ImMunoGeneTics Information system®

http://www.imgt.org

Systèmes d’information

-

Ontologies
Bases de connaissances

Bioinformatique (TC3 du MR2 BioMed)

Véronique Giudicelli
LIGM, IGH, UPR CNRS 1142
141 rue de la Cardonille
34296 Montpellier Cedex 5
tel: 04.34.35.99.28 fax: 04.34.35.99.01
e-mail: Veronique.Giudicelli@igh.cnrs.fr

02 octobre 2013

Système d’information

Ensemble des moyens

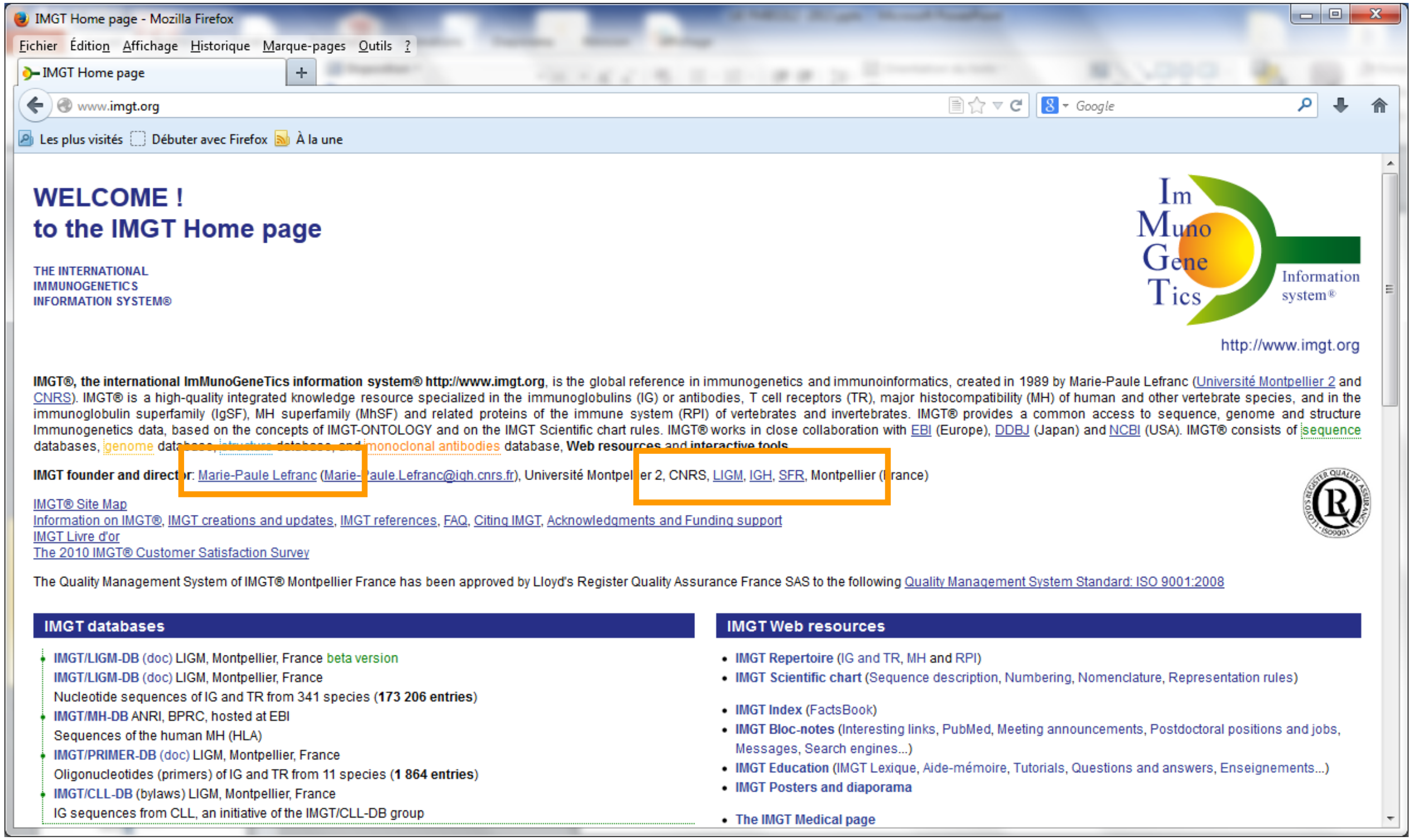
organisation (CNRS,…),
acteurs,
systèmes informatiques

nécessaires au traitement et à l’exploitation des informations dans le cadre d’objectifs définis

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

Source unique de connaissances en immunogénétique et immunoinformatique, au niveau international.

IMGT® est spécialisé dans les séquences, structures et données génétiques des immunoglobulines (IG), des récepteurs T (TR), du complexe majeur d'histocompatibilité (MHC) des vertébrés, des protéines des superfamilles IgSF et MhcSF, et des protéines apparentées du système immunitaire.



WELCOME !
to the IMGT Home page

THE INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

http://www.imgt.org

IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org, is the global reference in immunogenetics and immunoinformatics, created in 1989 by Marie-Paule Lefranc (Université Montpellier 2 and CNRS). IMGT® is a high-quality integrated knowledge resource specialized in the immunoglobulins (IG) or antibodies, T cell receptors (TR), major histocompatibility (MH) of human and other vertebrate species, and in the immunoglobulin superfamily (IgSF), MH superfamily (MhSF) and related proteins of the immune system (RPI) of vertebrates and invertebrates. IMGT® provides a common access to sequence, genome and structure Immunogenetics data, based on the concepts of IMGT-ONTOLOGY and on the IMGT Scientific chart rules. IMGT® works in close collaboration with EBI (Europe), DDBJ (Japan) and NCBI (USA). IMGT® consists of sequence databases, genome database, structure database, and monoclonal antibodies database, **Web resources and interactive tools**.

