http://www.imgt.org

Ontologies
Langage OWL

IMGT-ONTOLOGY

Application à la formalisation des concepts de classification et de description d'IMGT-ONTOLOGY avec l'éditeur Protégé

Véronique Giudicelli
LIGM, IGH, UPR CNRS 1142
141 rue de la Cardonille
34296 Montpellier Cedex 5
tel: 04.34.35.99.65 fax: 04.34.35.99.01
e-mail: Veronique.Giudicelli@igh.cnrs.fr

22-23 mai 2012

Plan

Ouverture compte Rez-UFR?

- Ontologies
- le Web sémantique
- Le langage OWL
- Exemples d'ontologies en biologie.

- Domaine d'application en immunogénétique:
IMGT, The international ImMunoGenetics information system

- Formalisation des concepts de classification
- Formalisation des concepts de description

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 connaissances dans un domaine particulier comme l'immunogénétique.
Une ontologie cherche à en représenter le sens (sémantique).

Nécessité de définir un vocabulaire standardisé utilisé dans le domaine de connaissance

Quelle est la signification des termes dans le contexte?

=> Terminologie + définition pour chaque terme.

Comment ces termes sont–ils organisés?
Comment sont-il reliés?

=> Propriétés expriment relations entre les connaissances

Ontologies

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

Ontologies

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

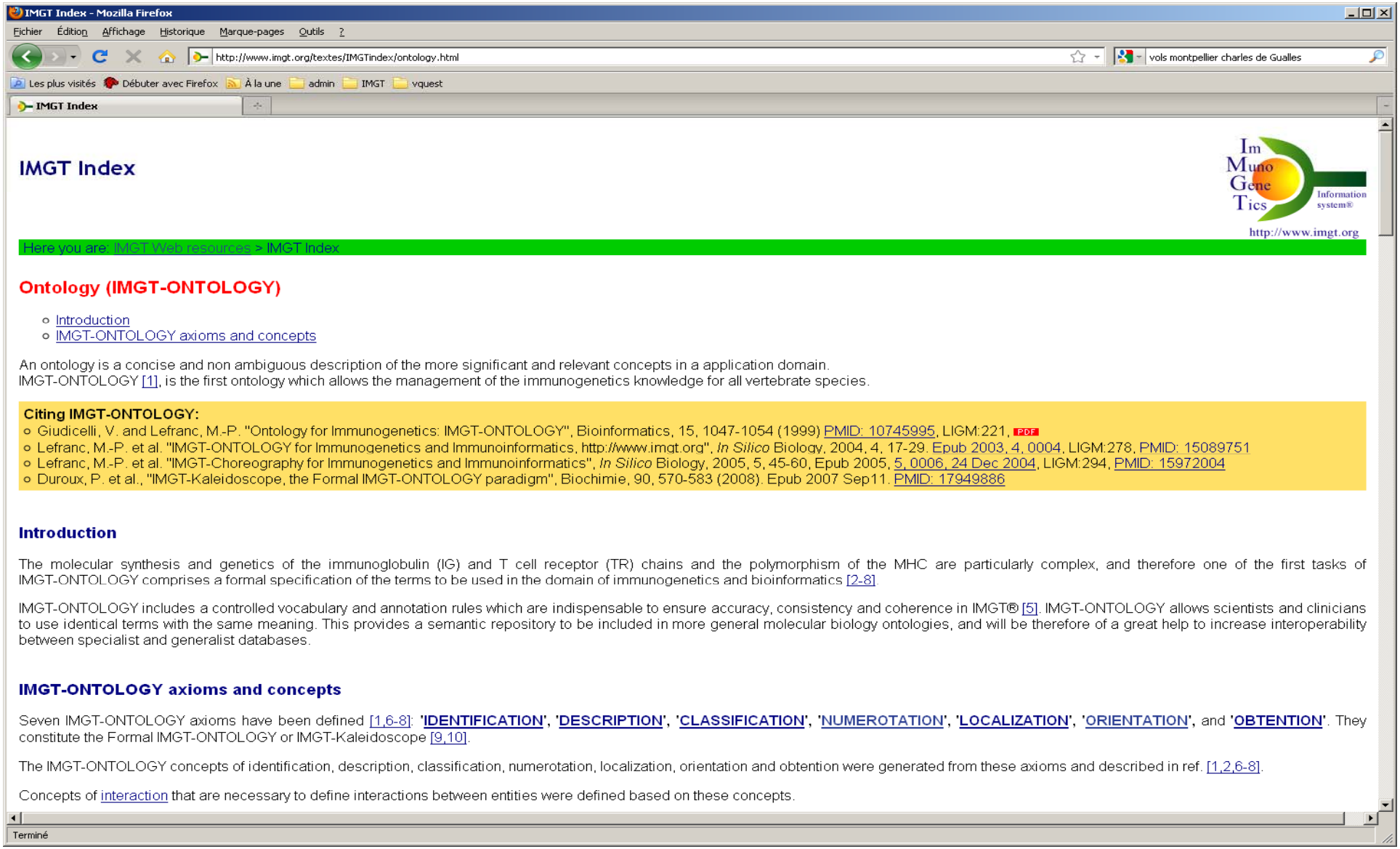
Bases de connaissances:

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.

World Wide Web

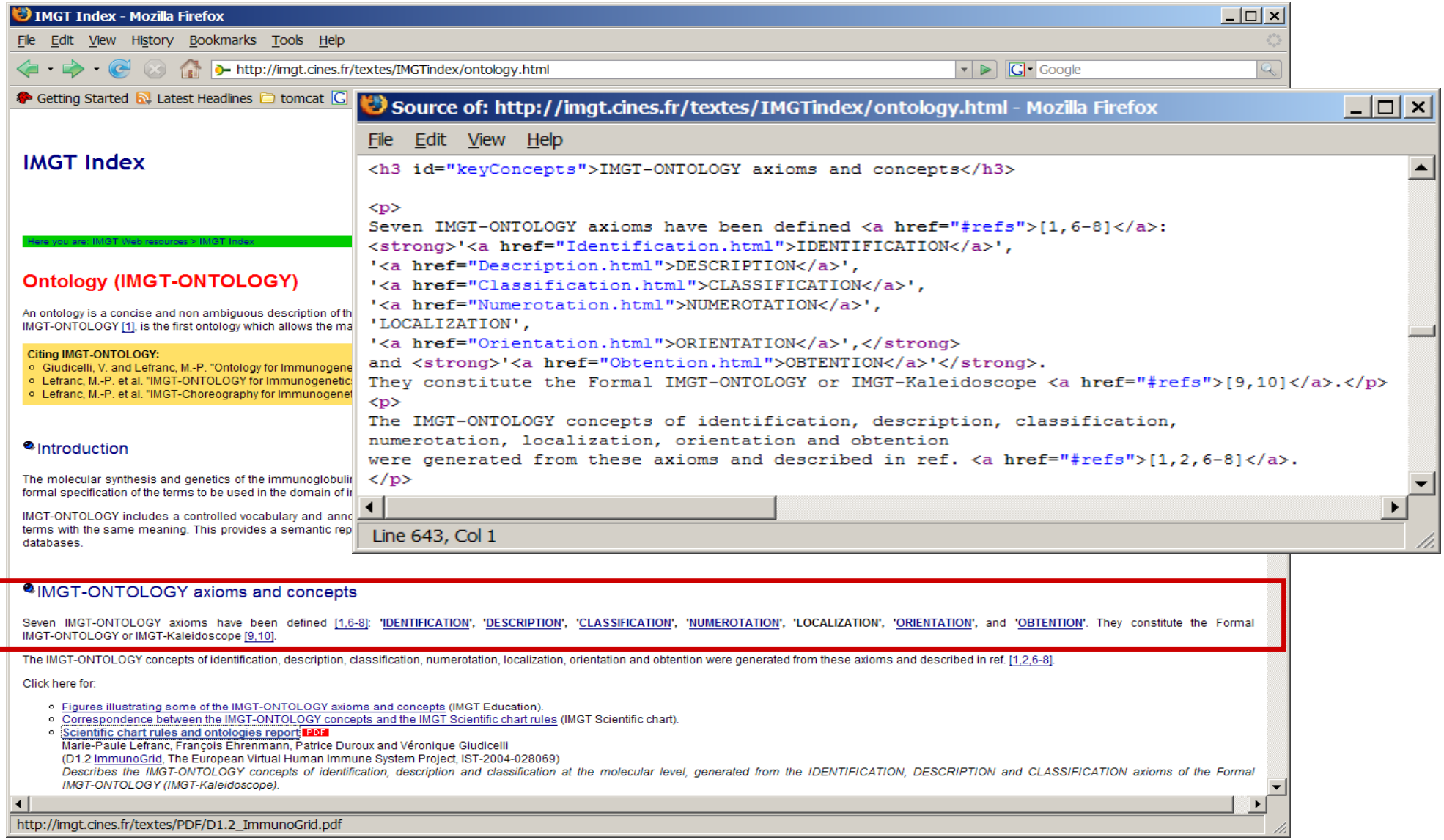
(Créé par Tim Berners Lee en 1990)

=> Pour partager des informations/connaissances et les relier



World Wide Web

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



World Wide Web

Limites

- **Balises HTML: la connaissance est encapsulée dans une couche de présentation**

- **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

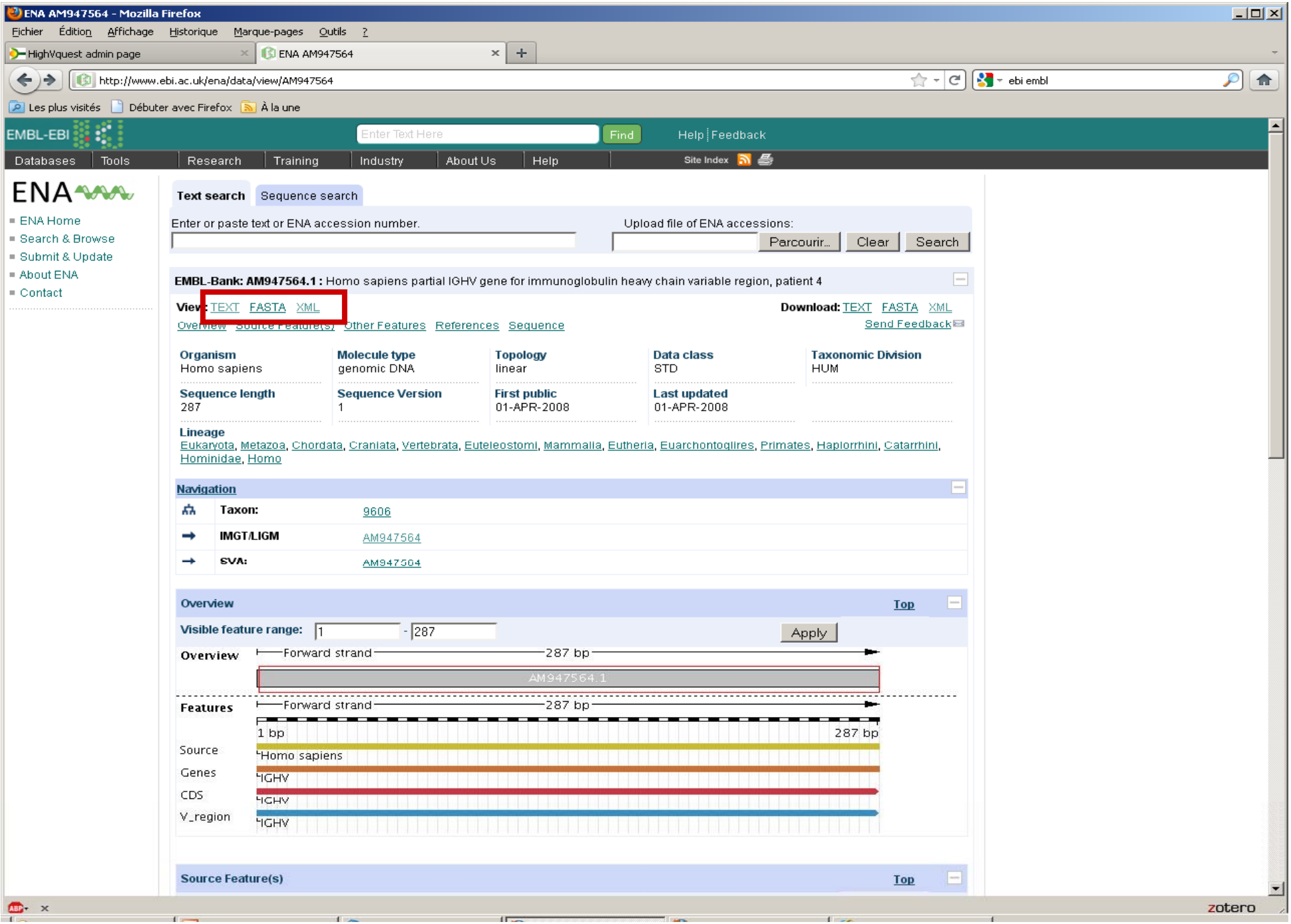
World Wide Web Consortium (W3C)

(créé en 1994)

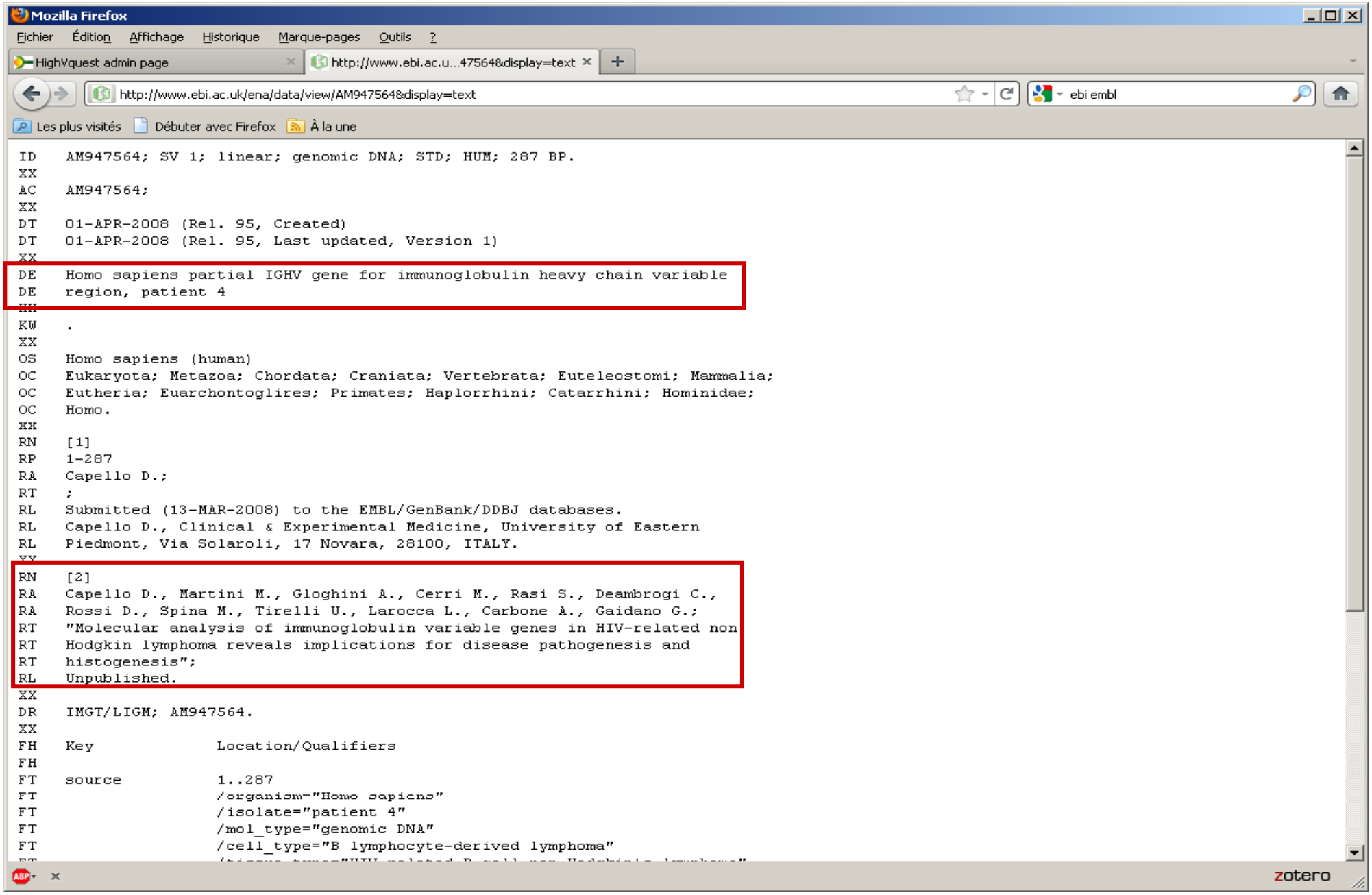
- Développement de technologies (spécifications, guides, logiciels et outils)
- Proposition des standards, pour la croissance et l'exploitation du Web

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

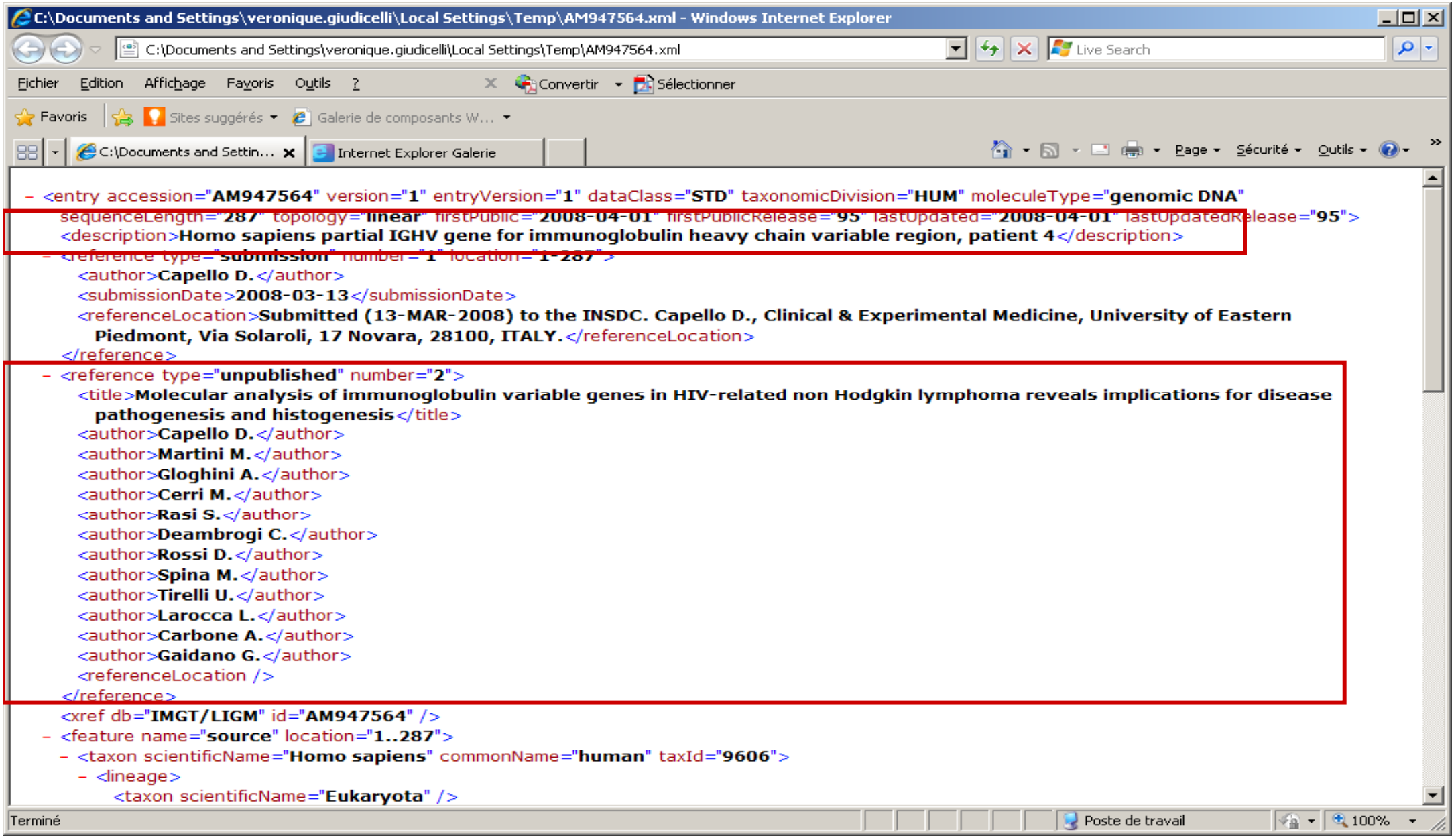
- XML devrait pouvoir être utilisé sans difficulté sur Internet
- Les documents XML devraient être lisibles par l'homme et raisonnablement clairs
- Il devrait être facile de créer des documents XML
- Il devrait être facile d'écrire des programmes traitant les documents XML


ENA AM947564 - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
HighVquest admin page
ENA AM947564
http://www.ebi.ac.uk/ena/data/view/AM947564
ebi embl
Les plus visités Débuter avec Firefox À la une
EMBL-EBI
Enter Text Here
Find
Help | Feedback
Databases Tools Research Training Industry About Us Help
Site Index
ENA
ENA Home
Search & Browse
Submit & Update
About ENA
Contact
Text search
Sequence search
Enter or paste text or ENA accession number.
Upload file of ENA accessions:
Parcourir...
Clear
Search
EMBL-Bank: AM947564.1 : Homo sapiens partial IGHV gene for immunoglobulin heavy chain variable region, patient 4
View: TEXT FASTA XML
Download: TEXT FASTA XML
Overview Source Feature(s) Other Features References Sequence
Send Feedback
Organism
Homo sapiens
Molecule type
genomic DNA
Topology
linear
Data class
STD
Taxonomic Division
HUM
Sequence length
287
Sequence Version
1
First public
01-APR-2008
Last updated
01-APR-2008
Lineage
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Euarchontoglires, Primates, Haplorrhini, Catarrhini, Hominidae, Homo
Navigation
Taxon: 9606
IMGT/LIGM AM947564
SVA: AM947564
Overview
Top
Visible feature range: 1 - 287
Apply
Overview
Forward strand
287 bp
AM947564.1
Features
Forward strand
287 bp
1 bp
287 bp
Source
Homo sapiens
Genes
IGHV
CDS
IGHV
V_region
IGHV
Source Feature(s)
Top
zotero

Séquence EMBL AM947564



Séquence AM947564 en format XML



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 URI comme schéma de nommage
  (pour éviter la confusion entre termes identiques mais différents sémantiquement selon le contexte)
- On peut tout décrire
- Grande souplesse quand à l'extensibilité

URI: Uniform Resource Identifier

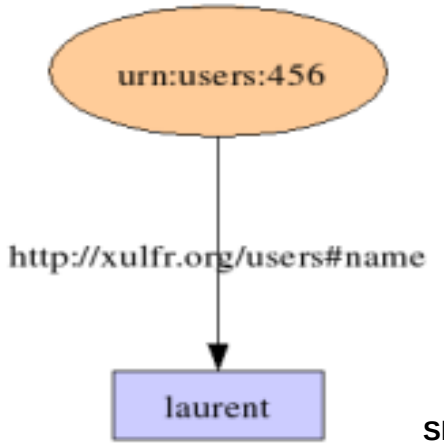
RDF

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"

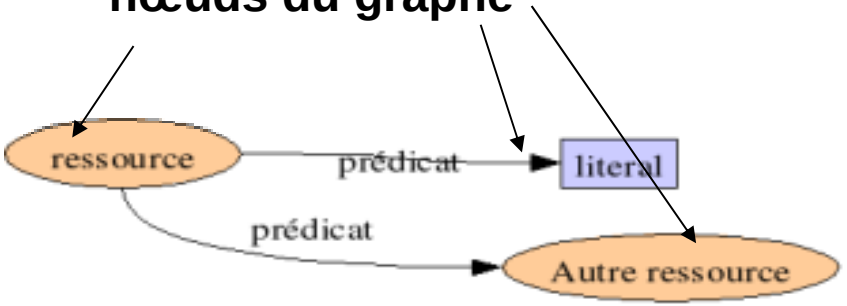
Sujet
(source, ressource)

Prédicat
(propriété)

Objet
(cible, valeur)



nœuds du graphe



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

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

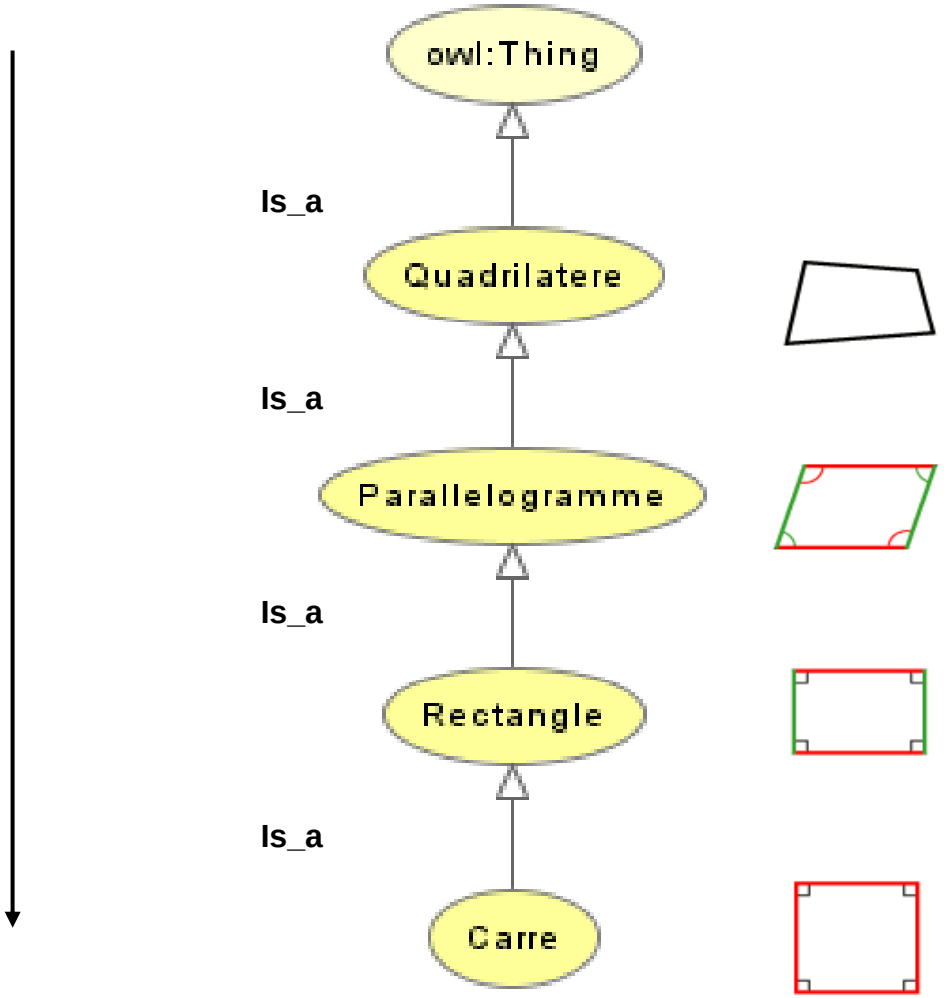
- **Un espace de nom, :**
- **L'entête <owl :Ontology> pour décrire l'ontologie**
- **La définition des classes, des propriétés et des instances**

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: Parallélogramme à angles droits

Def: rectangle dont les cotés sont égaux

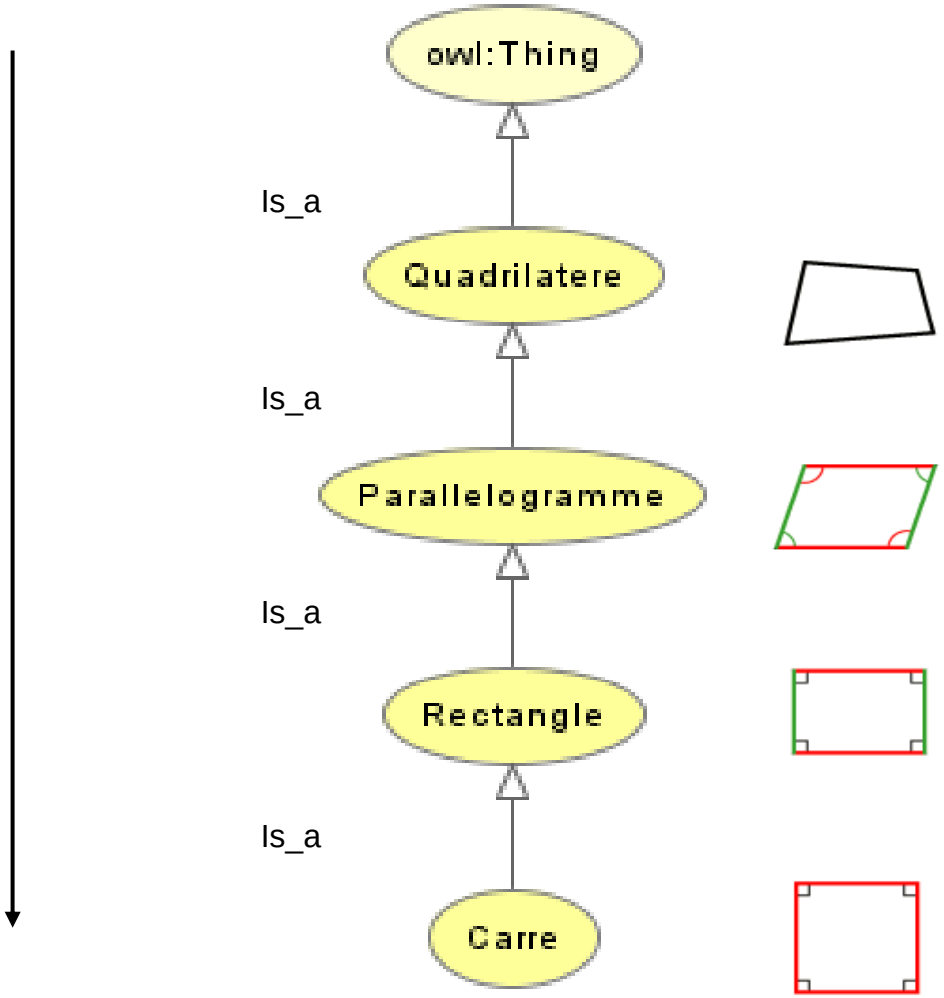
Exemple simple: les formes géométriques
représentation hiérarchique de concepts avec la relation « is_a »
(OWL, Web Ontology Language)

Notion de Classe :

Définit un groupe d'individus possédant des propriétés similaires.

Thing est la classe mère.

spécialisation



Def : polygone à 4 côtés

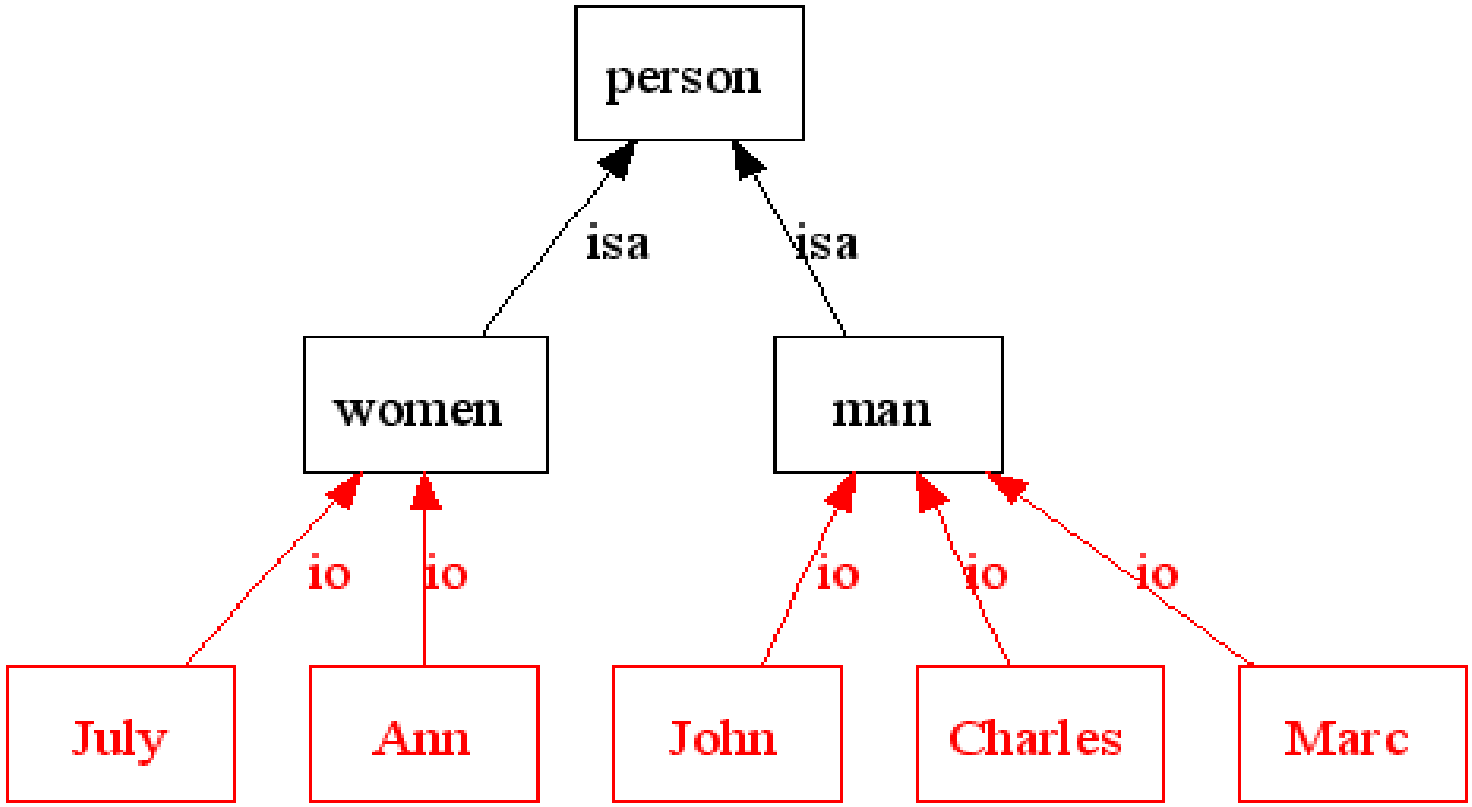
Def: quadrilatère (convexe) dont les côtés sont parallèles deux à deux

Def: Parallélogramme à angles droits

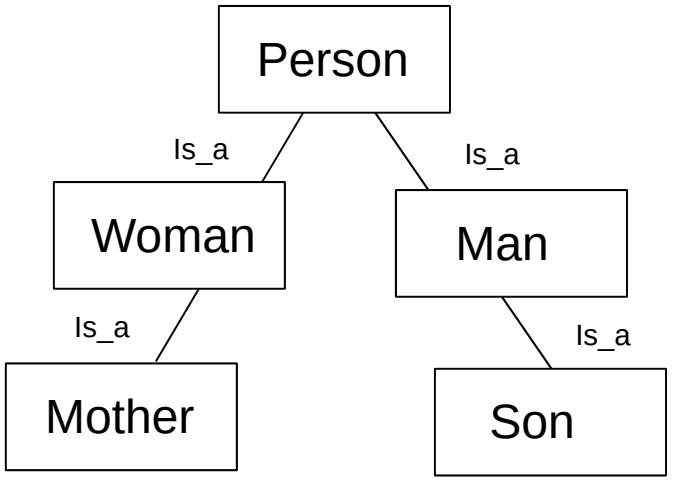
Def: rectangle dont les cotés sont égaux

Instances

Ce sont les « individus » qui peuplent les classes



Propriétés/Relations
(OWL, Web Ontology Language)



Existe-t-il des relations particulières
Autres que « is_a » :

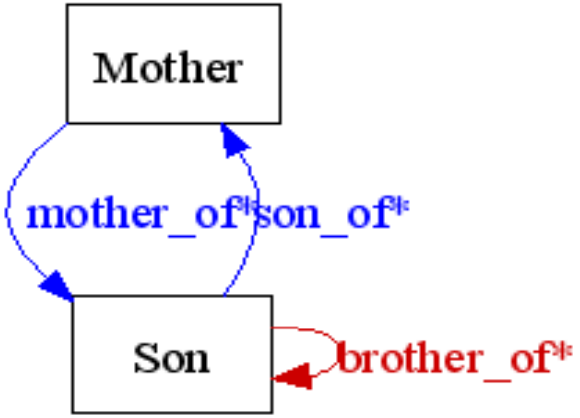
- fondamentales sur le plan sémantique
- caractéristiques des concepts/classes
- qui lient les concepts/classes entre eux?

