « Man »

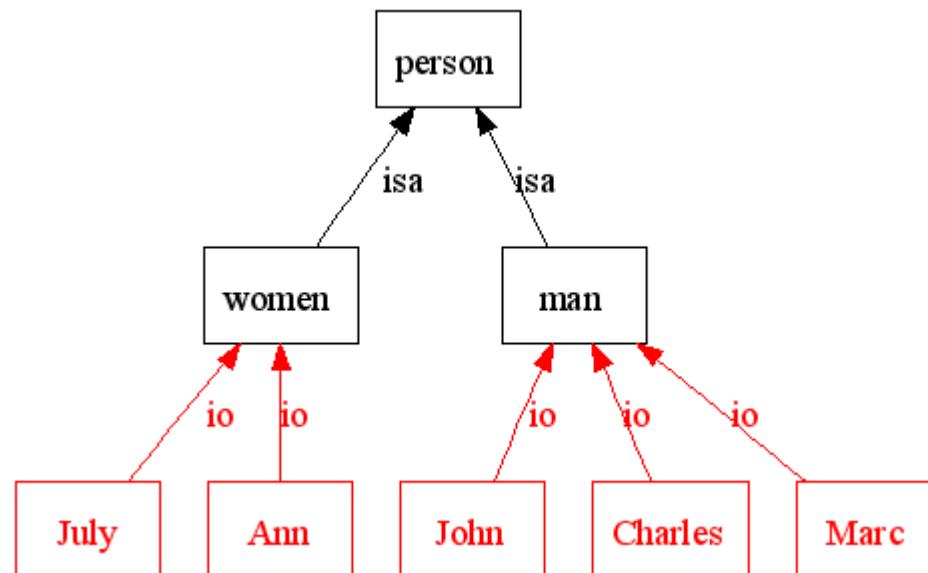




# Instances identiques ou différentes

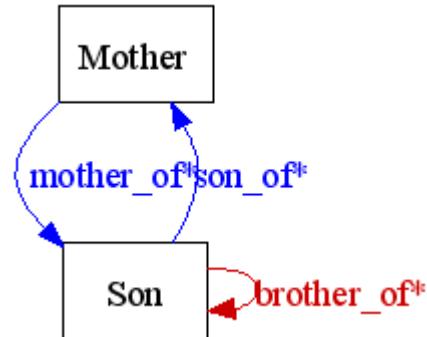
- OWL permet de déclarer que 2 instances sont les mêmes: «**sameAs**»
- Inversement on peut déclarer que 2 instances sont différentes: «**differentFrom**»

On peut déclarer les individus d'une même classe mutuellement différents les uns des autres: « **AllDifferent** »



# Propriétés

- Propriété d'objet (Object property) : relie une instance à une autre instance



- Propriété de type de données (Datatype property): relie une instance à des valeurs.

Son		
son_of	Instance	Mother
brother_of	Instance*	Son
age	Integer*	

OWL propose un héritage des propriétés (à exploiter pour les relations entre les motifs).

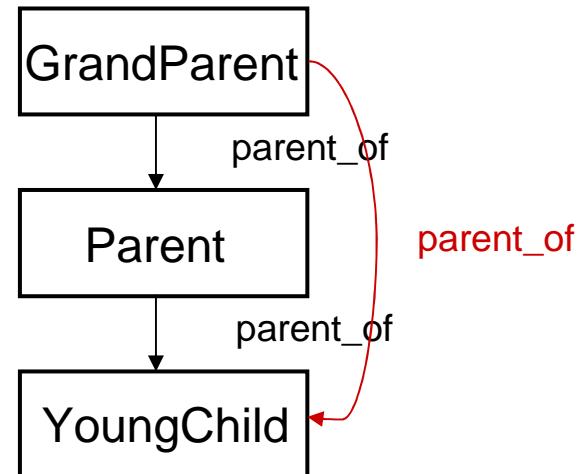
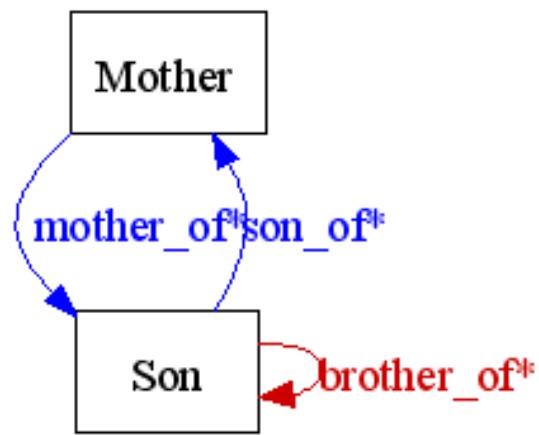
# Propriétés

restrictions globales

domain : classes pour lesquelles est définie la propriété,  
range : classes reliées par la propriété au domain.

domain	relation	range
Son	brother_of	Son
Mother	mother_of	Son
Son	son_of	Mother

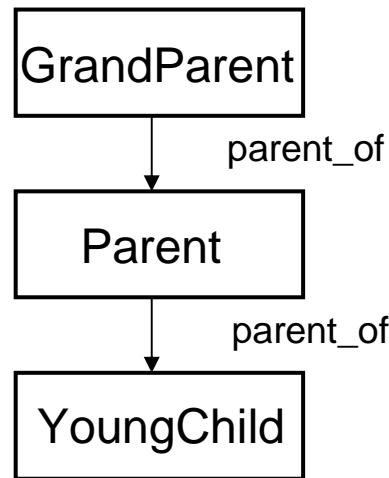
# Caractéristiques des Propriétés: symétriques, transitives, inverses et fonctionnelles



(familles de 3 générations avec de jeunes enfants)

Symétrique: brother\_of  
Transitive: parent\_of  
Inverse: son\_of / mother\_of  
Fonctionnelle : son\_of

# Restriction sur les Propriétés : déclaration au niveau des classes



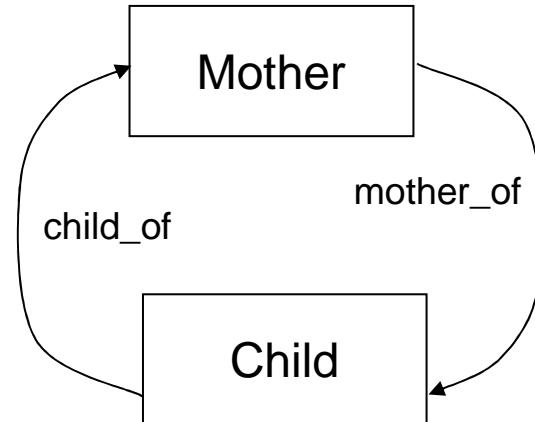
domain	relation	range
GrandParent, Parent	parent_of	Parent, YoungChild

**allValuesFrom:** au niveau d'une classe, permet de restreindre les classes liées par une relation. Exemple : un individu de la classe Parent ne peut être relié par la relation parent\_of qu'à un individu de la classe YoungChild

**someValuesFrom:** au niveau d'une classe, indique qu'un individu a au moins une relation avec un individu de la classe indiquée dans la restriction. Un individu de la classe GrandParent a au moins une relation avec un individu de la classe Parent

**hasValue:** au niveau d'une classe, permet de restreindre la relation à un individu donné

## Restriction sur les Propriétés : déclaration au niveau des classes (cardinalités)



Ex de familles nombreuses de moins de 10 enfants

- Une mère a au moins 3 enfants
- Une mère a au plus 10 enfants
- Un enfant a une et une seule mère

**minCardinality:** toute instance de la classe est liée par la propriété à au moins x individus Exemple: propriété mother\_of de la classe Mother : minCardinality=3

**maxCardinality:** toute instance de la classe est liée par la propriété à au plus x individus Exemple: propriété mother\_of de la classe Mother : maxCardinality=10

**cardinality:** toute instance de la classe est liée par la propriété à exactement x individus Exemple: propriété son\_of de la classe Mother : cardinality=1

# Les ontologies en Biologie

- On s'intéresse aux ontologies qui sont du domaine public.
- Leur nombre augmente régulièrement  
(besoin de définir, de partager).
- Elles couvrent des sujets et domaines différents.

NCBO Bioportal (<http://bioportal.bioontology.org/>)

OBO (Open Biological Ontologies) (<http://obo.sourceforge.net/>)

recensent les ontologies en biologie.

# Gene Ontology.

- GO a été créée en 1998. GO résulte d'une collaboration entre plusieurs bases de données: FlyBase (drosophile), the Saccharomyces Genome Database, et des base de données de génomes (homme et souris), etc.
- GO comprend 3 parties axées sur :
  - **la fonction moléculaire** (fonction des gènes exprimés ex: ATPase activity).
  - **le processus biologique** (rôles biologique généraux de fonctions moléculaire complexes ex: la mitose).
  - **les composants cellulaires** (structures subcellulaires, localisation des complexes macromoléculaires ex: le noyau, le télomère).

The Gene Ontology - Mozilla Firefox

http://www.geneontology.org/ gene ontology

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques Autres sites CNRS

WEB SEARCH

Y! Search

GO The Gene Ontology

the Gene Ontology

Search gene or protein name go!

Downloads Tools Documentation Projects About Contact

## Welcome to the Gene Ontology website!

The Gene Ontology project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases. The project provides a controlled vocabulary of terms for describing gene product characteristics and gene product annotation data from GO Consortium members, as well as tools to access and process this data. [Read more about the Gene Ontology...](#)

### Search the Gene Ontology Database

Search for genes, proteins or GO terms using AmiGO :

MITOSIS

gene or protein name  GO term or ID

AmiGO is the official GO browser and search engine. [Browse the Gene Ontology with AmiGO.](#)

The Gene Ontology project very much encourages input from the community into both the content of the GO and annotation using GO. We are very happy to work with others to ensure that the GO is both complete and accurate, and we also very much encourage communities to submit GO annotations for inclusion in the GO database. [Please contact us.](#)

Quick Links

- Tools
- AmiGO browser A
- OBO-Edit ontology editor
- Ontology downloads
- Annotation downloads
- Database downloads
- Documentation
- GO FAQ
- GO on SourceForge
- Contact GO

News

GO on Twitter t

Minutes for GO Consortium meeting now available (16 days ago) [News item](#)

Job vacancy: Gene Ontology developer at EBI (51 days ago) [News item](#)

Changes to the GO Consortium gene association file format (56

Done

AmiGO: Term Search Results - Mozilla Firefox

http://amigo.geneontology.org/cgi-bin/amigo/search.cgi?search\_query=MITOSIS&search\_constraint=term&ac

gene ontology

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques Autres sites CNRS

WEB SEARCH

Y! - AmiGO: Term Search Results

Perform an action with this page's selected terms... Go!

rel | Accession , Term | Ontology

<input type="checkbox"/>	GO:0007067 : <b>mitosis</b> [show def]	1746 gene products <a href="#">view in tree</a>	biological process
<input type="checkbox"/>	GO:0051337 : a <b>mitosis</b> [show def]	0 gene products <a href="#">view in tree</a>	biological process
<input type="checkbox"/>	GO:0007113 : endomitotic cell cycle [show def]	7 gene products <a href="#">view in tree</a>	biological process
	Query matches synonym "endo <b>mitosis</b> " [exact synonym]		
<input type="checkbox"/>	GO:0010458 : exit from <b>mitosis</b> [show def]	144 gene products <a href="#">view in tree</a>	biological process
<input type="checkbox"/>	GO:0007088 : regulation of <b>mitosis</b> [show def]	435 gene products <a href="#">view in tree</a>	biological process
<input type="checkbox"/>	GO:0055047 : generative cell <b>mitosis</b> [show def]	2 gene products <a href="#">view in tree</a>	biological process
<input type="checkbox"/>	GO:0045840 : positive regulation of <b>mitosis</b> [show def]	141 gene products <a href="#">view in tree</a>	biological process
	Query matches synonyms "up regulation of <b>mitosis</b> " [exact synonym], and 4 more		
<input type="checkbox"/>	GO:0000281 : cytokinesis after <b>mitosis</b> [show def]	37 gene products <a href="#">view in tree</a>	biological process
<input type="checkbox"/>	GO:0045839 : negative regulation of <b>mitosis</b> [show def]	142 gene products <a href="#">view in tree</a>	biological process
	Query matches synonyms "down regulation of <b>mitosis</b> " [exact synonym], and 3 more		
<input type="checkbox"/>	GO:0051727 : cell cycle switching, meiotic to mitotic cell cycle [show def]	0 gene products	biological process

Done

AmiGO: mitosis Details - Mozilla Firefox

http://amigo.geneontology.org/cgi-bin/amigo/term-details.cgi?term=GO:0007067&session\_id=7210amigo1272572568

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques Autres sites CNRS

WEB SEARCH

AmiGO: mitosis Details

CGD Bacillus subtilis

all : all [463707 gene products]

GO:0008150 : biological\_process [349949 gene products]

GO:0016043 : cellular component organization [26585 gene products]

GO:0006996 : organelle organization [13776 gene products]

GO:0048285 : organelle fission [1879 gene products]

GO:0000280 : nuclear division [1782 gene products]

GO:0007067 : mitosis [1746 gene products]

GO:0009987 : cellular process [187265 gene products]

GO:0007049 : cell cycle [7516 gene products]

GO:0022402 : cell cycle process [5185 gene products]

GO:0022403 : cell cycle phase [4193 gene products]

GO:0000279 : M phase [3561 gene products]

GO:0000087 : M phase of mitotic cell cycle [1888 gene products]

GO:0007067 : mitosis [1746 gene products]

GO:0007067 : mitosis [1746 gene products]

GO:0000278 : mitotic cell cycle [3414 gene products]

GO:0000087 : M phase of mitotic cell cycle [1888 gene products]

GO:0007067 : mitosis [1746 gene products]

GO:0022402 : cell cycle process [5185 gene products]

GO:0022403 : cell cycle phase [4193 gene products]

GO:0000279 : M phase [3561 gene products]

GO:0000087 : M phase of mitotic cell cycle [1888 gene products]

GO:0007067 : mitosis [1746 gene products]

GO:0007067 : mitosis [1746 gene products]

GO:0006996 : organelle organization [13776 gene products]

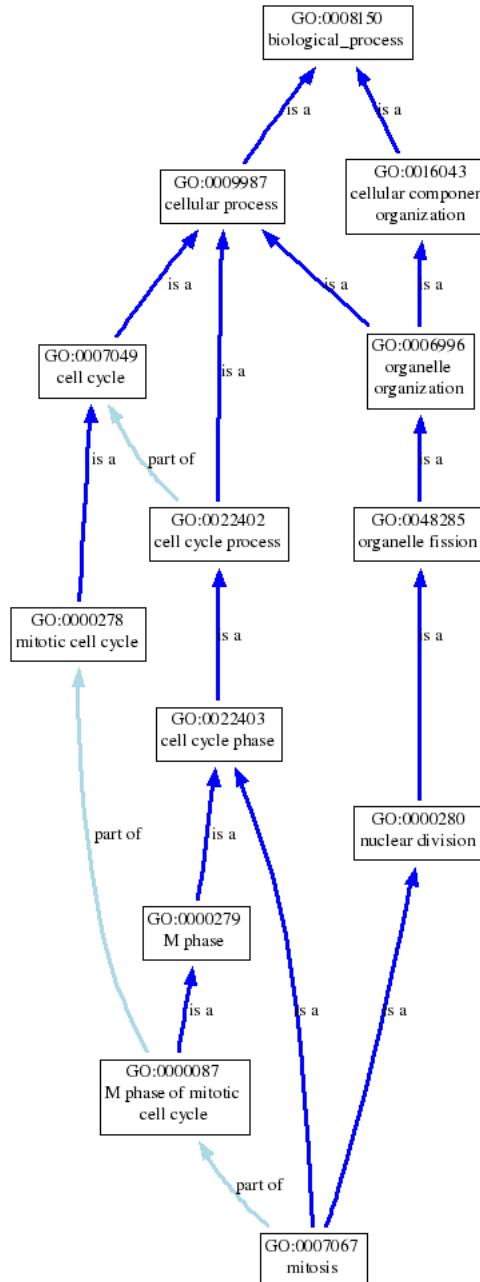
GO:0048285 : organelle fission [1879 gene products]