IMGT founder and direct r: Marie-Paule Lefranc (Marie-Paule.Lefranc@igh.cnrs.fr), Université Montpel er 2, CNRS, LIGM, IGH, SFR, Montpellier (France)

IMGT® Site Map
Information on IMGT®, IMGT creations and updates, IMGT references, FAQ, Citing IMGT, Acknowledgments and Funding support
IMGT Livre d'or
The 2010 IMGT® Customer Satisfaction Survey

The Quality Management System of IMGT® Montpellier France has been approved by Lloyd's Register Quality Assurance France SAS to the following Quality Management System Standard: ISO 9001:2008

IMGT databases

- **IMGT/LIGM-DB** (doc) LIGM, Montpellier, France beta version
- **IMGT/LIGM-DB** (doc) LIGM, Montpellier, France
  Nucleotide sequences of IG and TR from 341 species (**173 206 entries**)
- **IMGT/MH-DB** ANRI, BPRC, hosted at EBI
  Sequences of the human MH (HLA)
- **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 Web resources

- **IMGT Repertoire** (IG and TR, MH 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**

Système informatique (SI)

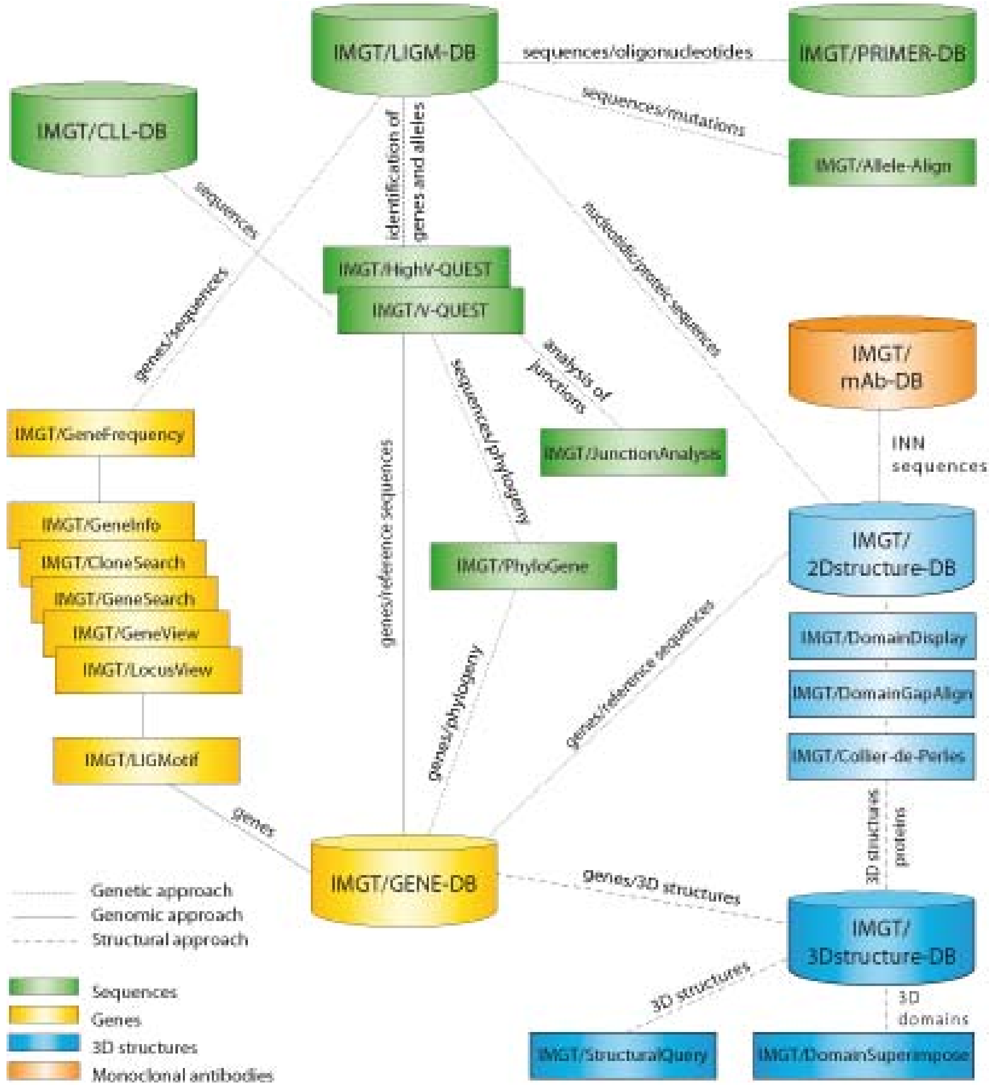
Partie automatisée d'un système d'information.

Le **système informatique** regroupe :

- Le matériel physique et câblage (PC, réseau,….)
- Le réseau (adresse IP, noms, …)
- Le(s) système(s) d'exploitation (Windows, Linux, …)
- Sauvegardes

- **Les applications :**
  bases de données,
  logiciels de développement,
  interfaces d'exploitation, …