Propriétés/Relations
(OWL, Web Ontology Language)

-Propriété : de type de données:
classe =>ensemble de valeurs (numériques, alphabétique,…)

- date de naissance
- numéro de sécurité sociale
…..

- Propriété d'objet : classe => classe

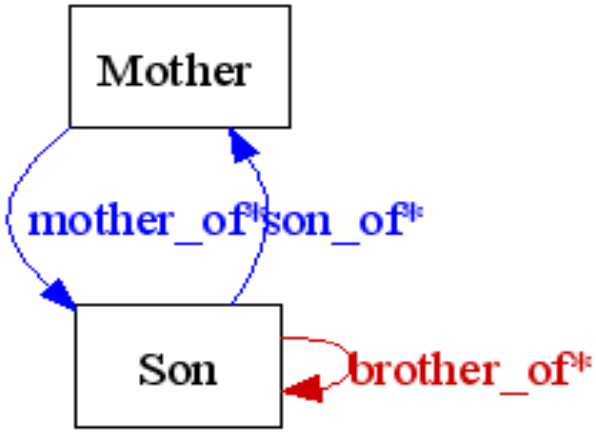


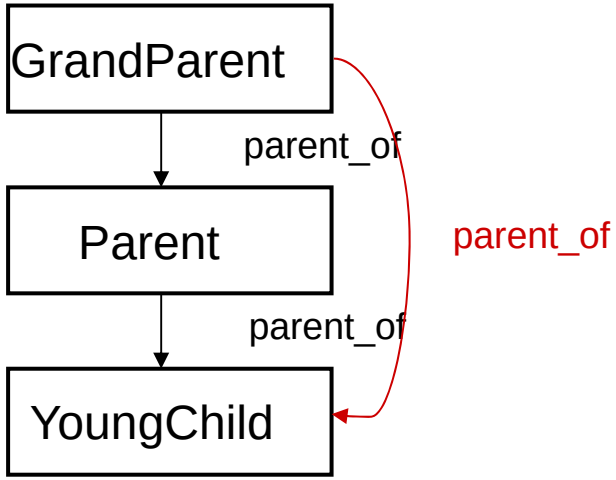
Définir et caractériser les Propriétés/Relations (OWL, Web Ontology Language)

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/relations: symétriques, transitives, inverses et fonctionnelles

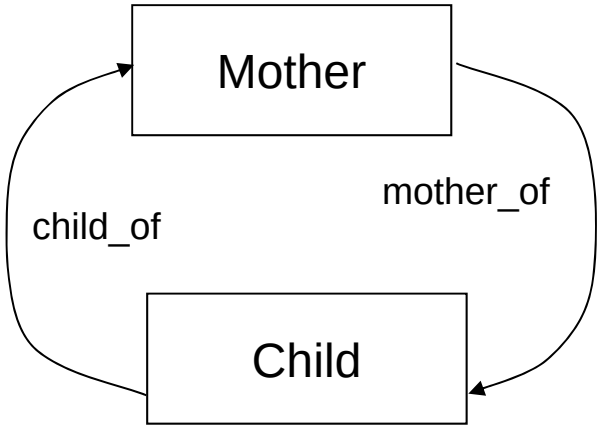




(familles de 3 générations avec de jeunes enfants)

Symétrique: brother_of
Transitive: parent_of
Inverse: son_of / mother_of
Fonctionnelle : son_of

Restriction sur les Propriétés/Relations : 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 sera relié par la propriété à au moins x individus Exemple: propriété mother_of de la classe Mother : minCardinality=3

maxCardinality: toute instance de la classe sera relié par la propriété à au plus x individus Exemple: propriété mother_of de la classe Mother : maxCardinality=10

cardinality: toute instance de la classe sera relié par la propriété à exactement x individus Exemple: propriété son_of de la classe Mother : cardinality=1

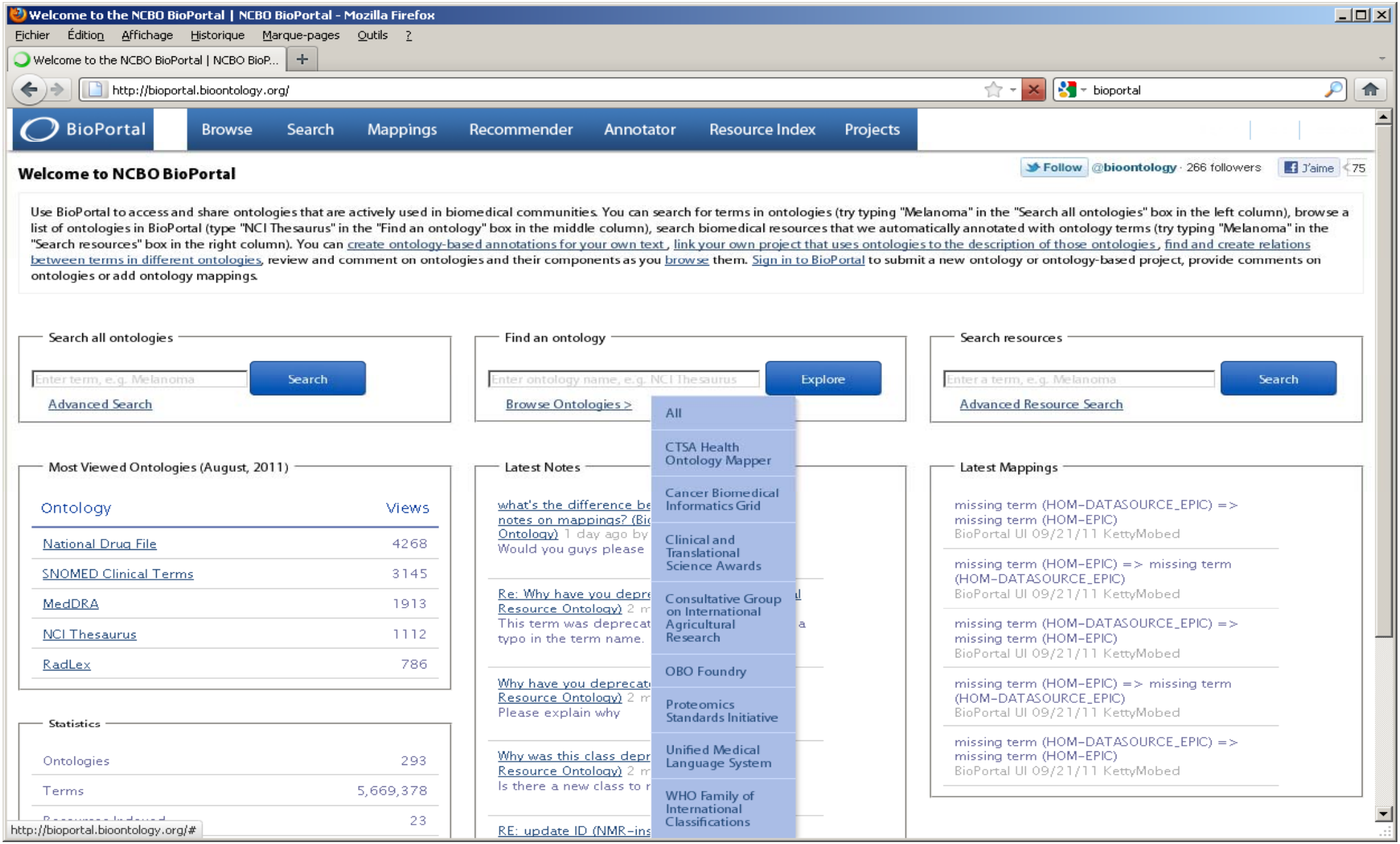
Les ontologies en Biologie

- 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.

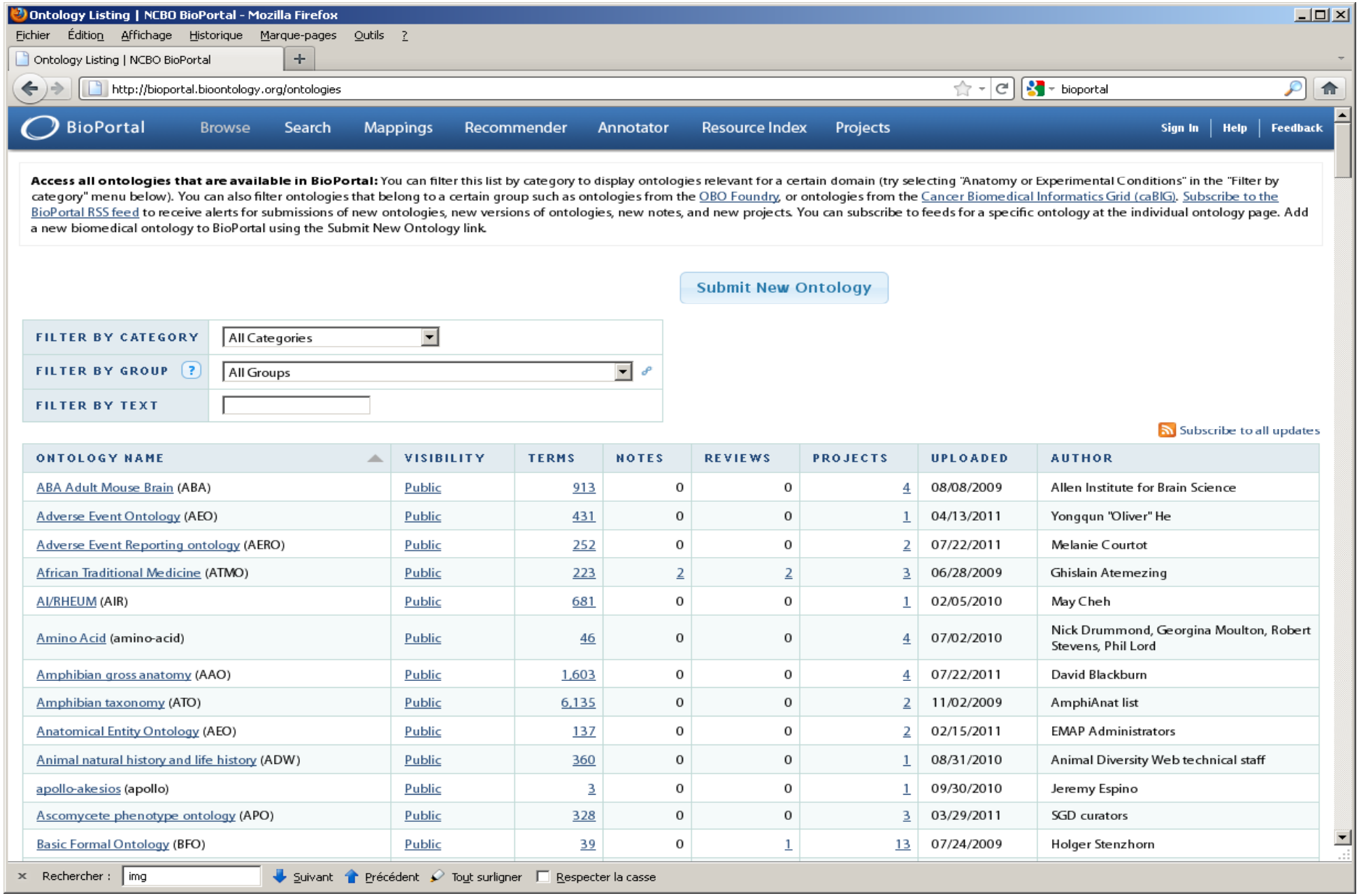
NCBO Bioportal (http://bioportal.bioontology.org/)
OBO (Open Biological Ontologies) (http://www.obofoundry.org/)

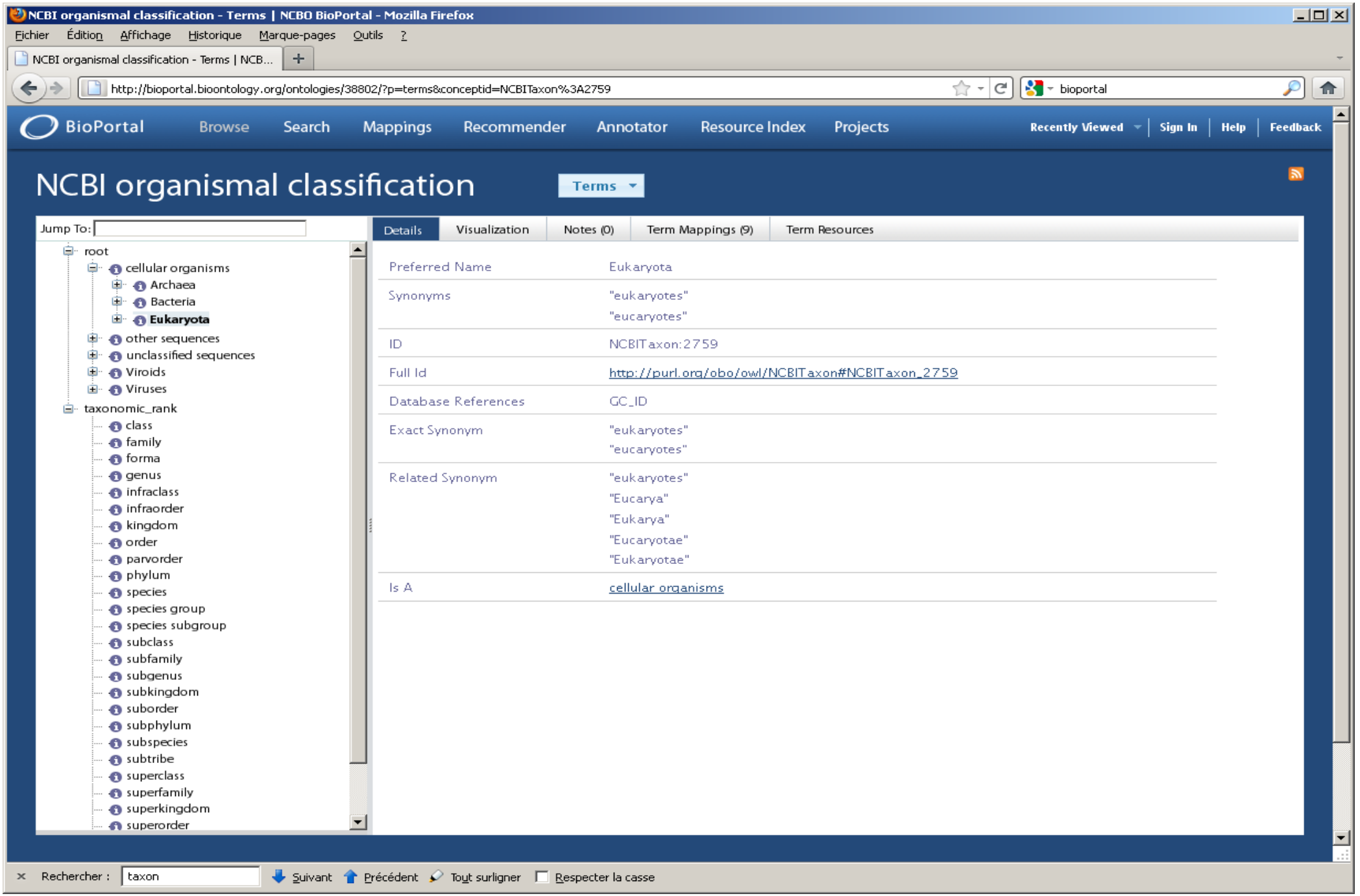
recensent les ontologies en biologie.

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

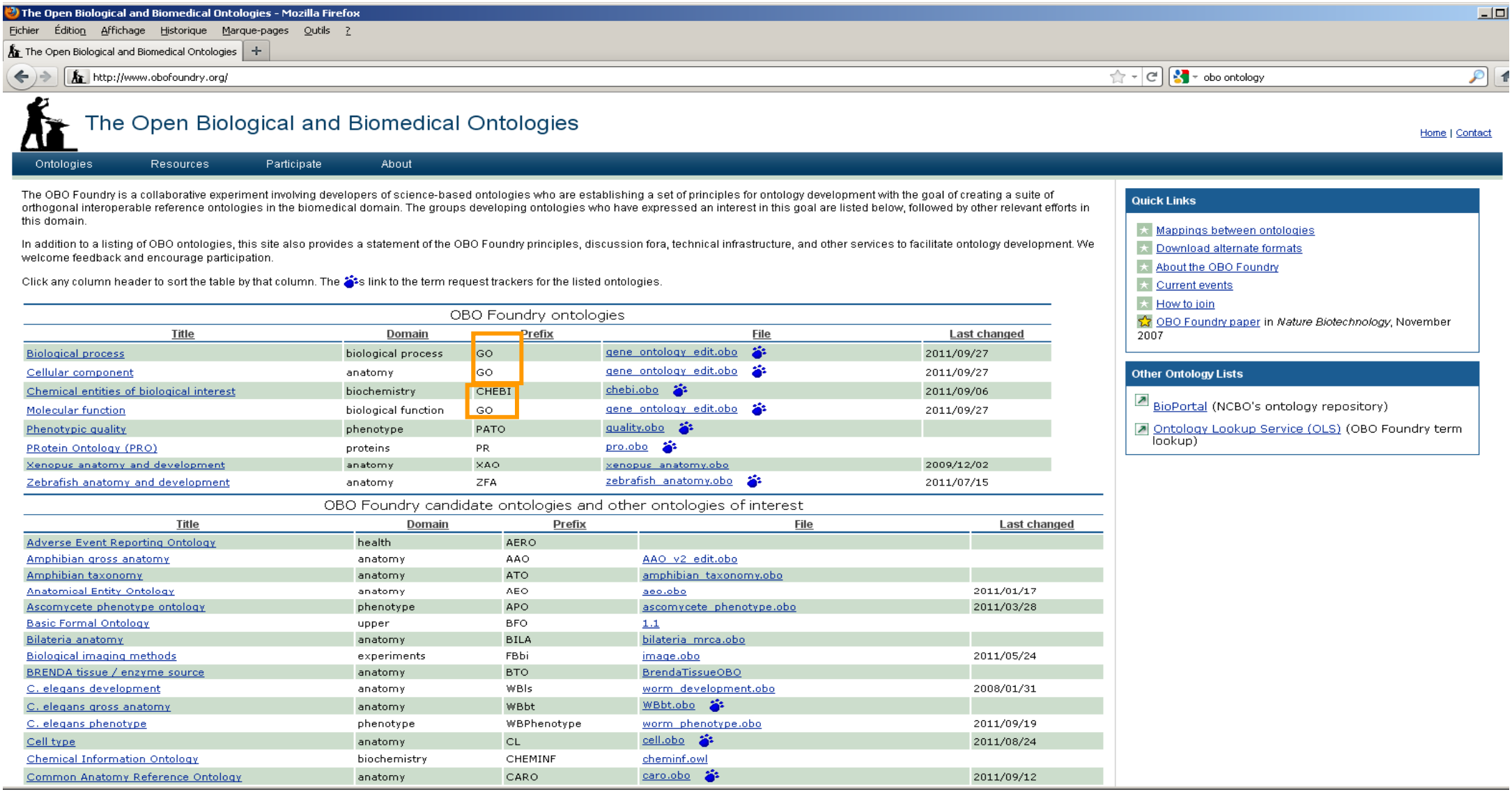


NCBO Bioportal

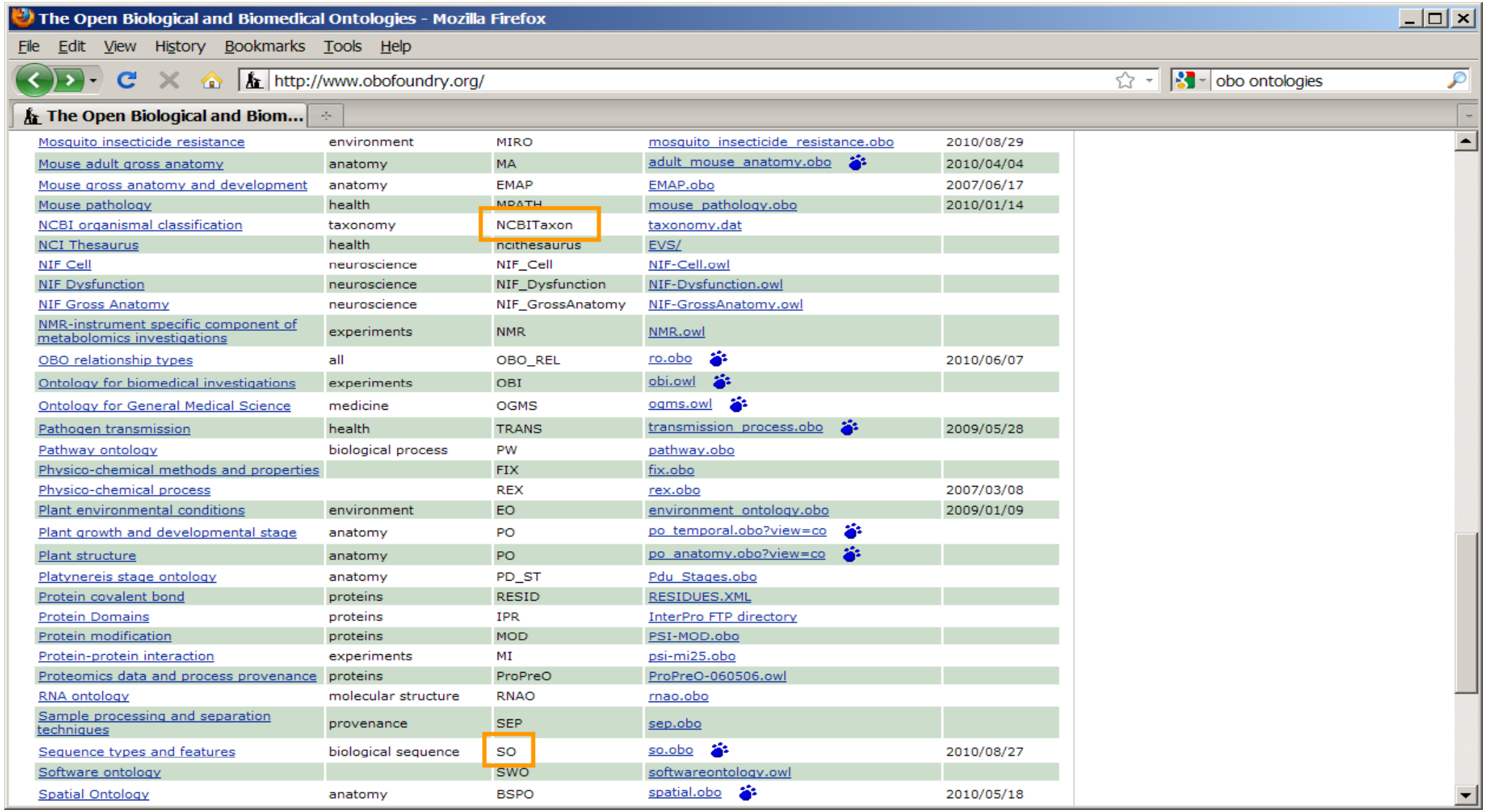



NCBI organismal classification - Terms | NCBO BioPortal - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
NCBI organismal classification - Terms | NCB...
http://bioportal.bioontology.org/ontologies/38802/?p=terms&conceptid=NCBITaxon%3A2759
bioportal
BioPortal
Browse
Search
Mappings
Recommender
Annotator
Resource Index
Projects
Recently Viewed
Sign In
Help
Feedback
NCBI organismal classification
Terms
Jump To:
root
cellular organisms
Archaea
Bacteria
Eukaryota
other sequences
unclassified sequences
Viroids
Viruses
taxonomic_rank
class
family
forma
genus
infraclass
infraorder
kingdom
order
parvorder
phylum
species
species group
species subgroup
subclass
subfamily
subgenus
subkingdom
suborder
subphylum
subspecies
subtribe
superclass
superfamily
superkingdom
superorder
Details
Visualization
Notes (0)
Term Mappings (9)
Term Resources
Preferred Name
Eukaryota
Synonyms
"eukaryotes"
"eucaryotes"
ID
NCBITaxon:2759
Full Id
http://purl.org/obo/owl/NCBITaxon#NCBITaxon_2759
Database References
GC_ID
Exact Synonym
"eukaryotes"
"eucaryotes"
Related Synonym
"eukaryotes"
"Eucarya"
"Eukarya"
"Eucaryotae"
"Eukaryotae"
Is A
cellular organisms
Rechercher :
taxon
Suivant
Précédent
Tout surligner
Respecter la casse

OBO (http://www.obofoundry.org/)



OBO (http://www.obofoundry.org/)



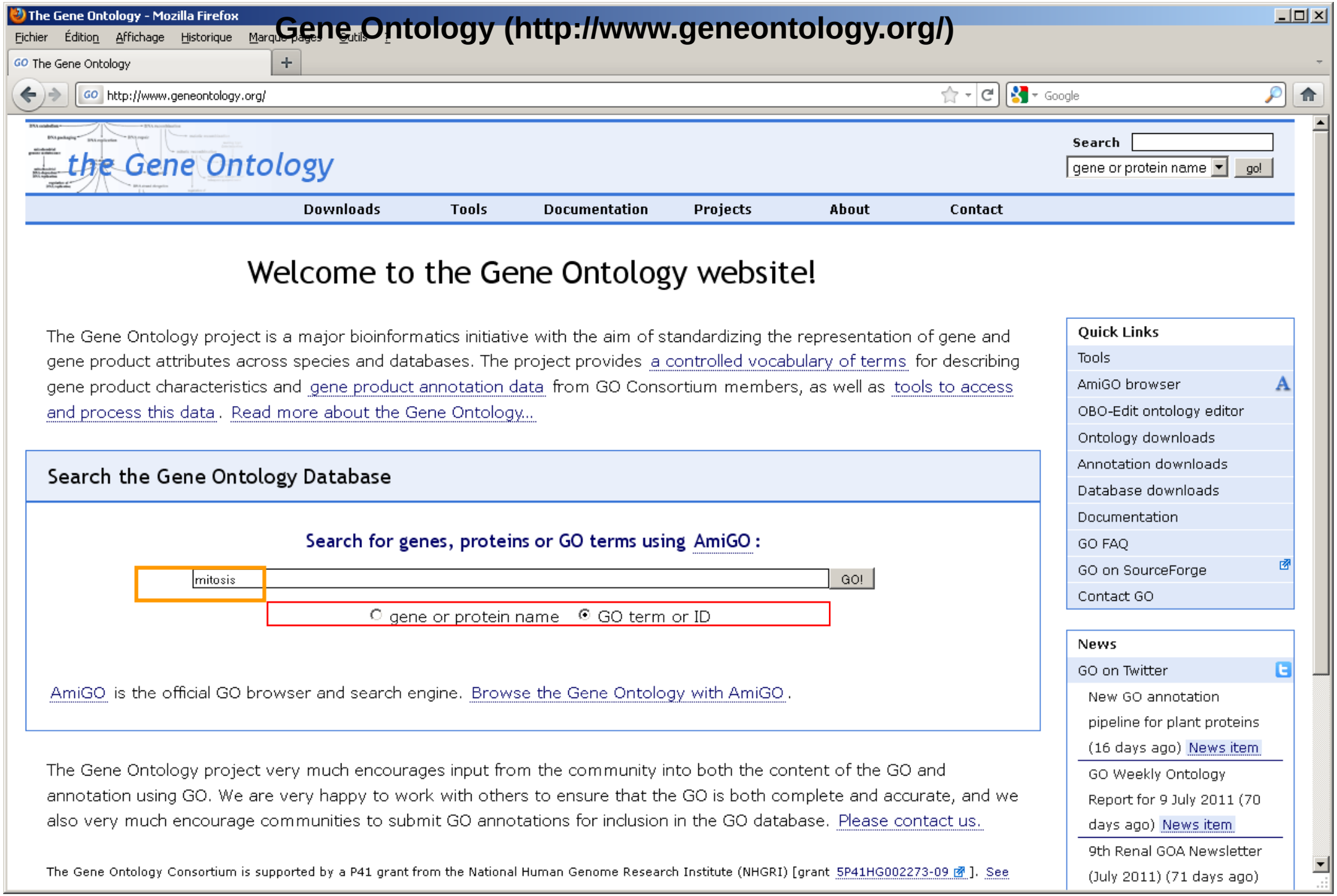
Mosquito insecticide resistance	environment	MIRO	mosquito_insecticide_resistance.obo	2010/08/29
Mouse adult gross anatomy	anatomy	MA	adult_mouse_anatomy.obo	2010/04/04
Mouse gross anatomy and development	anatomy	EMAP	EMAP.obo	2007/06/17
Mouse pathology	health	MPATH	mouse_pathology.obo	2010/01/14
NCBI organismal classification	taxonomy	NCBITaxon	taxonomy.dat	
NCI Thesaurus	health	ncithesaurus	EVS/	
NIF Cell	neuroscience	NIF_Cell	NIF-Cell.owl	
NIF Dysfunction	neuroscience	NIF_Dysfunction	NIF-Dysfunction.owl	
NIF Gross Anatomy	neuroscience	NIF_GrossAnatomy	NIF-GrossAnatomy.owl	
NMR-instrument specific component of metabolomics investigations	experiments	NMR	NMR.owl	
OBO relationship types	all	OBO_REL	ro.obo	2010/06/07
Ontology for biomedical investigations	experiments	OBI	obi.owl	
Ontology for General Medical Science	medicine	OGMS	ogms.owl	
Pathogen transmission	health	TRANS	transmission_process.obo	2009/05/28
Pathway ontology	biological process	PW	pathway.obo	
Physico-chemical methods and properties		FIX	fix.obo	
Physico-chemical process		REX	rex.obo	2007/03/08
Plant environmental conditions	environment	EO	environment_ontology.obo	2009/01/09
Plant growth and developmental stage	anatomy	PO	po_temporal.obo?view=co	
Plant structure	anatomy	PO	po_anatomy.obo?view=co	
Platynereis stage ontology	anatomy	PD_ST	Pdu_Stages.obo	
Protein covalent bond	proteins	RESID	RESIDUES.XML	
Protein Domains	proteins	IPR	InterPro FTP directory	
Protein modification	proteins	MOD	PSI-MOD.obo	
Protein-protein interaction	experiments	MI	psi-mi25.obo	
Proteomics data and process provenance	proteins	ProPreO	ProPreO-060506.owl	
RNA ontology	molecular structure	RNAO	rnao.obo	
Sample processing and separation techniques	provenance	SEP	sep.obo	
Sequence types and features	biological sequence	SO	so.obo	2010/08/27
Software ontology		SWO	softwareontology.owl	
Spatial Ontology	anatomy	BSPO	spatial.obo	2010/05/18

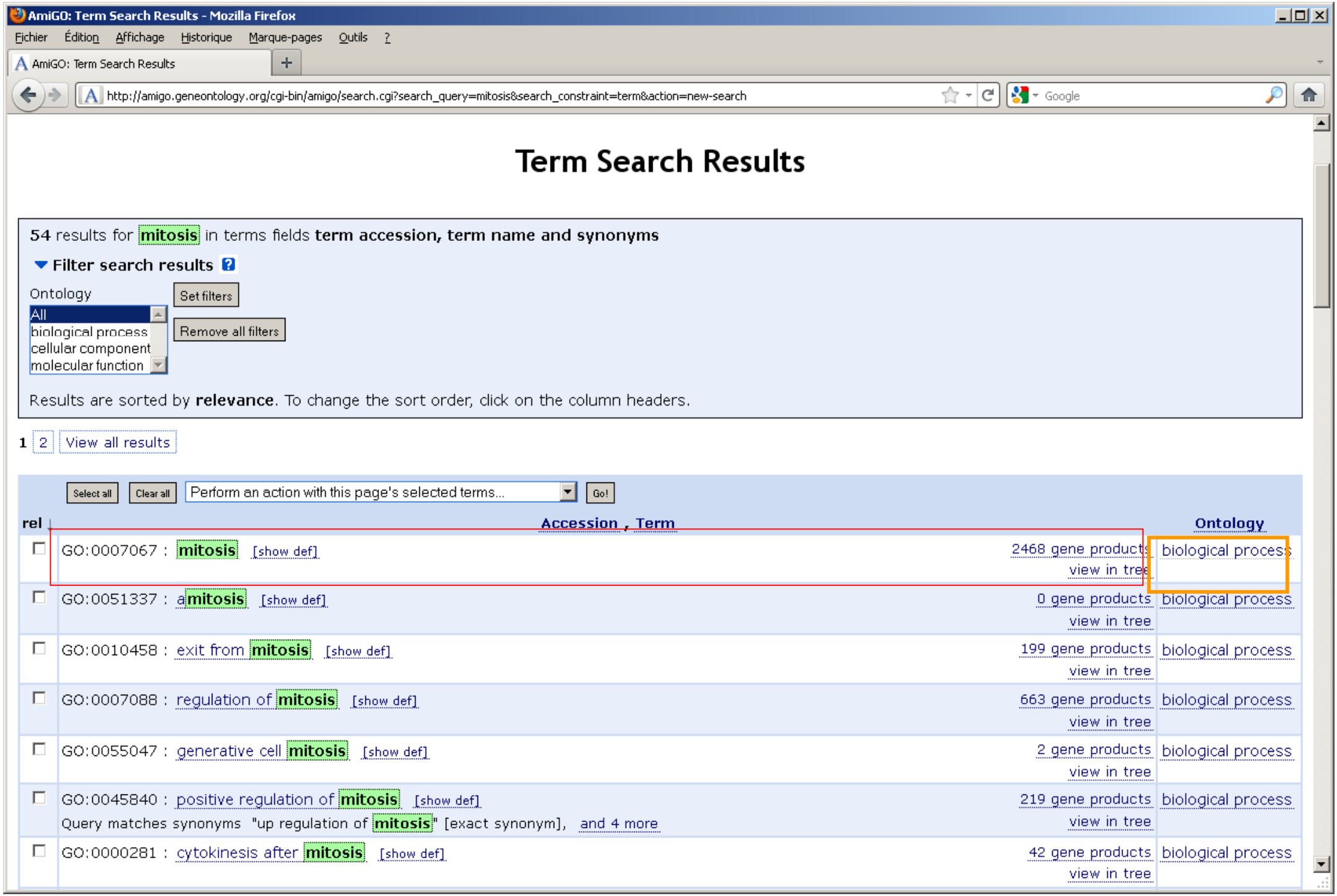
Gene Ontology.