GO:0000280 : nuclear division [1782 gene products]

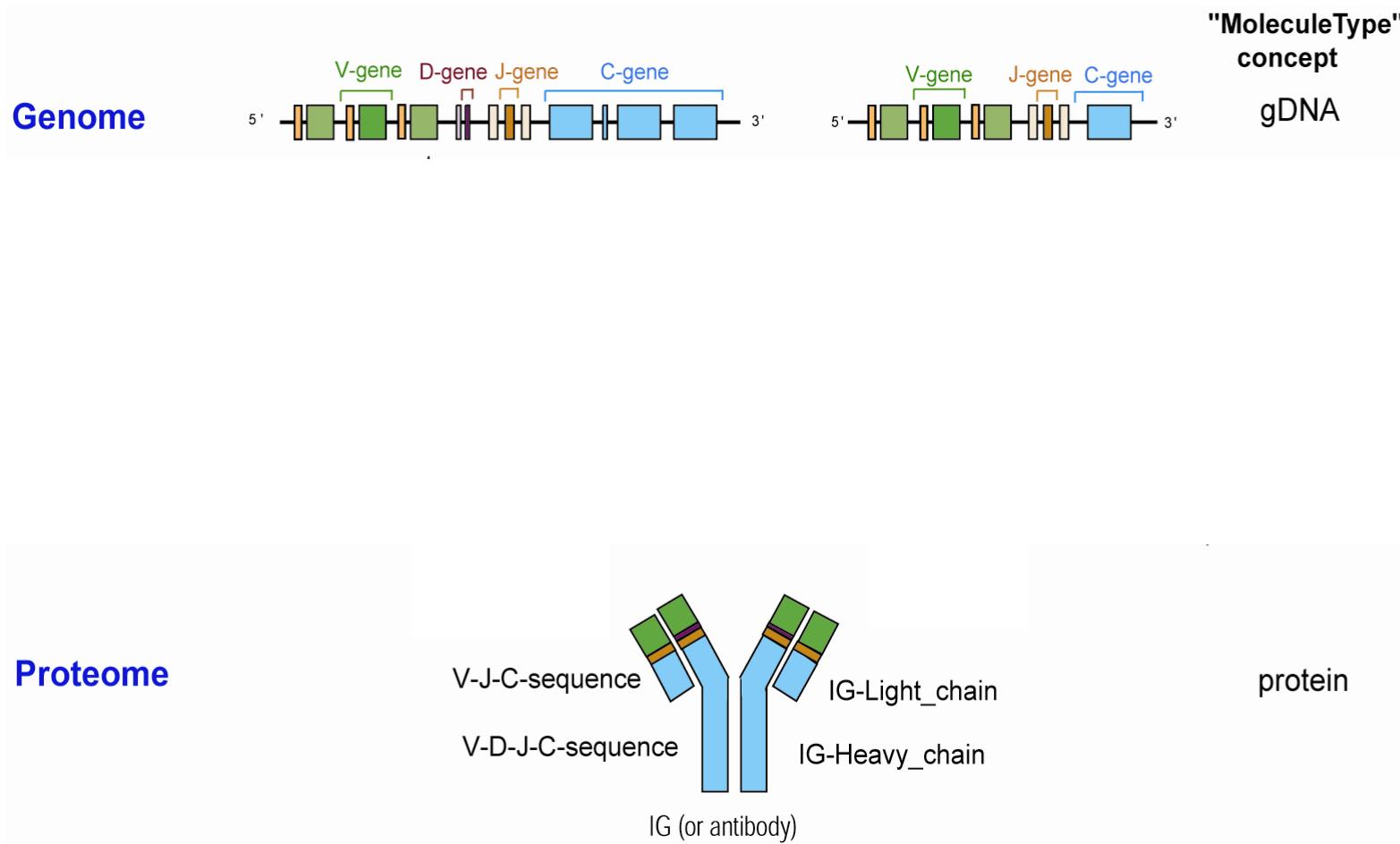
GO:0007067 : mitosis [1746 gene products]

Actions...  
Last action: Reset the tree  
Graphical View  
View in tree browser  
Download...  
OBO  
RDF-XML  
GraphViz dot

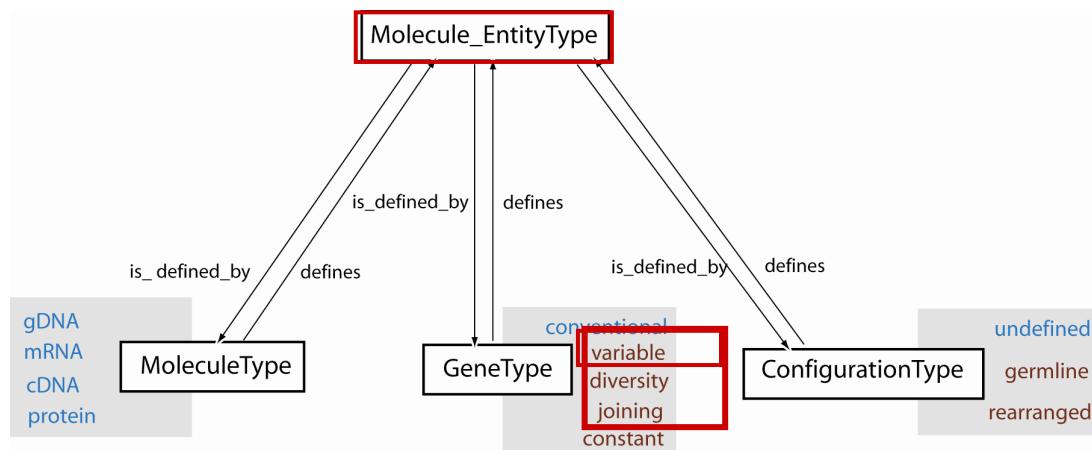
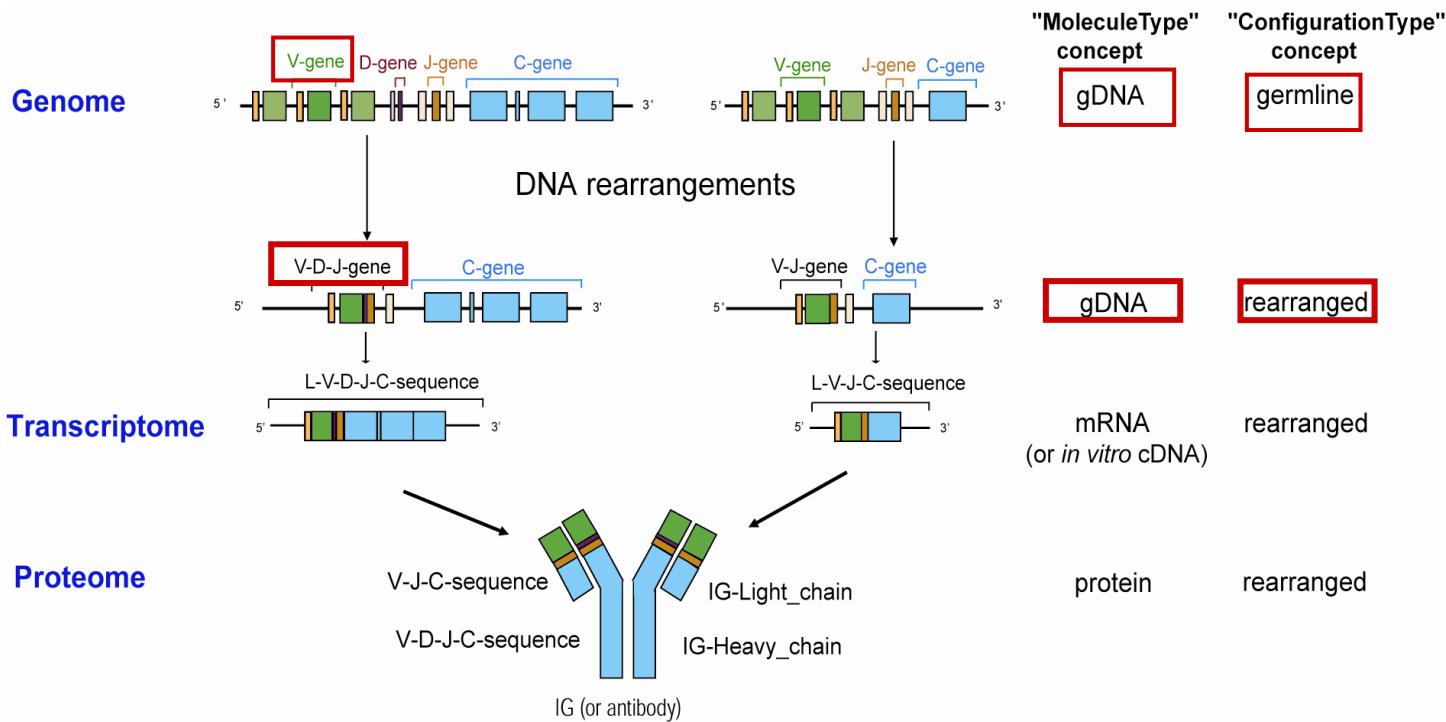
Done



# Synthèse des immunoglobulines et concepts d'identification d'IMGT-ONTOLOGY



# Synthèse des immunoglobulines et concepts d'identification d'IMGT-ONTOLOGY



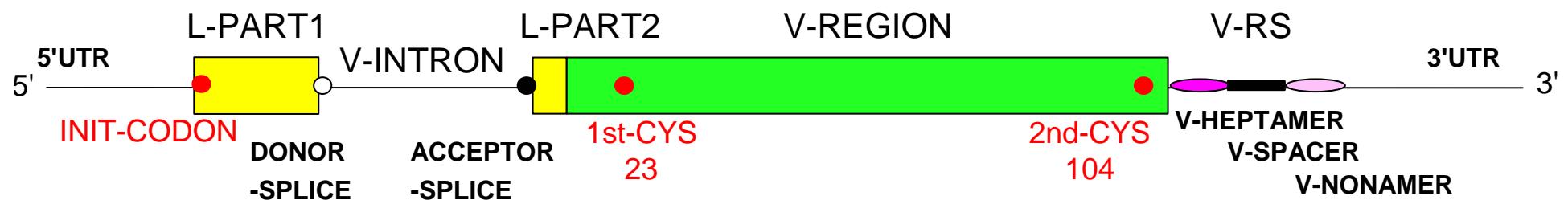
## **Concepts d'identification d'IMGT-ONTOLOGY**

**<http://bioportal.bioontology.org>**

# Les concepts de description de IMGT-ONTOLOGIE: La connaissance pour l'annotation des séquences

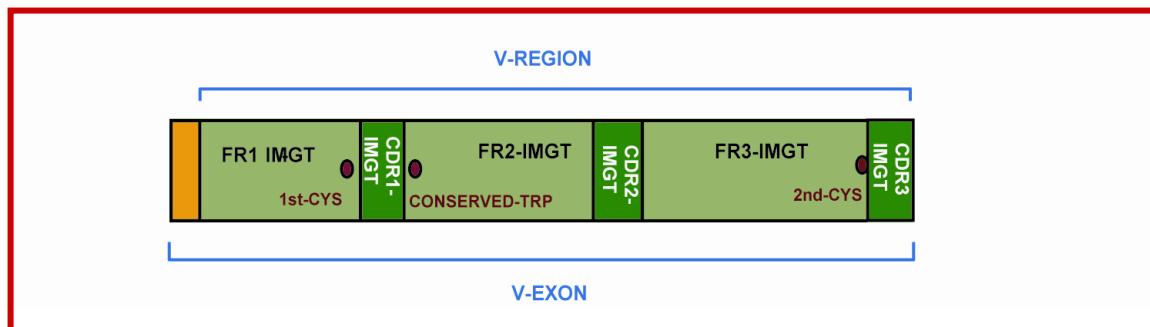
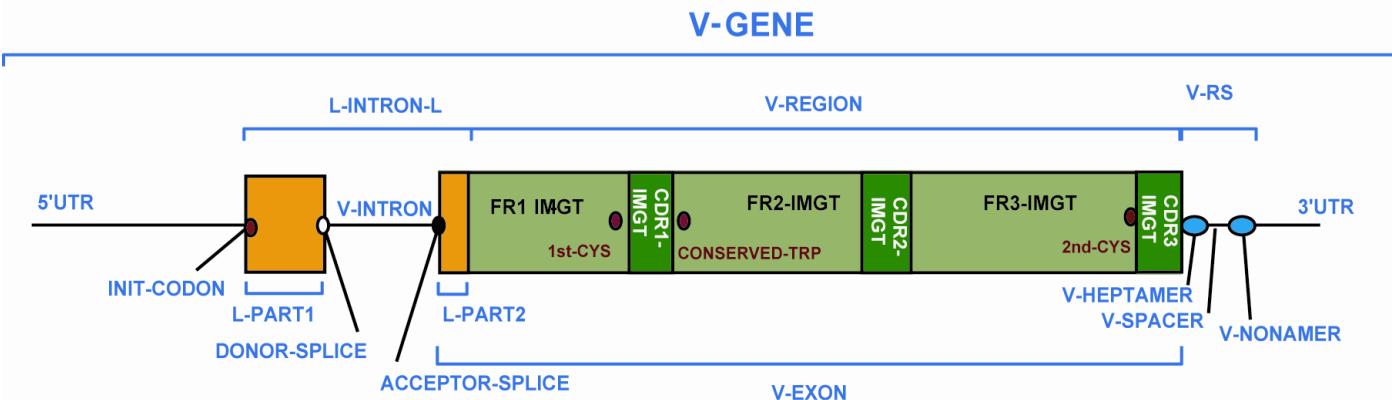
>X62106.0|HSV12|*Homo sapiens* VI-2 gene for immunoglobulin heavy chain

tgagagctcc	gttcctcacc	<b>atggactgga</b>	cctggaggat	cctcttcttg	gtggcagcag	60
<b>ccacaggtaa</b>	gaggctccct	agtcccagtg	atgagaaaaga	gattgagtcc	agtccaggga	120
gatctcatcc	acttctgtgt	tctctccapa	<b>ggagcccact</b>	cccaggtgca	gctggtgca	180
tctggggctg	aggtgaagaa	gcctggggcc	tcagtgaagg	tctcctgcaa	ggcttctgga	240
tacaccttca	ccggctacta	tatgcactgg	gtgcgacagg	cccctggaca	agggcttgag	300
tggatggat	ggatcaaccc	taacagtgg	ggcacaaact	atgcacagaa	gtttcagggc	360
agggtcacca	tgaccaggga	cacgtccatc	agcacagcct	acatggagct	gagcaggctg	420
agatctgacg	acacggccgt	gtattactgt	gcgagagaca	cagtgtgaaa	acccacatcc	480
<b>tgagggtgtc</b>	<b>agaaacccaa</b>	gggaggaggc	ag			