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

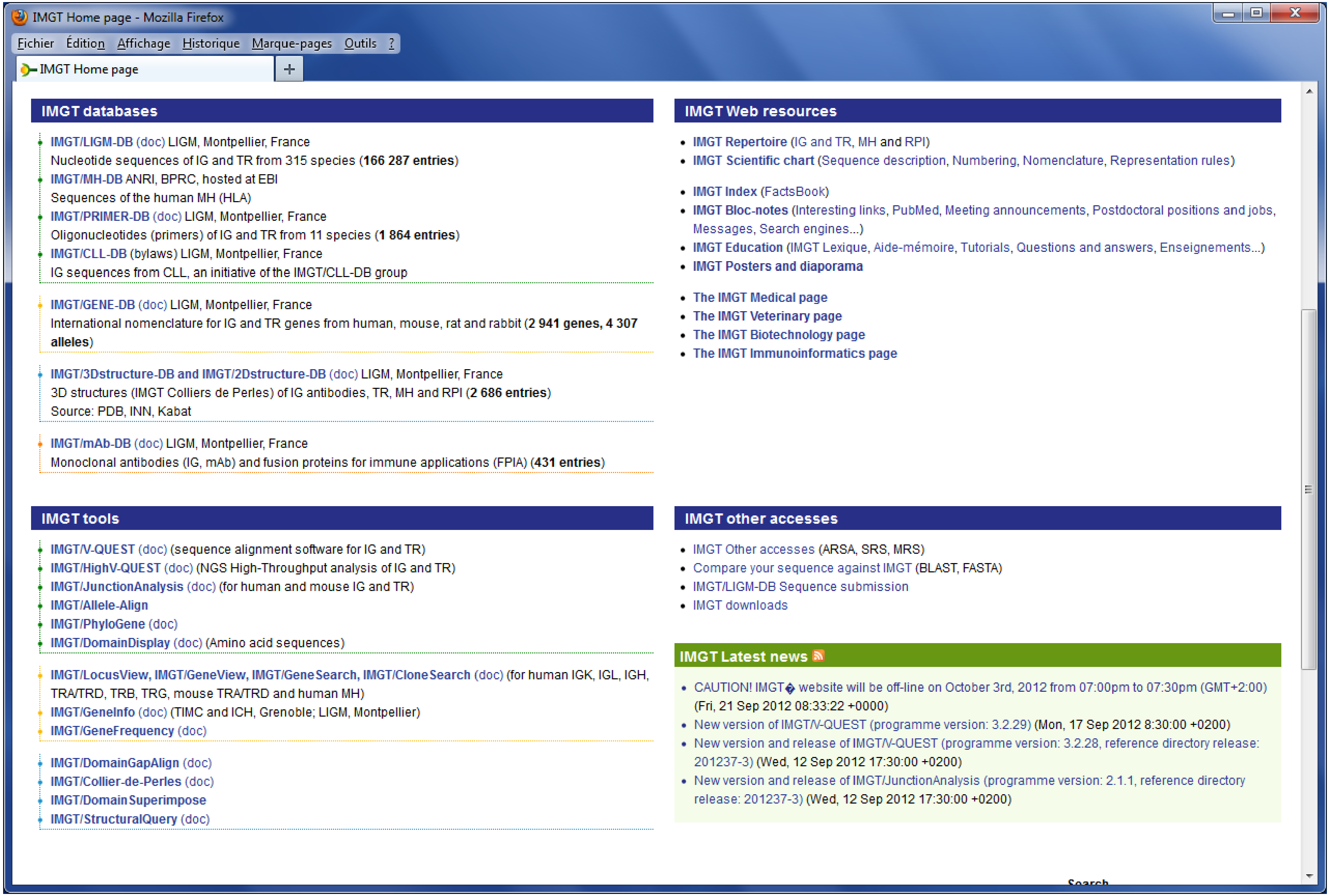


A unique integrated knowledge resource for :

- **IG**
- **TR**
- **MH**
- **IG superfamily (IgSF)**
- **MH superfamily (MhSF)**
- **related proteins of the immune system (RPI)**

7 Databases
15 online tools

Based on IMGT-ONTOLOGY, the first ontology for immunogenetics, and on the IMGT Scientific chart rules



IMGT databases

- IMGT/LIGM-DB (doc) LIGM, Montpellier, France
  Nucleotide sequences of IG and TR from 315 species (**166 287 entries**)
- IMGT/MH-DB ANRI, BPRC, hosted at EBI
  Sequences of the human MH (HLA)
- 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 (**2 941 genes, 4 307 alleles**)

- IMGT/3Dstructure-DB and IMGT/2Dstructure-DB (doc) LIGM, Montpellier, France
  3D structures (IMGT Colliers de Perles) of IG antibodies, TR, MH and RPI (**2 686 entries**)
  Source: PDB, INN, Kabat

- IMGT/mAb-DB (doc) LIGM, Montpellier, France
  Monoclonal antibodies (IG, mAb) and fusion proteins for immune applications (FPIA) (**431 entries**)

IMGT Web resources

- **IMGT Repertoire** (IG and TR, MH 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 tools

- **IMGT/V-QUEST** (doc) (sequence alignment software for IG and TR)
- **IMGT/HighV-QUEST** (doc) (NGS High-Throughput analysis of IG and TR)
- **IMGT/JunctionAnalysis** (doc) (for human and mouse 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 MH)
- **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 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

- CAUTION! IMGT� website will be off-line on October 3rd, 2012 from 07:00pm to 07:30pm (GMT+2:00) (Fri, 21 Sep 2012 08:33:22 +0000)
- New version of IMGT/V-QUEST (programme version: 3.2.29) (Mon, 17 Sep 2012 8:30:00 +0200)
- New version and release of IMGT/V-QUEST (programme version: 3.2.28, reference directory release: 201237-3) (Wed, 12 Sep 2012 17:30:00 +0200)
- New version and release of IMGT/JunctionAnalysis (programme version: 2.1.1, reference directory release: 201237-3) (Wed, 12 Sep 2012 17:30:00 +0200)

Search

Une application informatique distribuée

utilise plusieurs machines

L’architecture d’une application distribuée est donnée par:

- la description des machines utilisées,
- les fonctions de ces machines pour l’application:
  - fonction stockage des données
  - fonction traitements des données
  - fonction interface utilisateur