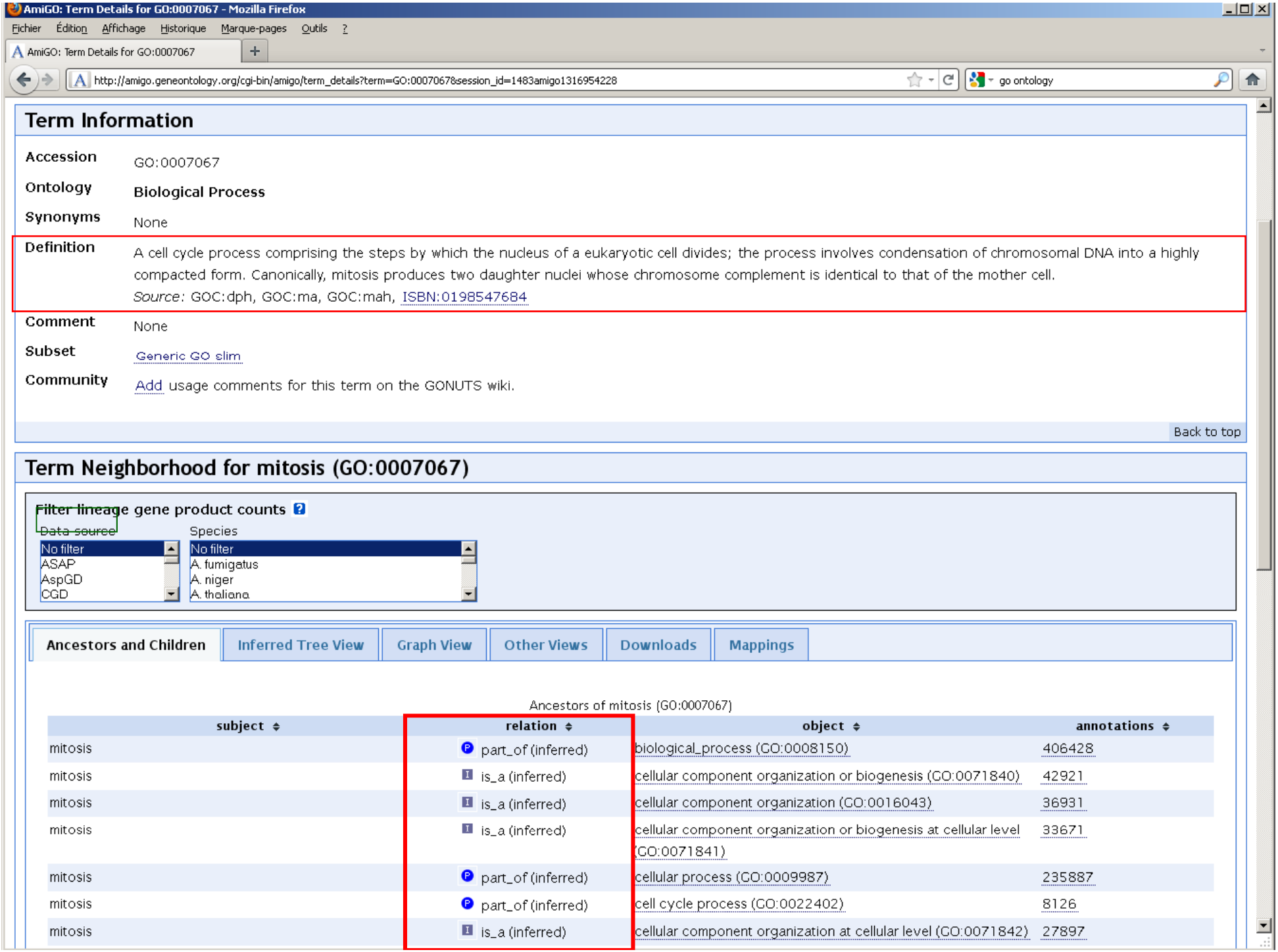
- GO a été créée en 1998. GO résulte d'une collaboration entre plusieurs bases de données: FlyBase (drosophile), the Saccharomyces Genome Database, et des bases de données de génomes (homme et souris), etc.

- 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 macromoleculaires ex: le noyau, le télomère).

Gene Ontology (http://www.geneontology.org/)




AmiGO: Term Search Results - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
AmiGO: Term Search Results
http://amigo.geneontology.org/cgi-bin/amigo/search.cgi?search_query=mitosis&search_constraint=term&action=new-search
Google
Term Search Results
54 results for mitosis in terms fields term accession, term name and synonyms
Filter search results
Ontology
All
biological process
cellular component
molecular function
Set filters
Remove all filters
Results are sorted by relevance. To change the sort order, click on the column headers.
1 2 View all results
Select all
Clear all
Perform an action with this page's selected terms...
Go!
rel
Accession , Term
Ontology
GO:0007067 : mitosis [show def]
2468 gene products
view in tree
biological process
GO:0051337 : amitosis [show def]
0 gene products
view in tree
biological process
GO:0010458 : exit from mitosis [show def]
199 gene products
view in tree
biological process
GO:0007088 : regulation of mitosis [show def]
663 gene products
view in tree
biological process
GO:0055047 : generative cell mitosis [show def]
2 gene products
view in tree
biological process
GO:0045840 : positive regulation of mitosis [show def]
Query matches synonyms "up regulation of mitosis" [exact synonym], and 4 more
219 gene products
view in tree
biological process
GO:0000281 : cytokinesis after mitosis [show def]
42 gene products
view in tree
biological process


AmiGO: Term Details for GO:0007067 - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
AmiGO: Term Details for GO:0007067
http://amigo.geneontology.org/cgi-bin/amigo/term_details?term=GO:0007067&session_id=1483amigo1316954228
go ontology
Term Information
Accession GO:0007067
Ontology Biological Process
Synonyms None
Definition A cell cycle process comprising the steps by which the nucleus of a eukaryotic cell divides; the process involves condensation of chromosomal DNA into a highly compacted form. Canonically, mitosis produces two daughter nuclei whose chromosome complement is identical to that of the mother cell.
Source: GOC:dph, GOC:ma, GOC:mah, ISBN:0198547684
Comment None
Subset Generic GO slim
Community Add usage comments for this term on the GONUTS wiki.
Back to top
Term Neighborhood for mitosis (GO:0007067)
Filter lineage gene product counts
Data source
No filter
ASAP
AspGD
CGD
Species
No filter
A. fumigatus
A. niger
A. thaliana
Ancestors and Children
Inferred Tree View
Graph View
Other Views
Downloads
Mappings
Ancestors of mitosis (GO:0007067)
subject
relation
object
annotations
mitosis
part_of (inferred)
biological_process (GO:0008150)
406428
mitosis
is_a (inferred)
cellular component organization or biogenesis (GO:0071840)
42921
mitosis
is_a (inferred)
cellular component organization (GO:0016043)
36931
mitosis
is_a (inferred)
cellular component organization or biogenesis at cellular level (GO:0071841)
33671
mitosis
part_of (inferred)
cellular process (GO:0009987)
235887
mitosis
part_of (inferred)
cell cycle process (GO:0022402)
8126
mitosis
is_a (inferred)
cellular component organization at cellular level (GO:0071842)
27897

AmiGO: Term Details for GO:0007067 - Mozilla Firefox

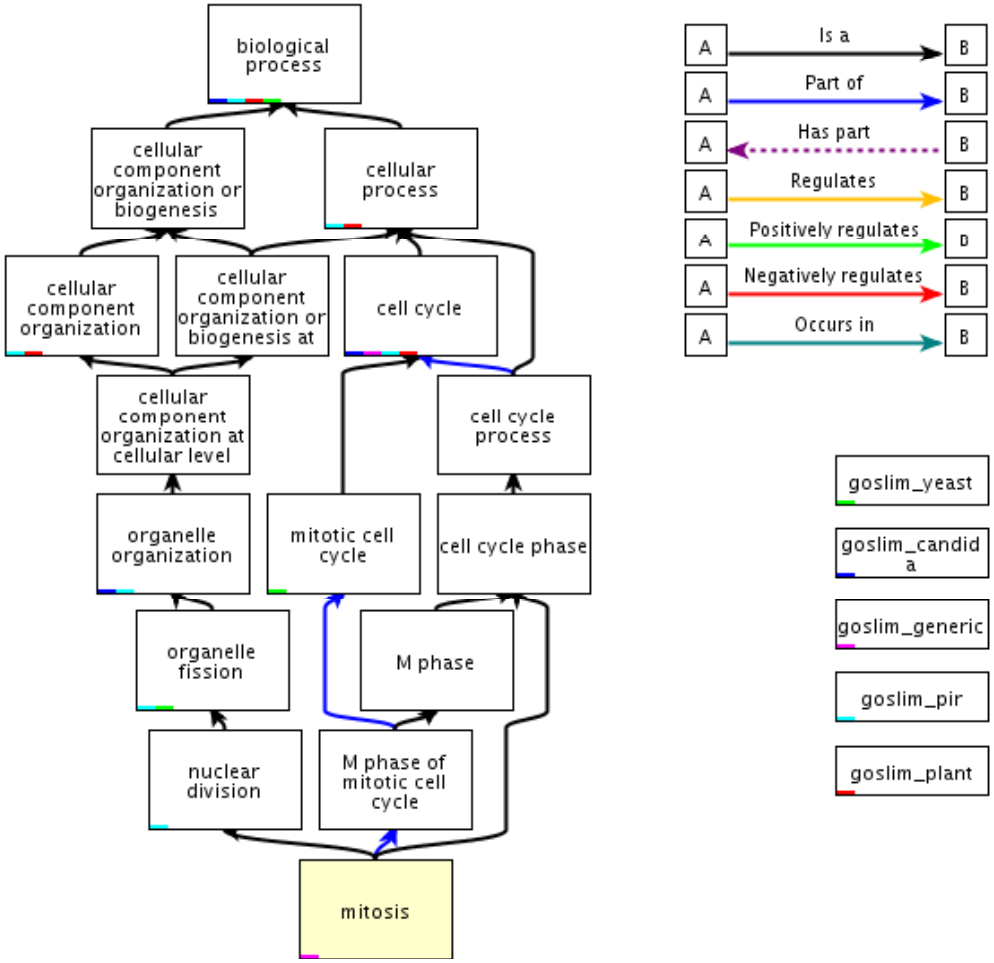
Fichier Édition Affichage Historique Marque-pages Outils ?

AmiGO: Term Details for GO:0007067

http://amigo.geneontology.org/cgi-bin/amigo/term_details?term=GO:0007067&session_id=1483amigo1316954228

go ontology

View this term in QuickGO.



goslim_yeast

goslim_candida

goslim_generic

goslim_pir

goslim_plant

QuickGO - http://www.ebi.ac.uk/QuickGO

Back to top

Sequence Ontology

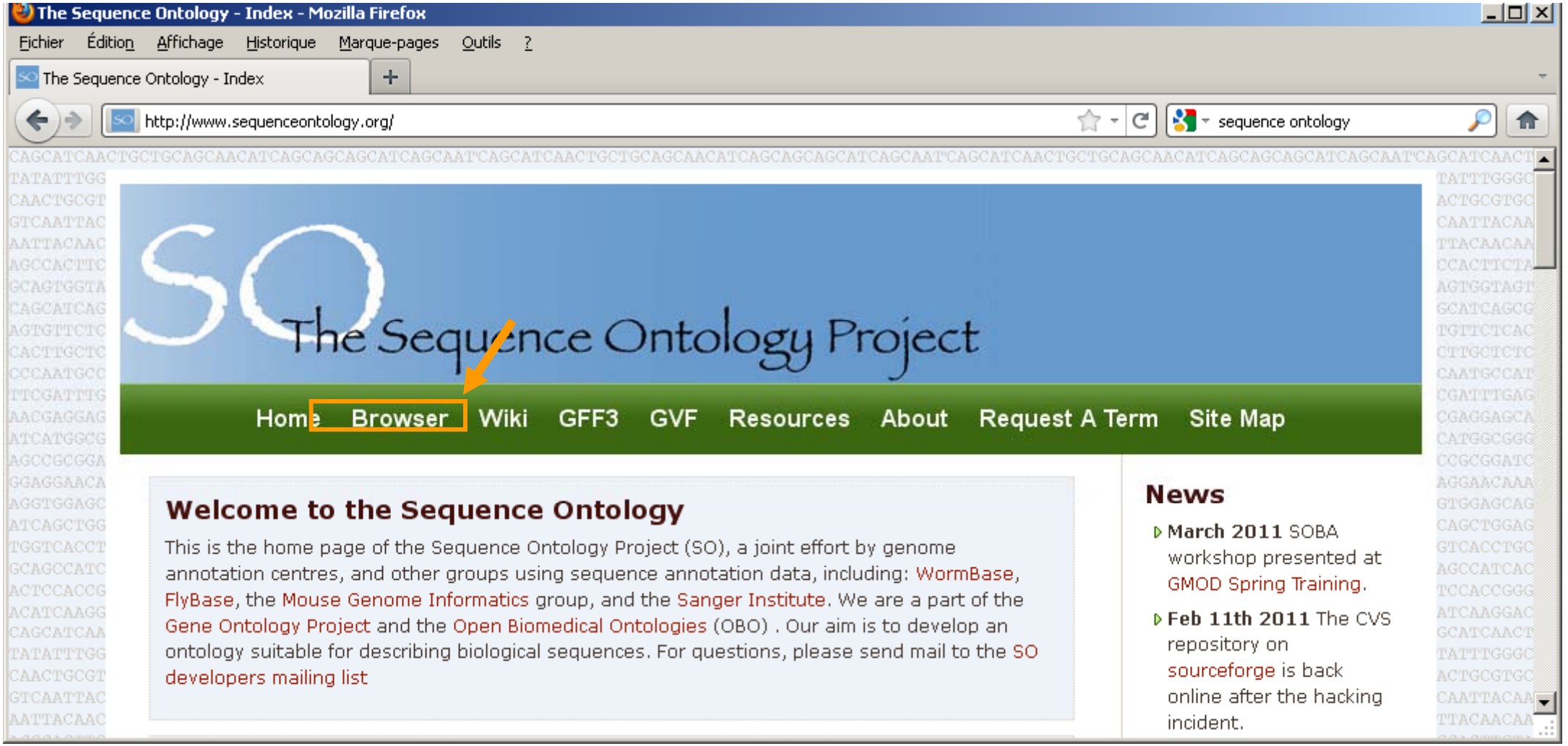
(http://www.sequenceontology.org/)

- Vocabulaire contrôlé pour l'annotation des séquences et l'annotation des génomes

- proposer une représentation structurée des annotations

- vocabulaire pour la description des mutations

Sequence Ontology:
une ontologie pour décrire les séquences biologiques



The MISO Sequence Ontology Browser - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
The MISO Sequence Ontology Browser
http://www.sequenceontology.org/browser/obob.cgi
sequence ontology
MISO
The Sequence Ontology Browser
Home Browser Wiki GFF3 Resources About Request A Term Site Map
Search for Terms: immmunoglobulin
Search
Select Release: CURRENT RELEASE
sequence_attribute
sequence_collection
sequence_feature
sequence_variant
Obsolete Terms
Relationship
Welcome to MISO, the Sequence Ontology Browser
Search for a SO term by entering a SO term name or synonym in the query box above;
Explore the structure of SO and browse for SO terms using the expandable, cascading tree on the left;
Go to the detail page for a term where you can:
Get details about a term, it's definition and relationships;
See graphical views of a term's place in the ontology and link to it's neighbors;
Export details about a term in a variety of formats;
And access and contribute detailed documentation about a term and it's history by linking through to the SO wiki
Copyright © 2008 Sequence Ontology. All rights reserved.
Valid XHTML 1.0 Strict

Le système d'information
IMGT, the international ImMunoGeneTics information system®
http://www.imgt.org

Immunoglobulin (IG)

T cell receptor (TR)

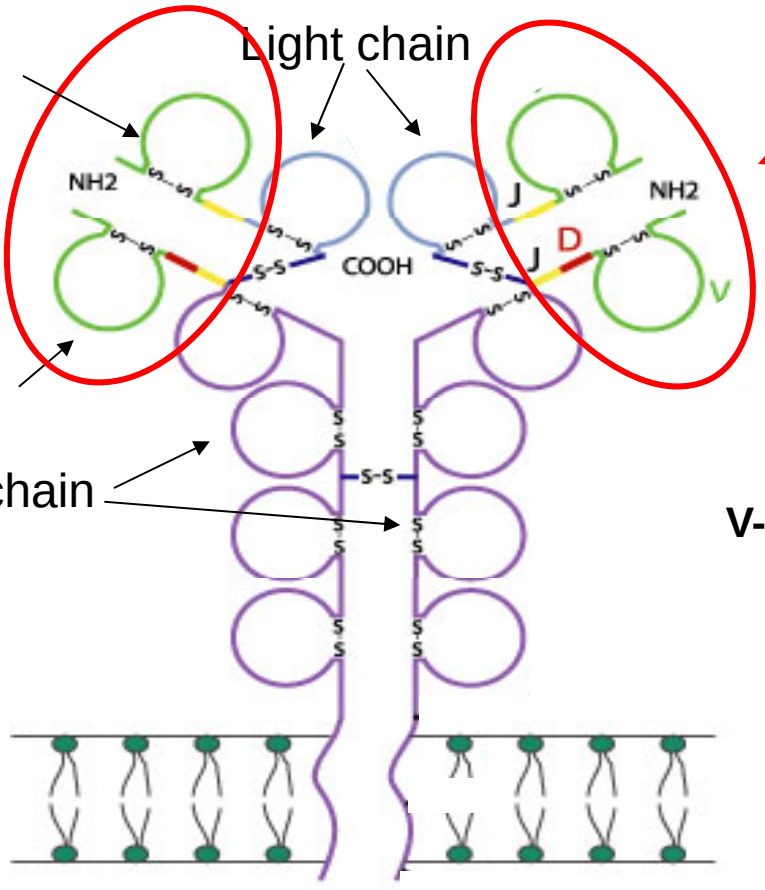
http://www.imgt.org

V-DOMAIN

V-J-REGION

V-DOMAIN

V-D-J-REGION

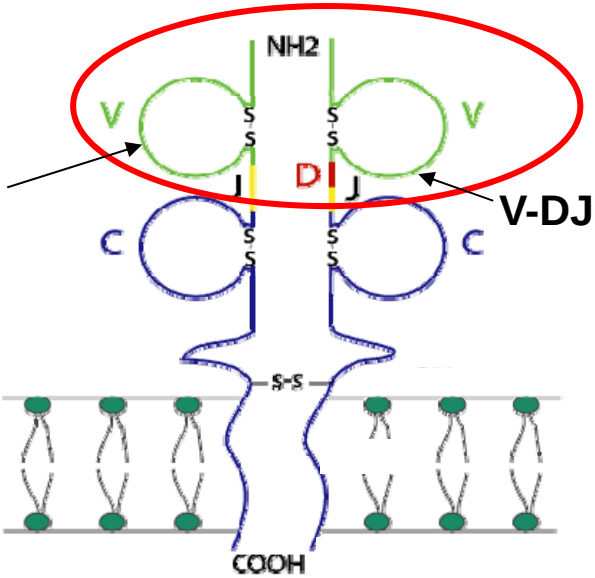


Membrane IgM

Contribution of the
2 V-DOMAINs
to the antigen binding site

Alpha - Beta

Gamma - Delta



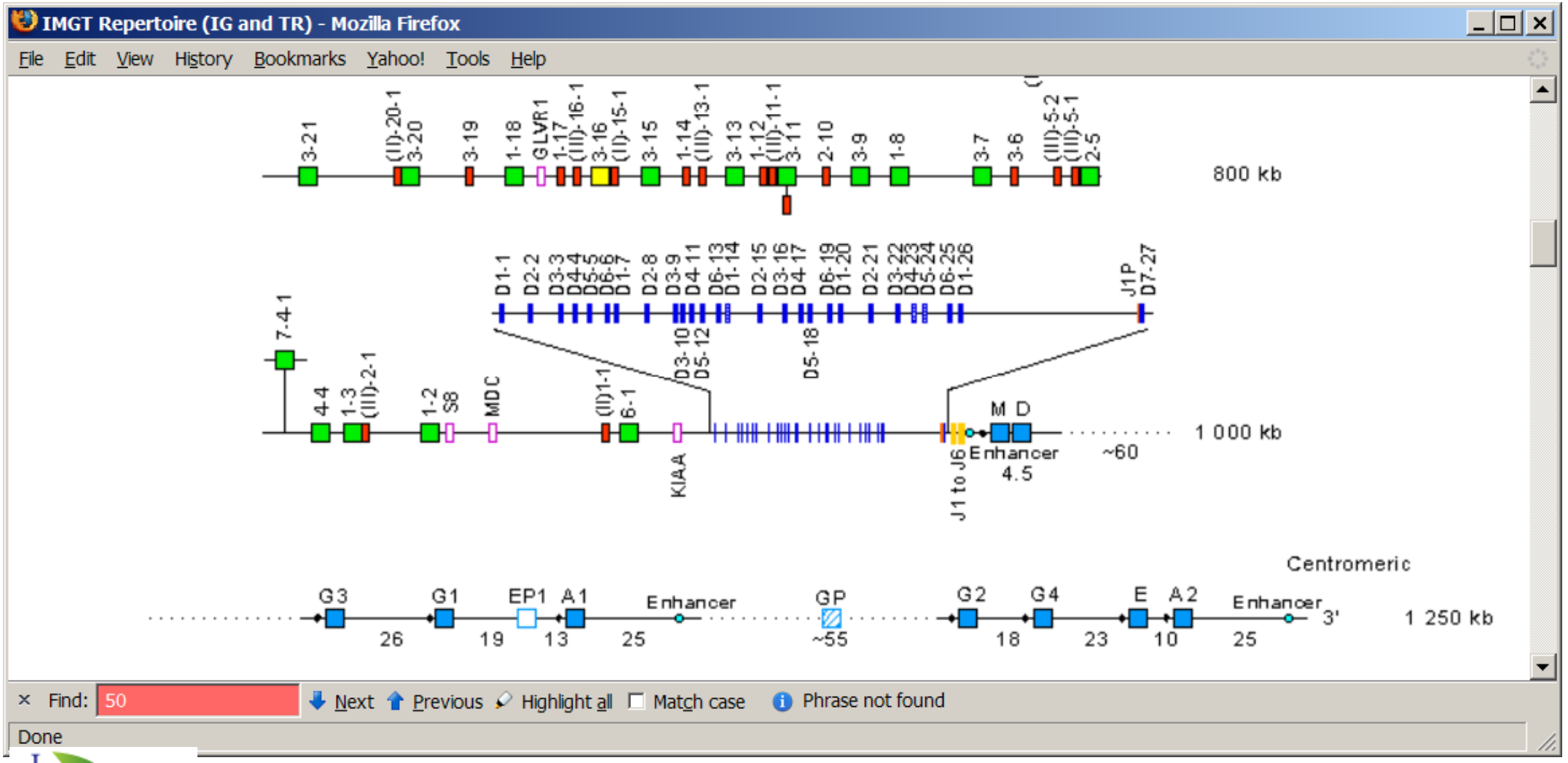
V-J-REGION

V-DJ-REGION

T cell receptor

The Immunoglobulin FactsBook, 2001

Part of Human *(Homo sapiens)* IGH locus on chromosome 14 (14q32.33)



Im Muno Gene Tics Information system®

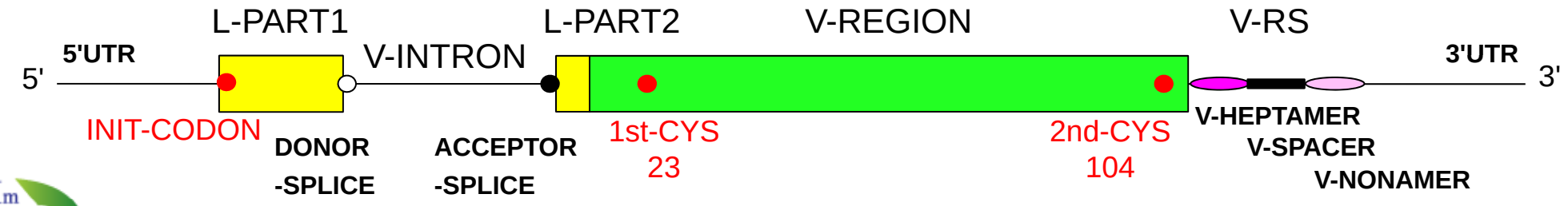
http://www.imgt.org

Séquences d'ADN génomique en configuration germline

1 - Exemple d'un V-GENE

>X62106.0|HSVI2|*Homo sapiens* VI-2 gene for immunoglobulin heavy chain

```
tgagagctcc gttcctcacc atggactgga cctggaggat cctcttcttg gtggcagcag      60
ccacaggtaa gaggctccct agtcccagtg atgagaaaga gattgagtcc agtccaggga     120
gatctcatcc acttctgtgt tctctccaca ggagcccact cccaggtgca gctggtgcag     180
tctggggctg aggtgaagaa gcctggggcc tcagtgaagg tctcctgcaa ggcttctgga     240
tacaccttca ccggctacta tatgcactgg gtgcgacagg cccctggaca agggcttgag     300
tggatgggat ggatcaaccc taacagtggt ggcacaaact atgcacagaa gtttcagggc     360
agggtcacca tgaccaggga cacgtccatc agcacagcct acatggagct gagcaggctg     420
agatctgacg acacggccgt gtattactgt gcgagagaca cagtgtgaaa acccacatcc     480
tgagggtgtc agaaacccaa gggaggaggc ag
```

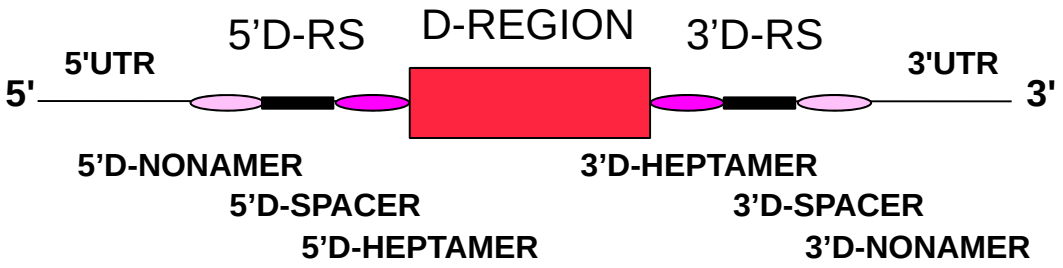


http://www.imgt.org

Séquences d'ADN génomique en configuration germline

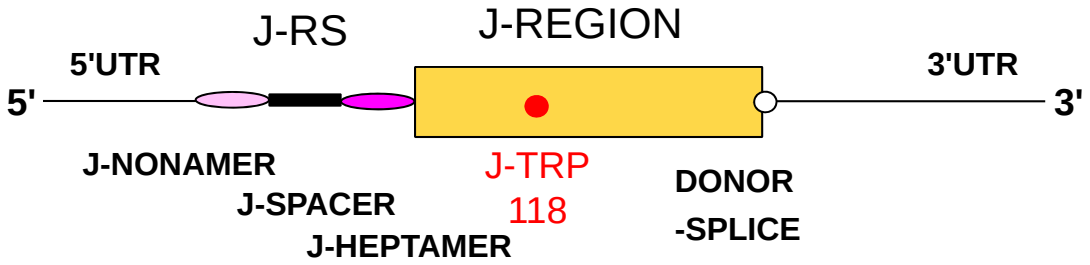
2 - Exemple d'un D-GENE

```
>J00256|IGHD7-27*01|Homo sapiens D-GENE
ccagccgcag ggtttttggc tgagctgaga accactgtgc taactgggga cacagtgatt      60
ggcagctcta caaaaaccat gctcccccgg g
```



3 - Exemple d'un J-GENE

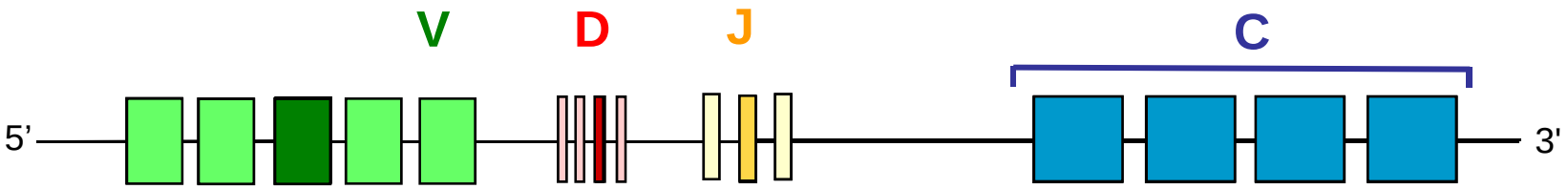
```
>J00256|IGHJ1*01|Homo sapiens J-GENE
accccgggct gtgggtttct gtgcccctgg ctcagggctg actcaccgtg gctgaatact      60
tccagcactg gggccagggc accctggtca ccgtctcctc aggtgagtct gctgtactgg     120
ggatagcggg gagccatgtg tactgggcca agcaagggct ttggcttcag                170
```



Im Muno Gene Tics Information system®
http://www.imgt.org

Séquences d'ADN et ADNc en configuration réarrangée

ADN génomique

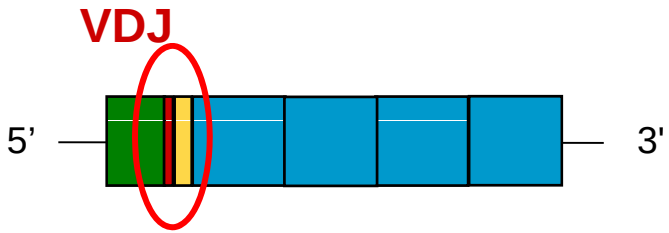


ADN réarrangé

5'

3'

ADNc/ARNm



Im Muno Gene Tics Information system®
http://www.imgt.org

Generation of V-D-J junction diversity: example of an IG heavy chain

http://www.imgt.org

3'V-REGION N1-REGION **D-REGION** N2-REGION **5'J-REGION**

tgtgcgaaa ga tacc agcatattgtg gtggtgactgctat tcc gatt acaactggttcg acccc**tgg**

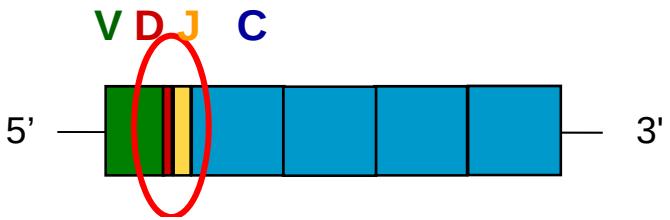
2nd-CYS

JUNCTION

J-TRP

C A P Y R G D T Y D Y S W

tgt gcg cca tac cgg ggt gac act tat gat tac tcc tgg

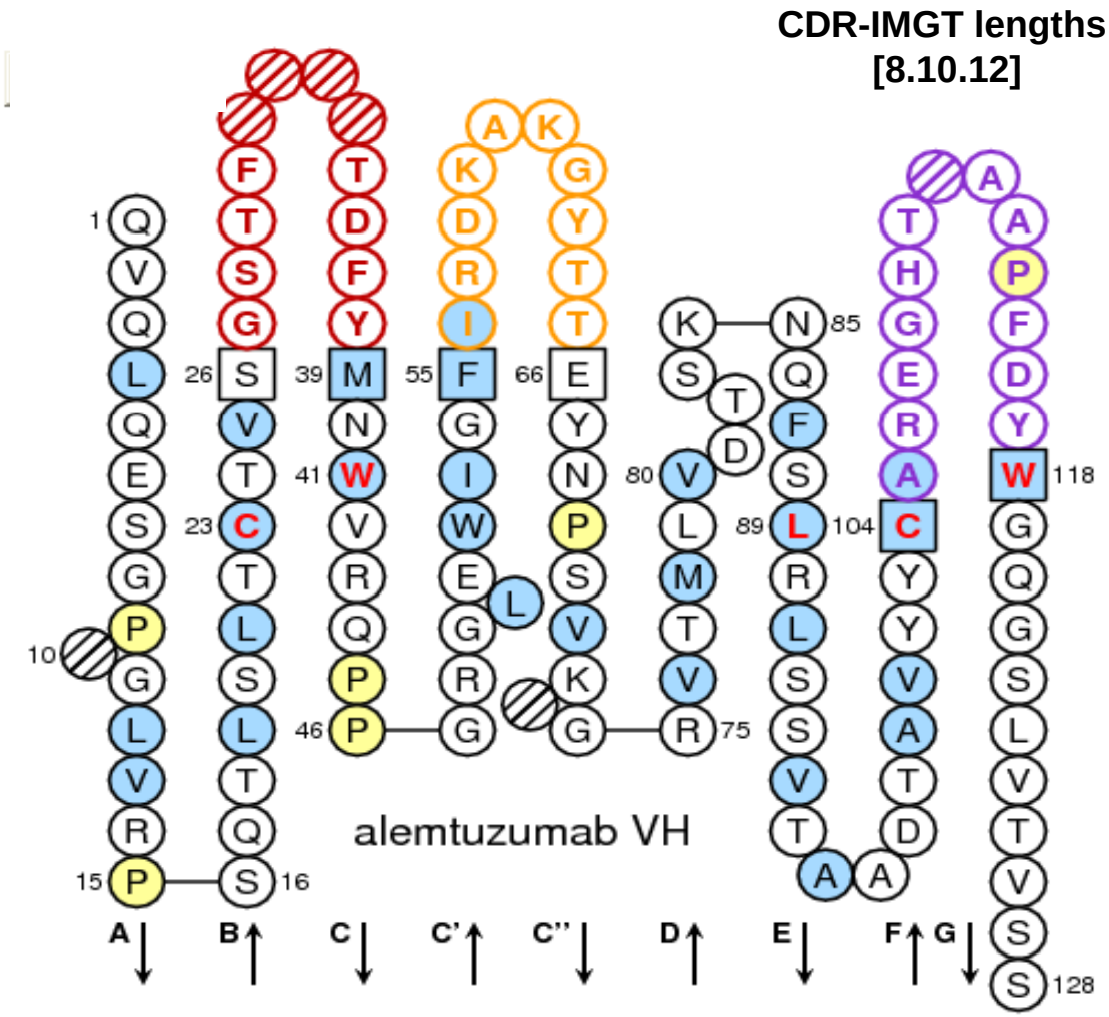


Représentation d'un domaine variable en «Collier de Perles»

http://www.imgt.org

- Conserved AA (and codons) are always at the **same positions**:
  - **23** 1st-CYS
  - **41** CONSERVED-TRP
  - **89** hydrophobic
  - **104** 2nd-CYS
  - **118** J-PHE, J-TRP

- Delimitation of the **FR-IMGT** and **CDR-IMGT** is standardized



Lefranc et al. Dev. Comp. Immunol. 27, 55-77 (2003)

Particularités des séquences IG et TR

http://www.imgt.org

- Elles sont nombreuses (2.10^{12} IG par individu)
- Une forte similitude
- De petites différences très significatives
- Nécessité de standardiser la nomenclature des gènes

IMGT, the international ImMunoGeneTics information system® http://www.imgt.org

Bases de données

Bases de données de séquences

IMGT/LIGM-DB
+ de 157.000 séquences nucléotidiques IG et TR de 280 espèces de vertébrés (sur le Web depuis 1995)

- **IMGT/PRIMER-DB**
- **IMGT/MHC-DB**

Base de données de gènes

- **IMGT/GENE-DB**

Base de données de structures 3D

- **IMGT/3Dstructure-DB**

Outils en ligne

Analyse de séquences

- **IMGT/V-QUEST**
- **IMGT/JunctionAnalysis**
- **IMGT/PhyloGene**
- **IMGT/Allele-Align**

Analyse de génomes

- **IMGT/GeneSearch**
- **IMGT/GeneView**
- **IMGT/LocusView**

Analyse de structures 3D

- **IMGT/StructuralQuery**

Ressources Web

Séquences

IMGT Repertoire:
- **Alignments of alleles**
- **Tables of alleles**
- **Protein displays, etc.**

Génomes

- **Chromosomal localizations**
- **Locus representations**
- **Gene tables, etc.**

Structures 2D et 3D

- **2D Colliers de Perles**
- **3D representations, etc.**

Page d'accueil IMGT (http://www.imgt.org)

http://www.imgt.org

Dans le domaine de l'immunogénétique

Problèmes et limites de GO:

- Comment se compose le site de reconnaissance d'un anticorps?