# Application: formalisation dans Protégé des labels qui composent Le V-EXON et de leur relations

(Source d'information: les informations de l'article de Biochimie)



Relation	Reciprocal relation (inverse)
"adjacent_at_its_5_prime_to"	"adjacent_at_its_3_prime_to"
"included_with_same_5_prime_in",	"includes_with_same_5_prime",
"included_with_same_3_prime_in",	"includes_with_same_3_prime",
"overlap_at_its_5_prime_with"	"overlap_at_its_3_prime_with"
"included_in"	"includes"

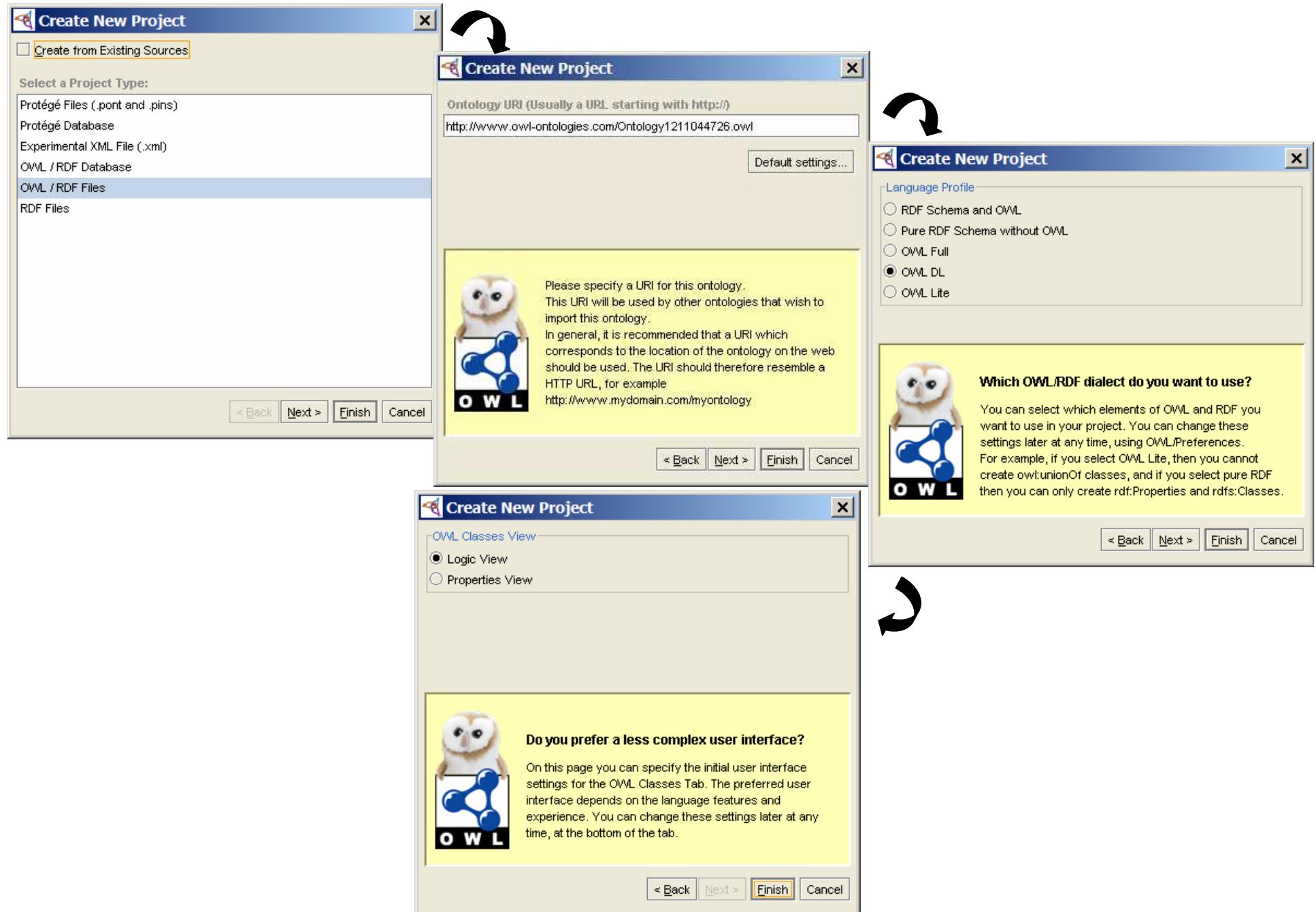
## Protégé (<http://protege.stanford.edu/>)

Protégé est un logiciel gratuit (JAVA), plate-forme open-source qui fournit une suite d'outils pour construire des bases de connaissance et des ontologies.

Protégé inclut de nombreux plugins pour la manipulation et la représentation d'ontologies dans différents formats.

L'éditeur Protégé-OWL permet aux utilisateurs de construire des ontologies pour le Web sémantique en OWL.

# Création d'une nouvelle ontologie avec Protege\_3.4\_beta



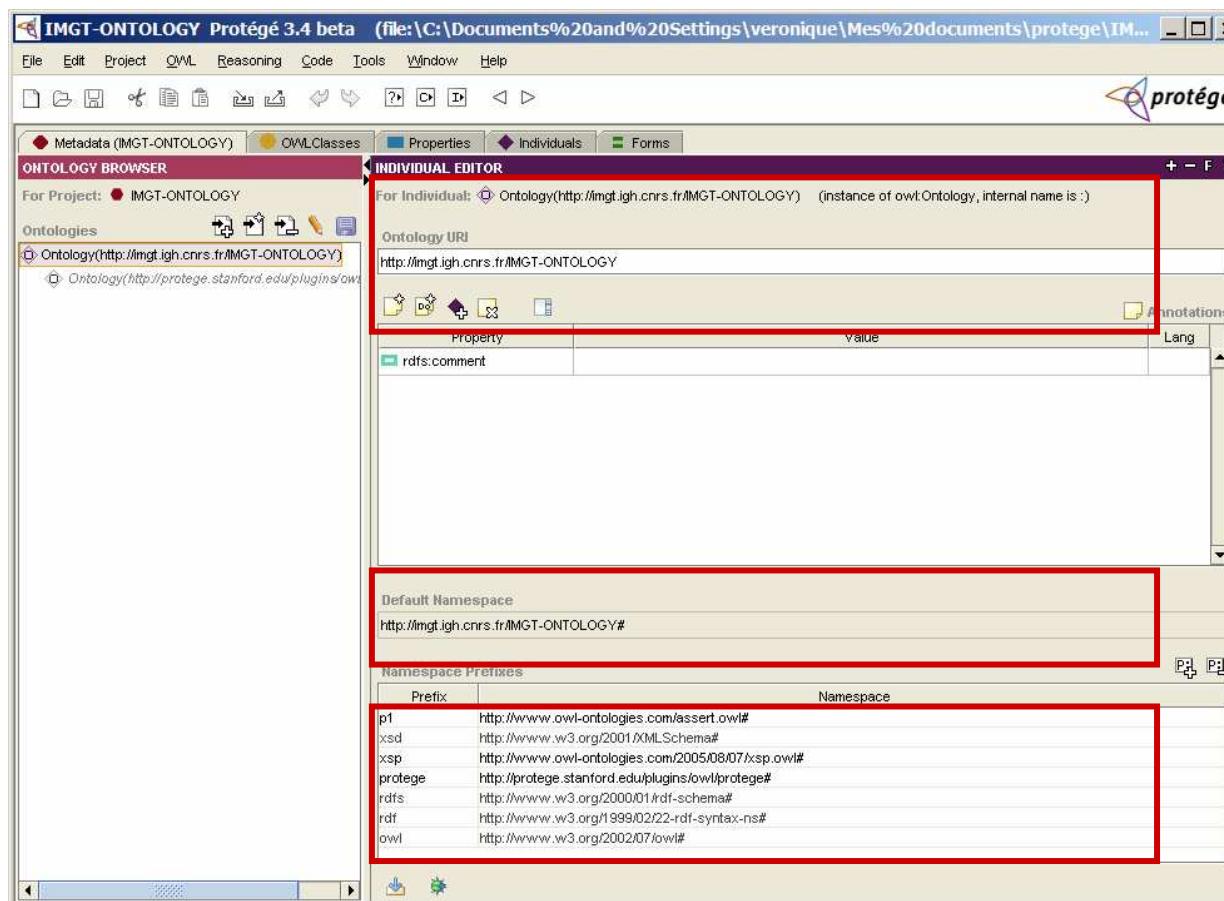
## Espace de nom

Chaque ontologie éditée avec protégé a son propre espace de nom:

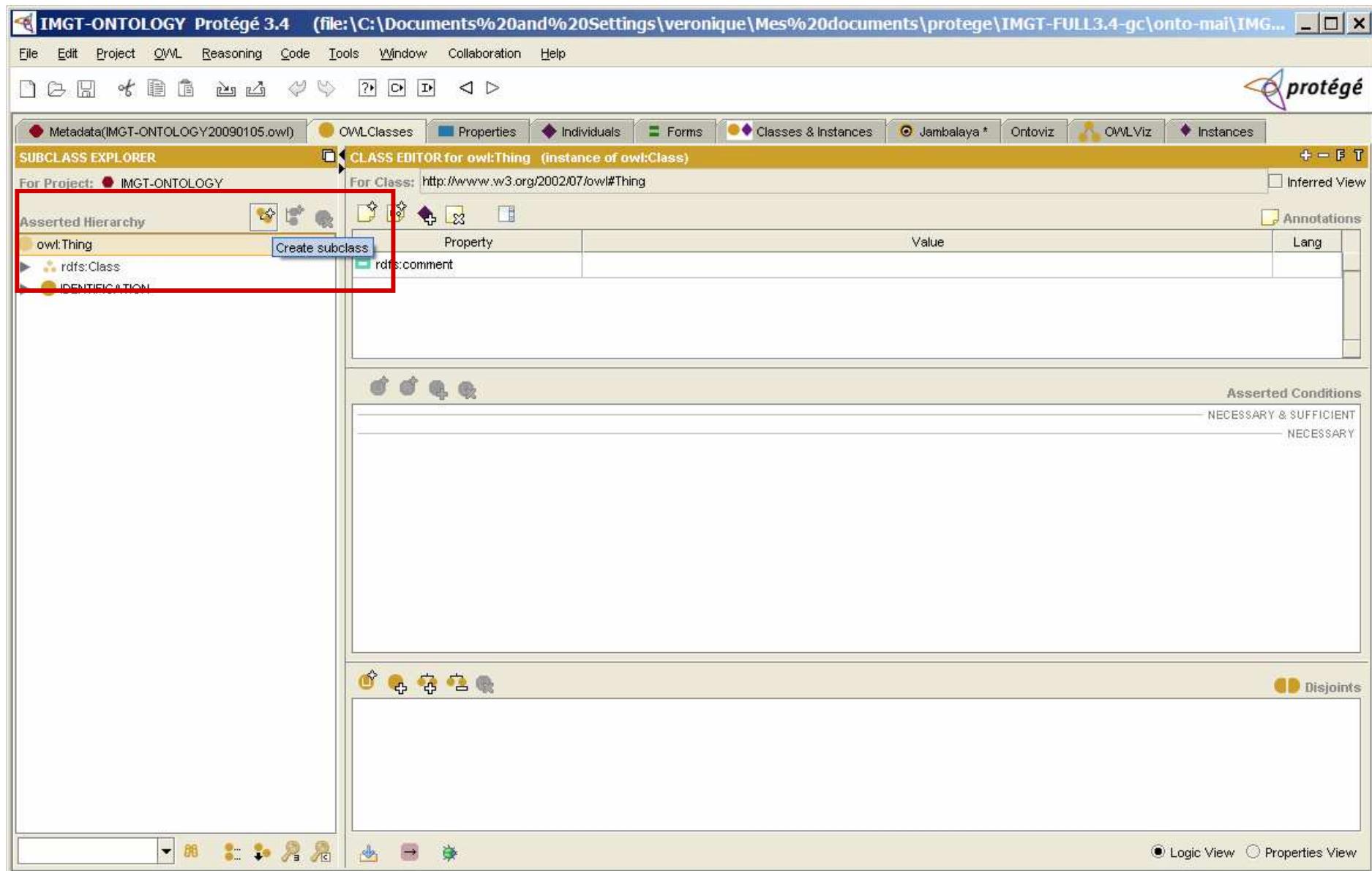
C'est le « default namespace »

C'est une chaîne de caractères qui préfixe les noms des classes, propriétés, instances afin d'assurer leur unicité (Unique Resource Identifiers URI).

=> Evite la confusion en cas de termes identiques utilisés avec des significations différentes en fonction du domaine de connaissance de l'ontologie



## Création de nouvelles classes avec Protégé (1)



## Création de nouvelles classes avec Protégé (2)

The screenshot shows the Protégé 3.4 interface for creating a new class. The project is titled "IMGT-ONTOLOGY".

**SUBCLASS EXPLORER:** Shows the asserted hierarchy with "owl:Thing" at the top, followed by "rdfs:Class", "IDENTIFICATION", and "DESCRIPTION".

**CLASS EDITOR for DESCRIPTION (instance of owl:Class):** This panel is highlighted with a red box.

**For Class:** <http://www.imgt.org/IMGT-ONTOLOGY1.owl#DESCRIPTION>

**Annotations:** A table showing annotations for the class:

Property	Value	Lang
rdfs:comment	commentaire ....	