IMGT databases

- **IMGT/LIGM-DB** (doc) LIGM, Montpellier, France
  Nucleotide sequences of IG and TR from 315 species (**166 287 entries**)
- **IMGT/MH-DB** ANRI, BPRC, hosted at EBI
  Sequences of the human MH (HLA)
- **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 (**2 941 genes, 4 307 alleles**)

- **IMGT/3Dstructure-DB and IMGT/2Dstructure-DB** (doc) LIGM, Montpellier, France
  3D structures (IMGT Colliers de Perles) of IG antibodies, TR, MH and RPI (**2 686 entries**)
  Source: PDB, INN, Kabat

- **IMGT/mAb-DB** (doc) LIGM, Montpellier, France
  Monoclonal antibodies (IG, mAb) and fusion proteins for immune applications (FPIA) (**431 entries**)

IMGT Web resources

- **IMGT Repertoire** (IG and TR, MH 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 tools

- **IMGT/V-QUEST** (doc) (sequence alignment software for IG and TR)
- **IMGT/HighV-QUEST** (doc) (NGS High-Throughput analysis of IG and TR)
- **IMGT/JunctionAnalysis** (doc) (for human and mouse 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 MH)
- **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 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

- CAUTION! IMGT� website will be off-line on October 3rd, 2012 from 07:00pm to 07:30pm (GMT+2:00) (Fri, 21 Sep 2012 08:33:22 +0000)
- New version of IMGT/V-QUEST (programme version: 3.2.29) (Mon, 17 Sep 2012 8:30:00 +0200)
- New version and release of IMGT/V-QUEST (programme version: 3.2.28, reference directory release: 201237-3) (Wed, 12 Sep 2012 17:30:00 +0200)
- New version and release of IMGT/JunctionAnalysis (programme version: 2.1.1, reference directory release: 201237-3) (Wed, 12 Sep 2012 17:30:00 +0200)

Les objectifs de IMGT/LIGM-DB

http://www.imgt.org

- Contenir toutes les séquences IG et TR des bases EMBL/GenBank/DDBJ

  + de 173.000 séquences de 341 espèces

- Expertiser les séquences en fonction des connaissances en immunogénétique

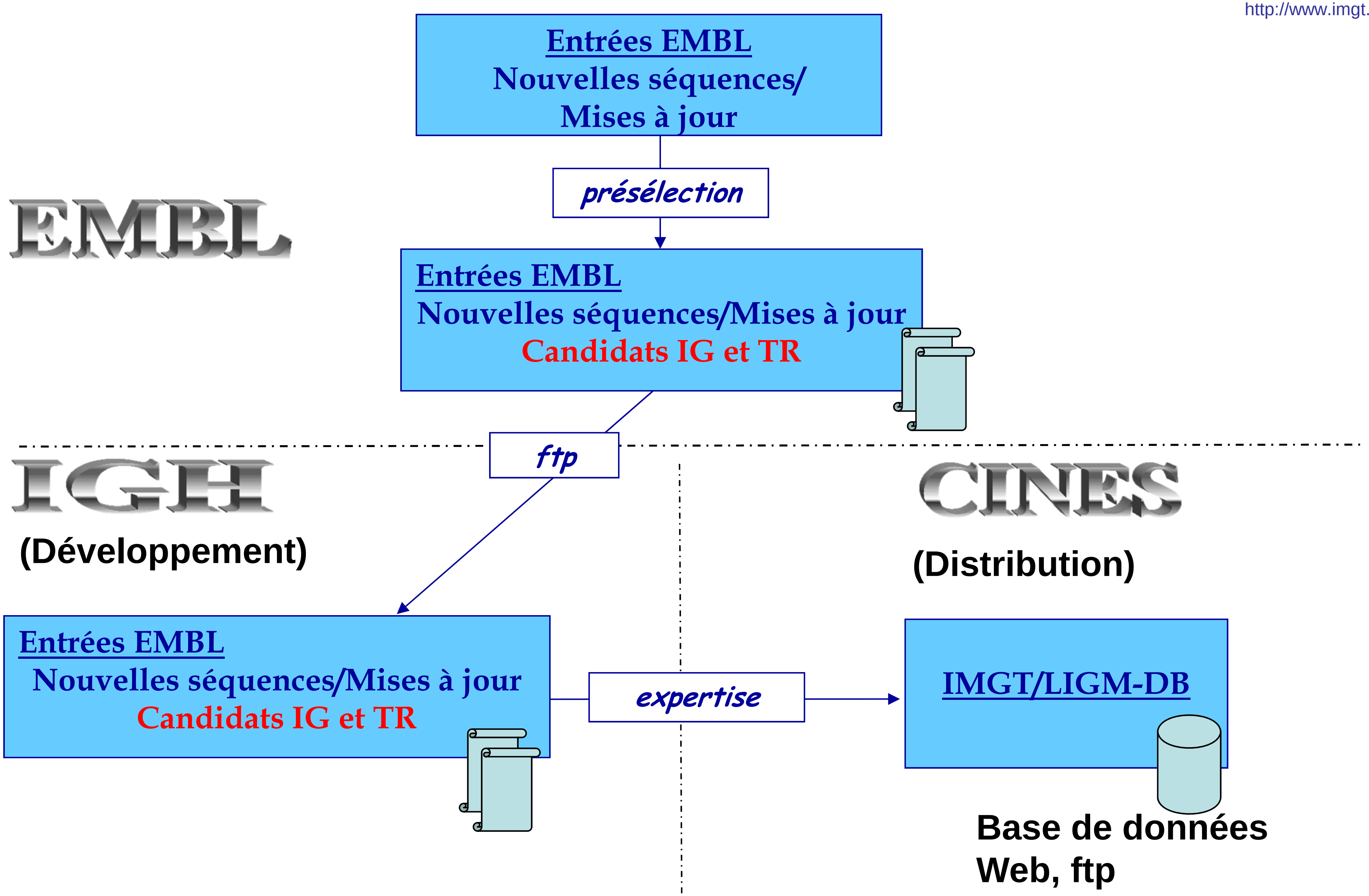
  - identification des gènes,
  - description des motifs (régions codantes, …)