- Quels sont les motifs constitutifs importants?

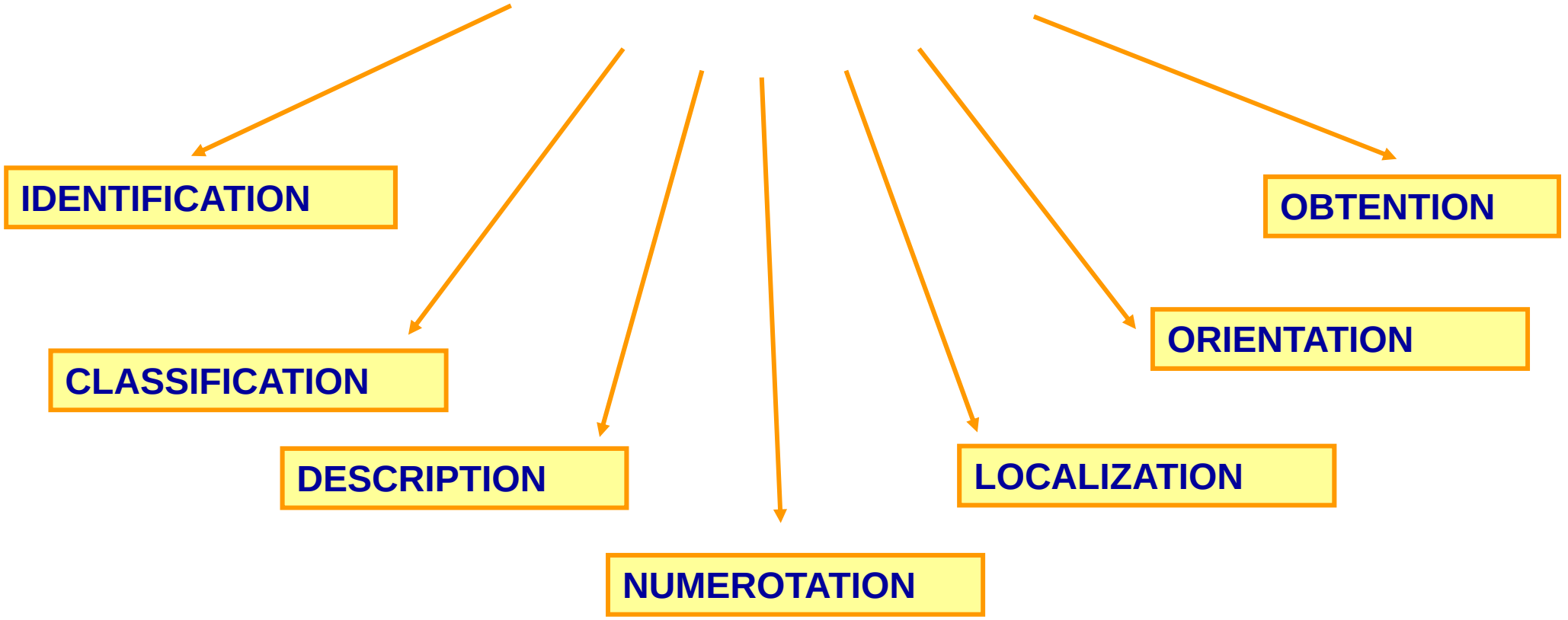
- Comment caractériser précisément des séquences d'IG (.. et des TR): identification, classification des gènes, description, numérotation des acides aminés, obtention ?

IMGT-ONTOLOGY

http://www.imgt.org

Pour standardiser, partager, réutiliser et représenter les connaissances en immunogénétique

IMGT-ONTOLOGY repose sur 7 axiomes:

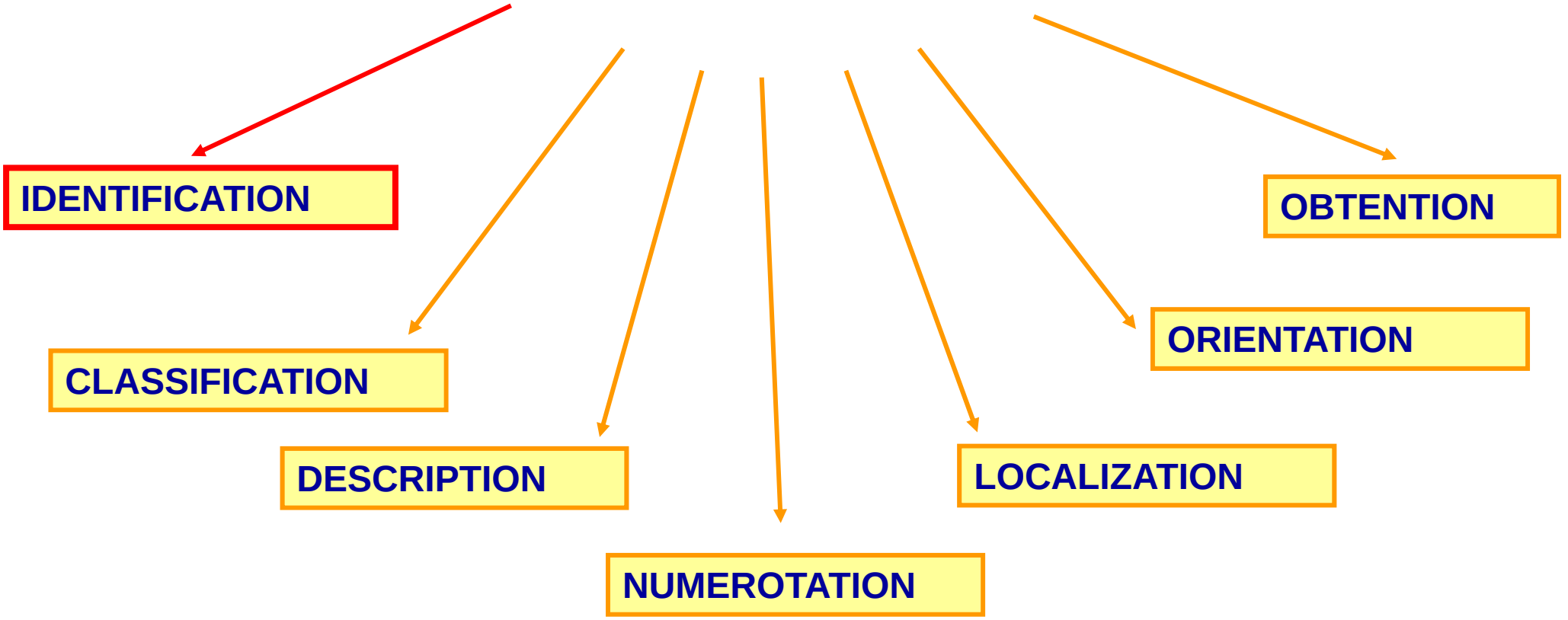


IMGT-ONTOLOGY

http://www.imgt.org

Pour standardiser, partager, réutiliser et représenter les connaissances en immunogénétique

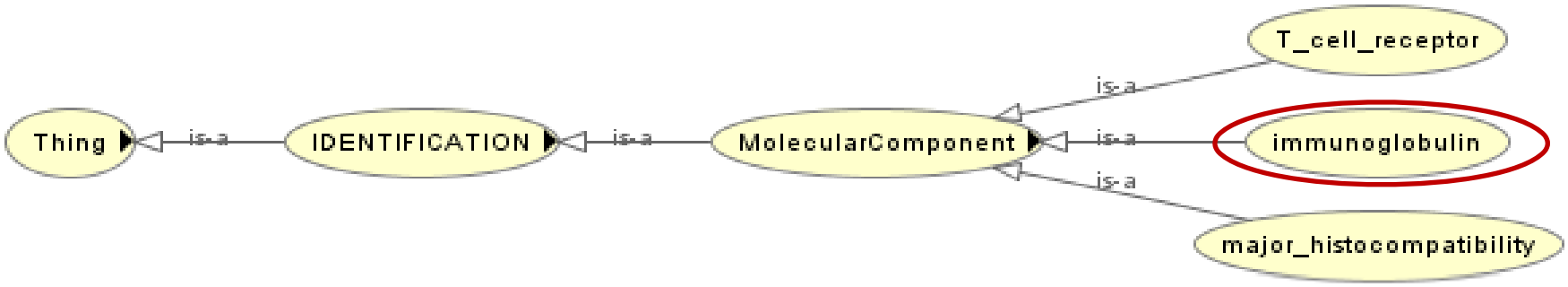
IMGT-ONTOLOGY repose sur 7 axiomes:



1 – Concepts d'identification d'IMGT-ONTOLOGY

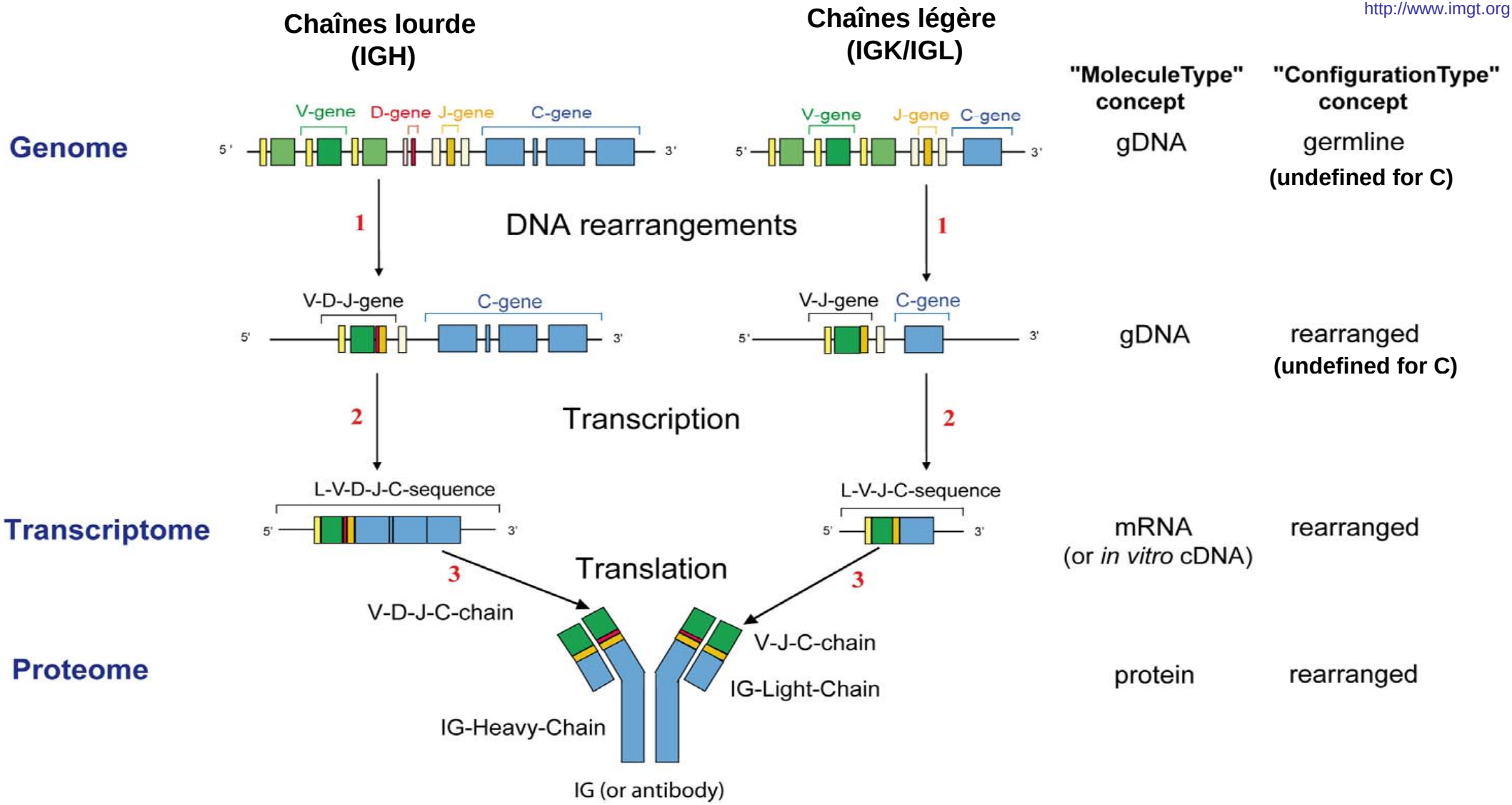
Les composants moléculaires (MolecularComponent) de la réponse immunitaire adaptative:

- **Les immunoglobulines**
- **Les récepteurs des cellules T**
- **Les molécules du complexe majeur d'histocompatibilité**



Synthèse des immunoglobulines

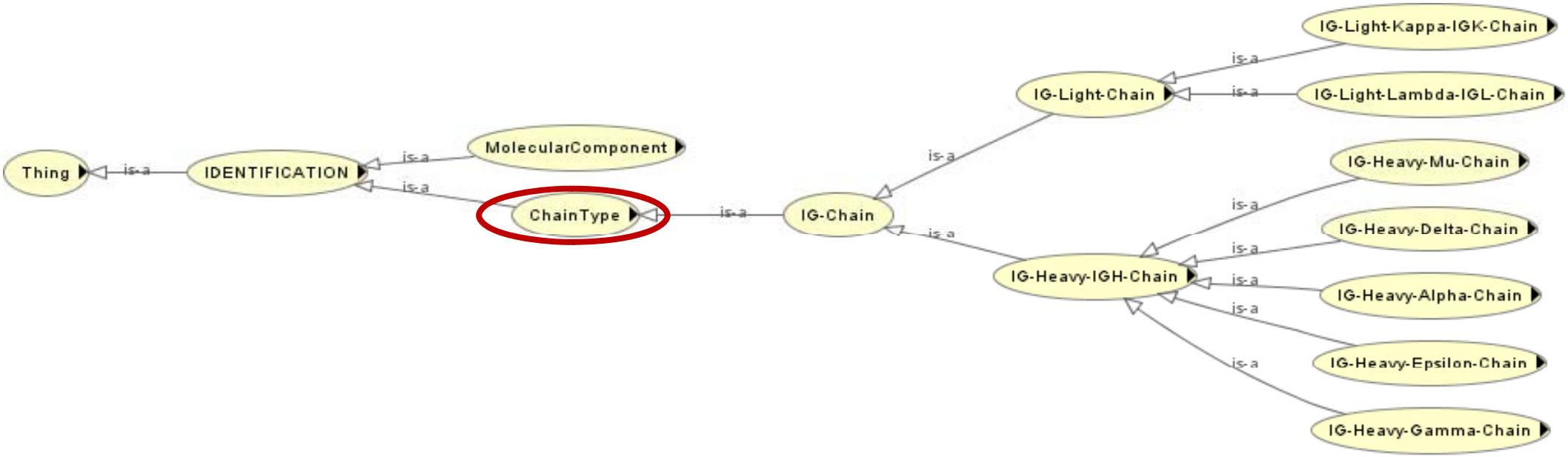
http://www.imgt.org



1 – Concepts d'identification d'IMGT-ONTOLOGY

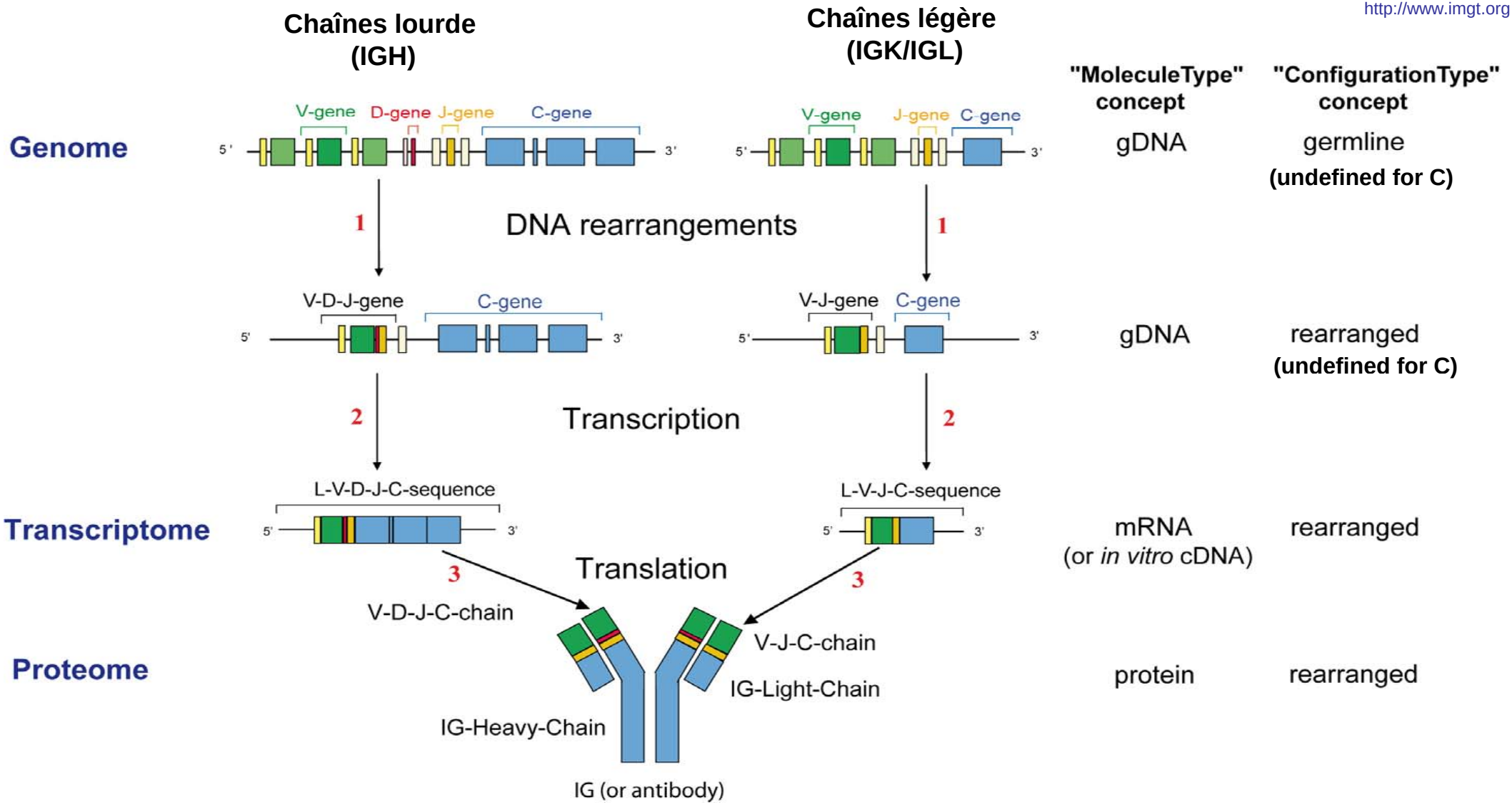
Les types de chaîne d'une immunoglobuline (ChainType) :

- Les chaînes lourdes (Alpha, Delta, Epsilon, Gamma, Mu)
- Les chaînes légères (Kappa, lambda)



Synthèse des immunoglobulines

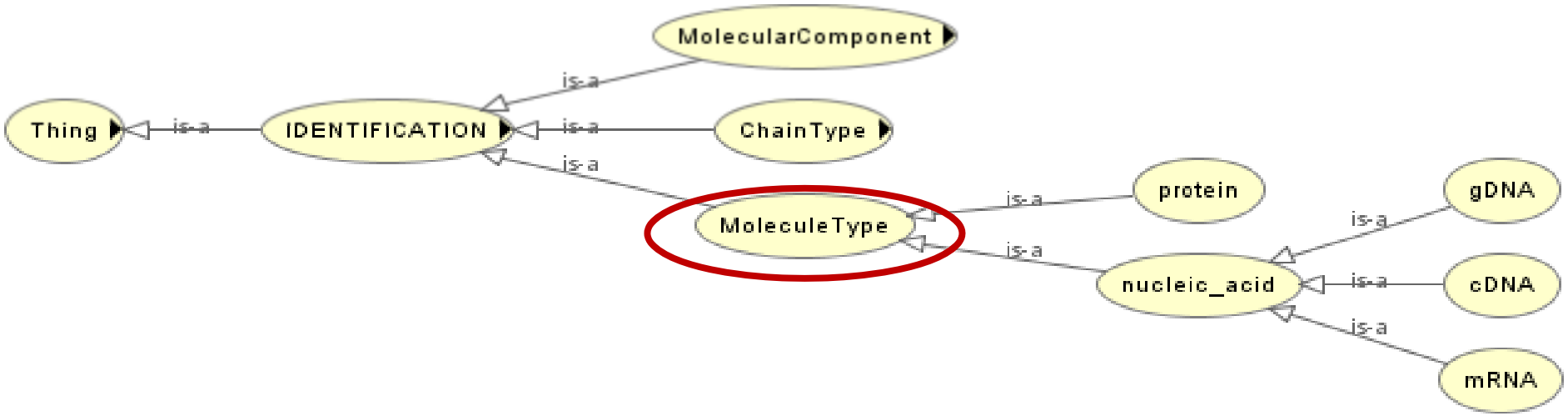
http://www.imgt.org



1 – Concepts d'identification d'IMGT-ONTOLOGY

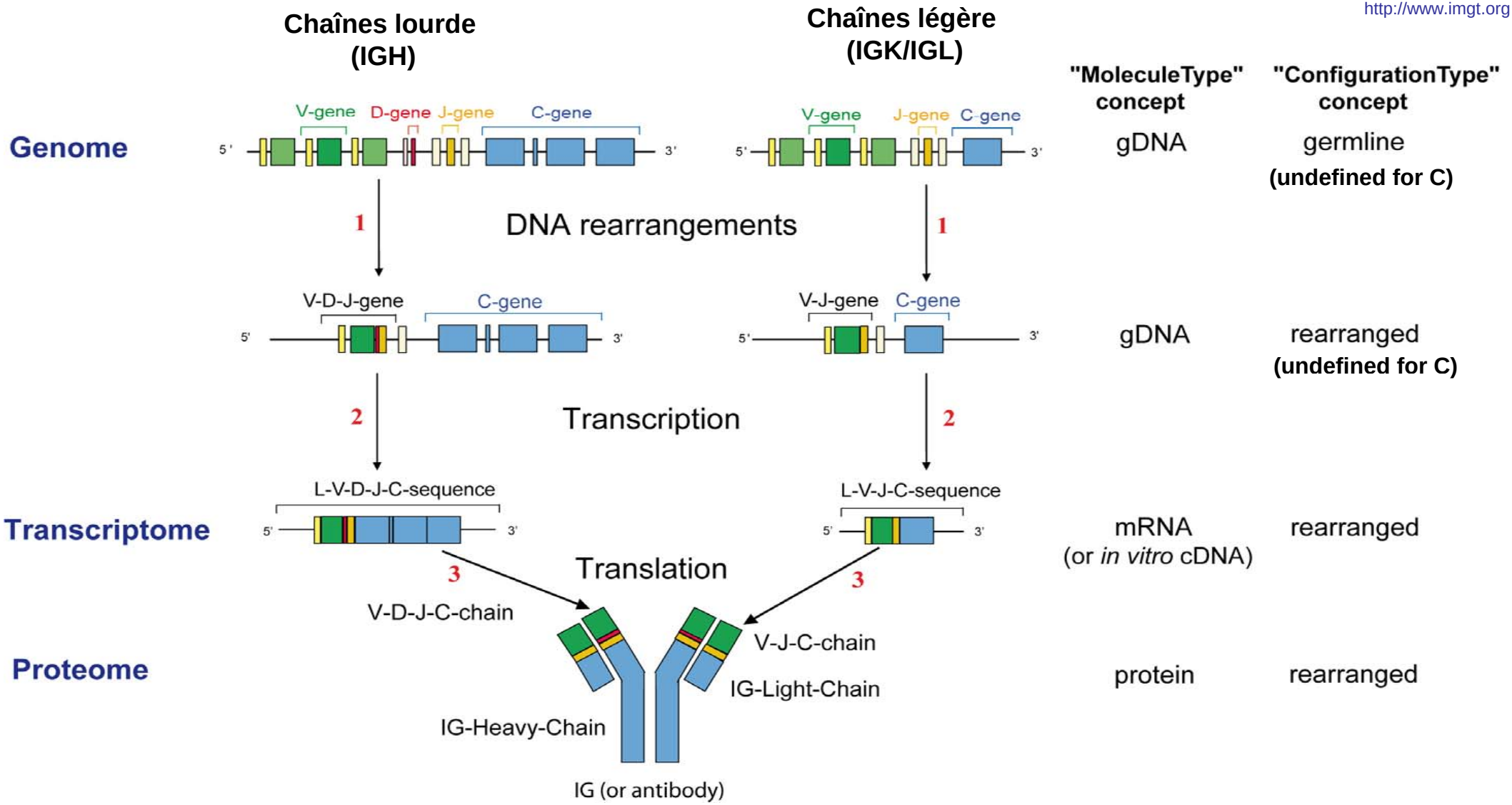
Les types de molécules (MoleculeType) :

- gDNA
- mRNA
- cDNA
- protein



Synthèse des immunoglobulines

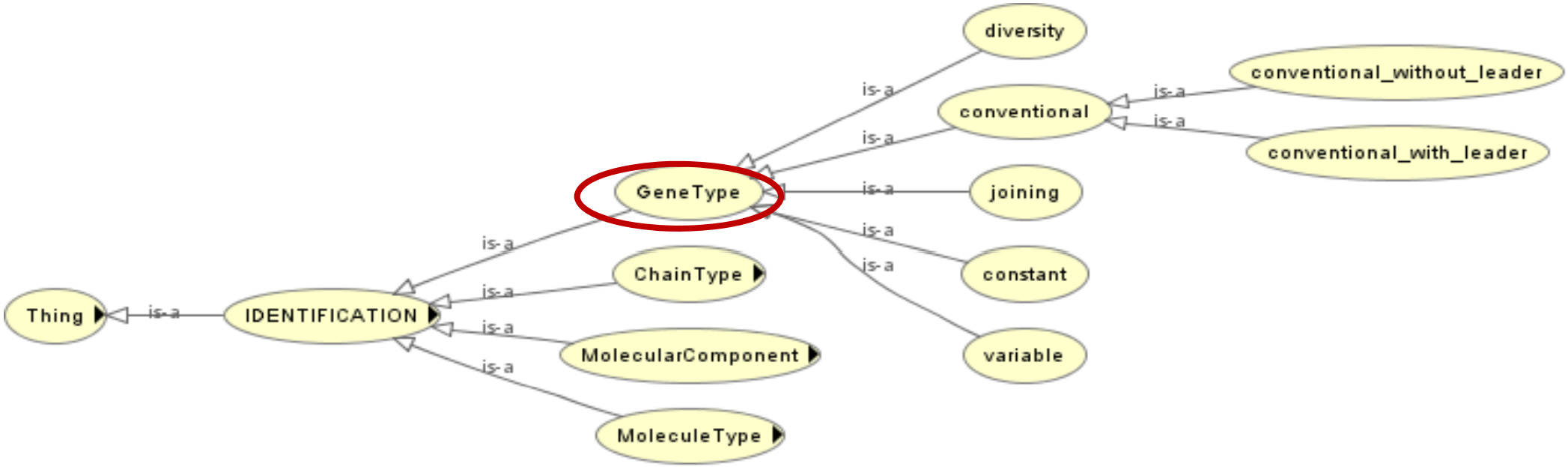
http://www.imgt.org



1 – Concepts d'identification d'IMGT-ONTOLOGY

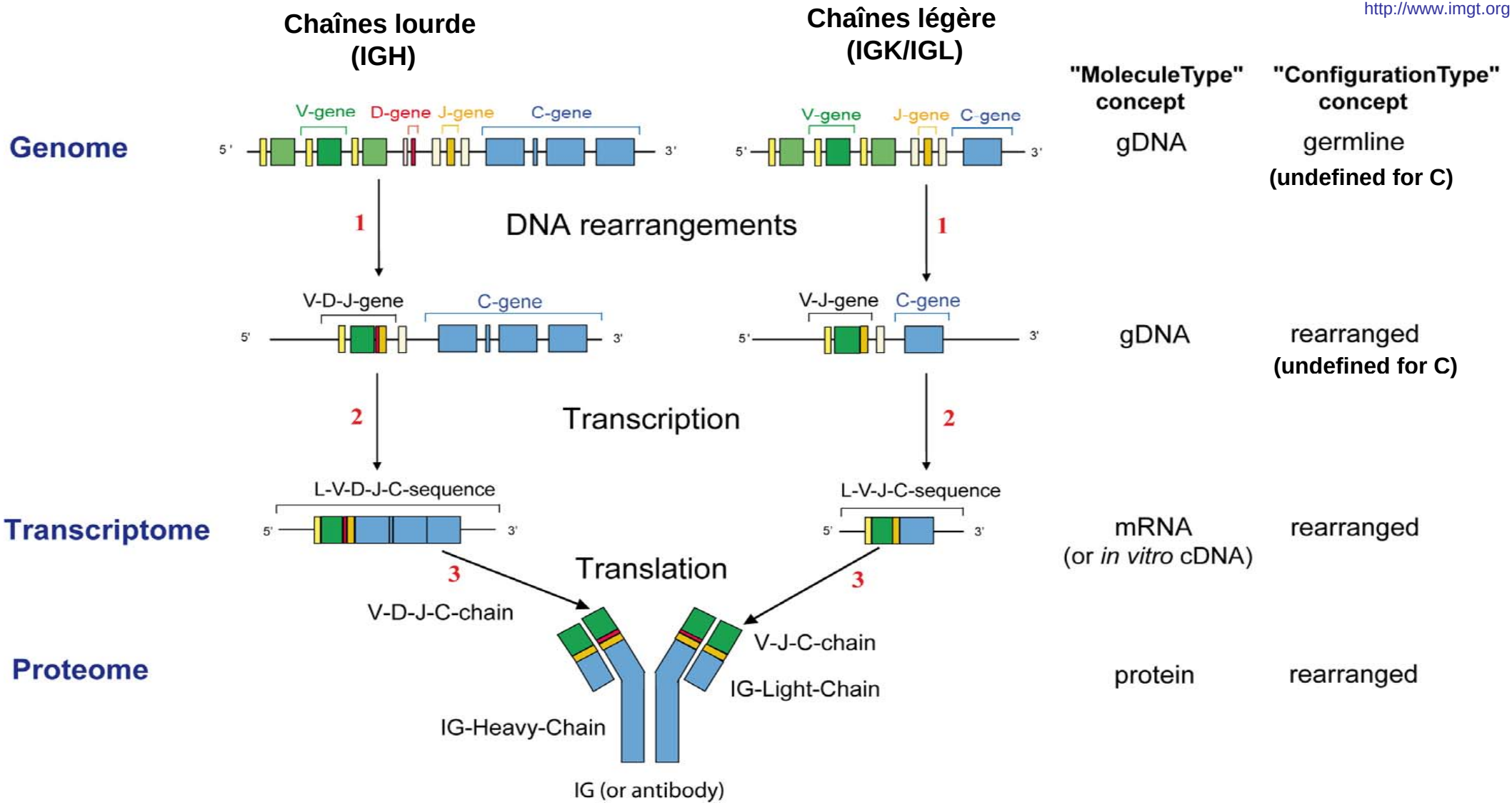
Les types de gène (GeneType) :

- variable
- diversity
- joining
- constant
- conventional (with or without leader)