**Asserted Conditions:** Shows conditions for "owl:Thing".

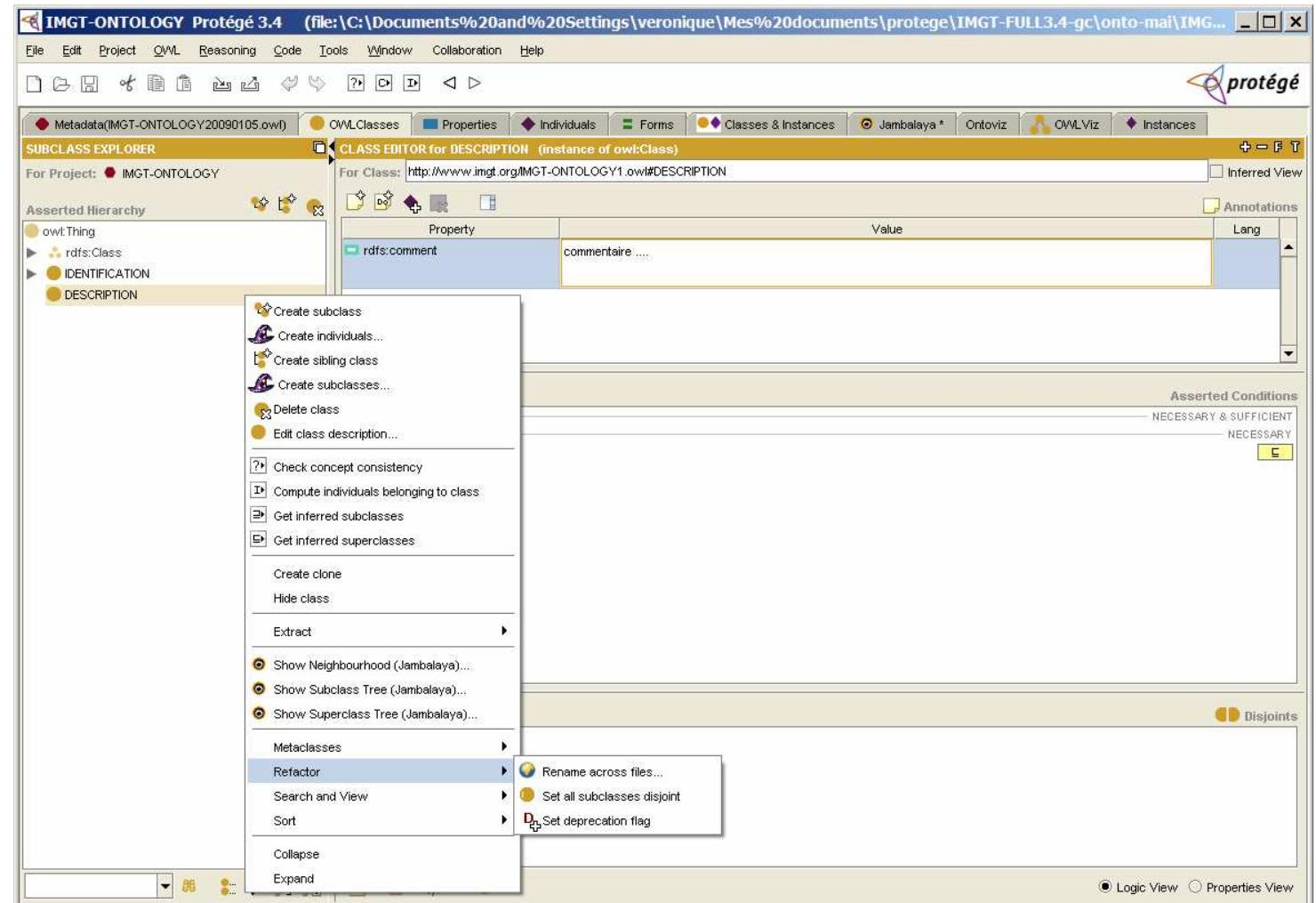
**Disjoins:** A large empty box highlighted with a red box.

**Bottom Navigation:** Logic View (selected) and Properties View.

## Classes disjointes

On peut soit exprimer que les sous classes d'une classe sont toutes disjointes

Click droit =>



Soit expliciter quelles sont les classes disjointes

The screenshot shows the Protégé 3.4 interface for editing an ontology. The main window title is "IMGT-ONTOLOGY Protégé 3.4". The menu bar includes File, Edit, Project, OWL, Reasoning, Code, Tools, Window, Collaboration, and Help. The toolbar contains various icons for file operations like Open, Save, and Print.

The left sidebar displays the "SUBCLASS EXPLORER" for the project "IMGT-ONTOLOGY". It shows the asserted hierarchy with nodes for owl:Thing, rdfs:Class, IDENTIFICATION, EntityType, MoleculeType, and GeneType. The MoleculeType node is currently selected and highlighted in orange.

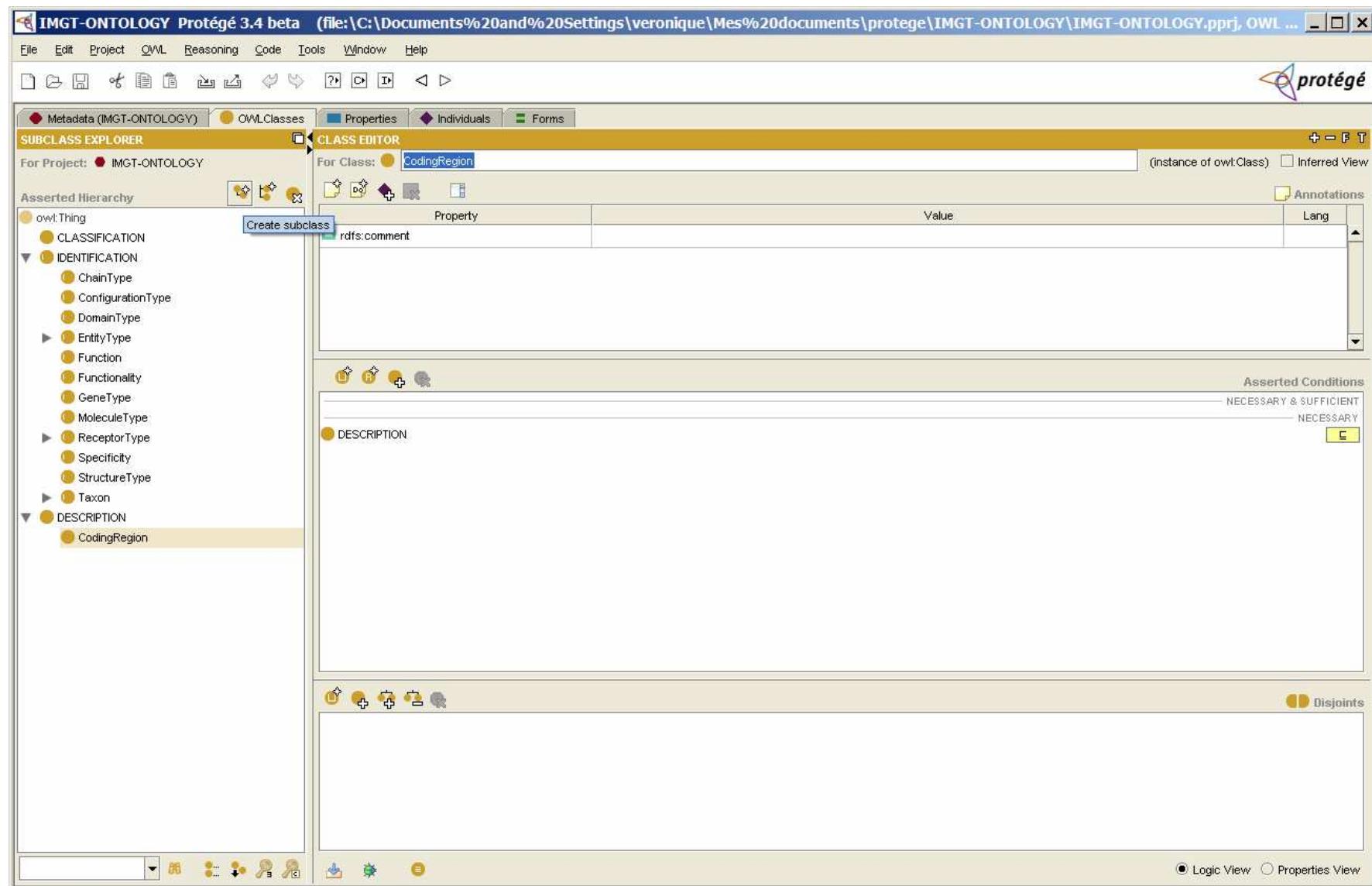
The central panel is the "CLASS EDITOR for MoleculeType (instance of owl:Class)". It shows the class URI "http://www.imgt.org/IMGT-ONTOLOGY1.owl#MoleculeType". The "Annotations" section contains a single entry for "rdfs:comment".

The bottom panel shows the "Asserted Conditions" for the MoleculeType class. It lists two conditions under the "NECESSARY & SUFFICIENT" heading:

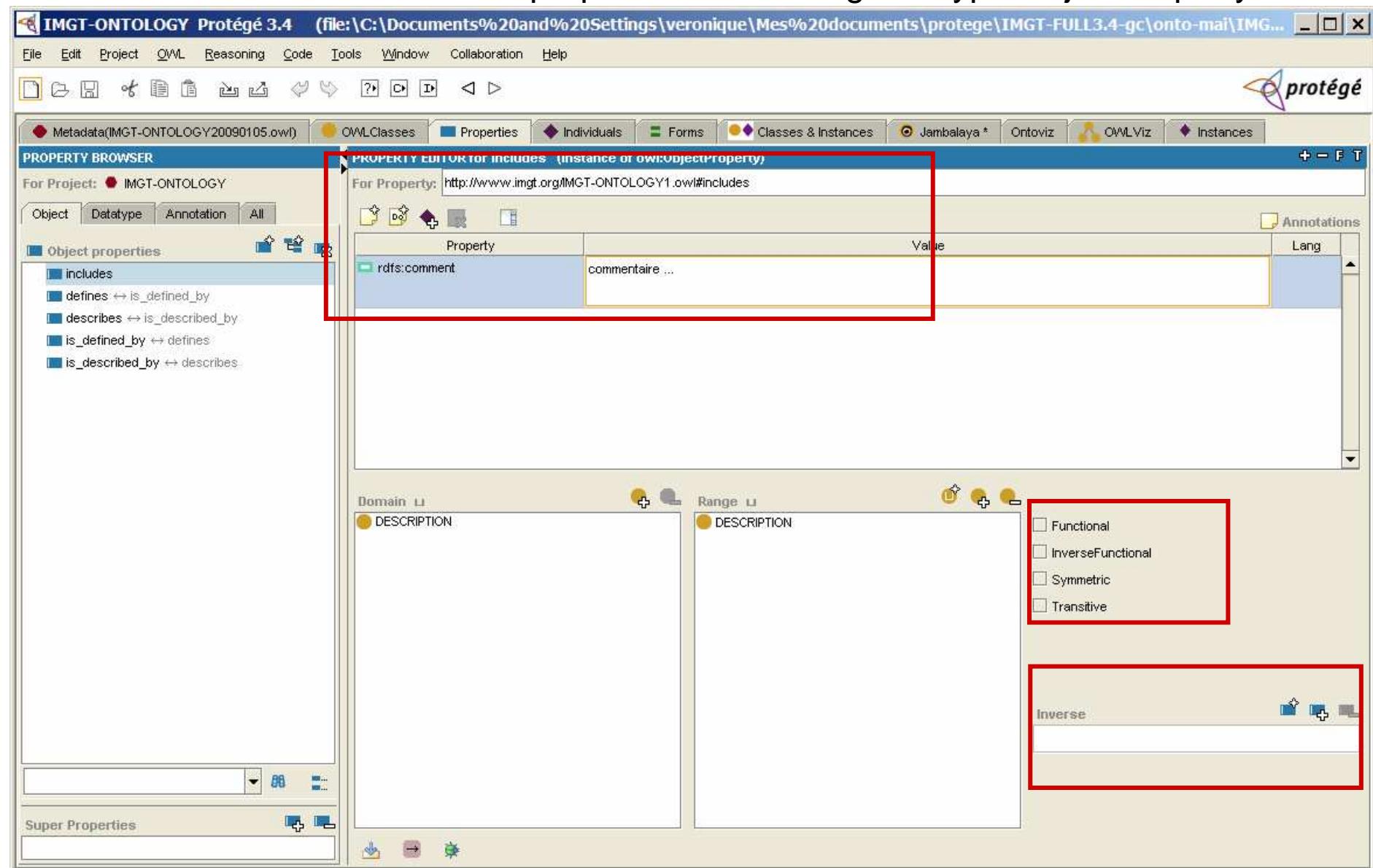
- defines **only** Molecule\_EntityType
- defines **some** Molecule\_EntityType

The bottom right corner of the main editor area is highlighted with a red box, indicating a specific region of interest.

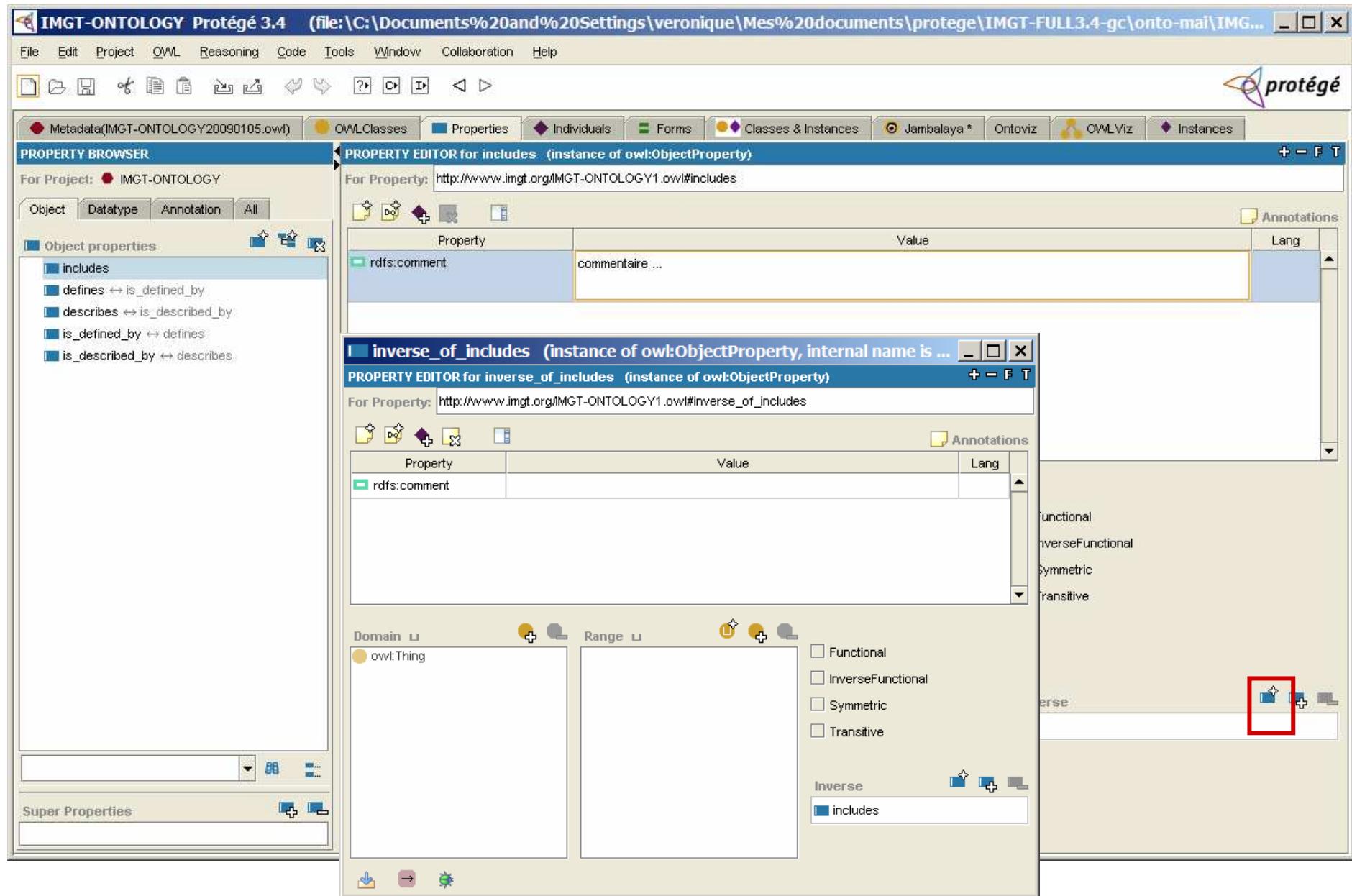
## Création de sous-classes avec Protégé