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

IMGT/LIGM-DB: application distribuée

Im Muno Gene Tics Information system®

http://www.imgt.org



EMBL

Entrées EMBL
Nouvelles séquences/
Mises à jour

présélection

Entrées EMBL
Nouvelles séquences/Mises à jour
Candidats IG et TR

ftp

IGH

(Développement)

Entrées EMBL
Nouvelles séquences/Mises à jour
Candidats IG et TR

expertise

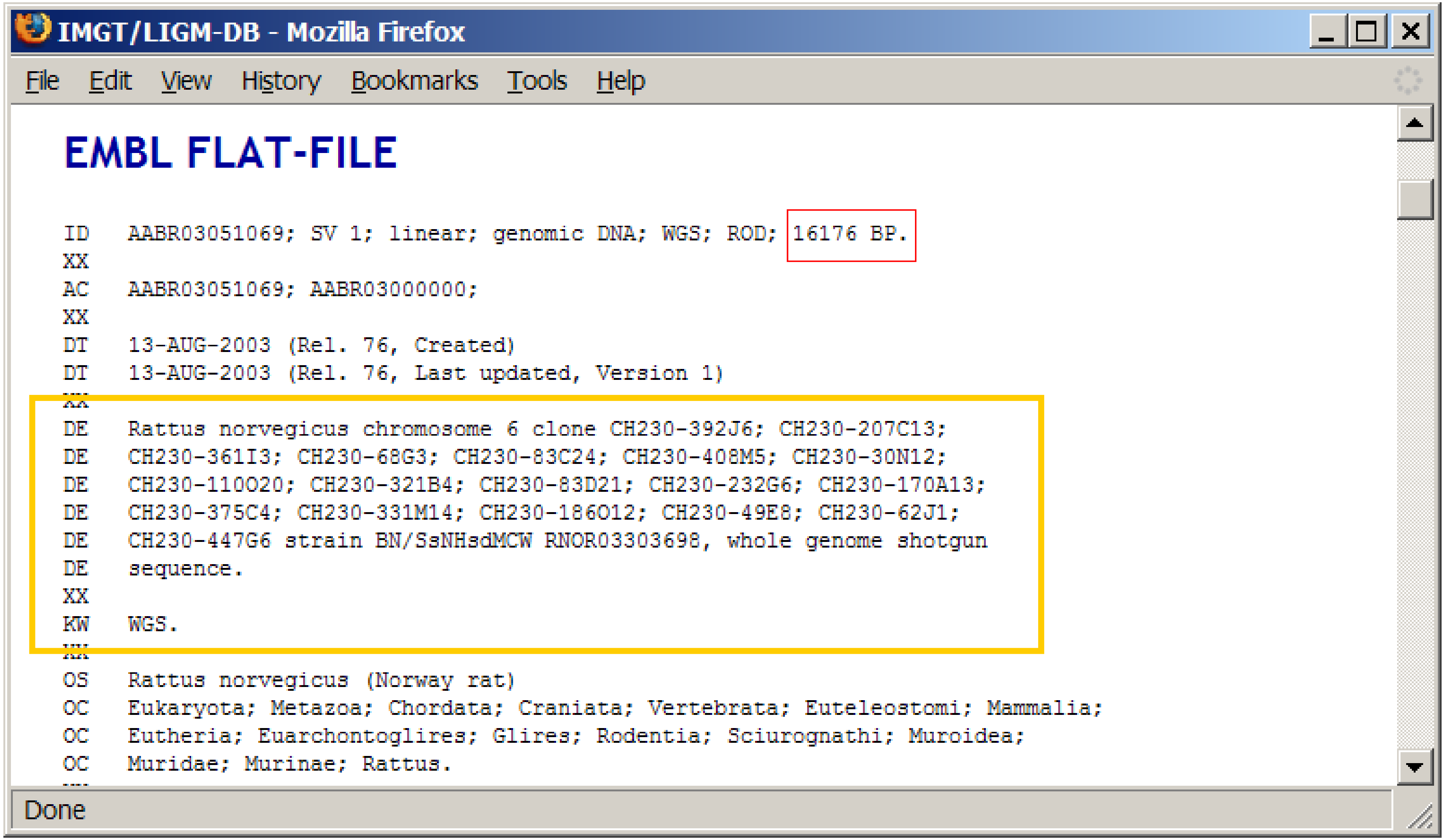
CINES

(Distribution)