Synthèse des immunoglobulines

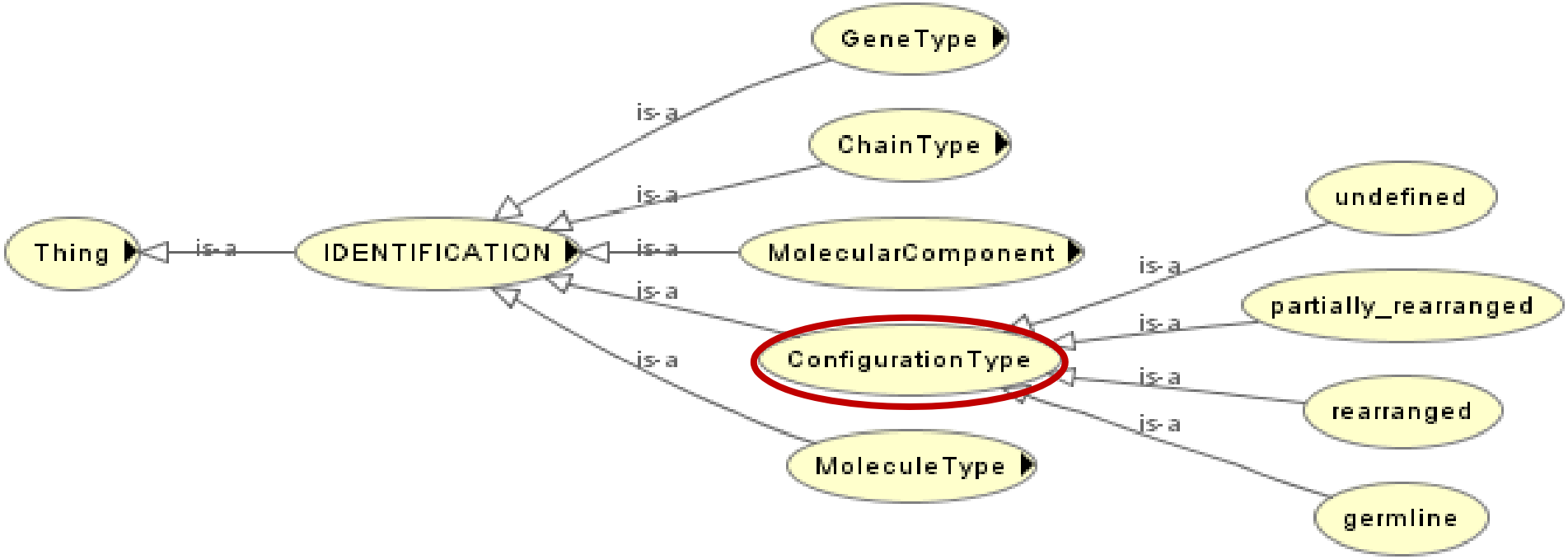
http://www.imgt.org



1 – Concepts d'identification d'IMGT-ONTOLOGY

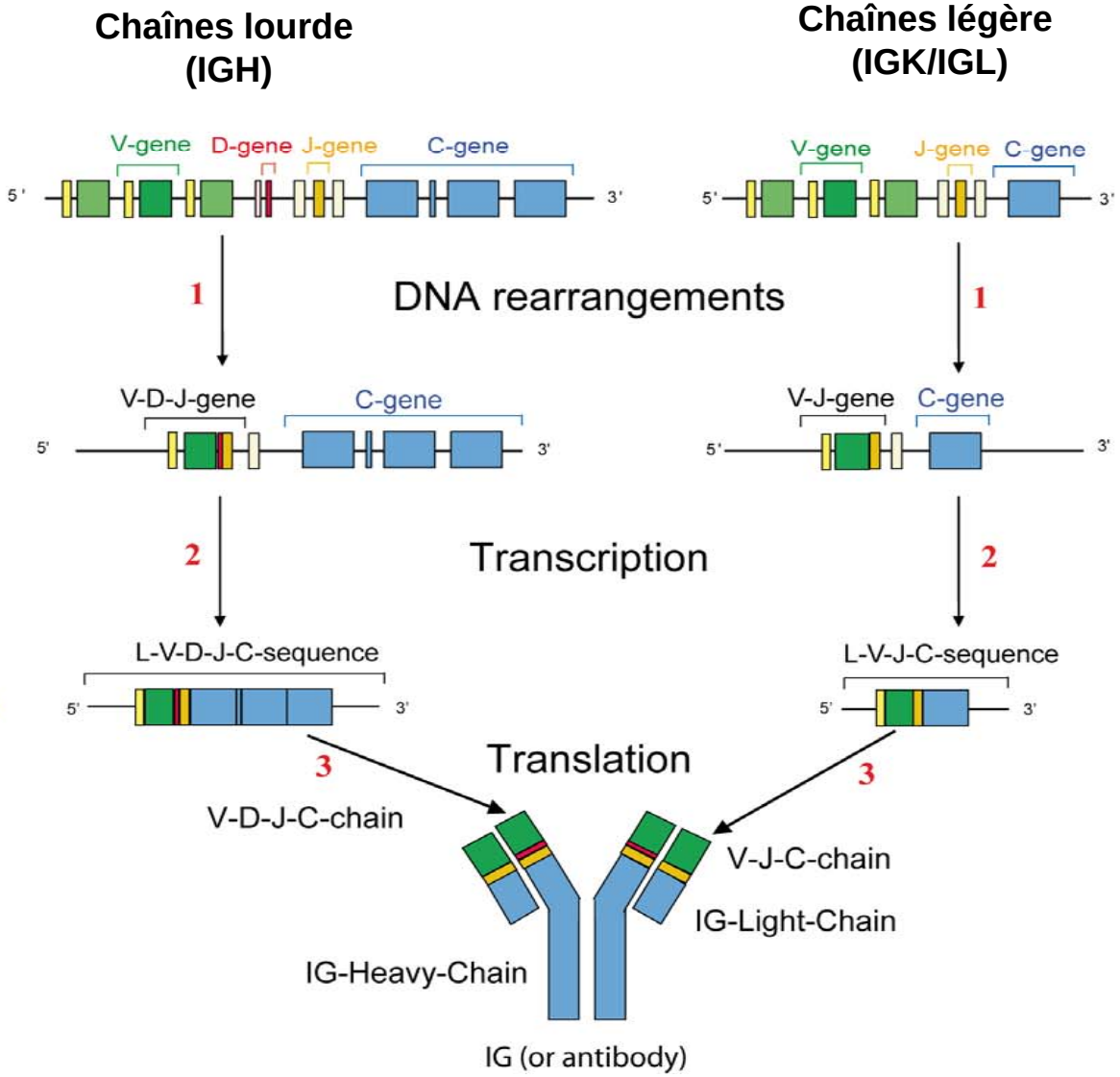
Les types de configuration (ConfigurationType) :

- germline
- rearranged
- partially-rearranged
- undefined



Synthèse des immunoglobulines

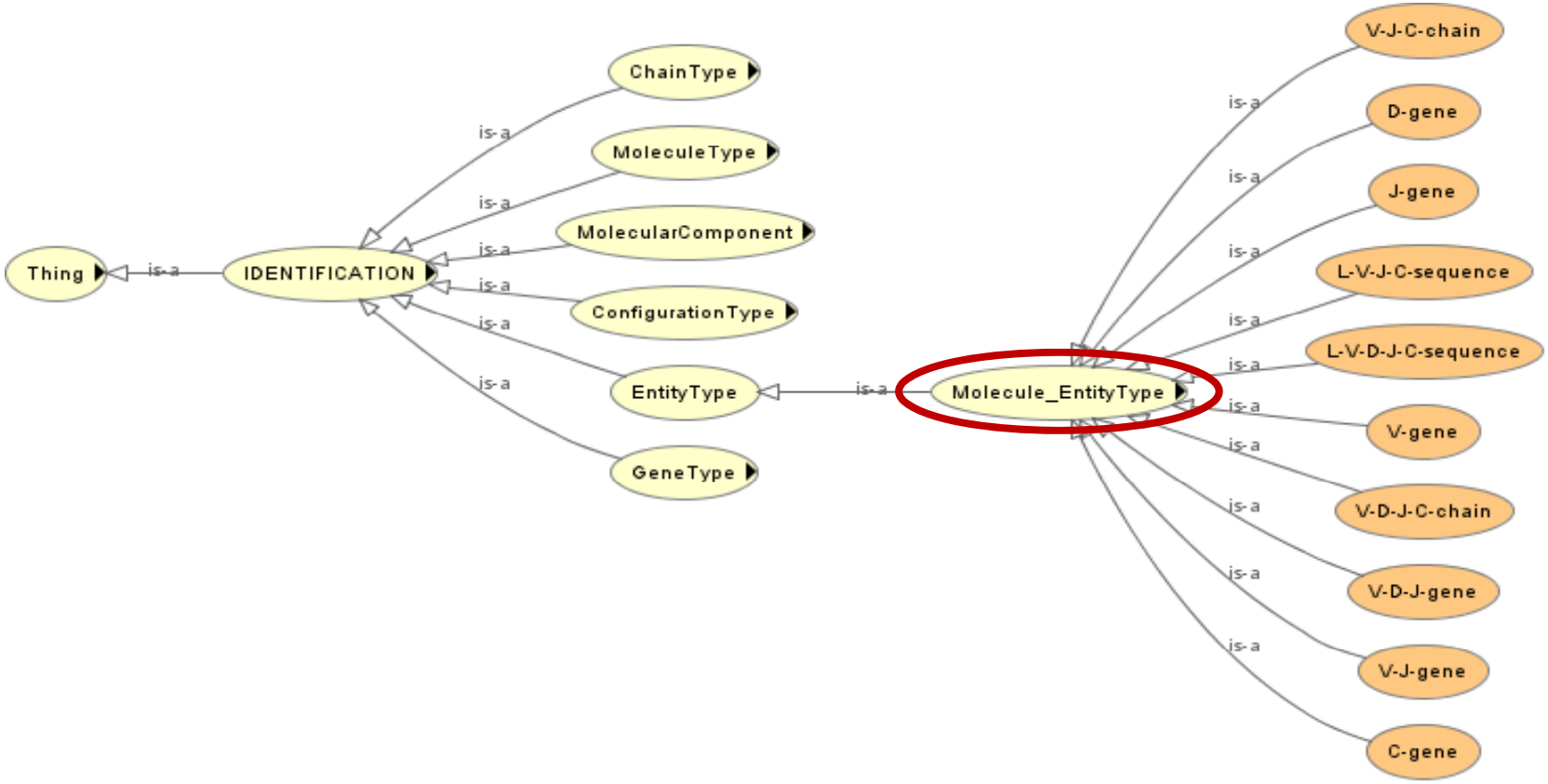
http://www.imgt.org



1 – Concepts d'identification d'IMGT-ONTOLOGY

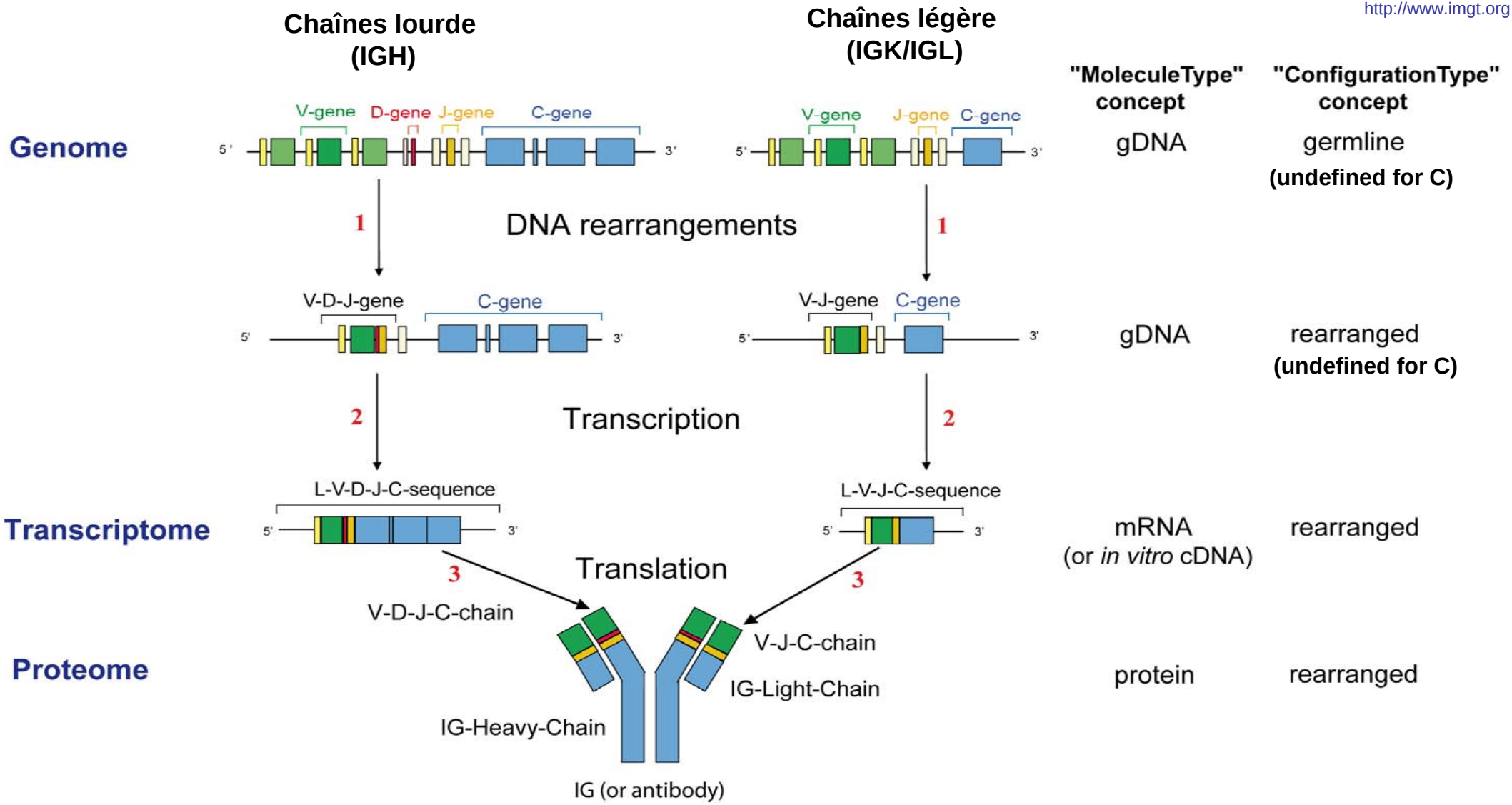
Les types d'Entités moléculaire (Molecule_EntityType) :

- V-gene, D-gene, J-gene, C-gene
- V-D-J-gene, V-J-gene
- L-V-J-C-sequence, L-V-D-J-C-sequence
- V-D-J-chain, V-J-chain



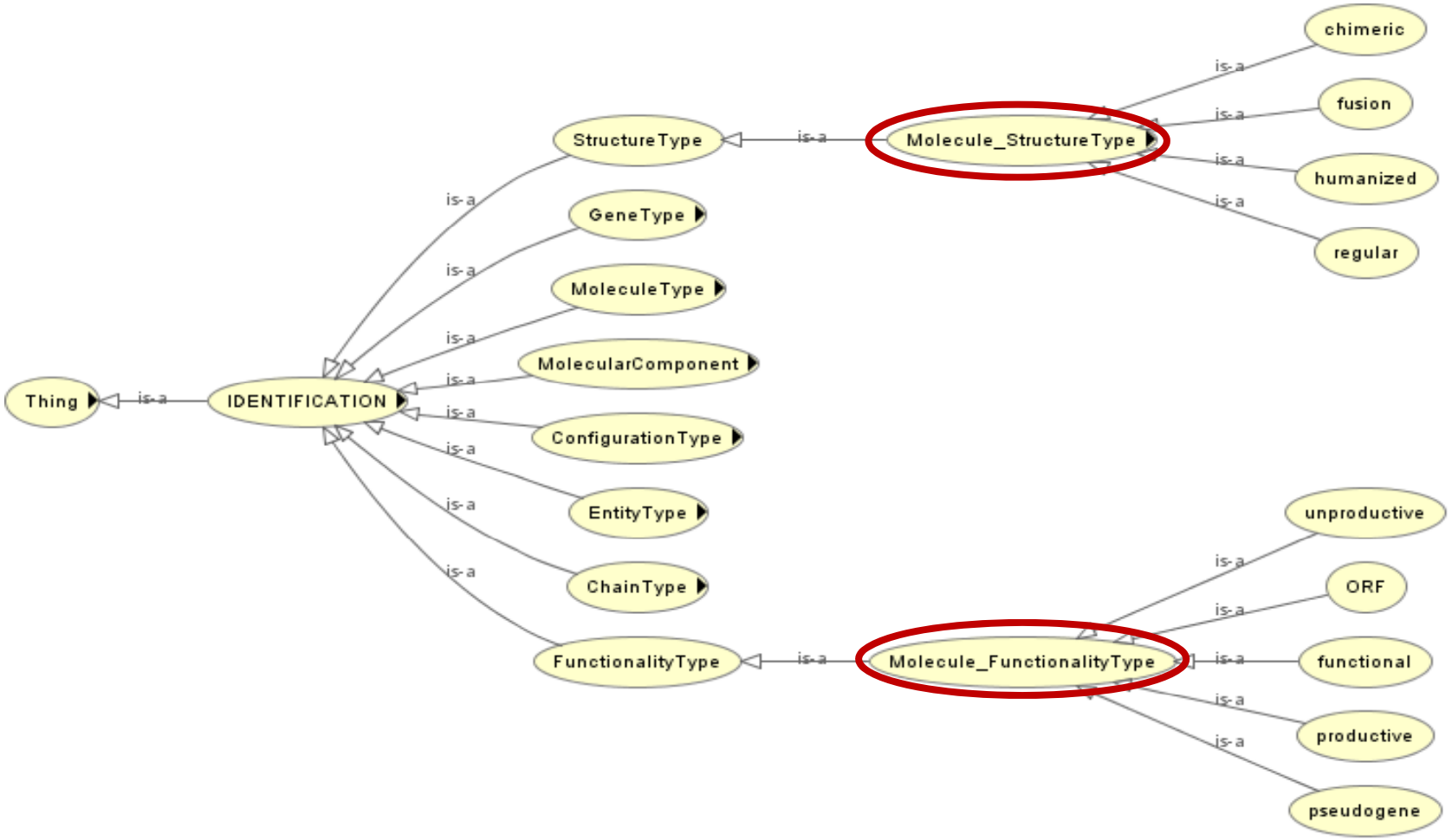
Synthèse des immunoglobulines

http://www.imgt.org

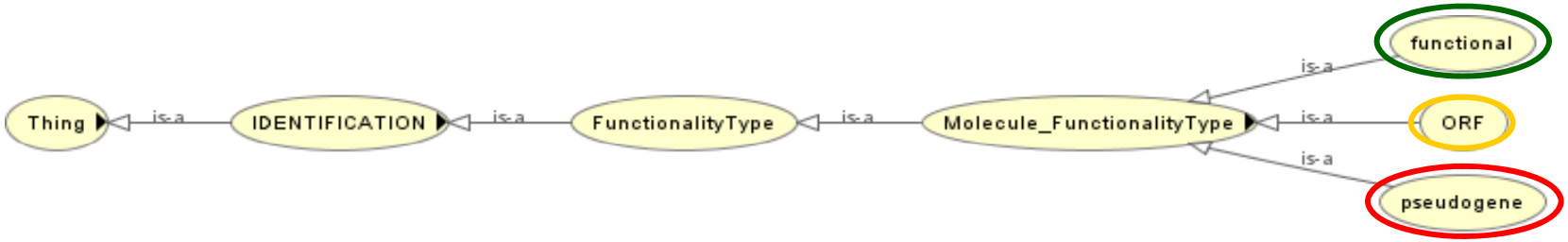


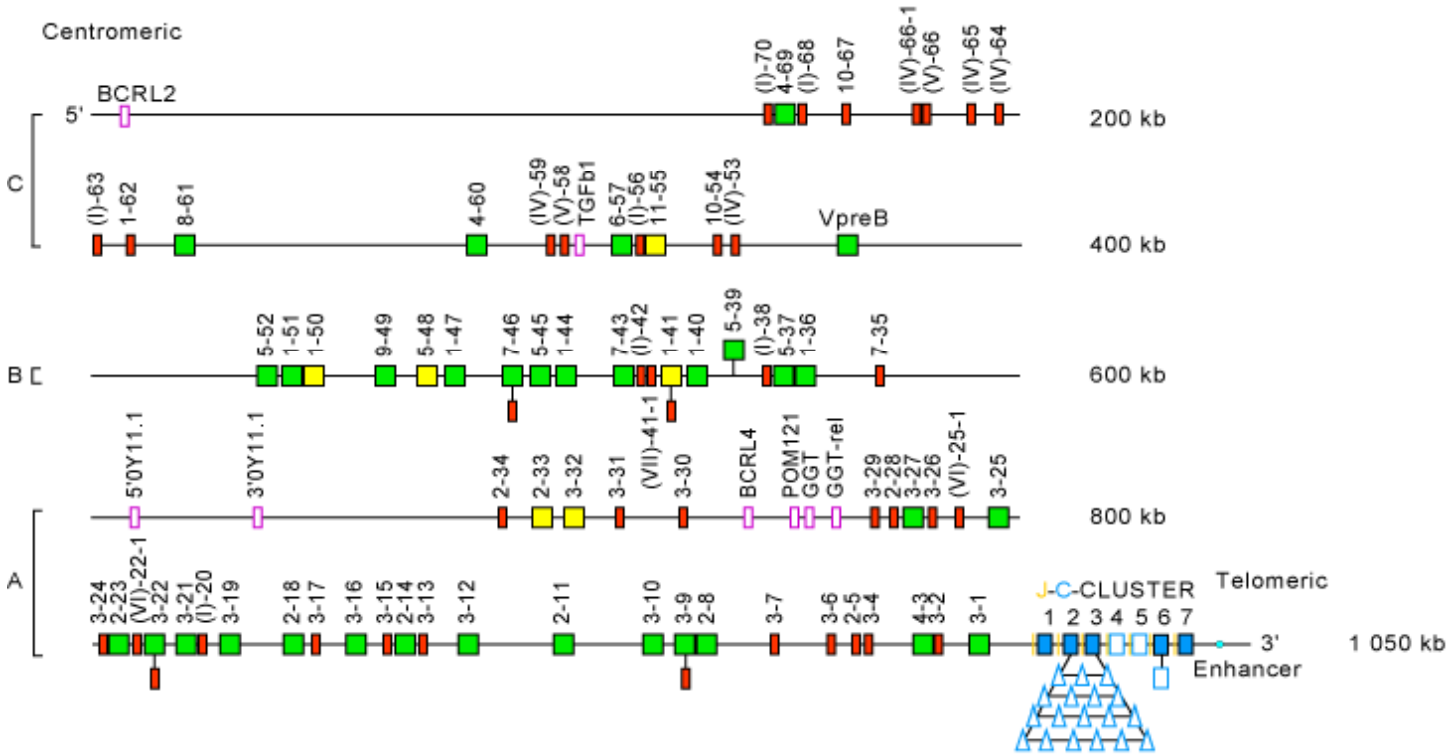
1 – Concepts d'identification d'IMGT-ONTOLOGY

Les types de structure (Molecule_StructureType) :
Les types de fonctionnalité (Molecule_FonctionnalityType) :



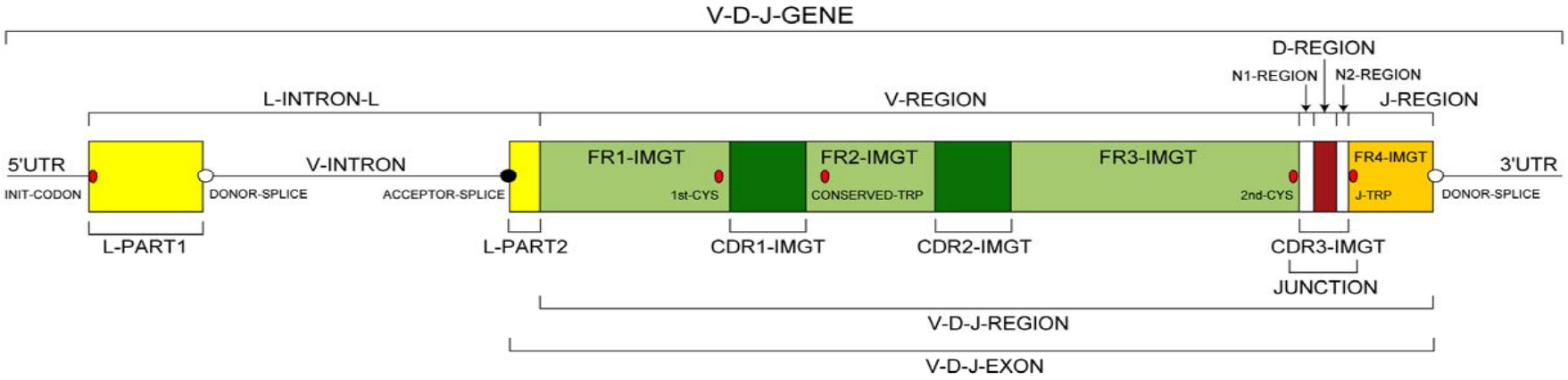
Les types de fonctionnalité pour les entités moléculaires en configuration germline ou undefined





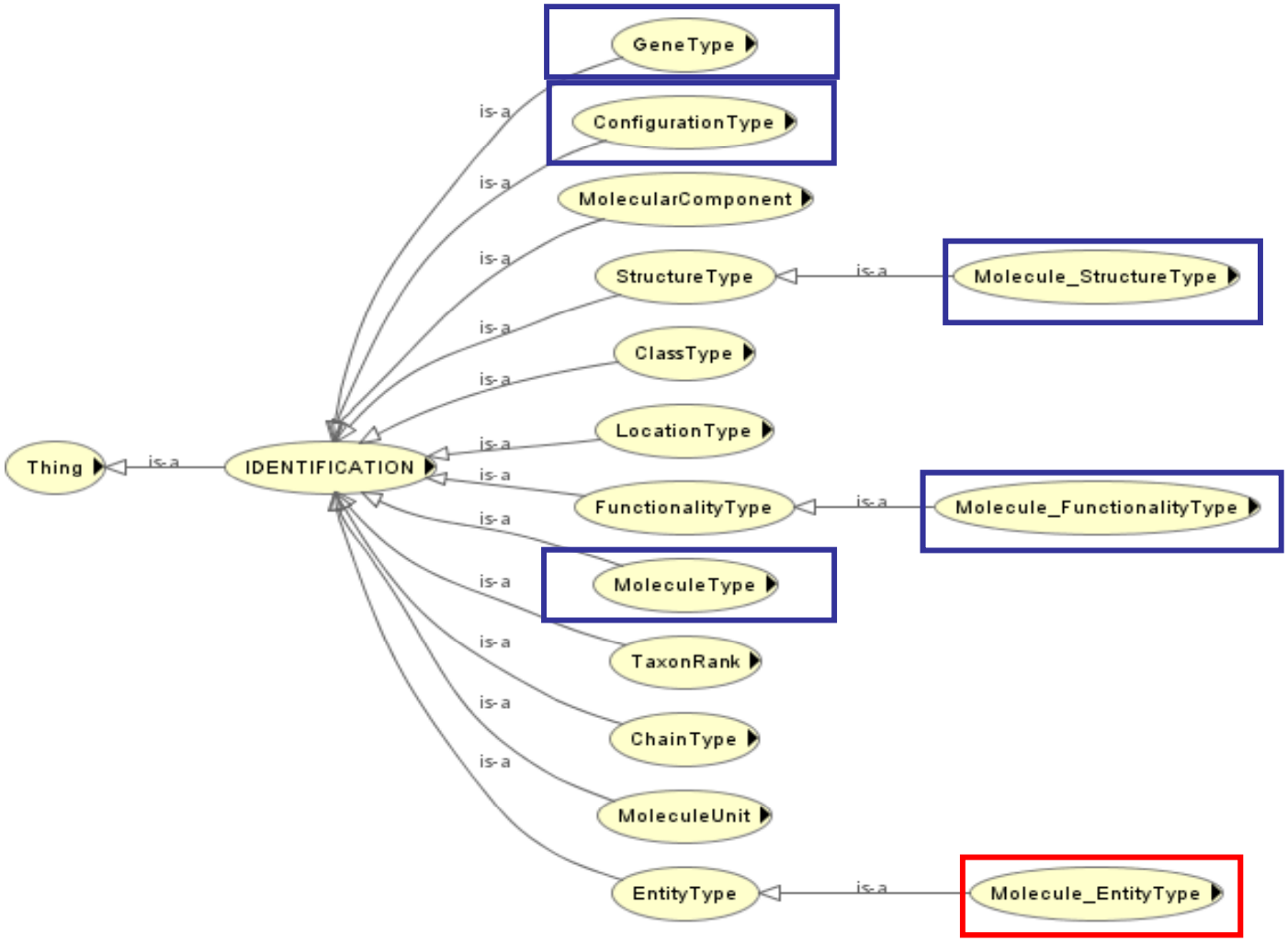
Human IGL locus (22q11.2)

Les types de fonctionnalité pour les entités moléculaires en configuration réarrangées



Concepts d'Indentification

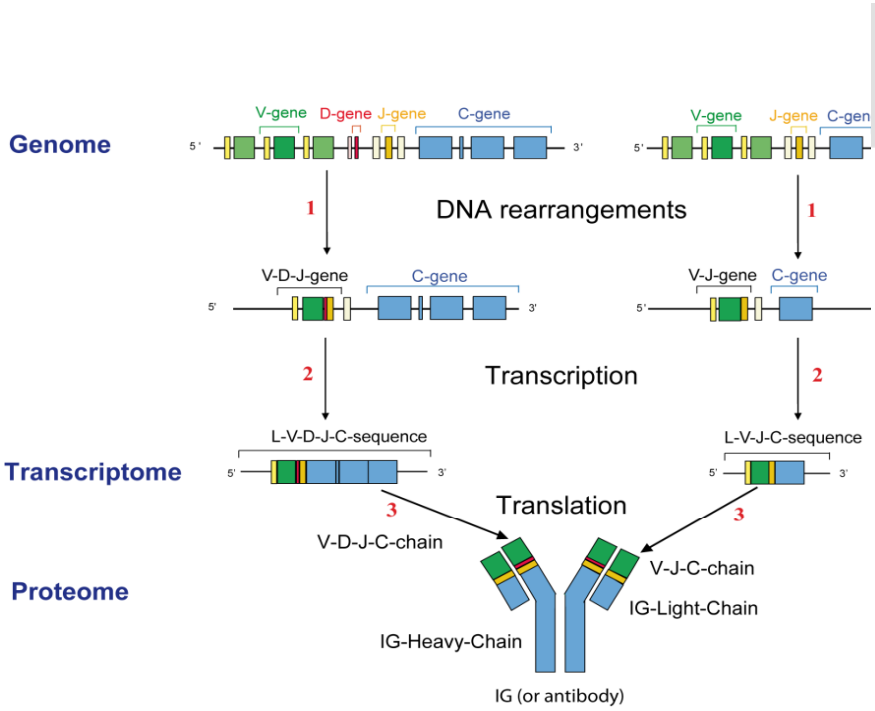
http://www.imgt.org

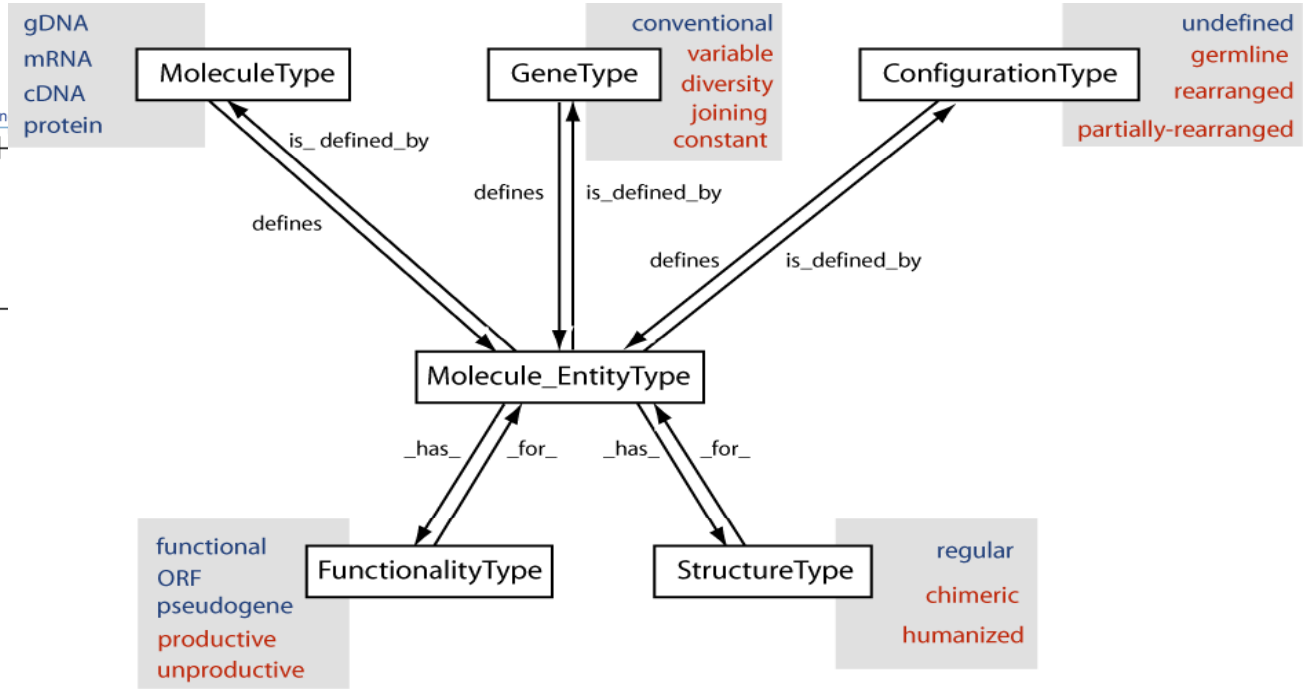


Quelles sont relations/contraintes sémantiques qui relient les concepts entre eux?