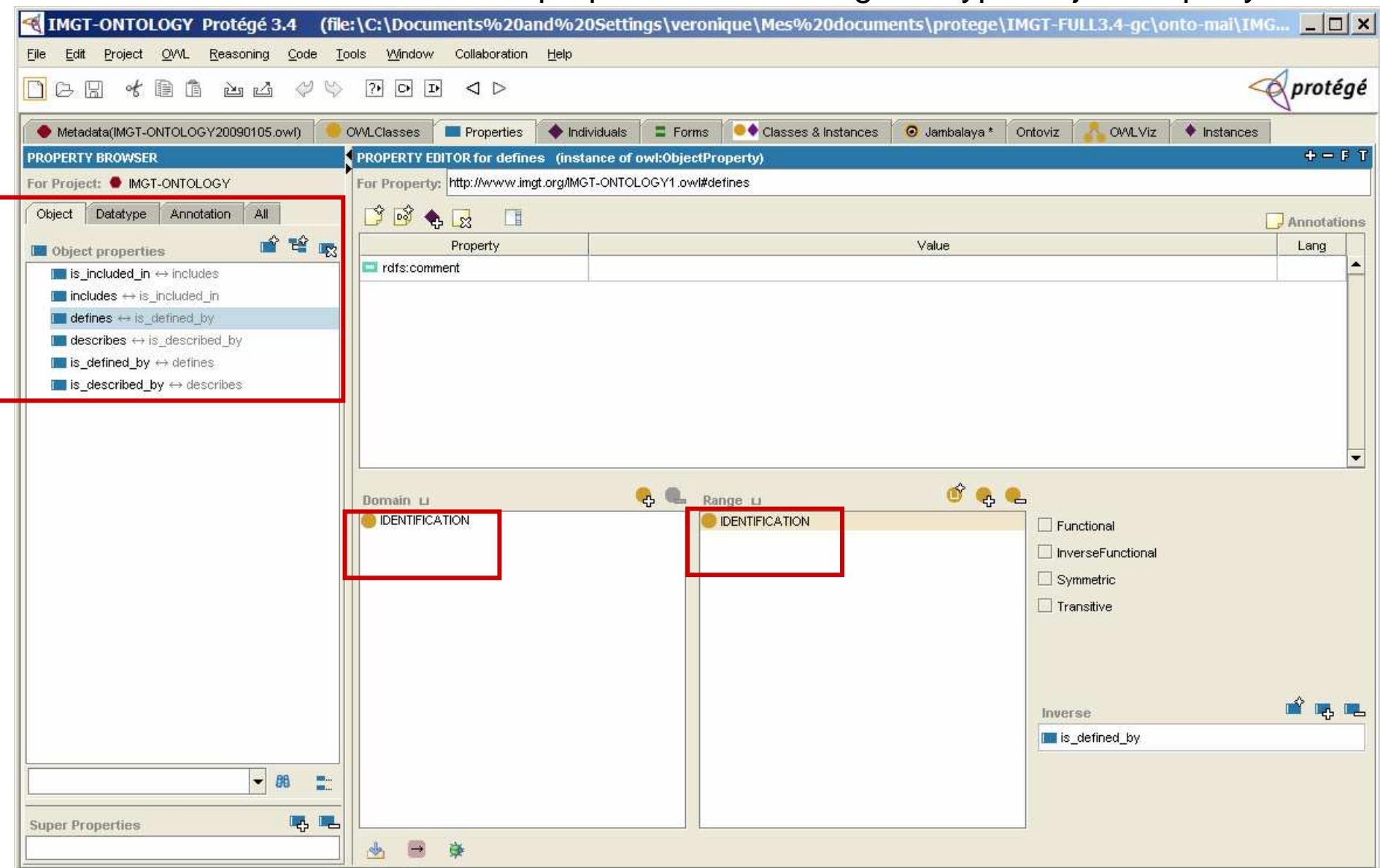
## Création d'une nouvelle propriété avec Protégé de type ObjectProperty



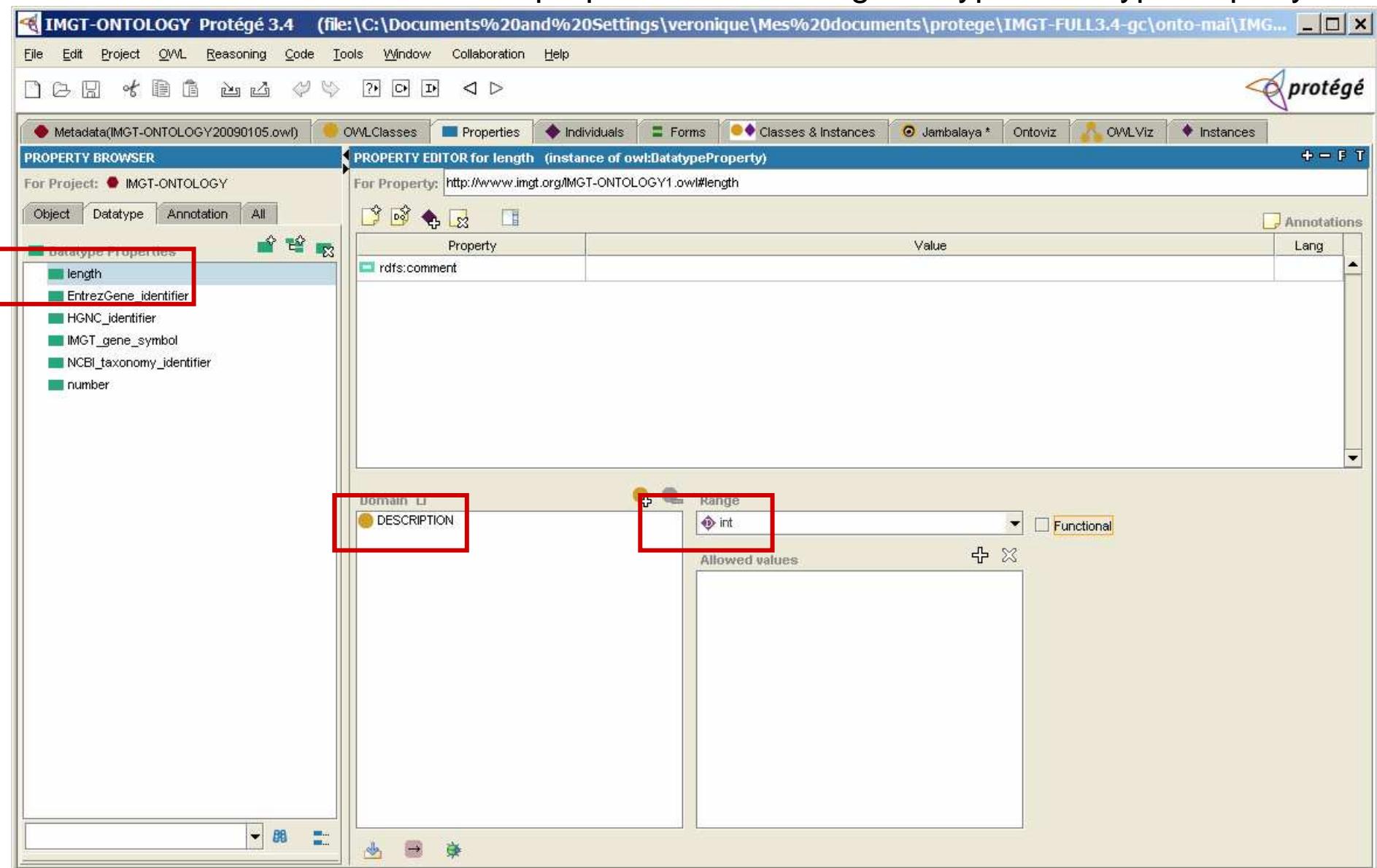
## Création d'une nouvelle propriété avec Protégé de type ObjectProperty



## Création d'une nouvelle propriété avec Protégé de type ObjectProperty



## Création d'une nouvelle propriété avec Protégé de type DataTypeProperty



# Restrictions sur les propriétés

The screenshot shows the Protégé 3.4 interface for editing an ontology. The main window displays the 'CLASS EDITOR for MoleculeType' (instance of owl:Class). The 'For Class:' field is set to <http://www.imgt.org/IMGT-ONTOLOGY1.owl#MoleculeType>. The 'Annotations' section contains a single entry: rdfs:comment. The 'Asserted Conditions' section shows two asserted conditions under the 'IDENTIFICATION' category:

- defines **only** Molecule\_EntityType
- defines **some** Molecule\_EntityType

The 'Disjoints' section lists three disjoint classes: EntityType, GeneType, and ConfigurationType.

Navigation icons at the bottom include: back, forward, search, and file operations.

# Restrictions sur les propriétés

The screenshot shows the Protégé 3.4 interface for editing an OWL class. The title bar indicates the project is "IMGT-ONTOLOGY" and the file path is "C:\Documents and Settings\veronique\Mes documents\protege\IMGT-FULL3.4-gc\onto-mai\IMG...". The menu bar includes File, Edit, Project, OWL, Reasoning, Code, Tools, Window, Collaboration, and Help.

The main window displays the "CLASS EDITOR for Molecule\_EntityType (instance of owl:Class)". The "For Class:" field specifies [http://www.imgt.org/IMGT-ONTOLOGY1.owl#Molecule\\_EntityType](http://www.imgt.org/IMGT-ONTOLOGY1.owl#Molecule_EntityType). The "Asserted Hierarchy" panel on the left shows the class hierarchy under the project "IMGT-ONTOLOGY", with "Molecule\_EntityType" selected. The "Annotations" section shows an annotation for "rdfs:comment". The "Asserted Conditions" section lists asserted conditions:

- EntityType
- is\_defined\_by **only** ConfigurationType or MoleculeType or GeneType
- is\_defined\_by **some** ConfigurationType
- is\_defined\_by **some** MoleculeType
- is\_defined\_by **some** GeneType

The bottom status bar shows "Logic View" selected. The interface includes various toolbars and panels for navigating and editing the ontology.

# Restrictions sur les propriétés

The screenshot shows the Protégé 3.4 interface for editing the IMGT-ONTOLOGY. The top menu bar includes File, Edit, Project, OWL, Reasoning, Code, Tools, Window, Collaboration, and Help. The toolbar below has icons for file operations like Open, Save, and Print, along with navigation and search tools.

The main workspace is divided into several panels:

- SUBCLASS EXPLORER:** Shows the asserted hierarchy starting from owl:Thing, including rdfs:Class, IDENTIFICATION, ConfigurationType, EntityType, Molecule\_EntityType, and various sequence types like AA-sequence, C-sequence, J-sequence, J-gene, V-gene, J-C-sequence, C-gene, V-J-C-sequence, D-J-gene, D-sequence, V-D-J-gene, gene, V-J-gene, L-V-sequence, nt-sequence, L-V-J-C-sequence, L-V-D-J-C-sequence, D-gene, and V-D-J-C-sequence.
- CLASS EDITOR for ConfigurationType (instance of owl:Class):** This panel is active. It displays the class <http://www.imgt.org/IMGT-ONTOLOGY1.owl#ConfigurationType>. The properties tab shows an rdfs:comment entry. The asserted conditions tab lists "IDENTIFICATION" with two restrictions: "defines only Molecule\_EntityType" and "defines some Molecule\_EntityType". The asserted disjointness tab lists EntityType, MoleculeType, and GeneType.
- Annotations:** A table for adding annotations to the class.
- Asserted Conditions:** A table for defining asserted conditions, currently showing three entries under NECESSARY.
- Disjoints:** A table for defining disjoint classes.

At the bottom, there are tabs for Logic View and Properties View, and a set of small icons for navigating between views.

# Tests et vérifications

The screenshot shows the Protégé 3.4 interface for editing the IMGT-ONTOLOGY. The main window displays the CLASS EDITOR for the class `intron` (instance of `owl:Class`). The left pane, titled "SUBCLASS EXPLORER", shows the asserted hierarchy, including `owl:Thing`, `rdfs:Class`, `IDENTIFICATION`, `EntityType`, `MoleculeType`, `GeneType`, `DESCRIPTION`, `CodingRegion`, and `intron`. The right pane contains tabs for "Annotations", "Asserted Conditions", and "Disjoints". The "Annotations" tab shows an `rdfs:comment` annotation. The "Asserted Conditions" tab lists a condition for `DESCRIPTION` with a "NECESSARY & SUFFICIENT" constraint. The "Disjoints" tab is currently empty. At the bottom, a "Test Results" table shows a warning for the `DESCRIPTION` class: "Missing disjoints on primitive subclasses: CodingRegion intron".

File Edit Project OWL Reasoning Code Tools Window Collaboration Help

protégé

SUBCLASS EXPLORER

For Project: IMGT-ONTOLOGY

Asserted Hierarchy

- owl:Thing
- rdfs:Class
- IDENTIFICATION
  - ConfigurationType
- EntityType
- MoleculeType
- GeneType
- DESCRIPTION
  - CodingRegion
  - intron

CLASS EDITOR for intron (instance of owl:Class)

For Class: <http://www.imgt.org/IMGT-ONTOLOGY1.owl#intron> Inferred View

Annotations

Property	Value	Lang
rdfs:comment		

Asserted Conditions

NECESSARY & SUFFICIENT NECESSARY

DESCRIPTION

Disjoints

Type Source Test Result

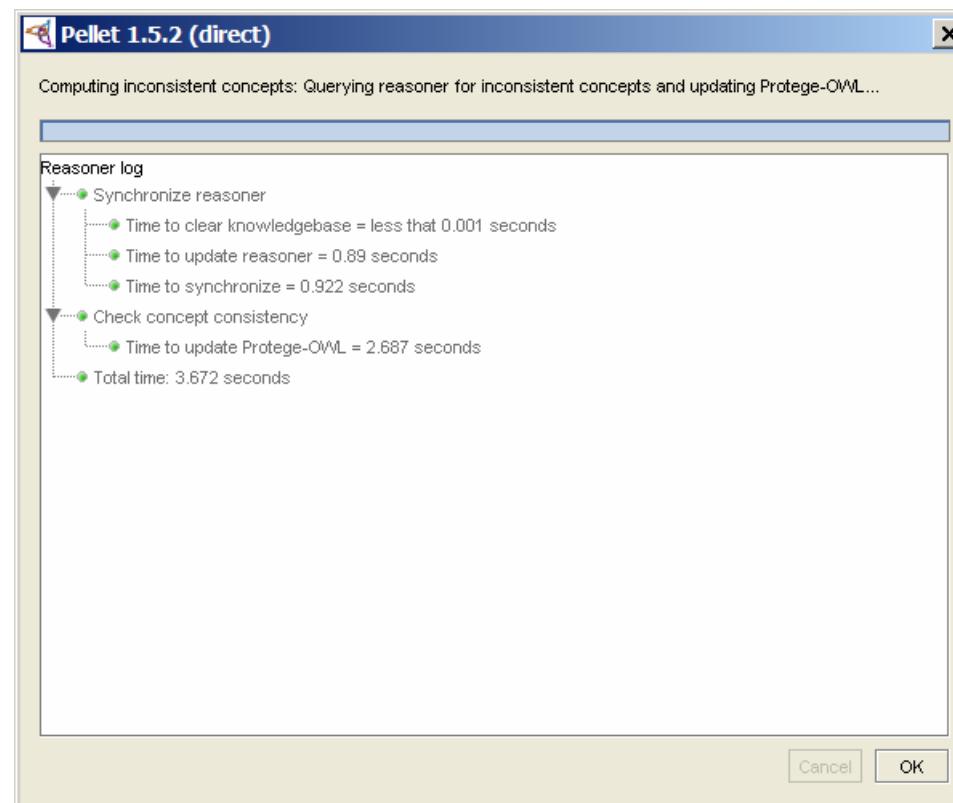
DESCRIPTION Missing disjoints on primitive subclasses: CodingRegion intron

Test Results

## Tests et vérifications

The screenshot shows the Protégé 3.4.4 interface with the IMGT-ONTOLOGY ontology loaded. The menu bar includes File, Edit, Project, OWL, Reasoning, Code, Tools, Window, Collaboration, and Help. The Reasoning tab is selected, showing options like None, DIG Reasoner, and Pellet 1.5.2 (direct). The main workspace displays the asserted hierarchy of the ontology, with 'imtg:J-gene' currently selected. The asserted hierarchy tree shows various classes such as imtg:Molecule\_EntityType, imtg:AA-chain, imtg:C-gene, etc. The central panel shows the properties and values for 'imtg:J-gene'. The asserted conditions panel lists various asserted conditions, and the disjoint panel lists classes that are disjoint with 'imtg:J-gene'. The bottom right corner shows tabs for Logic View (selected) and Properties View.