IMGT/LIGM-DB

Base de données
Web, ftp

A l’arrivée dans IMGT®

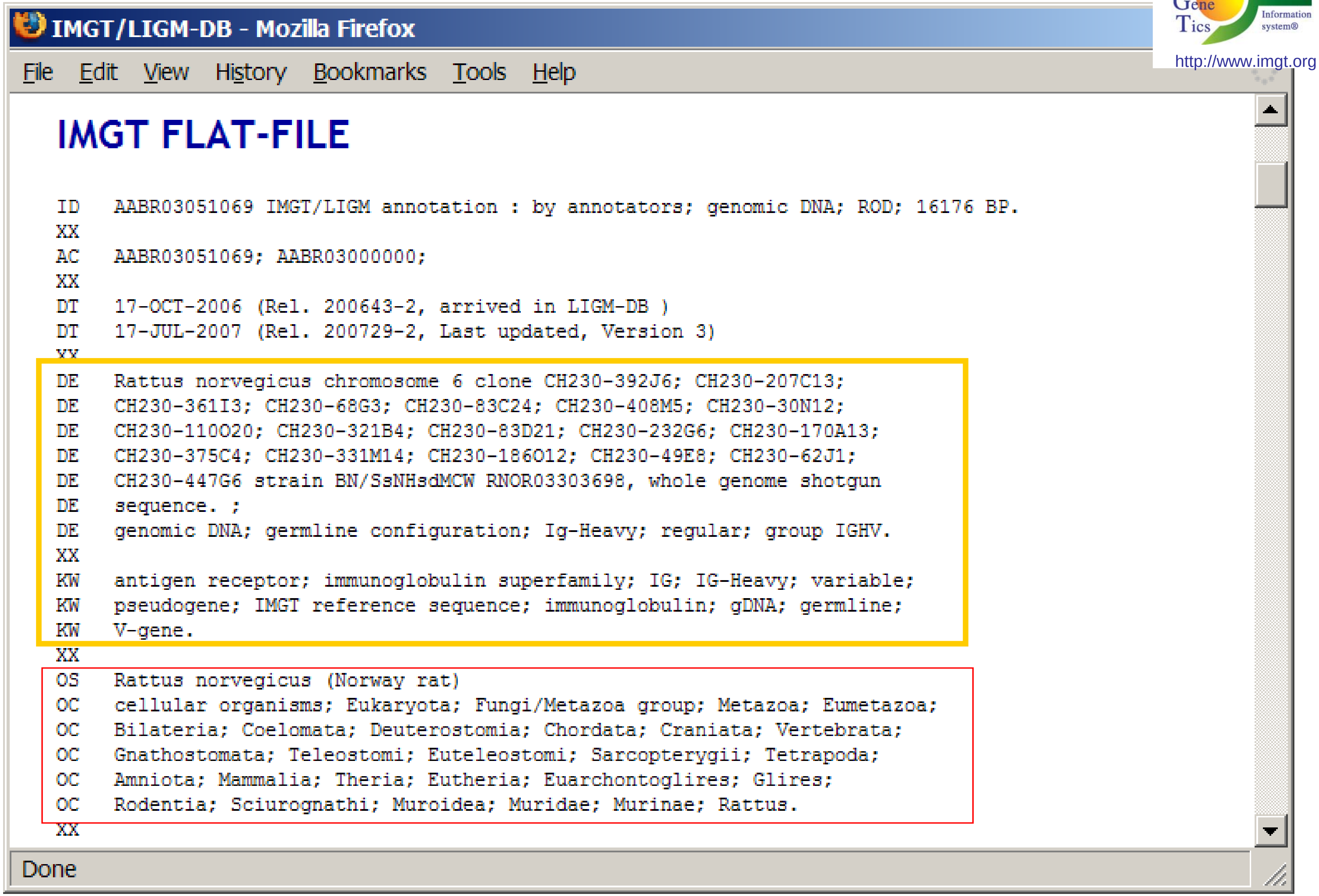


EMBL FLAT-FILE

```
ID   AABR03051069; SV 1; linear; genomic DNA; WGS; ROD; 16176 BP.
XX
AC   AABR03051069; AABR03000000;
XX
DT   13-AUG-2003 (Rel. 76, Created)
DT   13-AUG-2003 (Rel. 76, Last updated, Version 1)
XX
DE   Rattus norvegicus chromosome 6 clone CH230-392J6; CH230-207C13;
DE   CH230-361I3; CH230-68G3; CH230-83C24; CH230-408M5; CH230-30N12;
DE   CH230-110O20; CH230-321B4; CH230-83D21; CH230-232G6; CH230-170A13;
DE   CH230-375C4; CH230-331M14; CH230-186O12; CH230-49E8; CH230-62J1;
DE   CH230-447G6 strain BN/SsNHsdMCW RNOR03303698, whole genome shotgun
DE   sequence.
XX
KW   WGS.
XX
OS   Rattus norvegicus (Norway rat)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea;
OC   Muridae; Murinae; Rattus.
```