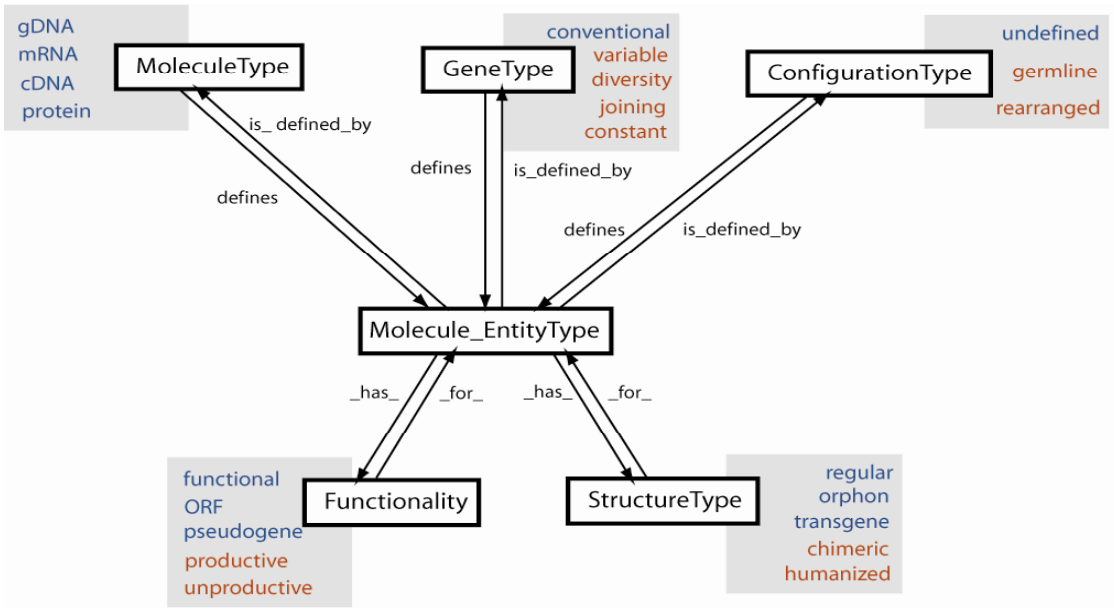
Relationships between Molecule_EntityType and the other IDENTIFICATION concepts

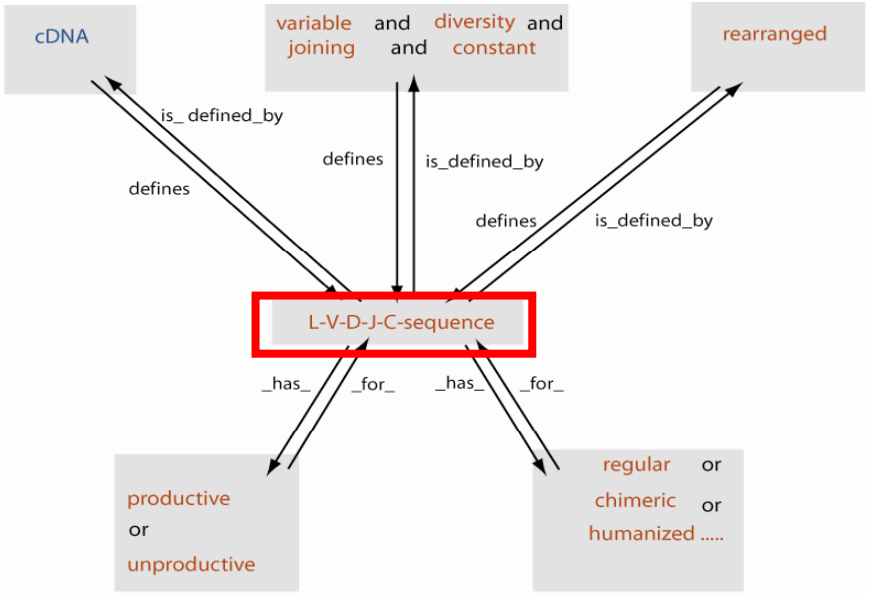
http://www.imgt.org

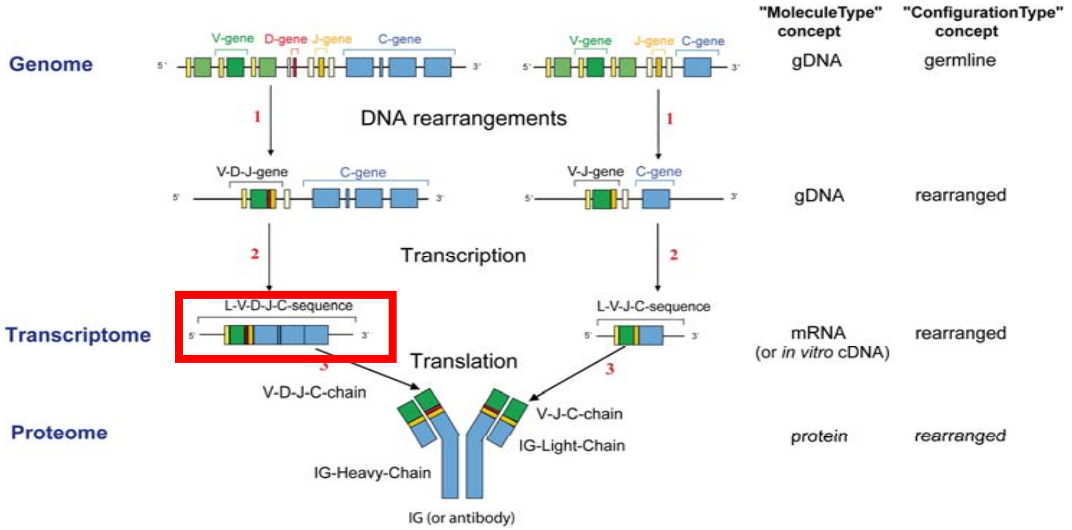


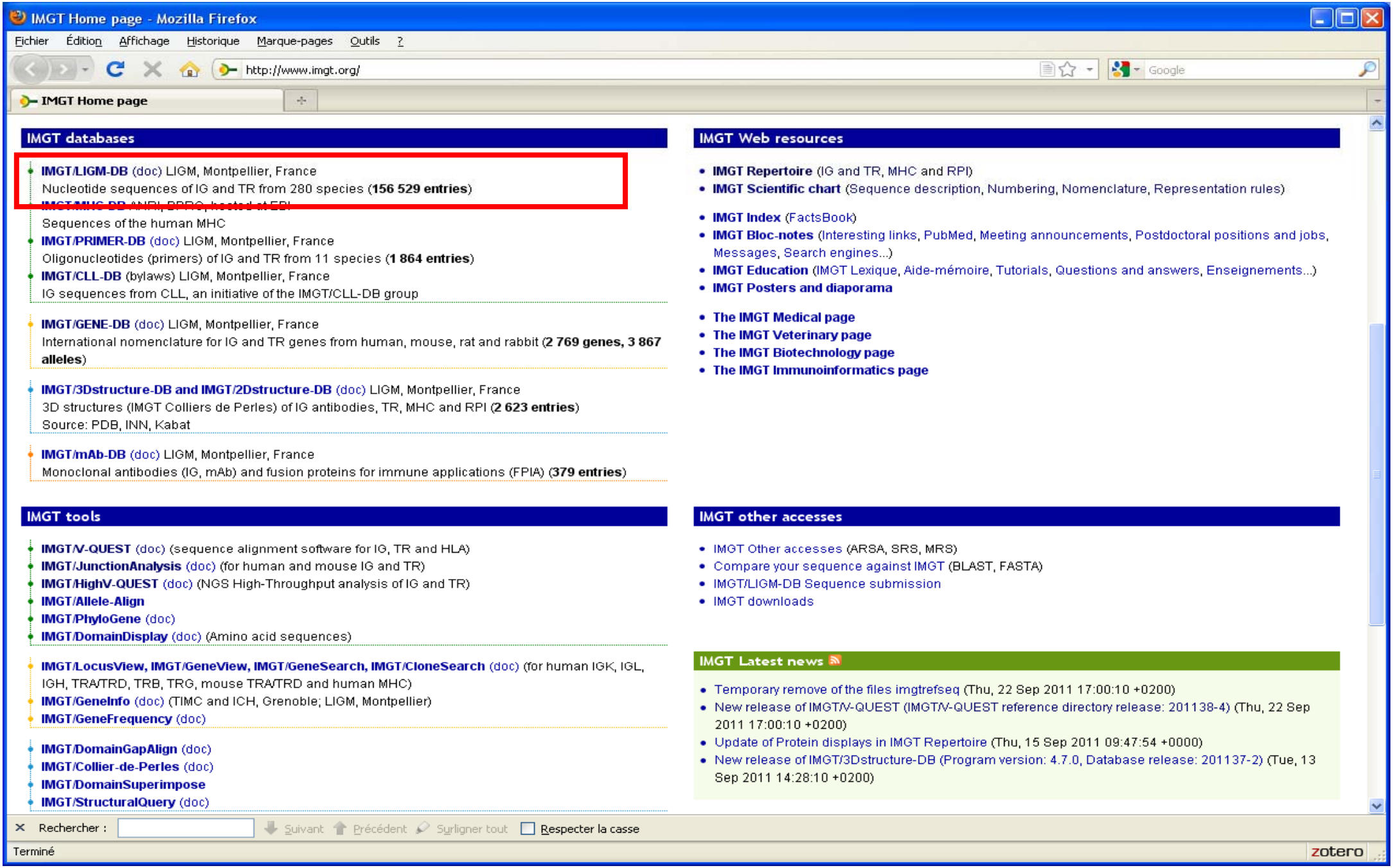


Relationships between the L-V-D-J-C-sequence concepts and the other IDENTIFICATION concepts








IMGT Home page - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
http://www.imgt.org/
Google
IMGT Home page
IMGT databases
IMGT/LIGM-DB (doc) LIGM, Montpellier, France
Nucleotide sequences of IG and TR from 280 species (156 529 entries)
IMGT/MHC-DB ANRI, PRBC, hosted at EBI
Sequences of the human MHC
IMGT/PRIMER-DB (doc) LIGM, Montpellier, France
Oligonucleotides (primers) of IG and TR from 11 species (1 864 entries)
IMGT/CLL-DB (bylaws) LIGM, Montpellier, France
IG sequences from CLL, an initiative of the IMGT/CLL-DB group
IMGT/GENE-DB (doc) LIGM, Montpellier, France
International nomenclature for IG and TR genes from human, mouse, rat and rabbit (2 769 genes, 3 867 alleles)
IMGT/3Dstructure-DB and IMGT/2Dstructure-DB (doc) LIGM, Montpellier, France
3D structures (IMGT Colliers de Perles) of IG antibodies, TR, MHC and RPI (2 623 entries)
Source: PDB, INN, Kabat
IMGT/mAb-DB (doc) LIGM, Montpellier, France
Monoclonal antibodies (IG, mAb) and fusion proteins for immune applications (FPIA) (379 entries)
IMGT tools
IMGT/V-QUEST (doc) (sequence alignment software for IG, TR and HLA)
IMGT/JunctionAnalysis (doc) (for human and mouse IG and TR)
IMGT/HighV-QUEST (doc) (NGS High-Throughput analysis of IG and TR)
IMGT/Allele-Align
IMGT/PhyloGene (doc)
IMGT/DomainDisplay (doc) (Amino acid sequences)
IMGT/LocusView, IMGT/GeneView, IMGT/GeneSearch, IMGT/CloneSearch (doc) (for human IGK, IGL, IGH, TRA/TRD, TRB, TRG, mouse TRA/TRD and human MHC)
IMGT/GeneInfo (doc) (TIMC and ICH, Grenoble; LIGM, Montpellier)
IMGT/GeneFrequency (doc)
IMGT/DomainGapAlign (doc)
IMGT/Collier-de-Perles (doc)
IMGT/DomainSuperimpose
IMGT/StructuralQuery (doc)
IMGT Web resources
IMGT Repertoire (IG and TR, MHC and RPI)
IMGT Scientific chart (Sequence description, Numbering, Nomenclature, Representation rules)
IMGT Index (FactsBook)
IMGT Bloc-notes (Interesting links, PubMed, Meeting announcements, Postdoctoral positions and jobs, Messages, Search engines...)
IMGT Education (IMGT Lexique, Aide-mémoire, Tutorials, Questions and answers, Enseignements...)
IMGT Posters and diaporama
The IMGT Medical page
The IMGT Veterinary page
The IMGT Biotechnology page
The IMGT Immunoinformatics page
IMGT other accesses
IMGT Other accesses (ARSA, SRS, MRS)
Compare your sequence against IMGT (BLAST, FASTA)
IMGT/LIGM-DB Sequence submission
IMGT downloads
IMGT Latest news
Temporary remove of the files imgtrefseq (Thu, 22 Sep 2011 17:00:10 +0200)
New release of IMGT/V-QUEST (IMGT/V-QUEST reference directory release: 201138-4) (Thu, 22 Sep 2011 17:00:10 +0200)
Update of Protein displays in IMGT Repertoire (Thu, 15 Sep 2011 09:47:54 +0000)
New release of IMGT/3Dstructure-DB (Program version: 4.7.0, Database release: 201137-2) (Tue, 13 Sep 2011 14:28:10 +0200)
Rechercher :
Suivant Précédent Surligner tout Respecter la casse
Terminé
zotero

Les objectifs de IMGT/LIGM-DB

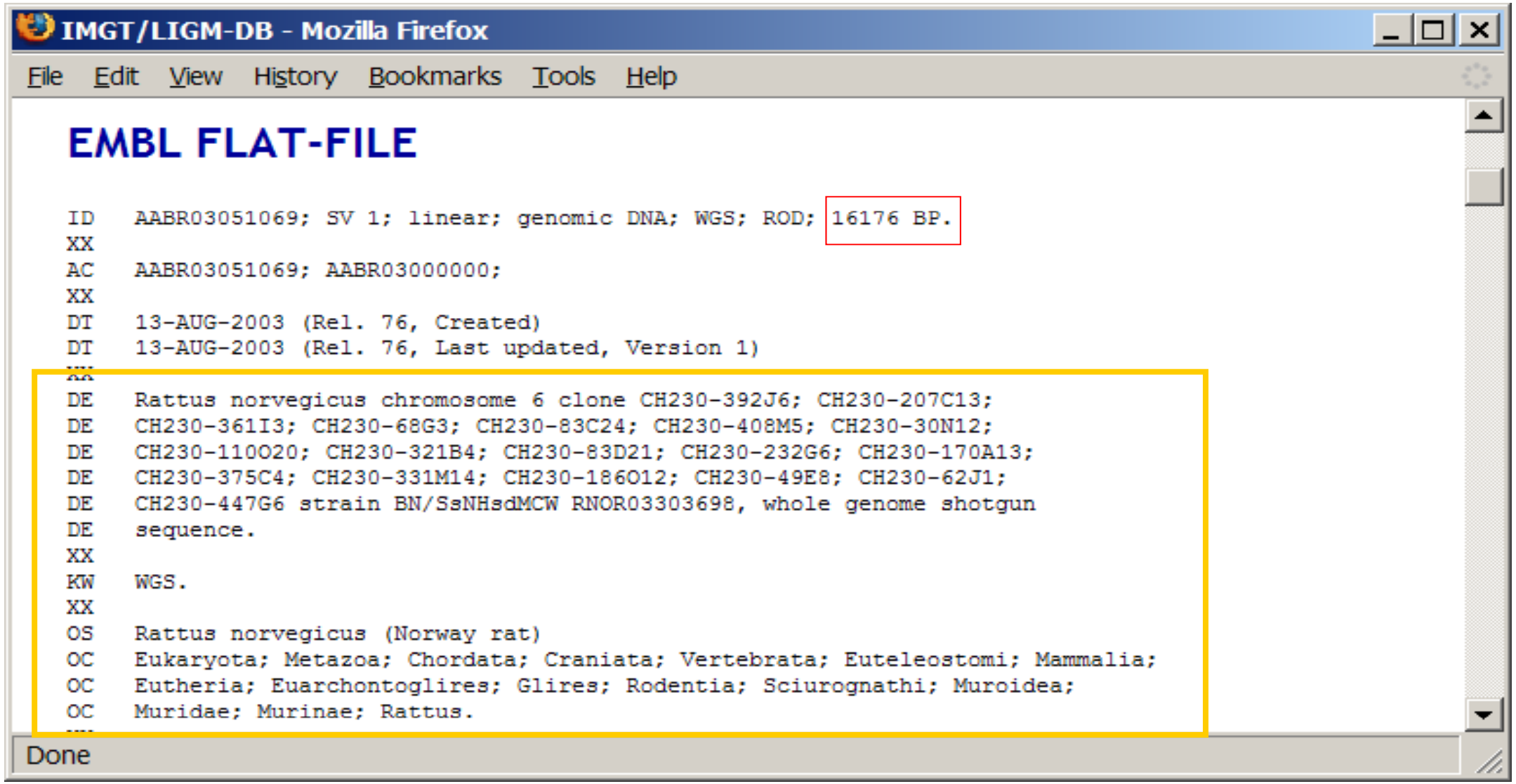
http://www.imgt.org

- Contenir toutes les séquences IG et TR des bases EMBL/GenBank/DDBJ
  + de 157.000 séquences de 280 espèces

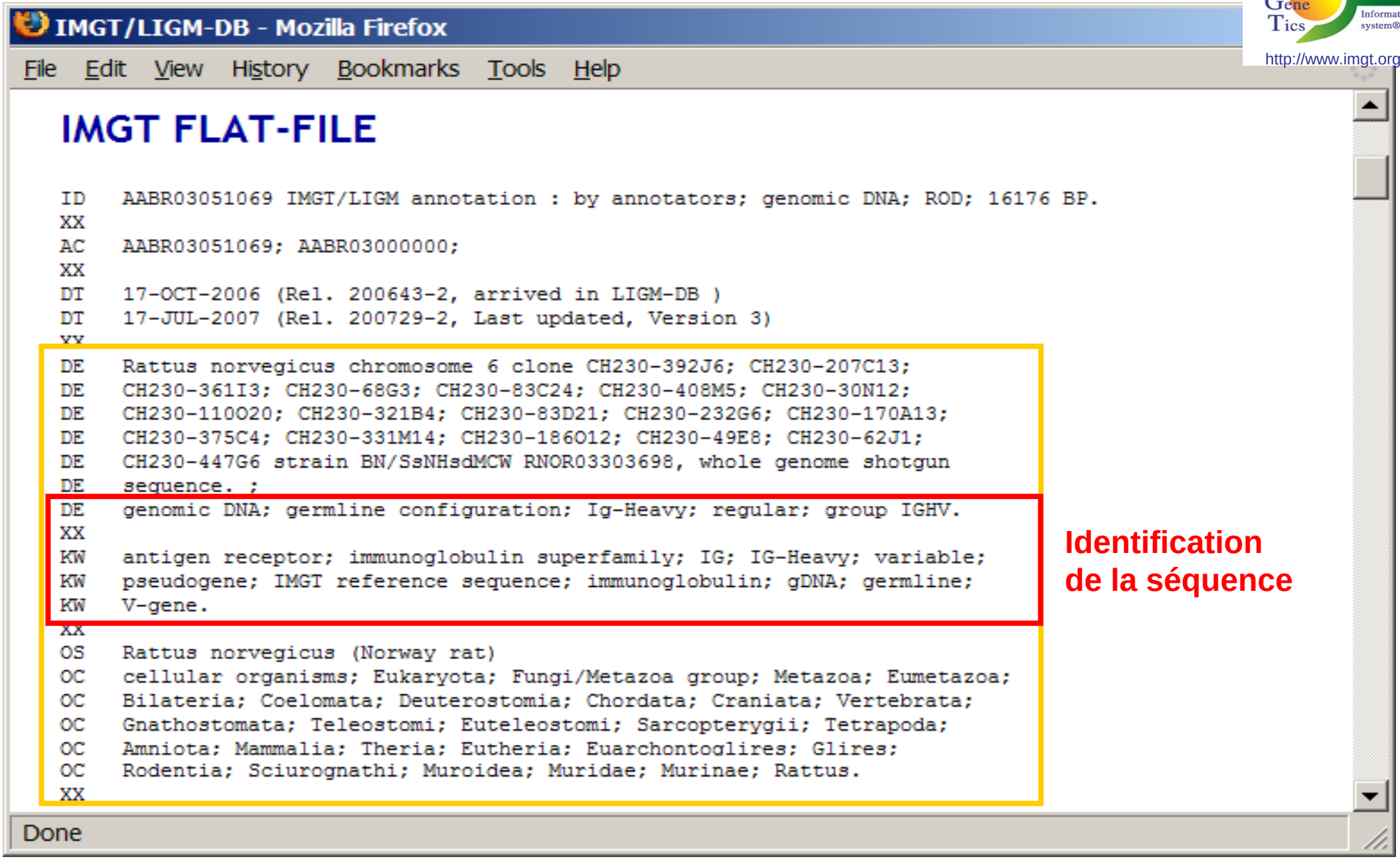
- Expertiser les séquences en fonction des connaissances en immunogénétique
  - identification des gènes,
  - description des motifs (régions codantes, …)

- Suivi et intégration des nouvelles connaissances dans le système d'information IMGT®

A l'arrivée dans IMGT®



Après expertise dans IMGT®

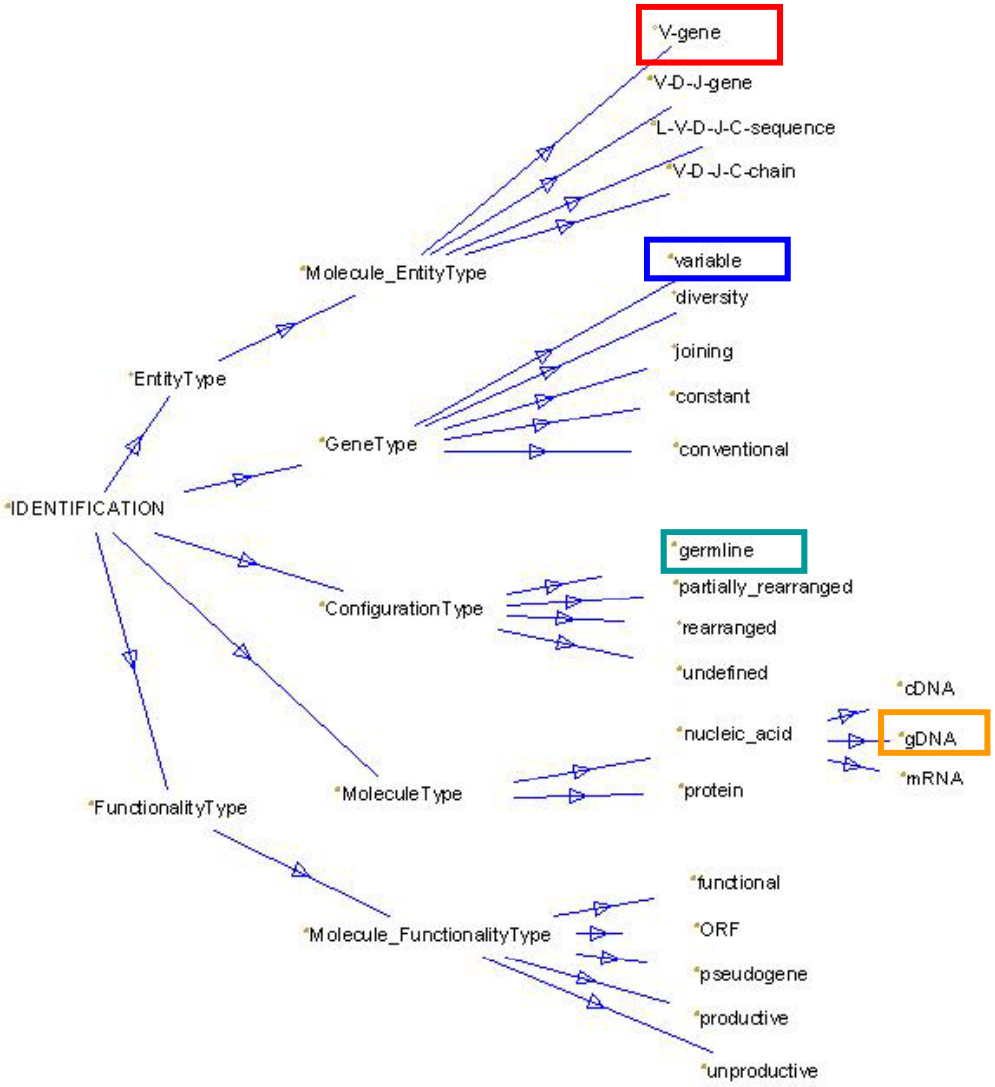


Identification de la séquence

Im Muno Gene Tics Information system®

http://www.imgt.org

Result: the formalization of the relationships between concepts of IDENTIFICATION allows to represent the rules of the dependency between keywords used in IMGT® databases



IMGT/LIGM-DB flatfile

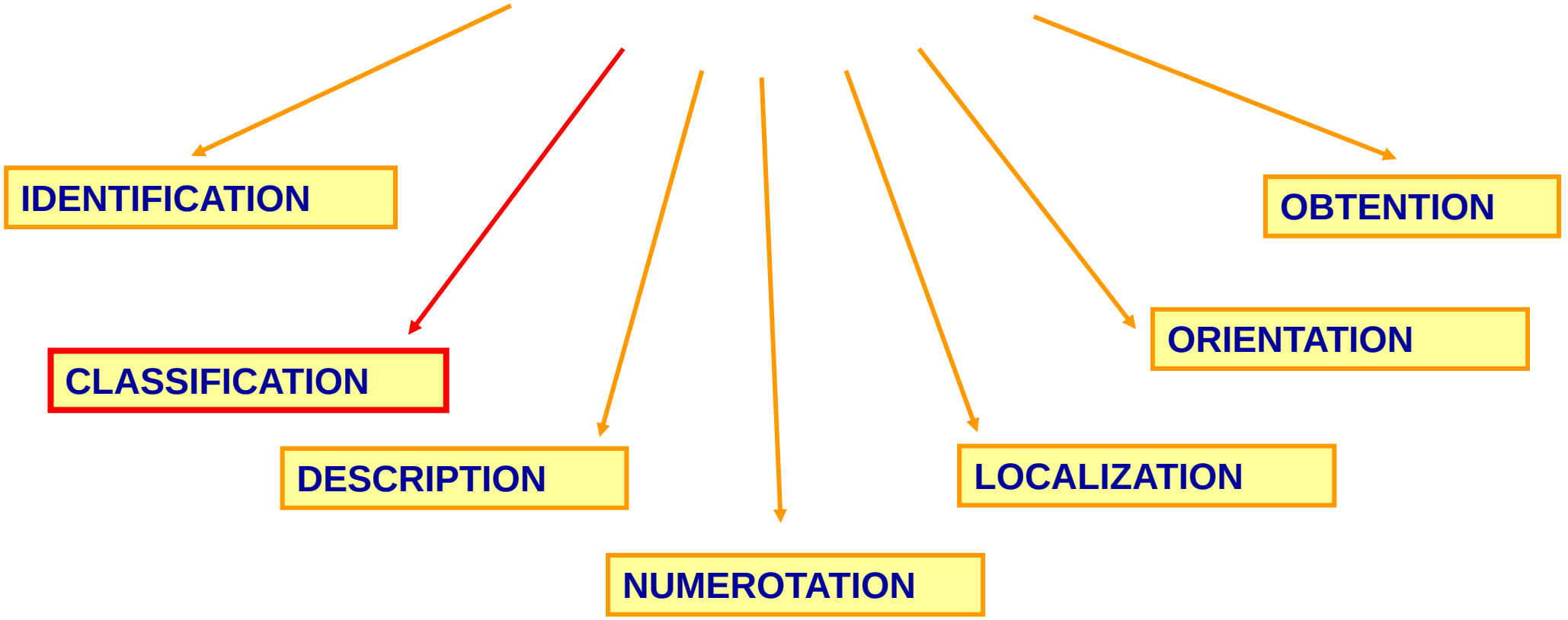
```
ID   X07448 IMGT/LIGM annotation : by annotators; genomic DNA; HUM; 618 BP.
XX
AC   X07448;
XX
DT   15-MAY-1995 (Rel. 2, arrived in LIGM-DB )
DT   20-OCT-2008 (Rel. 200843-1, Last updated, Version 10)
XX
DE   Human V35 gene for Ig heavy chain  ;
DE   genomic DNA; germline configuration; Ig-Heavy; regular; functionality
DE   functional; group IGHV; subgroup HV1.
XX
KW   antigen receptor; Immunoglobulin superfamily (IgSF);
KW   immunoglobulin (IG); IG-Heavy; variable; IMGT reference sequence; gDNA;
KW   germline; functional; V-gene.
XX
OS   Homo sapiens (human)
OC   cellular organisms; Eukaryota; Fungi/Metazoa group; Metazoa; Eumetazoa;
OC   Bilateria; Coelomata; Deuterostomia; Chordata; Craniata; Vertebrata;
OC   Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Tetrapoda;
OC   Amniota; Mammalia; Theria; Eutheria; Euarchontoglires; Primates;
OC   Haplorrhini; Simiiformes; Catarrhini; Hominoidea; Hominidae;
OC   Homo/Pan/Gorilla group; Homo.
XX
RN   [1]
RP   1-618
RX   PUBMED; 2841108.
RA   Matsuda F., Lee K.H., Nakai S., Sato T., Kodaira M., Zong S.Q., Ohno H.,
RA   Fukuhara S., Honjo T.;
RT   "Dispersed localization of D segments in the human immunoglobulin
RT   heavy-chain locus";
RL   EMBO J. 7(4):1047-1051(1988).
XX
```

IMGT-ONTOLOGY

http://www.imgt.org

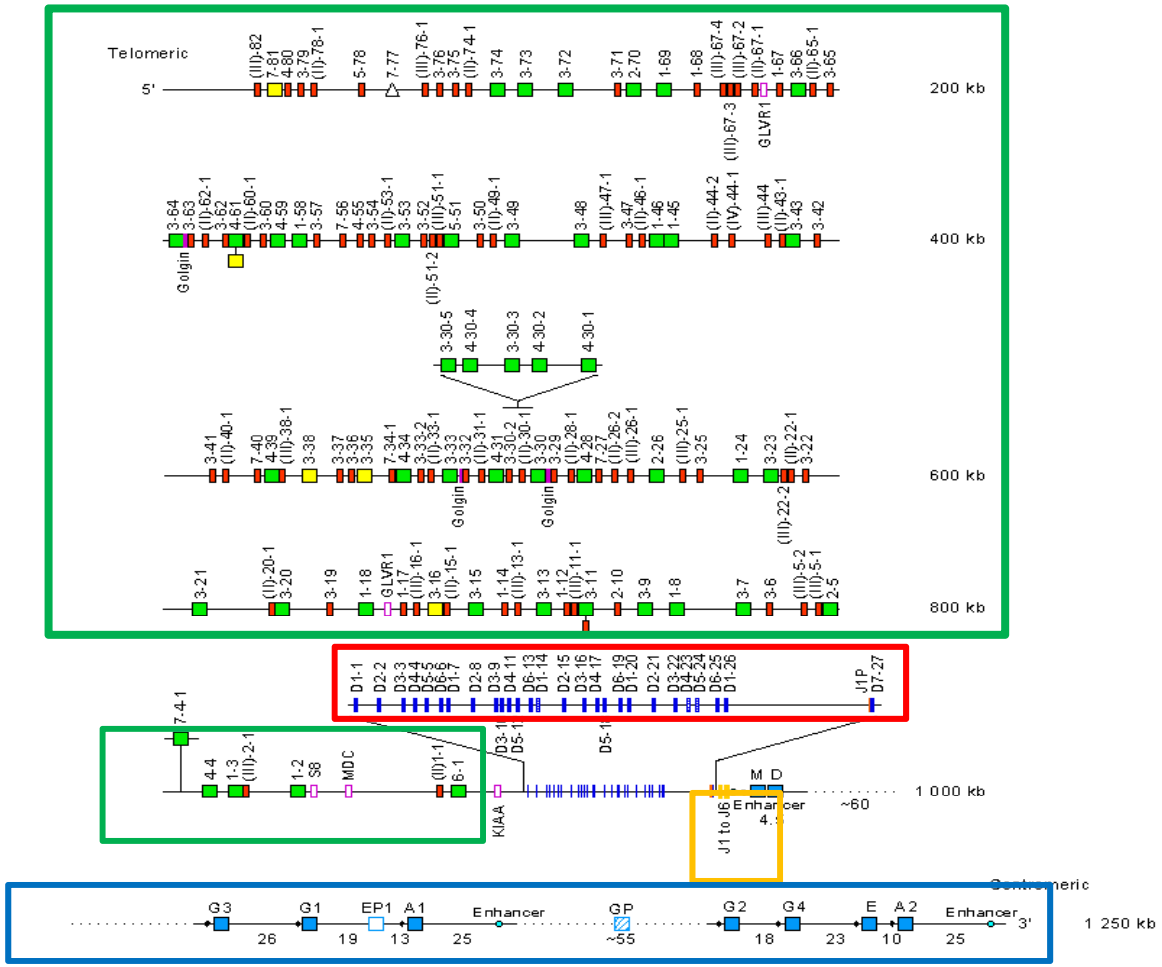
Pour standardiser, partager, réutiliser et représenter les connaissances en immunogénétique

IMGT-ONTOLOGY repose sur 7 axiomes:



2 – Concepts de classification d'IMGT-ONTOLOGY

Group: A group is a set of genes which share the same 'gene type' and participate potentially to the synthesis of a polypeptide of the same 'chain type'.