## Tests et vérifications

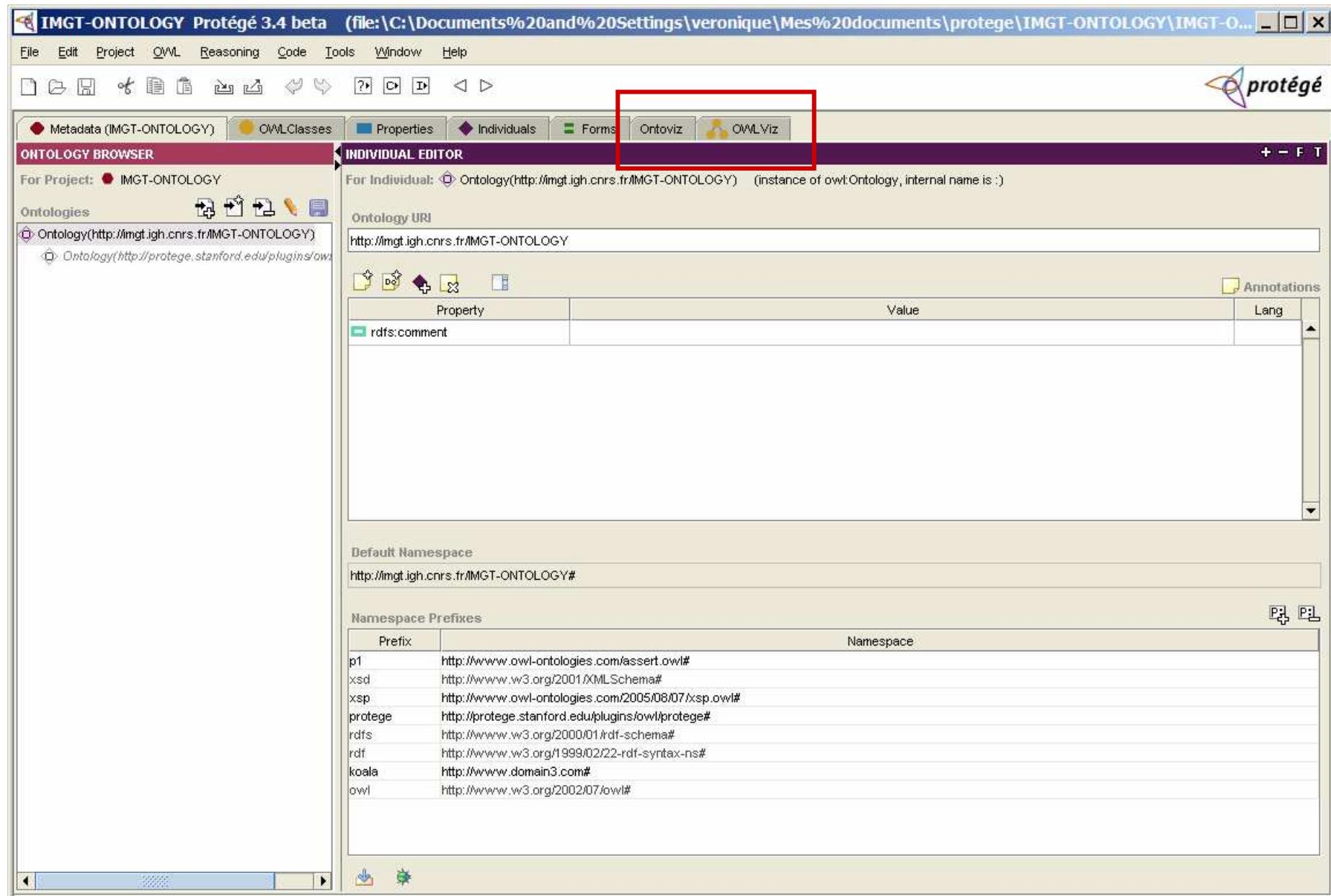


# Création d'instances avec Protégé

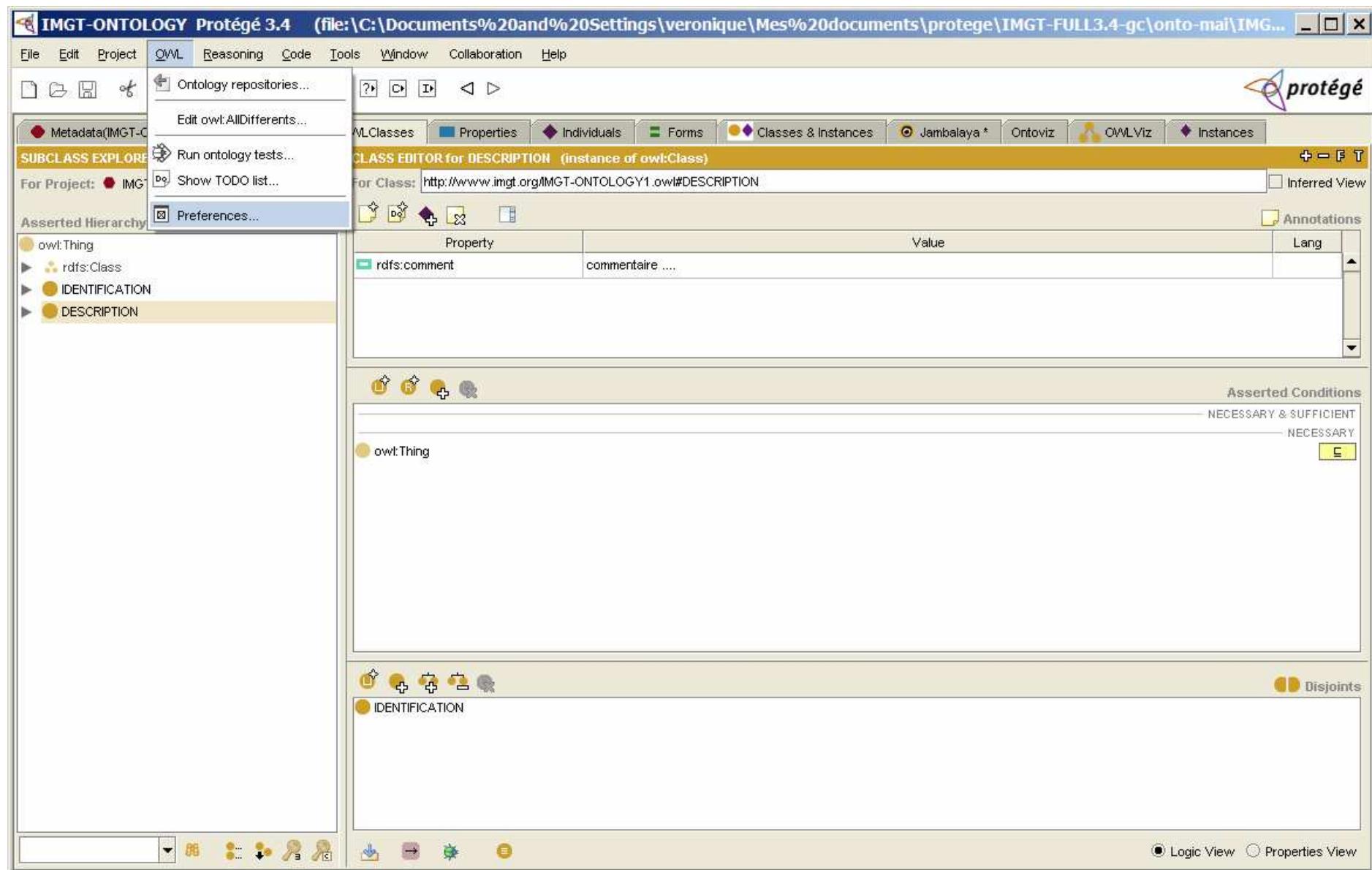
The screenshot shows the Protégé 3.4.4 interface with the following panels:

- CLASS BROWSER:** Shows the class hierarchy for the project "IMGT-ONTOLOGY". The selected class is "imgt:J-gene".
- INSTANCE BROWSER:** Shows asserted instances for the class "imgt:J-gene". One instance, "imgt:J-gene\_1", is selected.
- INDIVIDUAL EDITOR:** Displays properties and values for the individual "imgt:J-gene\_1".
  - Properties:** rdfs:isDefinedBy, dc:identifier, rdfs:comment, rdfs:label.
  - Annotations:** A table with columns "Property", "Value", and "Lang".
  - Relationships:** imgt:has\_ (with value "w.imgur.org/IMGT-ONTOLOGY/functional") and imgt:is\_defined\_by (with values "www.imgur.org/IMGT-ONTOLOGY/DNA", "www.imgur.org/IMGT-ONTOLOGY/germline", and "www.imgur.org/IMGT-ONTOLOGY/joining").

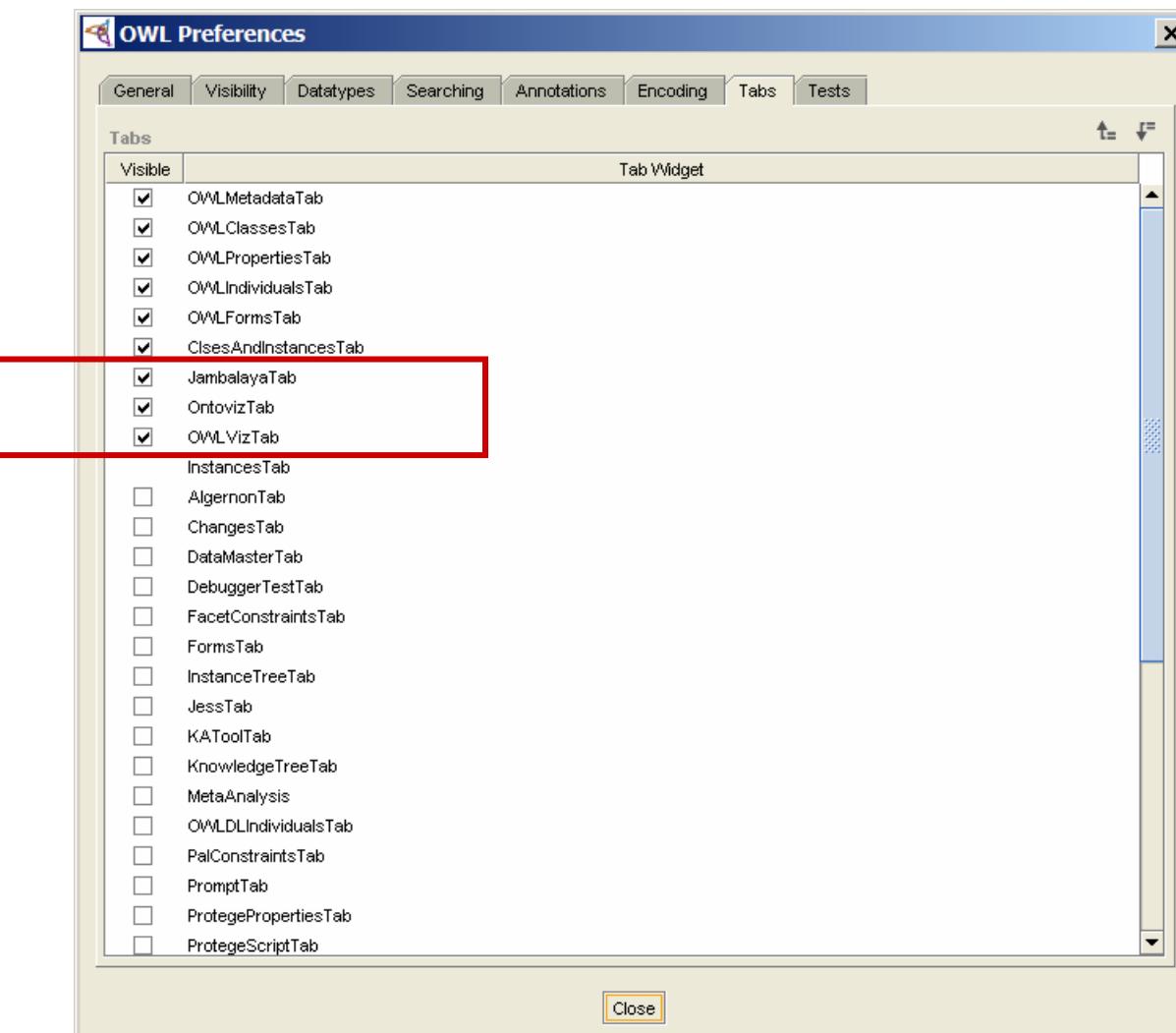
# Représentation graphique de l'ontologie