Après expertise dans IMGT®

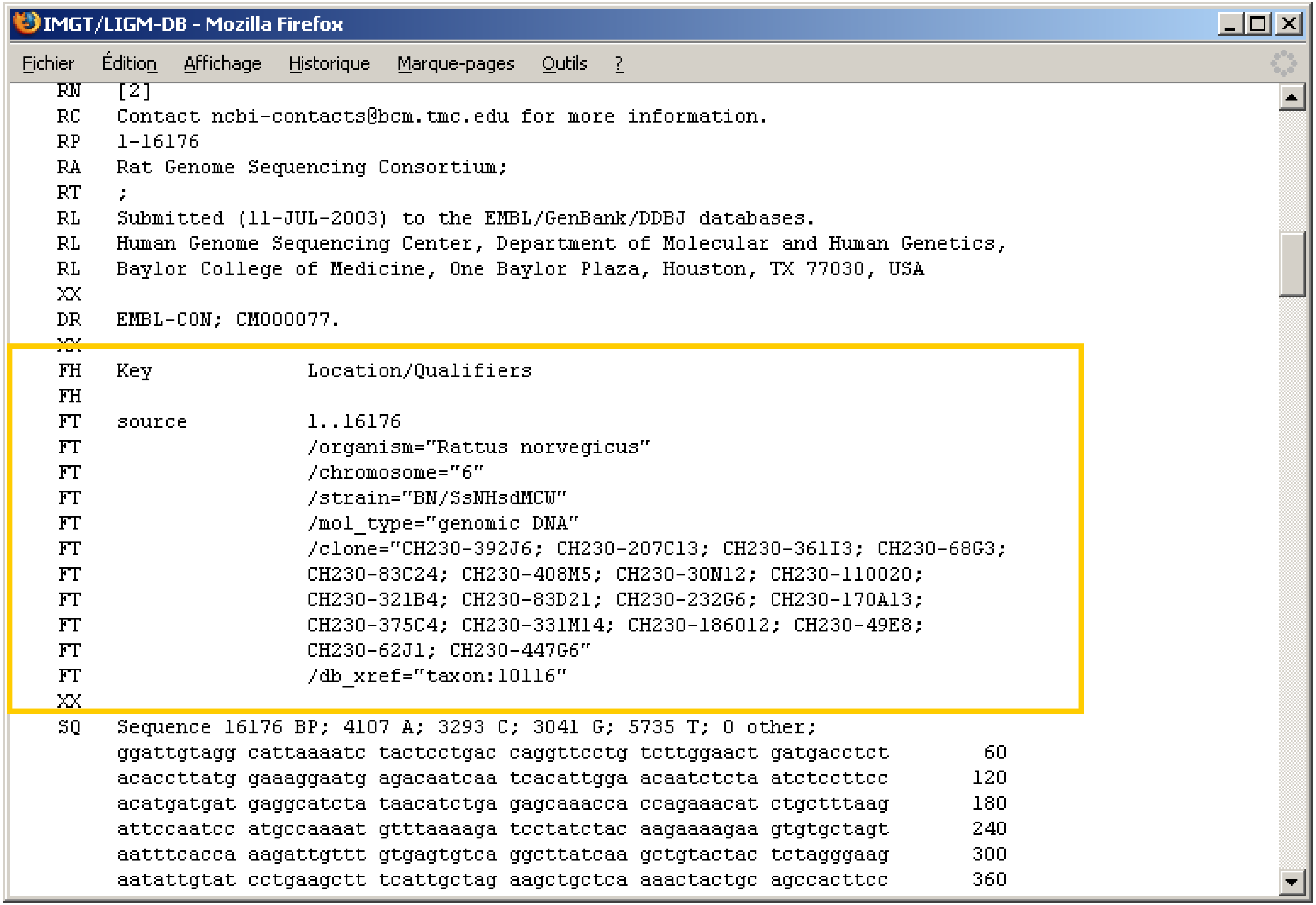
http://www.imgt.org



IMGT FLAT-FILE

```
ID   AABR03051069 IMGT/LIGM annotation : by annotators; genomic DNA; ROD; 16176 BP.
XX
AC   AABR03051069; AABR03000000;
XX
DT   17-OCT-2006 (Rel. 200643-2, arrived in LIGM-DB )
DT   17-JUL-2007 (Rel. 200729-2, Last updated, Version 3)
XX
DE   Rattus norvegicus chromosome 6 clone CH230-392J6; CH230-207C13;
DE   CH230-361I3; CH230-68G3; CH230-83C24; CH230-408M5; CH230-30N12;
DE   CH230-110O20; CH230-321B4; CH230-83D21; CH230-232G6; CH230-170A13;
DE   CH230-375C4; CH230-331M14; CH230-186O12; CH230-49E8; CH230-62J1;
DE   CH230-447G6 strain BN/SsNHsdMCW RNOR03303698, whole genome shotgun
DE   sequence. ;
DE   genomic DNA; germline configuration; Ig-Heavy; regular; group IGHV.
XX
KW   antigen receptor; immunoglobulin superfamily; IG; IG-Heavy; variable;
KW   pseudogene; IMGT reference sequence; immunoglobulin; gDNA; germline;
KW   V-gene.
XX
OS   Rattus norvegicus (Norway rat)
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; Glires;
OC   Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
XX
```