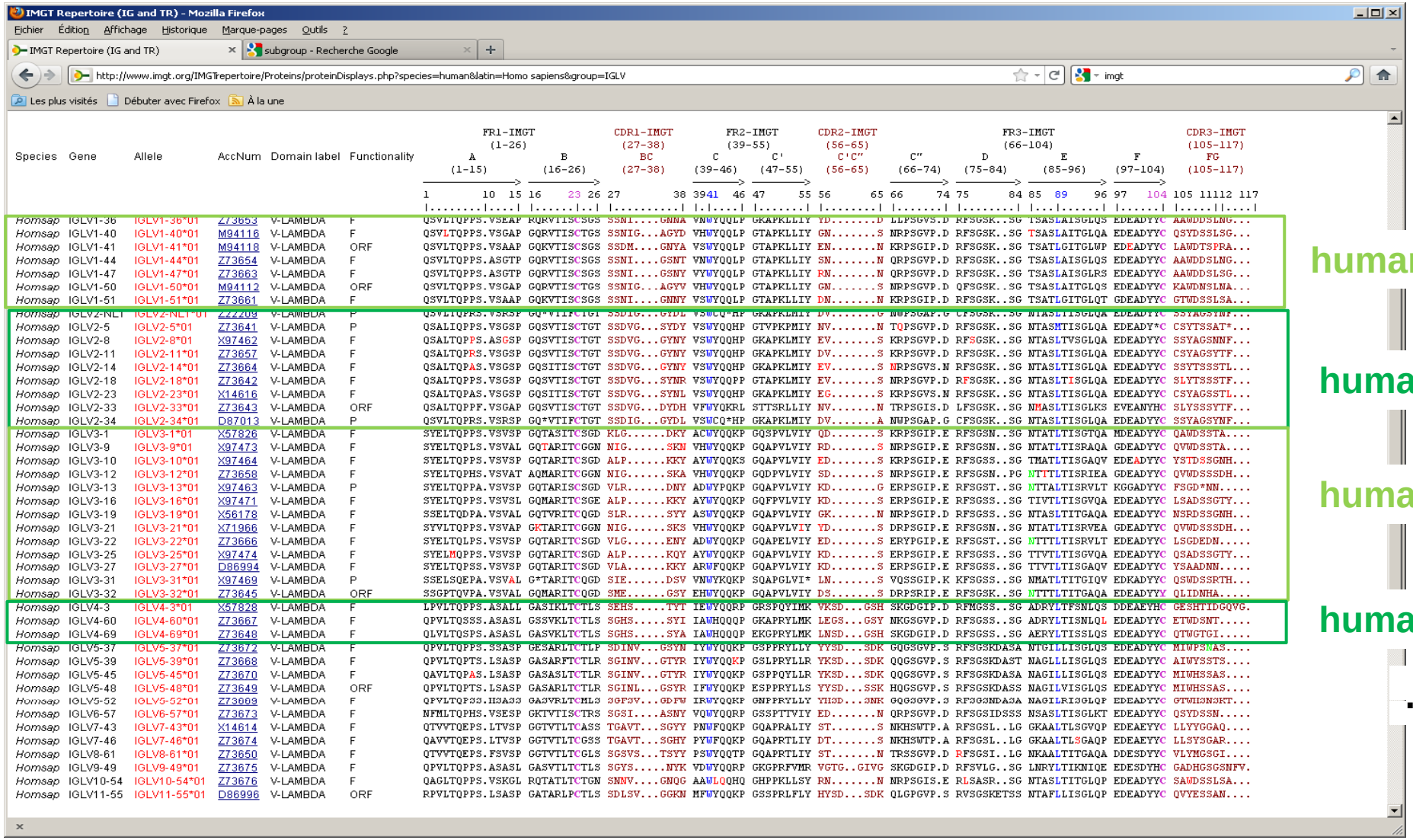


Human IGH locus on chromosome 14 (14q32.33)

2 – Concepts de classification d'IMGT-ONTOLOGY

Subgroup:

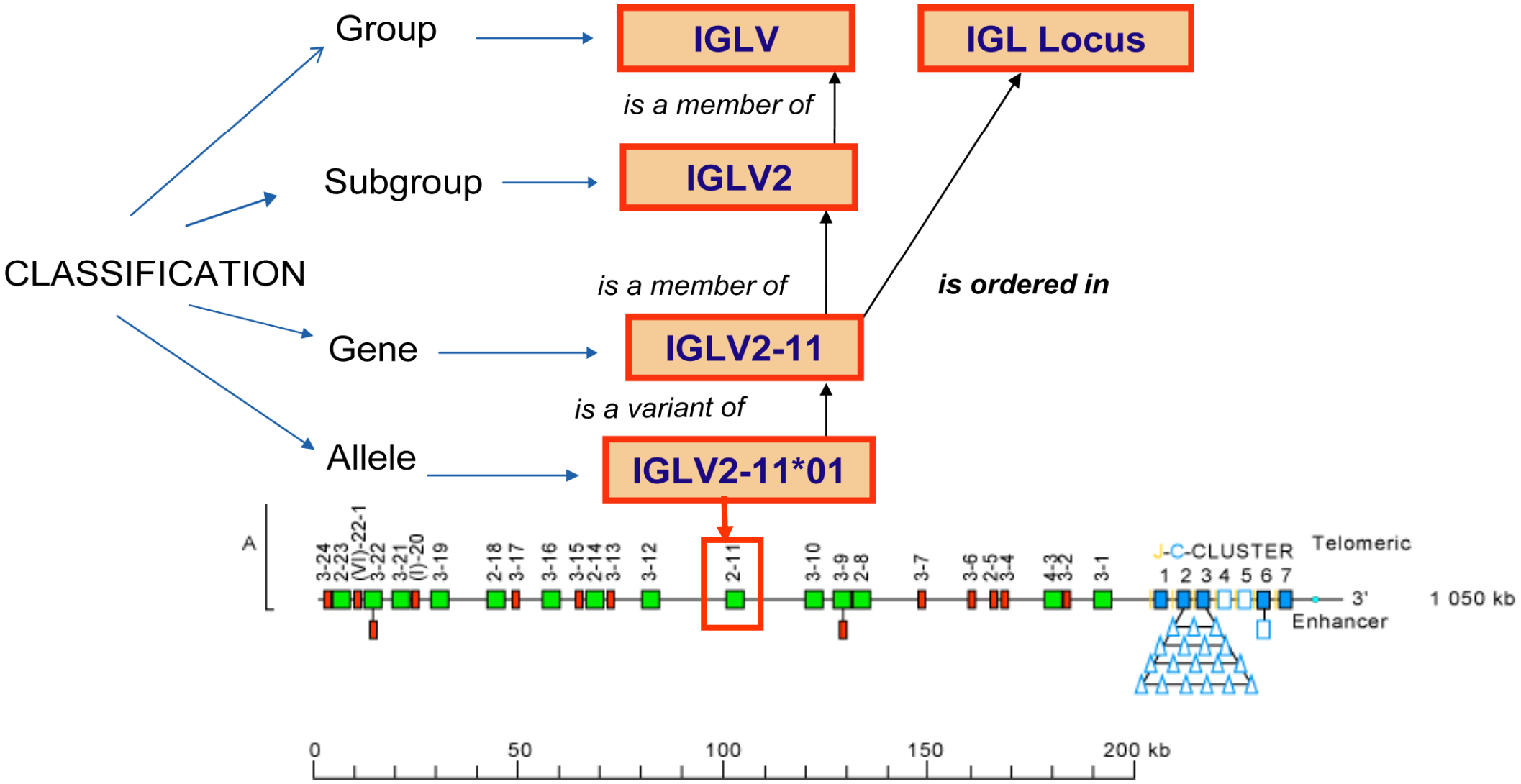
A subgroup is a set of IG or TR genes (C-GENE, V-GENE, D-GENE , J-GENE) which belong to the same group, in a given species, and which share at least 75% identity at the nucleotide level (in the germline configuration for V, D, and J).



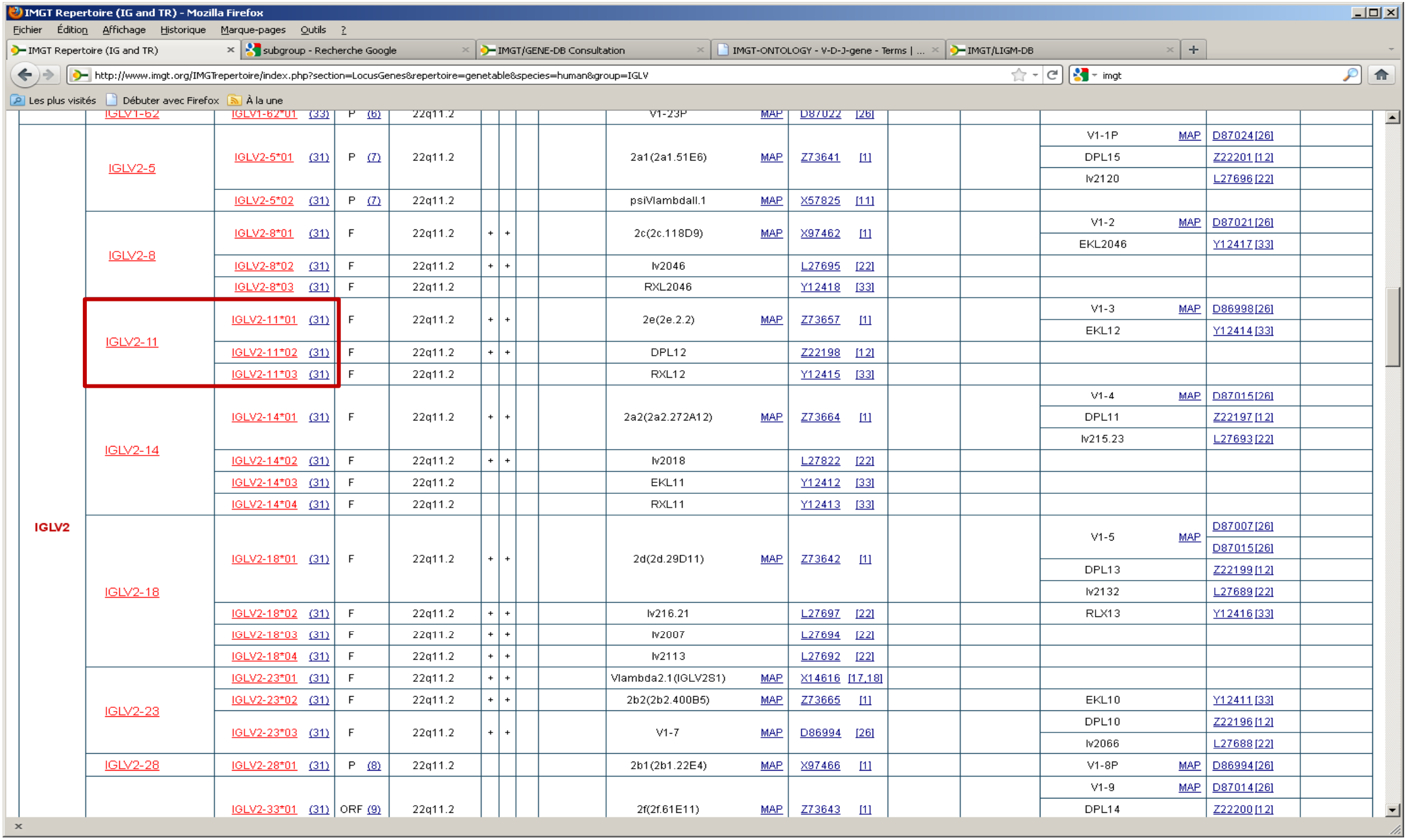
Concepts of CLASSIFICATION: IMGT® gene nomenclature

http://www.imgt.org

Human IGL locus (22q11.2)



Gene table: Human *(Homo sapiens)* IGLV



IMGT Repertoire (IG and TR) - Mozilla Firefox

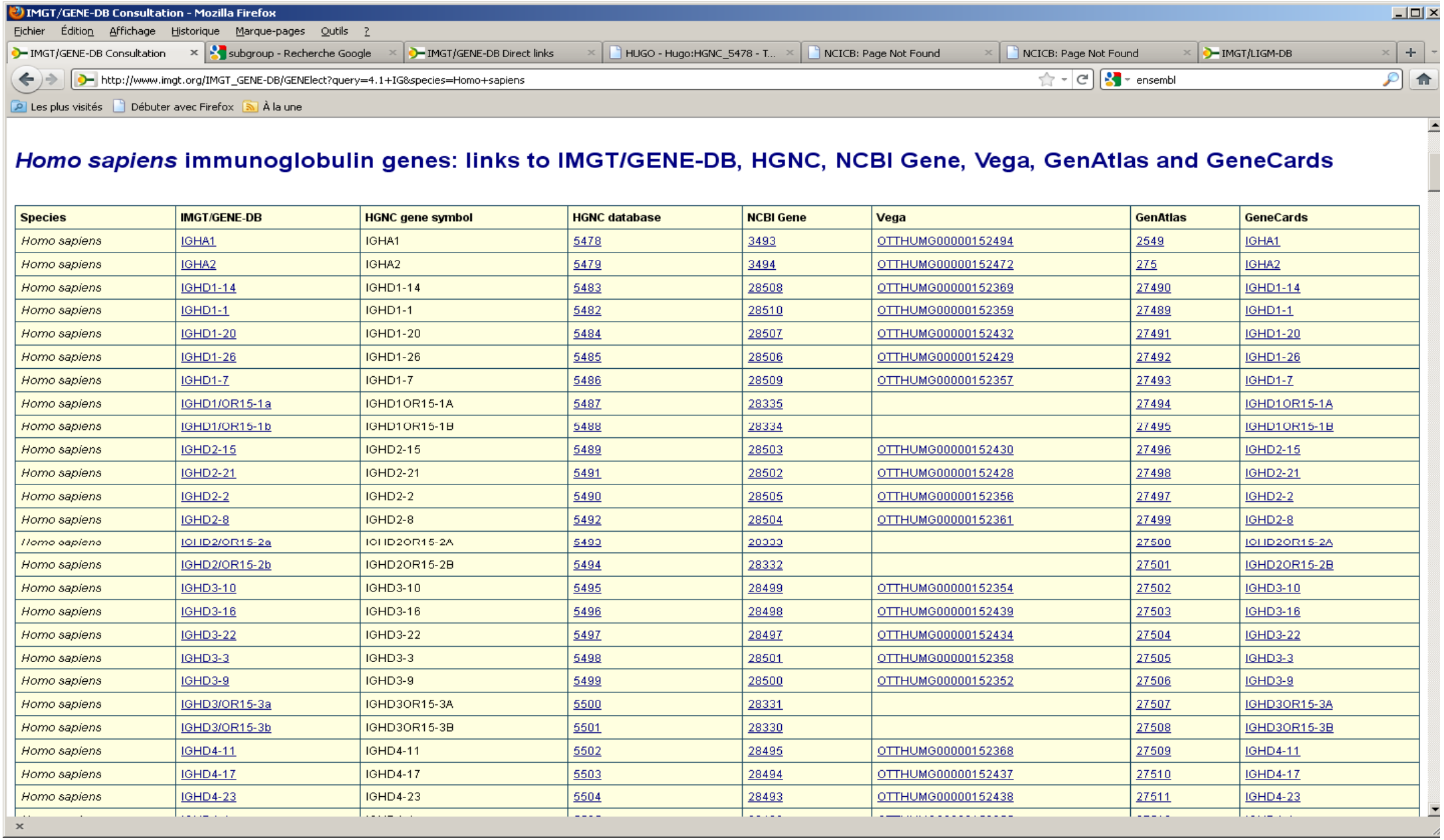
Fichier Édition Affichage Historique Marque-pages Outils ?

IMGT Repertoire (IG and TR) | subgroup - Recherche Google | IMGT/GENE-DB Consultation | IMGT-ONTOLOGY - V-D-J-gene - Terms | ... | IMGT/LIGM-DB

http://www.imgt.org/IMGTrepertoire/index.php?section=LocusGenes&repertoire=genetable&species=human&group=IGLV

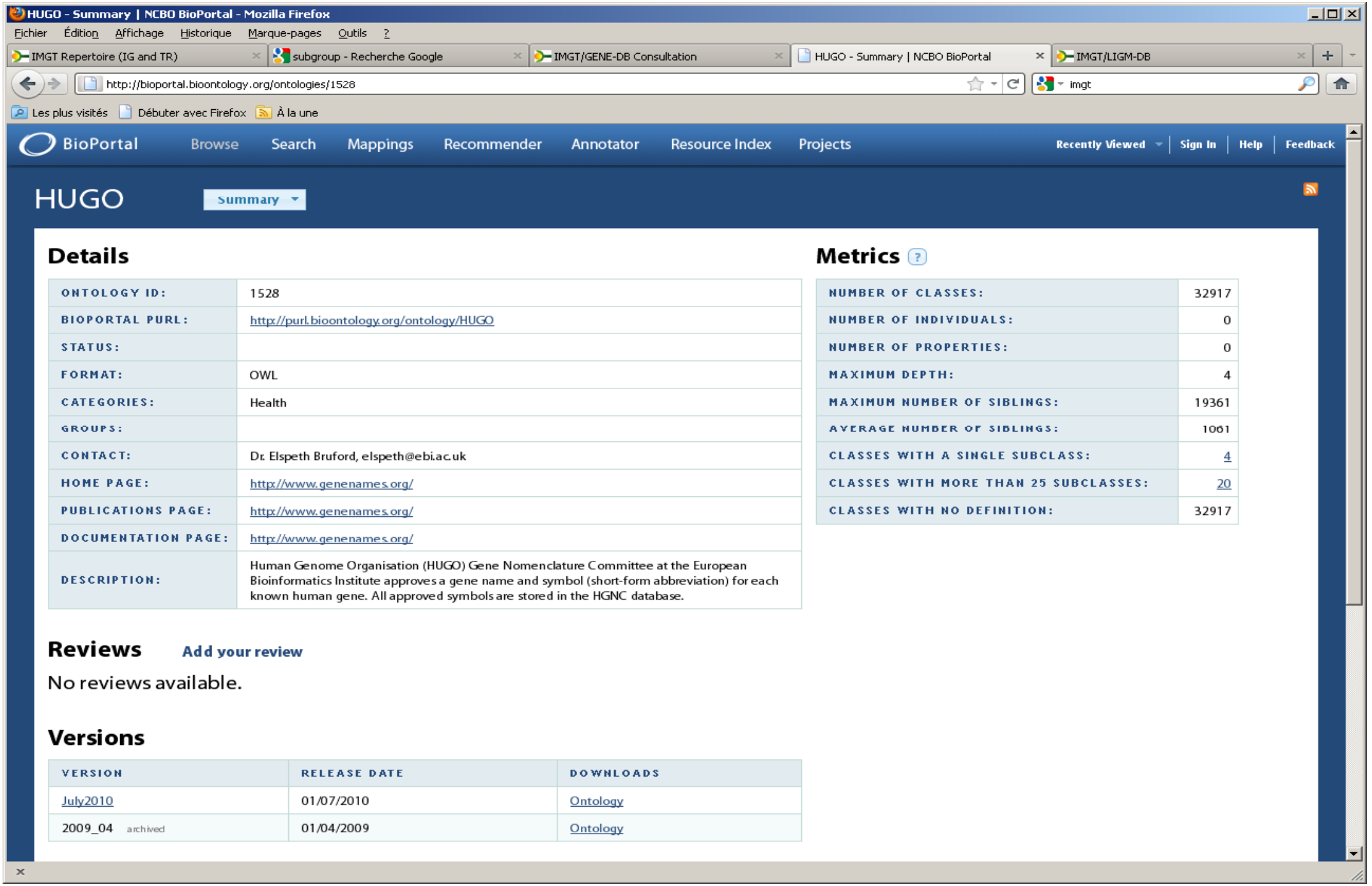
Les plus visités Débuter avec Firefox À la une

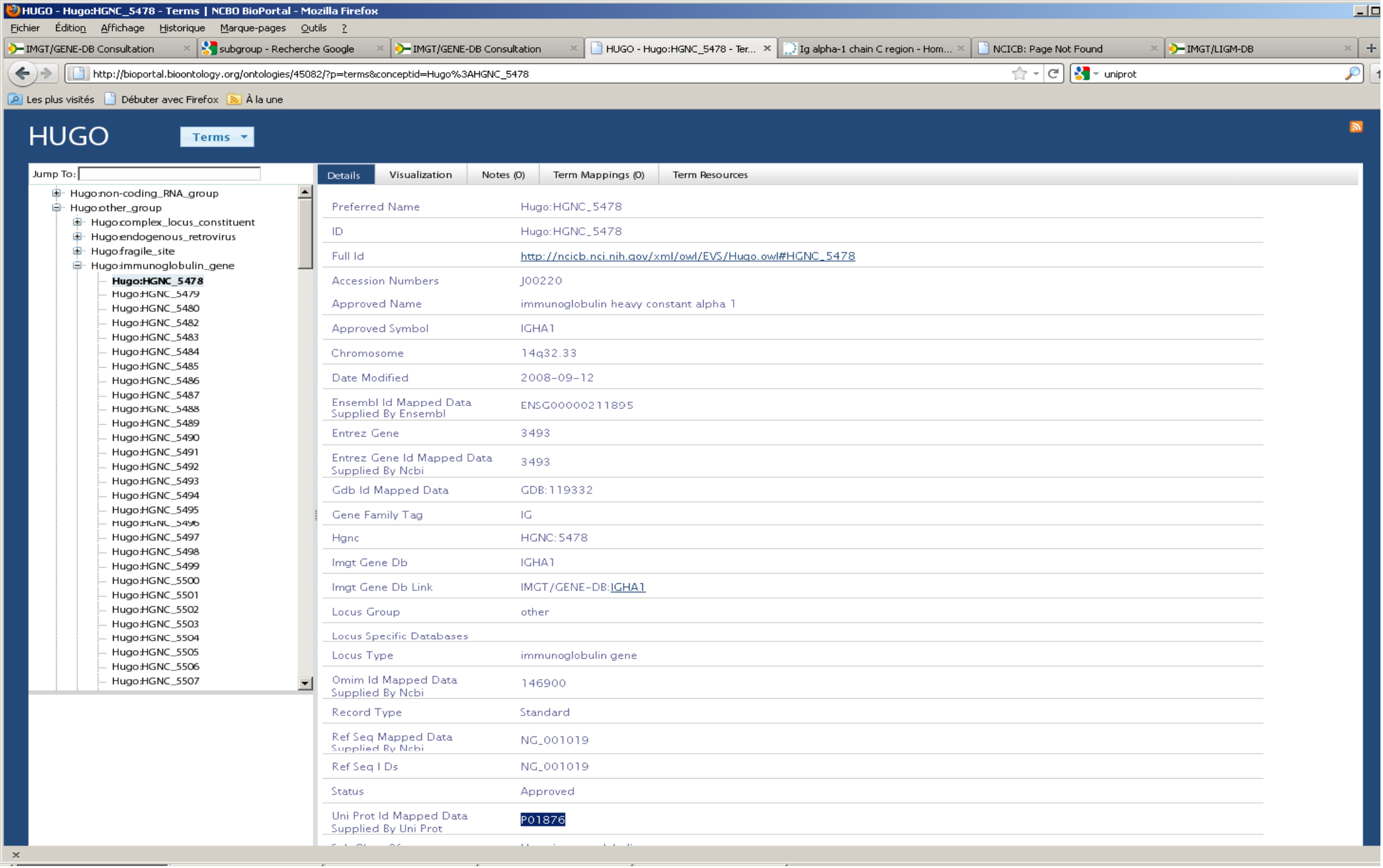
	IGLV1-62	IGLV1-62*01 (33)	P (6)	22q11.2					V1-23P	MAP	D87022 [26]					
IGLV2	IGLV2-5	IGLV2-5*01 (31)	P (7)	22q11.2					2a1(2a1.51E6)	MAP	Z73641 [1]			V1-1P	MAP	D87024 [26]
														DPL15		Z22201 [12]
														lv2120		L27696 [22]
		IGLV2-5*02 (31)	P (7)	22q11.2					psiVlambdaII.1	MAP	X57825 [11]					
	IGLV2-8	IGLV2-8*01 (31)	F	22q11.2	+	+			2c(2c.118D9)	MAP	X97462 [1]			V1-2	MAP	D87021 [26]
														EKL2046		Y12417 [33]
		IGLV2-8*02 (31)	F	22q11.2	+	+			lv2046		L27695 [22]					
		IGLV2-8*03 (31)	F	22q11.2					RXL2046		Y12418 [33]					
	IGLV2-11	IGLV2-11*01 (31)	F	22q11.2	+	+			2e(2e.2.2)	MAP	Z73657 [1]			V1-3	MAP	D86998 [26]
														EKL12		Y12414 [33]
		IGLV2-11*02 (31)	F	22q11.2	+	+			DPL12		Z22198 [12]					
		IGLV2-11*03 (31)	F	22q11.2					RXL12		Y12415 [33]					
	IGLV2-14	IGLV2-14*01 (31)	F	22q11.2	+	+			2a2(2a2.272A12)	MAP	Z73664 [1]			V1-4	MAP	D87015 [26]
														DPL11		Z22197 [12]
														lv215.23		L27693 [22]
		IGLV2-14*02 (31)	F	22q11.2	+	+			lv2018		L27822 [22]					
		IGLV2-14*03 (31)	F	22q11.2					EKL11		Y12412 [33]					
		IGLV2-14*04 (31)	F	22q11.2					RXL11		Y12413 [33]					
	IGLV2-18	IGLV2-18*01 (31)	F	22q11.2	+	+			2d(2d.29D11)	MAP	Z73642 [1]			V1-5	MAP	D87007 [26]
																D87015 [26]
														DPL13		Z22199 [12]
														lv2132		L27689 [22]
		IGLV2-18*02 (31)	F	22q11.2	+	+			lv216.21		L27697 [22]			RLX13		Y12416 [33]
		IGLV2-18*03 (31)	F	22q11.2	+	+			lv2007		L27694 [22]					
		IGLV2-18*04 (31)	F	22q11.2	+	+			lv2113		L27692 [22]					
	IGLV2-23	IGLV2-23*01 (31)	F	22q11.2	+	+			Vlambda2.1(IGLV2S1)	MAP	X14616 [17,18]					
		IGLV2-23*02 (31)	F	22q11.2	+	+			2b2(2b2.400B5)	MAP	Z73665 [1]			EKL10		Y12411 [33]
		IGLV2-23*03 (31)	F	22q11.2	+	+			V1-7	MAP	D86994 [26]			DPL10		Z22196 [12]
														lv2066		L27688 [22]
	IGLV2-28	IGLV2-28*01 (31)	P (8)	22q11.2					2b1(2b1.22E4)	MAP	X97466 [1]			V1-8P	MAP	D86994 [26]
		IGLV2-33*01 (31)	ORF (9)	22q11.2					2f(2f.61E11)	MAP	Z73643 [1]			V1-9	MAP	D87014 [26]
														DPL14		Z22200 [12]



Homo sapiens immunoglobulin genes: links to IMGT/GENE-DB, HGNC, NCBI Gene, Vega, GenAtlas and GeneCards

Species	IMGT/GENE-DB	HGNC gene symbol	HGNC database	NCBI Gene	Vega	GenAtlas	GeneCards
Homo sapiens	IGHA1	IGHA1	5478	3493	OTTHUMG00000152494	2549	IGHA1
Homo sapiens	IGHA2	IGHA2	5479	3494	OTTHUMG00000152472	275	IGHA2
Homo sapiens	IGHD1-14	IGHD1-14	5483	28508	OTTHUMG00000152369	27490	IGHD1-14
Homo sapiens	IGHD1-1	IGHD1-1	5482	28510	OTTHUMG00000152359	27489	IGHD1-1
Homo sapiens	IGHD1-20	IGHD1-20	5484	28507	OTTHUMG00000152432	27491	IGHD1-20
Homo sapiens	IGHD1-26	IGHD1-26	5485	28506	OTTHUMG00000152429	27492	IGHD1-26
Homo sapiens	IGHD1-7	IGHD1-7	5486	28509	OTTHUMG00000152357	27493	IGHD1-7
Homo sapiens	IGHD1/OR15-1a	IGHD1OR15-1A	5487	28335		27494	IGHD1OR15-1A
Homo sapiens	IGHD1/OR15-1b	IGHD1OR15-1B	5488	28334		27495	IGHD1OR15-1B
Homo sapiens	IGHD2-15	IGHD2-15	5489	28503	OTTHUMG00000152430	27496	IGHD2-15
Homo sapiens	IGHD2-21	IGHD2-21	5491	28502	OTTHUMG00000152428	27498	IGHD2-21
Homo sapiens	IGHD2-2	IGHD2-2	5490	28505	OTTHUMG00000152356	27497	IGHD2-2
Homo sapiens	IGHD2-8	IGHD2-8	5492	28504	OTTHUMG00000152361	27499	IGHD2-8
Homo sapiens	IGHD2/OR15-2a	IGHD2OR15-2A	5493	28333		27500	IGHD2OR15-2A
Homo sapiens	IGHD2/OR15-2b	IGHD2OR15-2B	5494	28332		27501	IGHD2OR15-2B
Homo sapiens	IGHD3-10	IGHD3-10	5495	28499	OTTHUMG00000152354	27502	IGHD3-10
Homo sapiens	IGHD3-16	IGHD3-16	5496	28498	OTTHUMG00000152439	27503	IGHD3-16
Homo sapiens	IGHD3-22	IGHD3-22	5497	28497	OTTHUMG00000152434	27504	IGHD3-22
Homo sapiens	IGHD3-3	IGHD3-3	5498	28501	OTTHUMG00000152358	27505	IGHD3-3
Homo sapiens	IGHD3-9	IGHD3-9	5499	28500	OTTHUMG00000152352	27506	IGHD3-9
Homo sapiens	IGHD3/OR15-3a	IGHD3OR15-3A	5500	28331		27507	IGHD3OR15-3A
Homo sapiens	IGHD3/OR15-3b	IGHD3OR15-3B	5501	28330		27508	IGHD3OR15-3B
Homo sapiens	IGHD4-11	IGHD4-11	5502	28495	OTTHUMG00000152368	27509	IGHD4-11
Homo sapiens	IGHD4-17	IGHD4-17	5503	28494	OTTHUMG00000152437	27510	IGHD4-17
Homo sapiens	IGHD4-23	IGHD4-23	5504	28493	OTTHUMG00000152438	27511	IGHD4-23


HUGO - Summary | NCBO BioPortal - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
IMGT Repertoire (IG and TR)
subgroup - Recherche Google
IMGT/GENE-DB Consultation
HUGO - Summary | NCBO BioPortal
IMGT/LIGM-DB
http://bioportal.bioontology.org/ontologies/1528
imgt
Les plus visités Débuter avec Firefox À la une
BioPortal Browse Search Mappings Recommender Annotator Resource Index Projects Recently Viewed Sign In Help Feedback
HUGO Summary
Details
ONTOLOGY ID: 1528
BIOPORTAL PURL: http://purl.bioontology.org/ontology/HUGO
STATUS:
FORMAT: OWL
CATEGORIES: Health
GROUPS:
CONTACT: Dr. Elspeth Bruford, elspeth@ebi.ac.uk
HOME PAGE: http://www.genenames.org/
PUBLICATIONS PAGE: http://www.genenames.org/
DOCUMENTATION PAGE: http://www.genenames.org/
DESCRIPTION: Human Genome Organisation (HUGO) Gene Nomenclature Committee at the European Bioinformatics Institute approves a gene name and symbol (short-form abbreviation) for each known human gene. All approved symbols are stored in the HGNC database.
Metrics
NUMBER OF CLASSES: 32917
NUMBER OF INDIVIDUALS: 0
NUMBER OF PROPERTIES: 0
MAXIMUM DEPTH: 4
MAXIMUM NUMBER OF SIBLINGS: 19361
AVERAGE NUMBER OF SIBLINGS: 1061
CLASSES WITH A SINGLE SUBCLASS: 4
CLASSES WITH MORE THAN 25 SUBCLASSES: 20
CLASSES WITH NO DEFINITION: 32917
Reviews Add your review
No reviews available.
Versions
VERSION RELEASE DATE DOWNLOADS
July2010 01/07/2010 Ontology
2009_04 archived 01/04/2009 Ontology

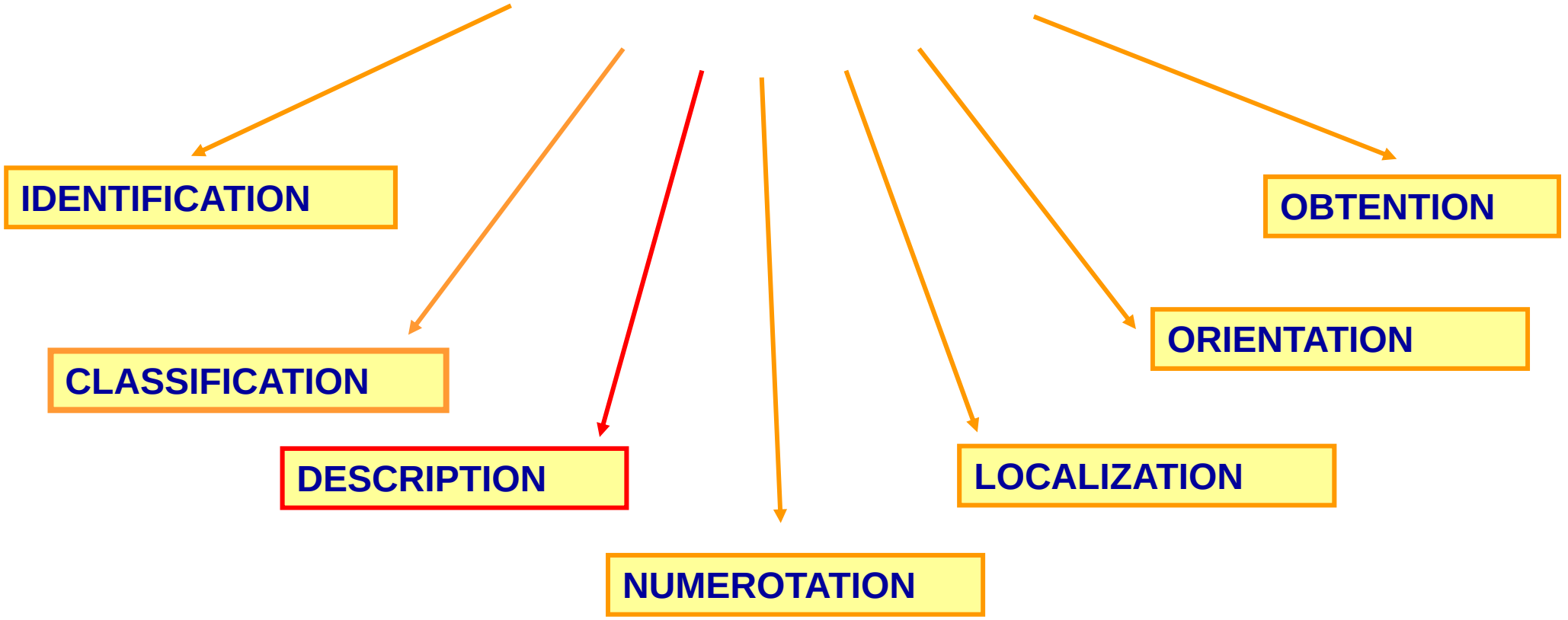

HUGO - Hugo:HGNC_5478 - Terms | NCBO BioPortal - Mozilla Firefox
Fichier Édition Affichage Historique Marque-pages Outils ?
IMGT/GENE-DB Consultation
subgroup - Recherche Google
IMGT/GENE-DB Consultation
HUGO - Hugo:HGNC_5478 - Ter...
Ig alpha-1 chain C region - Hom...
NCICB: Page Not Found
IMGT/LIGM-DB
http://bioportal.bioontology.org/ontologies/45082/?p=terms&conceptid=Hugo%3AHGNC_5478
uniprot
Les plus visités
Débuter avec Firefox
À la une
HUGO
Terms
Jump To:
Hugo:non-coding_RNA_group
Hugo:other_group
Hugo:complex_locus_constituent
Hugo:endogenous_retrovirus
Hugo:fragile_site
Hugo:immunoglobulin_gene
Hugo:HGNC_5478
Hugo:HGNC_5479
Hugo:HGNC_5480
Hugo:HGNC_5482
Hugo:HGNC_5483
Hugo:HGNC_5484
Hugo:HGNC_5485
Hugo:HGNC_5486
Hugo:HGNC_5487
Hugo:HGNC_5488
Hugo:HGNC_5489
Hugo:HGNC_5490
Hugo:HGNC_5491
Hugo:HGNC_5492
Hugo:HGNC_5493
Hugo:HGNC_5494
Hugo:HGNC_5495
Hugo:HGNC_5497
Hugo:HGNC_5498
Hugo:HGNC_5499
Hugo:HGNC_5500
Hugo:HGNC_5501
Hugo:HGNC_5502
Hugo:HGNC_5503
Hugo:HGNC_5504
Hugo:HGNC_5505
Hugo:HGNC_5506
Hugo:HGNC_5507
Details
Visualization
Notes (0)
Term Mappings (0)
Term Resources
Preferred Name Hugo:HGNC_5478
ID Hugo:HGNC_5478
Full Id http://ncicb.nci.nih.gov/xml/owl/EVS/Hugo.owl#HGNC_5478
Accession Numbers J00220
Approved Name immunoglobulin heavy constant alpha 1
Approved Symbol IGHA1
Chromosome 14q32.33
Date Modified 2008-09-12
Ensembl Id Mapped Data Supplied By Ensembl ENSG00000211895
Entrez Gene 3493
Entrez Gene Id Mapped Data Supplied By Ncbi 3493
Gdb Id Mapped Data GDB:119332
Gene Family Tag IG
Hgnc HGNC:5478
Imgt Gene Db IGHA1
Imgt Gene Db Link IMGT/GENE-DB:IGHA1
Locus Group other
Locus Specific Databases
Locus Type immunoglobulin gene
Omim Id Mapped Data Supplied By Ncbi 146900
Record Type Standard
Ref Seq Mapped Data Supplied By Ncbi NG_001019
Ref Seq I Ds NG_001019
Status Approved
Uni Prot Id Mapped Data Supplied By Uni Prot P01876