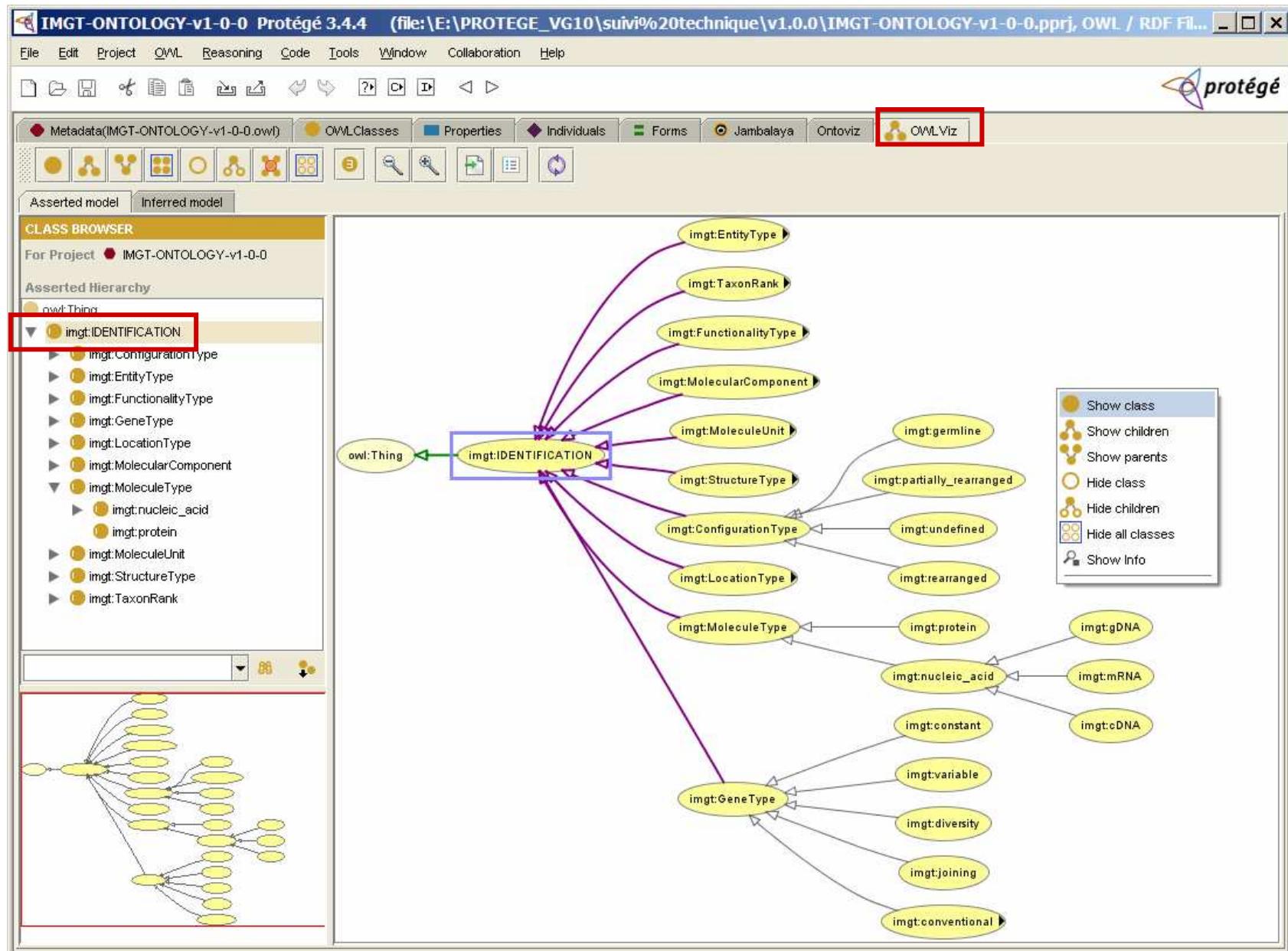
# Représentation graphique de l'ontologie



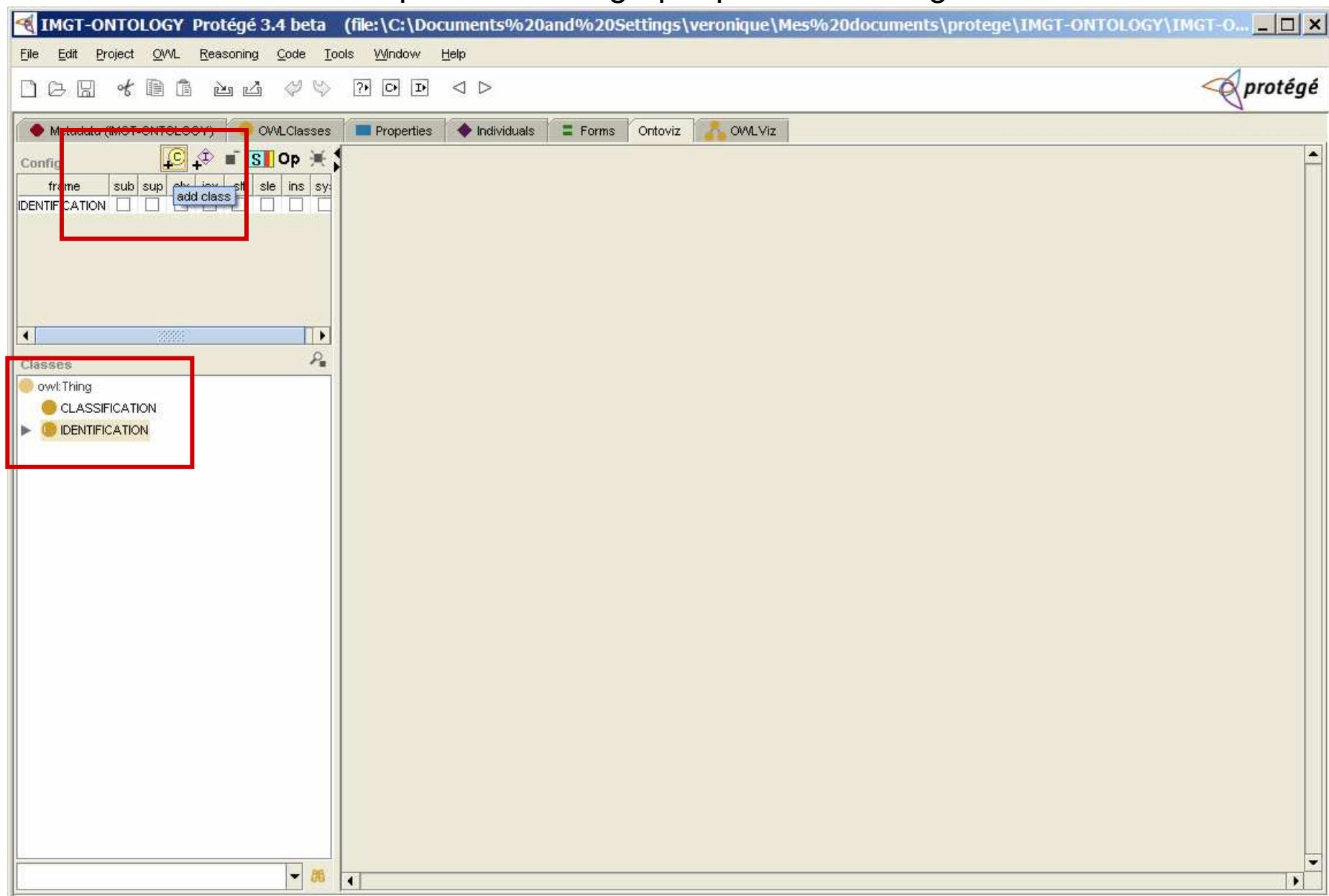
# Représentation graphique de l'ontologie



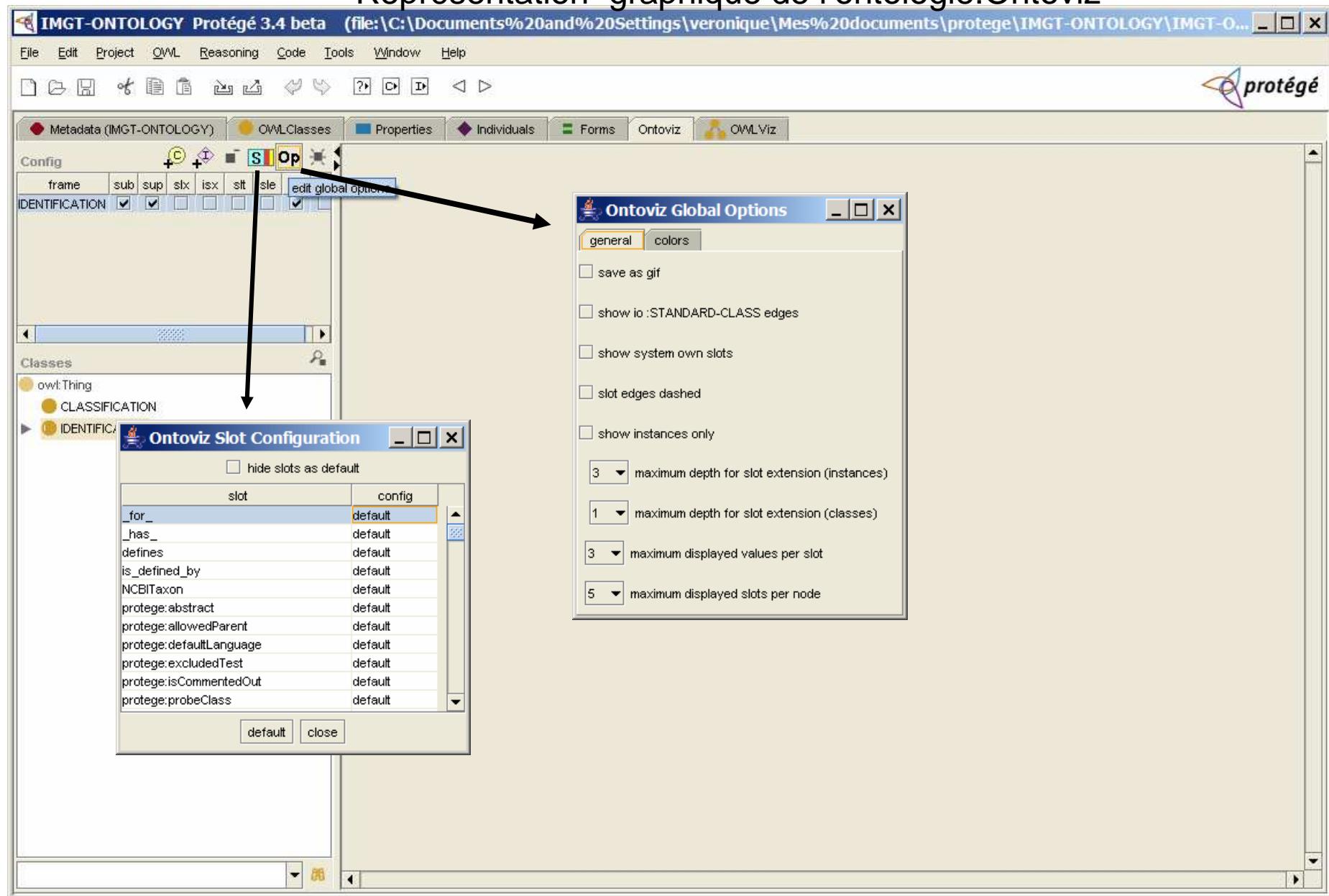
## Représentation graphique de l'ontologie: OWLviz



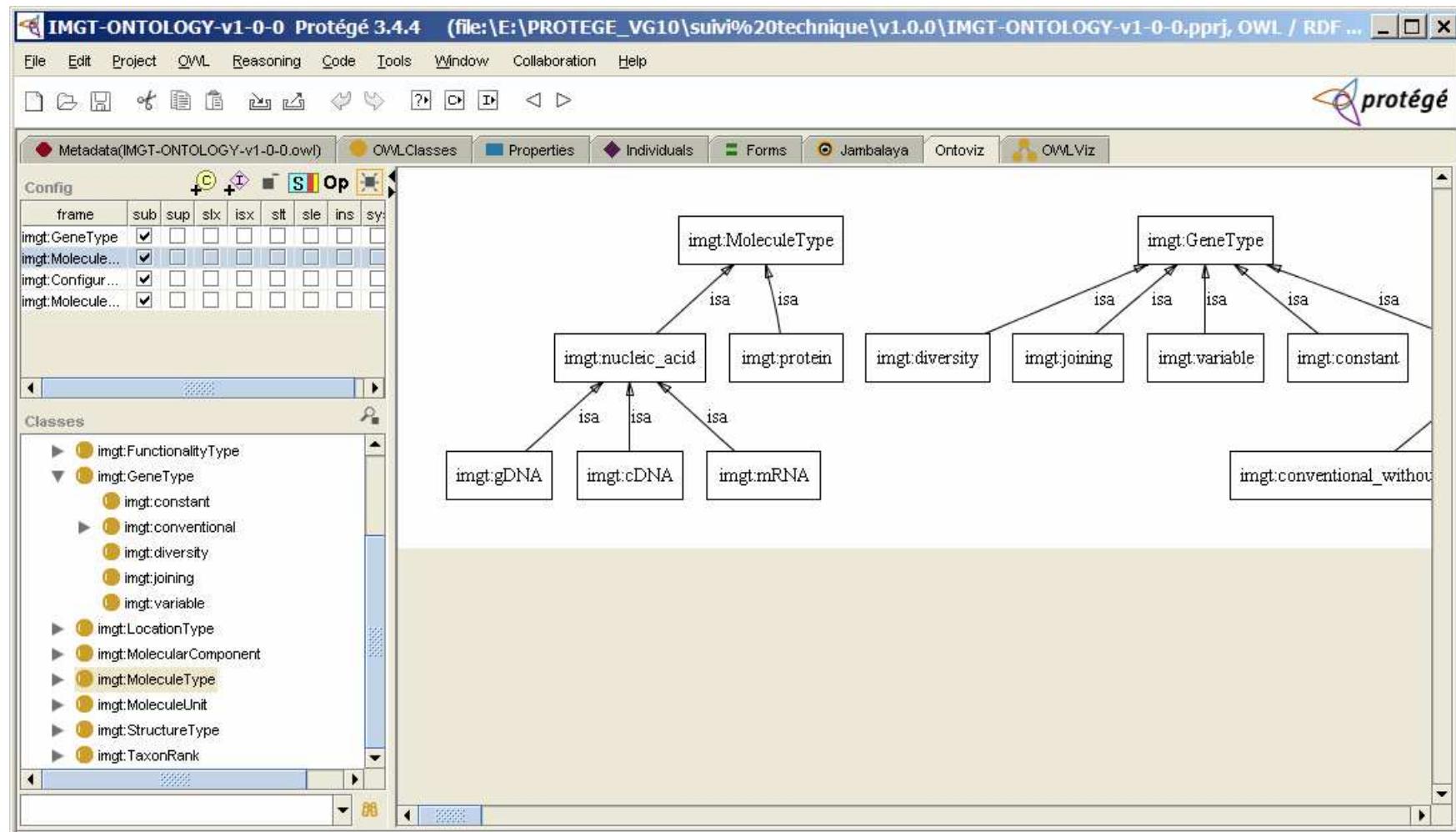
## Représentation graphique de l'ontologie:Ontoviz



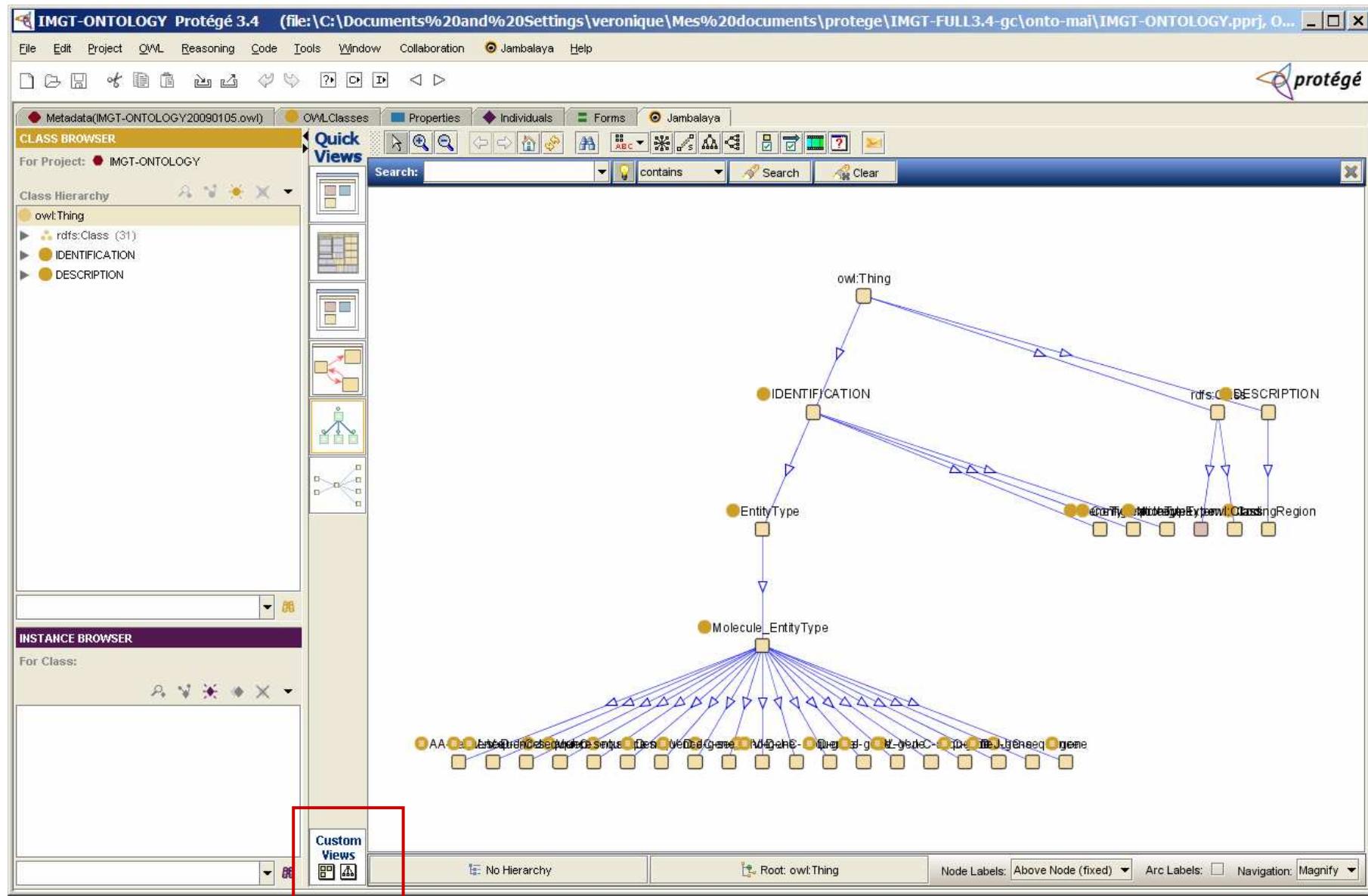
## Représentation graphique de l'ontologie:Ontoviz



# Représentation graphique de l'ontologie: Ontoviz



# Représentation graphique de l'ontologie:Jambalaya



**Quick Views - IMGT-ONTOLOGY-v1-0-0**

New Quick View Delete Delete All Defaults Run Quick View Import Export ?



Nested View (Default)


Nested Treemap


Nested Composite View


Class & Individual Tree


Class Tree


Domain/Range


New Quick View


test

### test Details

Name:  Display?

Layout:  Icon: 

Label mode:

Select the node types of interest:

Defined Class  
 Enumeration  
 Individual  
 Logical Operation  
 Primitive Class  
 RDFS Class  
 Restriction

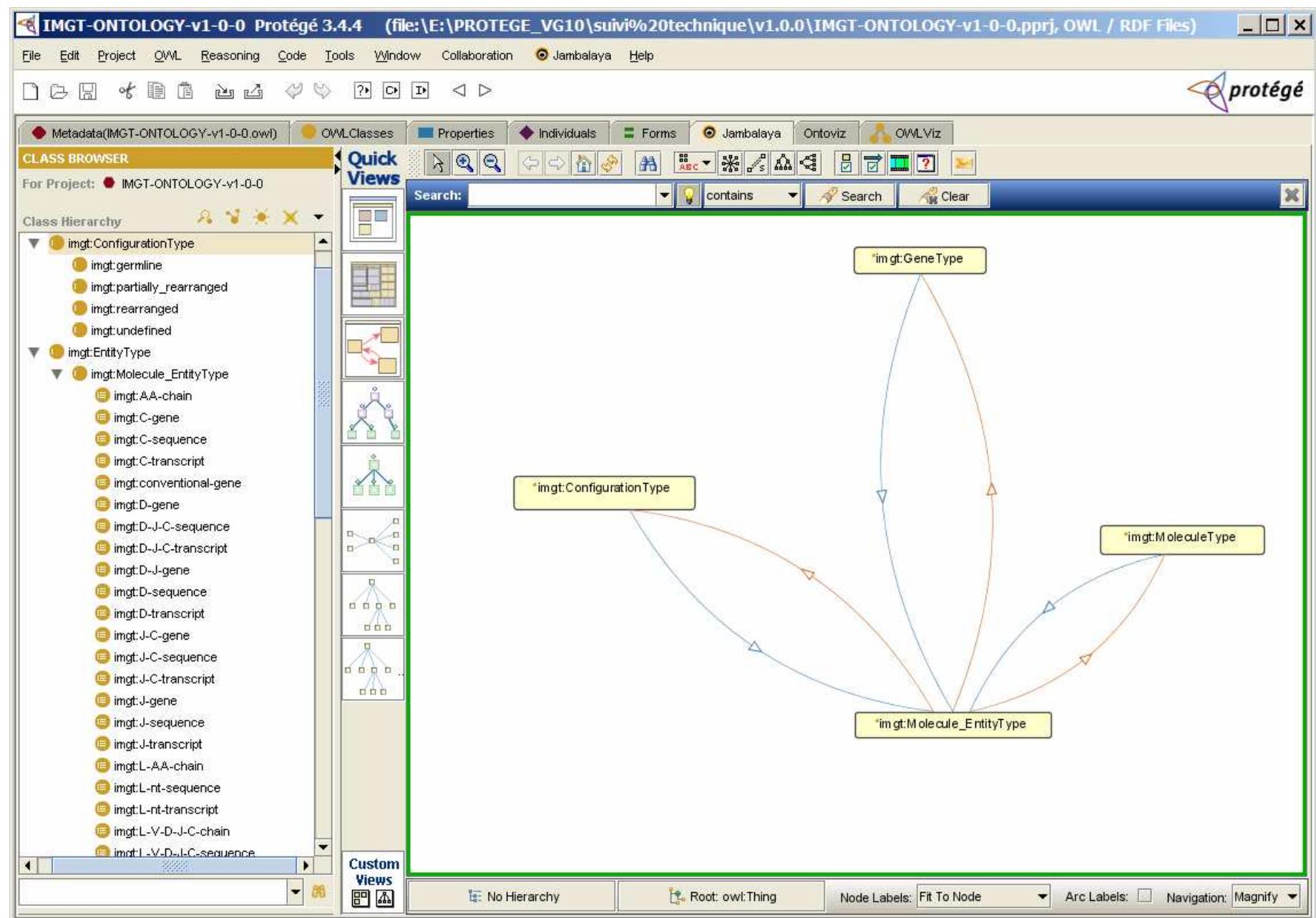
Select the arc types of interest:

has instance  
 has subclass  
 imgt:\_for\_  
 imgt:\_for\_(Domain>Range)  
 imgt:\_for\_(Necessary and Sufficient)  
 imgt:\_for\_(Necessary)  
 imgt:\_has\_  
 imgt:\_has\_(Domain>Range)  
 imgt:\_has\_(Necessary and Sufficient)  
 imgt:\_has\_(Necessary)  
 imgt:defines  
 imgt:defines(Domain>Range)  
 imgt:defines(Necessary and Sufficient)  
 imgt:defines(Necessary)  
 imgt:is\_defined\_by  
 imgt:is\_defined\_by(Domain>Range)  
 imgt:is\_defined\_by(Necessary and Sufficient)  
 imgt:is\_defined\_by(Necessary)

**New Quick Vi...** New Quick

Name:  New Quick

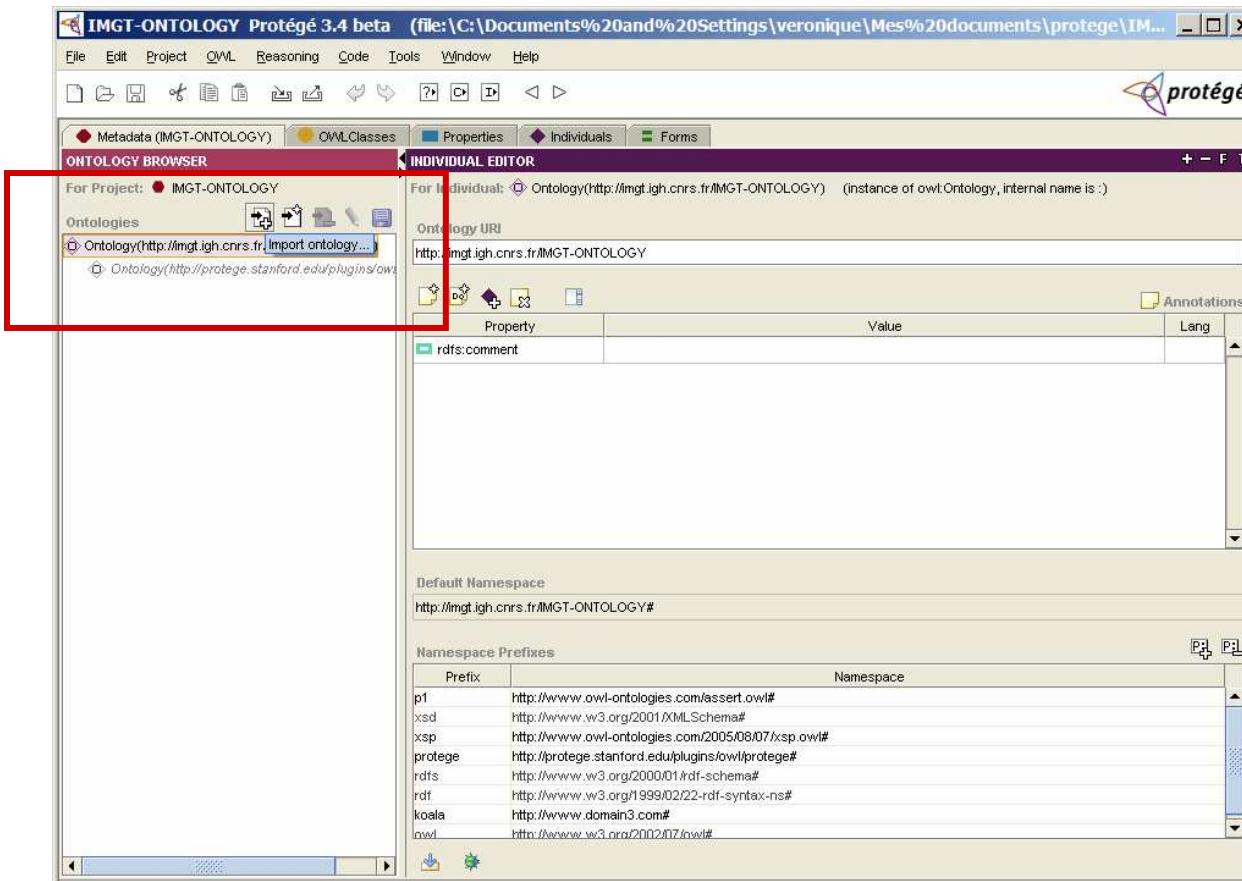
Nested Quick View  
 Flat Quick View

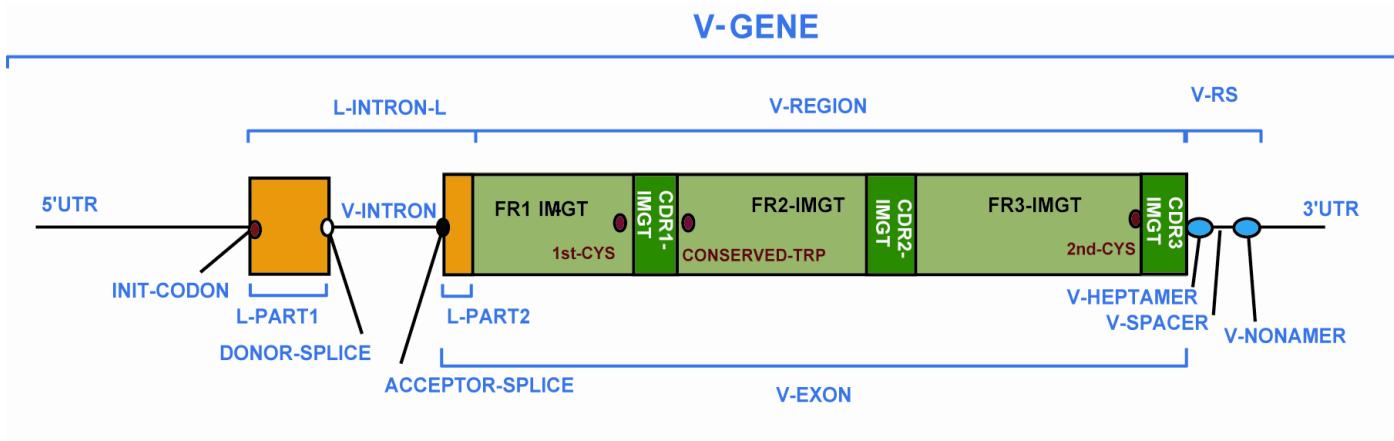


## Importation d'ontologies

Il est possible d'utiliser les classes les propriétés et les instances de L'ontologie importée, d'en étendre la description

⇒ Attention, la référence aux classes, propriétés et individus d'une autre Ontologie en utilisant l'espace de nom est différent de l'importation complète D'ontologies.





- 1- faire l'inventaire des termes (qui correspondent à autant de concepts)
- 2- les classer dans 2 catégories: codants ou non codants
- 3- créer l'arborescence des concepts
- 4- les visualiser avec OWLViz, OntoViz, Jambalaya

Relation	Reciprocal relation (inverse)
"adjacent_at_its_5_prime_to"	"adjacent_at_its_3_prime_to"
"included_with_same_5_prime_in",	"includes_with_same_5_prime",
"included_with_same_3_prime_in",	"includes_with_same_3_prime",
"overlap_at_its_5_prime_with"	"overlap_at_its_3_prime_with"
"included_in"	"includes"

- 5- créer les relations (et leurs caractéristiques)
- 6- définir les restrictions sur les relations
- 7- faire les tests de consistance
- 8- créer les instances (pour la séquence X62106)