IMGT-ONTOLOGY

http://www.imgt.org

Pour standardiser, partager, réutiliser et représenter les connaissances en immunogénétique

IMGT-ONTOLOGY repose sur 7 axiomes:

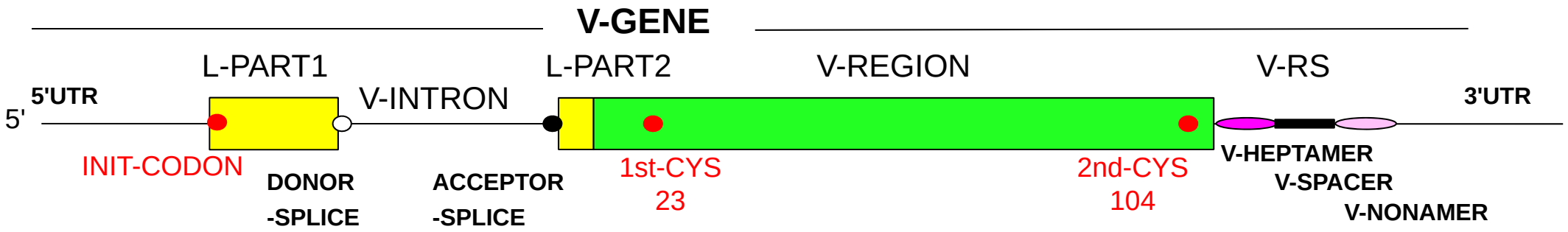


3- Concepts de description: décrire les motifs constitutifs

1 – Description d'une entité V-gene

>X62106.0|HSVI2|*Homo sapiens* VI-2 gene for immunoglobulin heavy chain

```
tgagagctcc gttcctcacc atggactgga cctggaggat cctcttcttg gtggcagcag      60
ccacaggtaa gaggctccct agtcccagtg atgagaaaga gattgagtcc agtccaggga     120
gatctcatcc acttctgtgt tctctccaca ggagcccact cccaggtgca gctggtgcag     180
tctggggctg aggtgaagaa gcctggggcc tcagtgaagg tctcctgcaa ggcttctgga     240
tacaccttca ccggctacta tatgcactgg gtgcgacagg cccctggaca agggcttgag     300
tggatgggat ggatcaaccc taacagtggt ggcacaaact atgcacagaa gtttcagggc     360
agggtcacca tgaccaggga cacgtccatc agcacagcct acatggagct gagcaggctg     420
agatctgacg acacggccgt gtattactgt gcgagagaca cagtgtgaaa tgaaa acccacatcc
tgagggtgc agaaaacccaa gggaggaggc ag
```



Exemple de « Fichier à plat » (Flat file) de IMGT/LIGM-DB

http://www.imgt.org

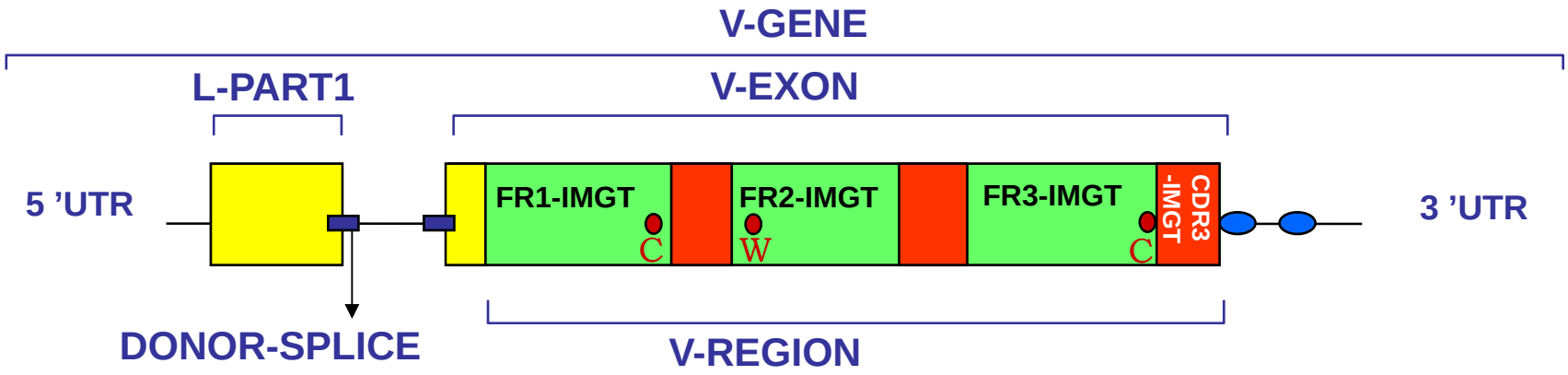
IMGT/LIGM-DB - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

```
FH
FT   V-GENE              1..650
FT                       /db_xref="taxon:9606"
FT                       /chromosome="14 (q32)"
FT                       /cell_line="CGM1 (Human Epstein-Barr virus-transformed
FT                       cell line)"
FT                       /clone="Y103 YAC"
FT                       /allele="IGHV4-4*07"
FT                       /map="155Kb upstream form JH region"
FT                       /gene="IGHV4-4"
FT                       /organism="Homo sapiens"
FT                       /mol_type="genomic DNA"
FT   5'UTR               1..90
FT   L-INTRON-L          91..228
FT   L-PART1             91..136
FT                       /protein_id="CAA44022.1"
FT                       /translation="MKHLWFFLLLVAAPR"
FT   INIT-CODON          91..93
FT   DONOR-SPLICE        136..138
FT   V-INTRON            137..217
FT   ACCEPTOR-SPLICE     215..219
FT   V-EXON              218..521
FT                       /protein_id="CAA44022.1"
FT                       /codon_start=3
FT                       /translation="VLSQVQLQESGPGLVKPSETLSLTCTVSGGSISSYYWSWIR
FT                       QPAGKGLEWIGRIYTSGSTNYNPSLKSRVTMSVDTSKNQFSLKLSSVTAADTAVY
FT                       YCAR"
FT   L-PART2             218..228
FT                       /codon_start=3
FT                       /translation="VLS"
FT   V-REGION            229..521
FT                       /allele="IGHV4-4*07"
FT                       /gene="IGHV4-4"
FT                       /CDR_length="[8.7.2]"
FT                       /translation="QVQLQESGPGLVKPSETLSLTCTVSGGSISSYYWSWIRQPA
FT                       GKGLEWIGRIYTSGSTNYNPSLKSRVTMSVDTSKNQFSLKLSSVTAADTAVYYCA
FT                       R"
FT   FR1-IMGT            229..303
FT                       /AA_IMGT="1 to 26, AA 10 is missing"
FT                       /translation="QVQLQESGPGLVKPSETLSLTCTVS"
FT   1st-CYS             292..294
FT   CDR1-IMGT           304..327
FT                       /AA_IMGT="27 to 34"
FT                       /translation="GGSISSYY"
FT   FR2-IMGT            328..378
```

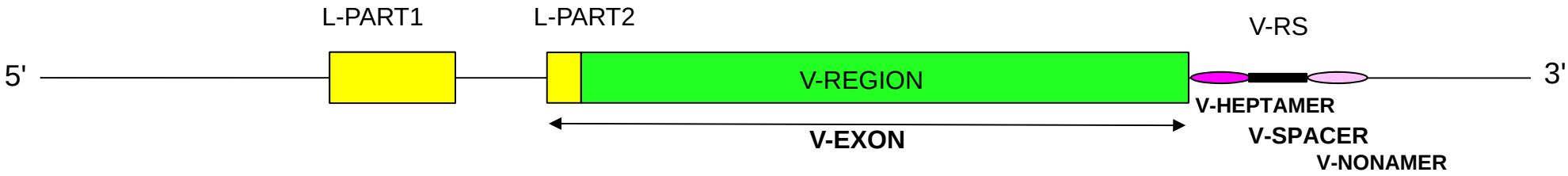
3- Concepts de description: les relations

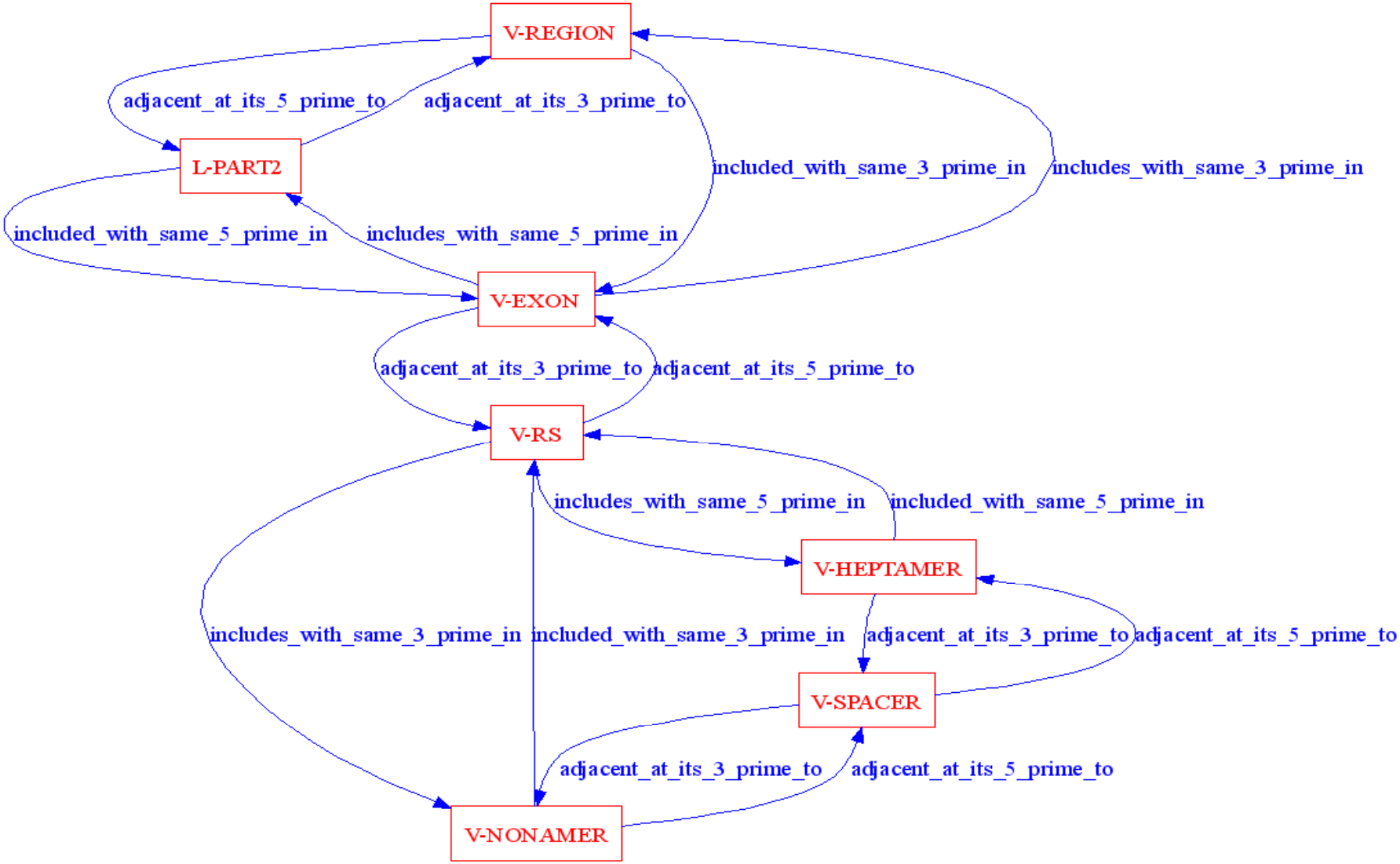
http://www.imgt.org



Label 1	Label 2	Label relations
V-GENE	V-EXON	Includes/included_in
FR3-IMGT	CDR3-IMGT	Adjacent _at_its_5_prime_to/ Adjacent _at_its_3_prime_to
L-PART1	DONOR-SPLICE	Overlaps_in_5_prime/ Overlaps_in_3_prime
V-REGION	FR1-IMGT	Includes_with_same 5_prime/ included_with_same_5_prime_in
V-REGION	CDR3-IMGT	Includes_with_same 3_prime/ included_with_same_3_prime_in

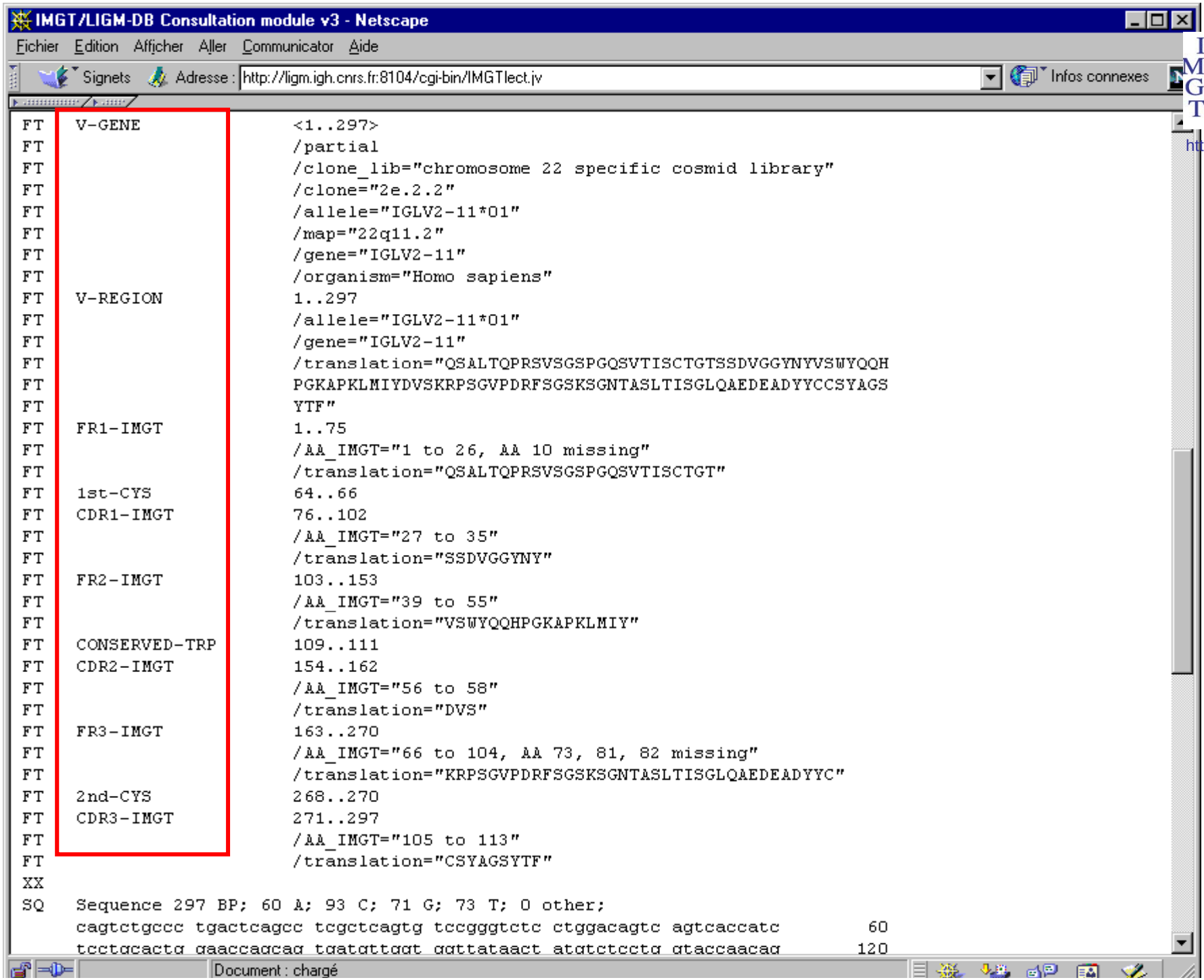
V-GENE





Im Muno Gene Tics Information system®

http://imgt.cines.fr

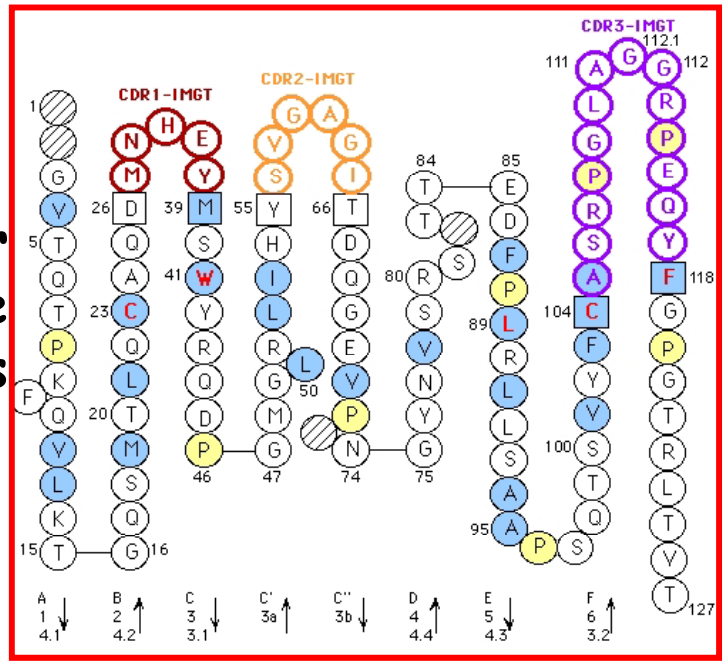


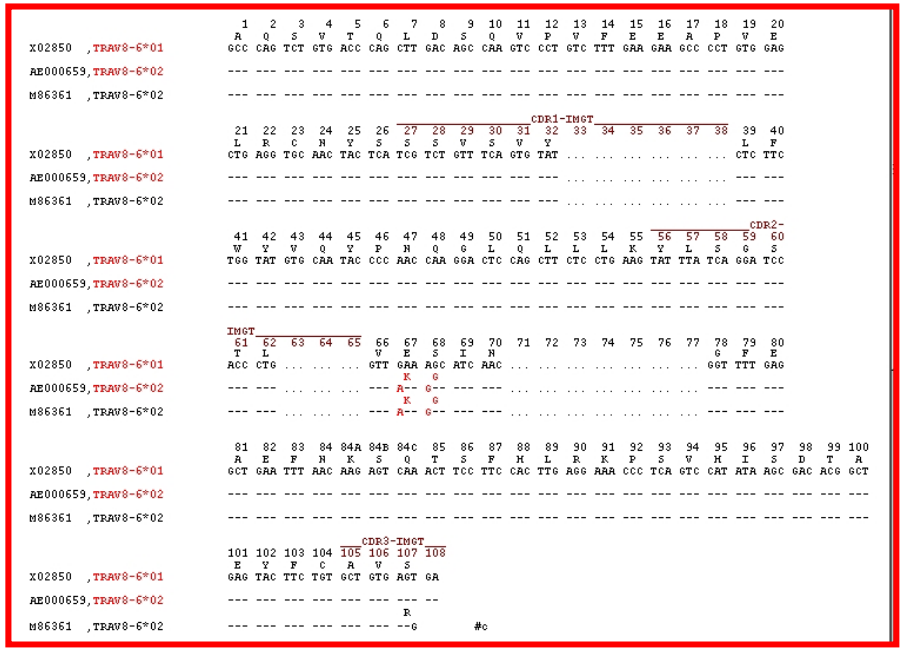
Im Muno Gene Tics Information system®
http://www.imgt.org

"NUMEROTATION"

http://www.imgt.org

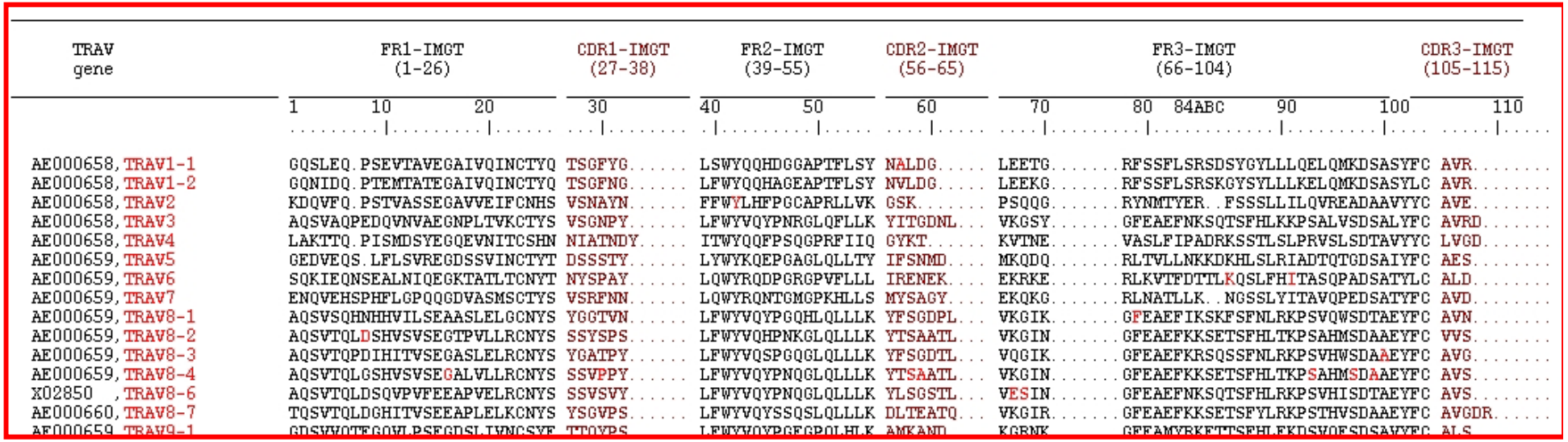
Collier De Perles





Alignment of alleles

Protein Display

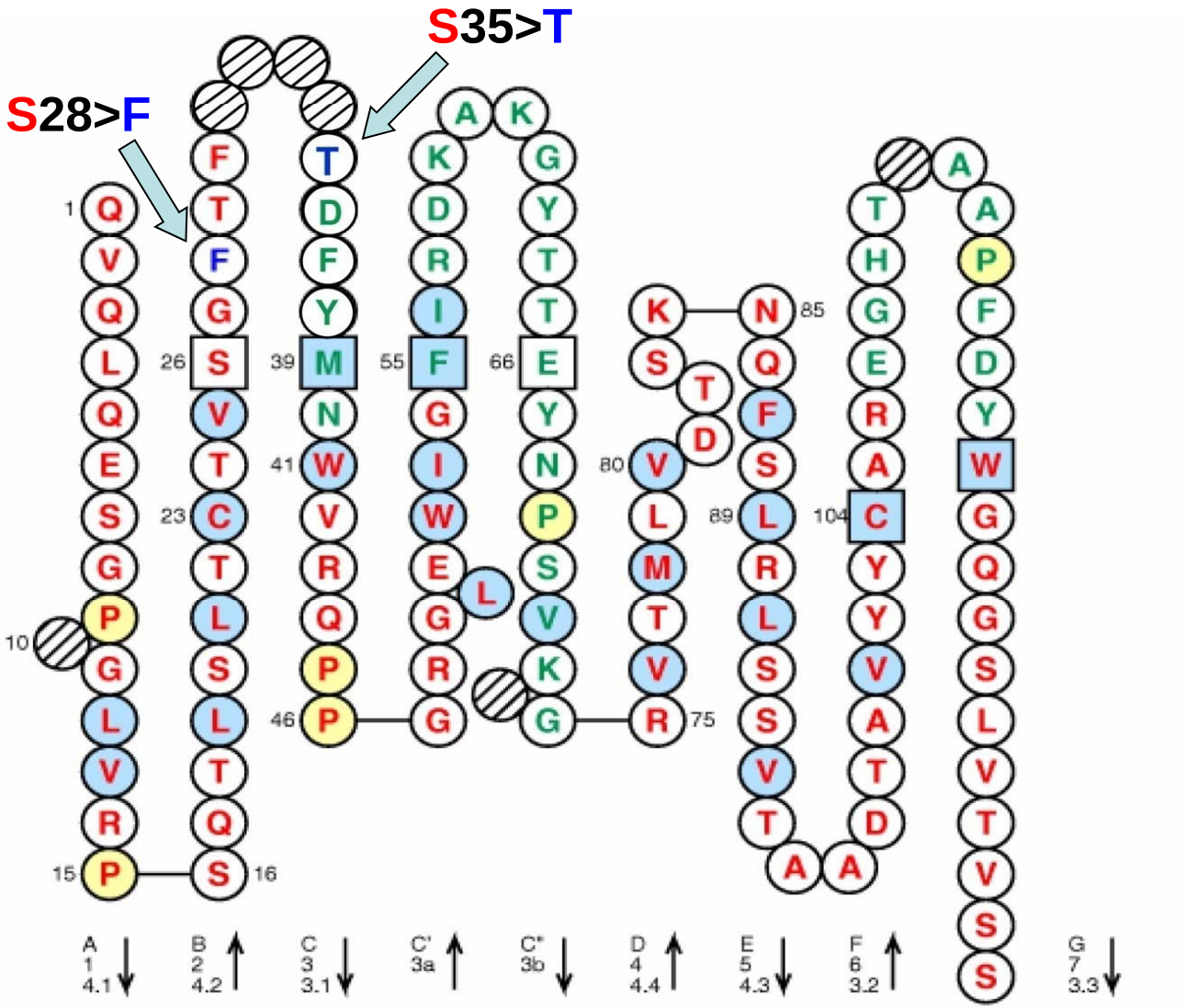


Antibody humanization and engineering

Alemtuzumab (CAMPATH®)

http://www.imgt.org

2 mutations:



VH domain
[8.10.12]

human
rat

IMGT® http://www.imgt.org

Annotation and IMGT-ONTOLOGY concepts

IDENTIFICATION

IMGT/LIGM-DB Consultation module v3 - Mozilla Firefox

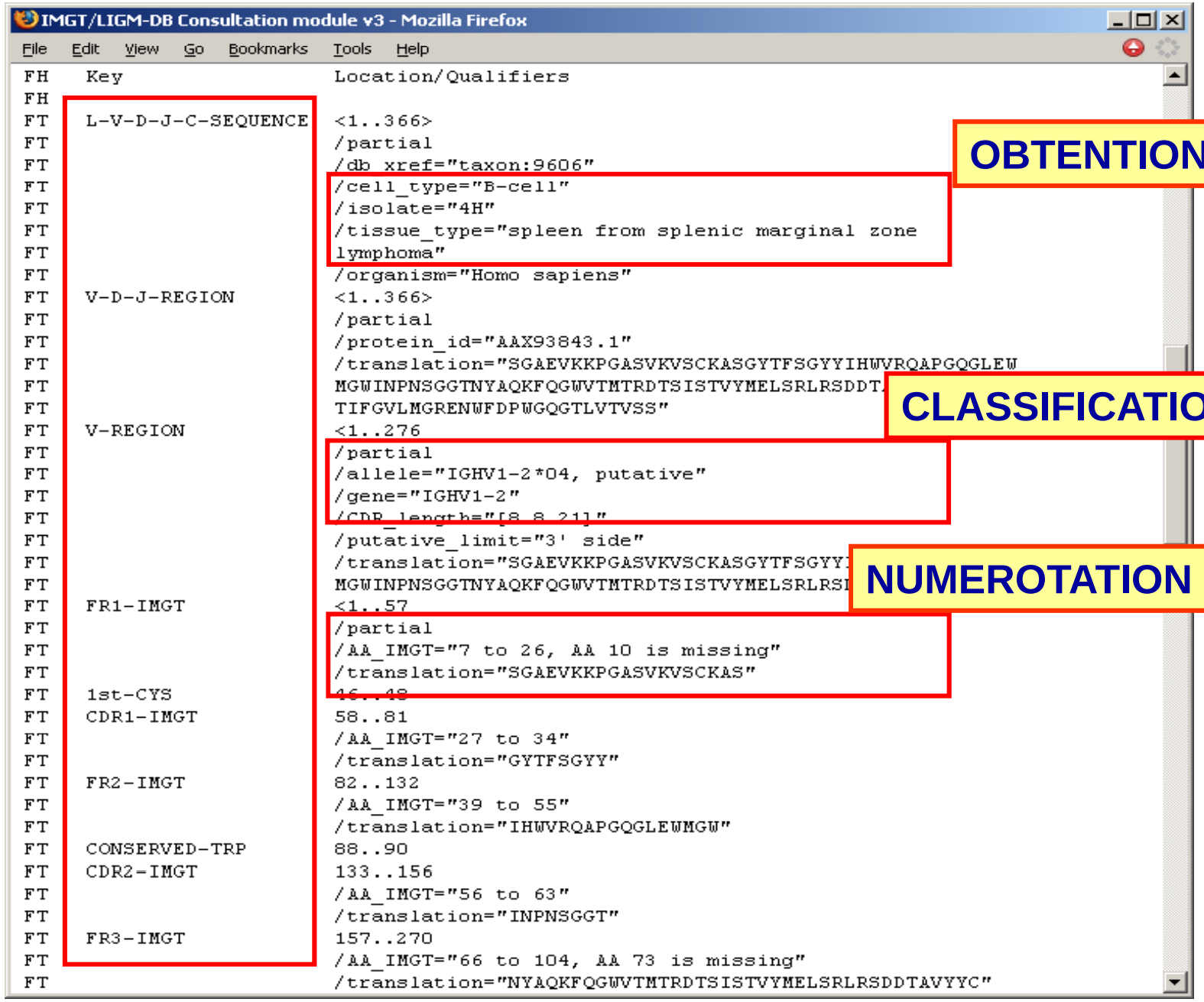
File Edit View Go Bookmarks Tools Help

```
ID   AY998750 IMGT/LIGM annotation : automatic; mRNA; HUM; 366 BP.
XX
AC   AY998750;
XX
DT   25-MAY-2005 (Rel. 200521-3, arrived in LIGM-DB )
DT   25-AUG-2005 (Rel. 200534-4, Last updated, Version 3)
XX
DE   Homo sapiens isolate 4H immunoglobulin heavy chain variable region (IGHV1)
DE   mRNA, IGHV1-2*04 allele, partial cds.
DE   mRNA; rearranged configuration; Ig-Heavy; regular; functionality
DE   productive; group IGHV; subgroup HV1.
XX
KW   antigen receptor; immunoglobulin superfamily; Ig; Ig-Heavy; variable;
KW   diversity; joining; immunoglobulin.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;
OC   Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-366
RX   PUBMED; 15706403.
RA   Stamatopoulos K., Belessi C., Papadaki T., Kalagiakou E., Stavroyianni N.,
RA   Douka V., Afendaki S., Saloum R., Parasi A., Anagnostou D., Laoutaris N.,
RA   Fassas A., Anagnostopoulos A.;
RT   "Immunoglobulin Heavy- And Light-chain Repertoire in Splenic Marginal Zone
RT   Lymphoma";
RL   (er) Mol. Med. (2005) In press
XX
RN   [2]
RP   1-366
RA   Stamatopoulos K., Belessi C., Papadaki T., Kalagiakou E., Stavroyianni N.,
RA   Douka V., Afendaki S., Saloum R., Parasi A., Anagnostou D., Laoutaris N.,
RA   Fassas A., Anagnostopoulos A.;
RT   ;
RL   Submitted (30-MAR-2005) to the EMBL/GenBank/DDBJ databases.
RL   Hematology Department and HCT Unit, Papanikolaou General Hospital,
RL   Thessaloniki 57010, Greece
XX
DR   EMBL; AY998750.
XX
FH   Key                 Location/Qualifiers
FH
FT   L-V-D-J-C-SEQUENCE  <1..366>
```

Im Muno Gene Tics Information system®

http://www.imgt.org

Annotation and IMGT-ONTOLOGY concepts



Im Muno Gene Tics Information system®

http://www.imgt.org

IMGT-ONTOLOGY

http://www.imgt.org

IMGT Scientific chart :

IMGT-ONTOLOGY en *langage naturel* pour les biologistes

http://imgt.cines.fr/textes/IMGTScientificChart/

En cours : édition en langage OWL

Version v1.0.0 sur le site BioPortal

(http://bioportal.bioontology.org/visualize/42685)

Les atouts de IMGT-ONTOLOGY

http://www.imgt.org

1 - Pour les utilisateurs:

- Définir un vocabulaire précis et spécifique du domaine qui permette de décrire en détail toutes les caractéristiques des IG et des TR
- Proposer ce vocabulaire standardisé comme critères de sélection dans la base

Les atouts de IMGT-ONTOLOGY

http://www.imgt.org

2 - Pour les annotateurs du laboratoire:

- Faciliter la communication au sein de l'équipe
- Faciliter la formation du nouveau personnel
- Etablir les règles d'annotation qui expriment les dépendances entre les termes du vocabulaire

Les atouts de IMGT-ONTOLOGY

http://www.imgt.org

3 - Automatisation des procédures:

- Alléger le travail des annotateurs
- Développer des outil d'annotation

Les atouts de IMGT-ONTOLOGY

http://www.imgt.org

4 – Intégration des nouvelles connaissances en immunogénétique

- Modifier les règles existantes
- Appliquer les modifications aux données préexistantes

Im
Muno
Gene
Tics
Information
system®
http://www.imgt.org