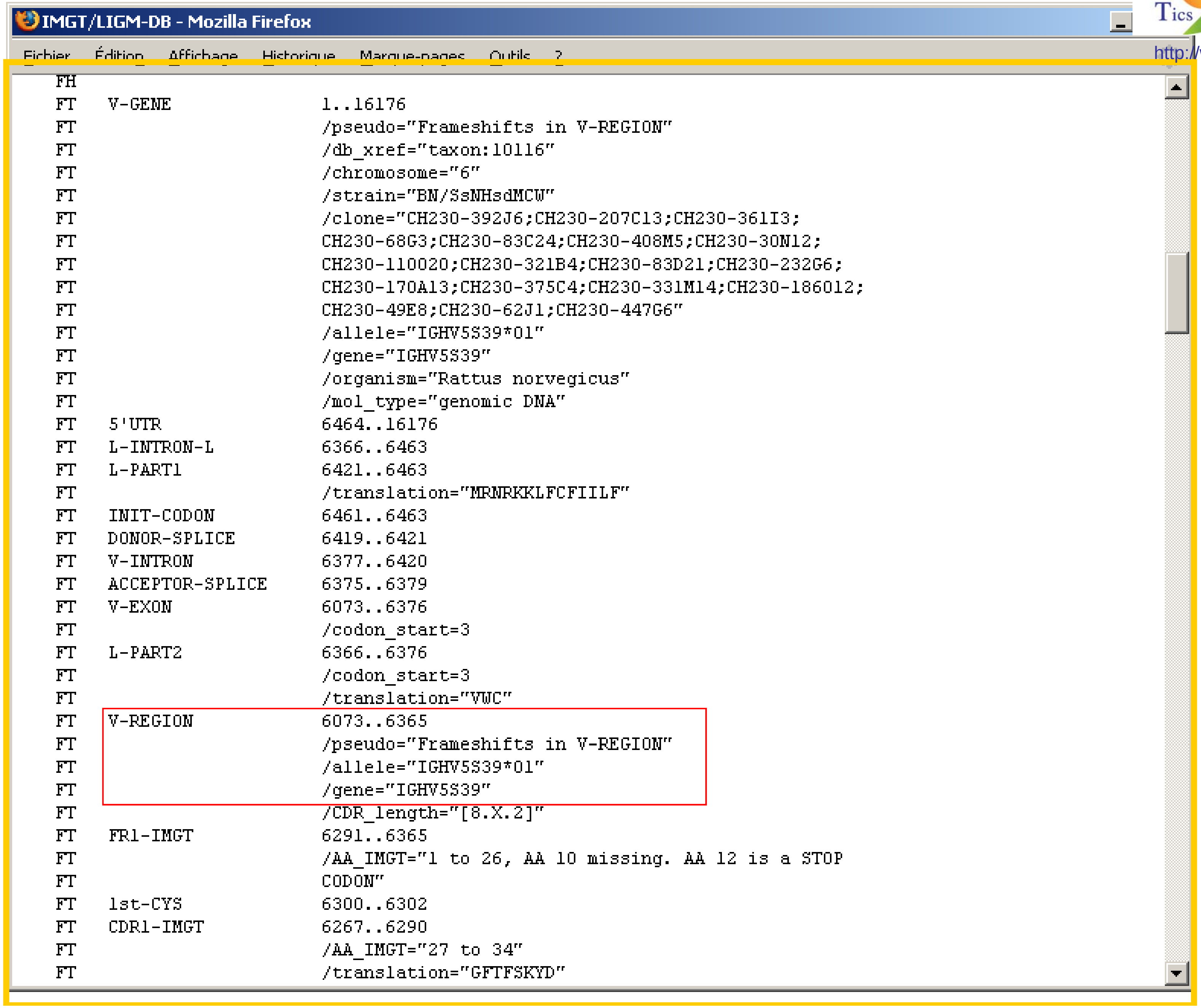
A l’arrivée dans IMGT®



```
RN   [2]
RC   Contact ncbi-contacts@bcm.tmc.edu for more information.
RP   1-16176
RA   Rat Genome Sequencing Consortium;
RT   ;
RL   Submitted (11-JUL-2003) to the EMBL/GenBank/DDBJ databases.
RL   Human Genome Sequencing Center, Department of Molecular and Human Genetics,
RL   Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
XX
DR   EMBL-CON; CM000077.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..16176
FT                   /organism="Rattus norvegicus"
FT                   /chromosome="6"
FT                   /strain="BN/SsNHsdMCW"
FT                   /mol_type="genomic DNA"
FT                   /clone="CH230-392J6; CH230-207C13; CH230-361I3; CH230-68G3;
FT                   CH230-83C24; CH230-408M5; CH230-30N12; CH230-110020;
FT                   CH230-321B4; CH230-83D21; CH230-232G6; CH230-170A13;
FT                   CH230-375C4; CH230-331M14; CH230-186012; CH230-49E8;
FT                   CH230-62J1; CH230-447G6"
FT                   /db_xref="taxon:10116"
XX
SQ   Sequence 16176 BP; 4107 A; 3293 C; 3041 G; 5735 T; 0 other;
     ggattgtagg cattaaaatc tactcctgac caggttcctg tcttggaact gatgacctct        60
     acaccttatg gaaaggaatg agacaatcaa tcacattgga acaatctcta atctccttcc       120
     acatgatgat gaggcatcta taacatctga gagcaaacca ccagaaacat ctgctttaag       180
     attccaatcc atgccaaaat gtttaaaaga tcctatctac aagaaaagaa gtgtgctagt       240
     aatttcacca aagattgttt gtgagtgtca ggcttatcaa gctgtactac tctagggaag       300
     aatattgtat cctgaagctt tcattgctag aagctgctca aaactactgc agccacttcc       360
```

Après expertise dans IMGT®

http://www.imgt.org



```
FH
FT   V-GENE              1..16176
FT                       /pseudo="Frameshifts in V-REGION"
FT                       /db_xref="taxon:10116"
FT                       /chromosome="6"
FT                       /strain="BN/SsNHsdMCW"
FT                       /clone="CH230-392J6;CH230-207C13;CH230-361I3;
FT                       CH230-68G3;CH230-83C24;CH230-408M5;CH230-30N12;
FT                       CH230-110020;CH230-321B4;CH230-83D21;CH230-232G6;
FT                       CH230-170A13;CH230-375C4;CH230-331M14;CH230-186012;
FT                       CH230-49E8;CH230-62J1;CH230-447G6"
FT                       /allele="IGHV5S39*01"
FT                       /gene="IGHV5S39"
FT                       /organism="Rattus norvegicus"
FT                       /mol_type="genomic DNA"
FT   5'UTR               6464..16176
FT   L-INTRON-L          6366..6463
FT   L-PART1             6421..6463
FT                       /translation="MRNRKKLFCFIILF"
FT   INIT-CODON          6461..6463
FT   DONOR-SPLICE        6419..6421
FT   V-INTRON            6377..6420
FT   ACCEPTOR-SPLICE     6375..6379
FT   V-EXON              6073..6376
FT                       /codon_start=3
FT   L-PART2             6366..6376
FT                       /codon_start=3
FT                       /translation="VWC"
FT   V-REGION            6073..6365
FT                       /pseudo="Frameshifts in V-REGION"
FT                       /allele="IGHV5S39*01"
FT                       /gene="IGHV5S39"
FT                       /CDR_length="[8.X.2]"
FT   FR1-IMGT            6291..6365
FT                       /AA_IMGT="1 to 26, AA 10 missing. AA 12 is a STOP
FT                       CODON"
FT   1st-CYS             6300..6302
FT   CDR1-IMGT           6267..6290
FT                       /AA_IMGT="27 to 34"
FT                       /translation="GFTFSKYD"
```

http://www.imgt.org

Administration et contrôle de données dans un système d'information: exemple de IMGT/LIGM-DB

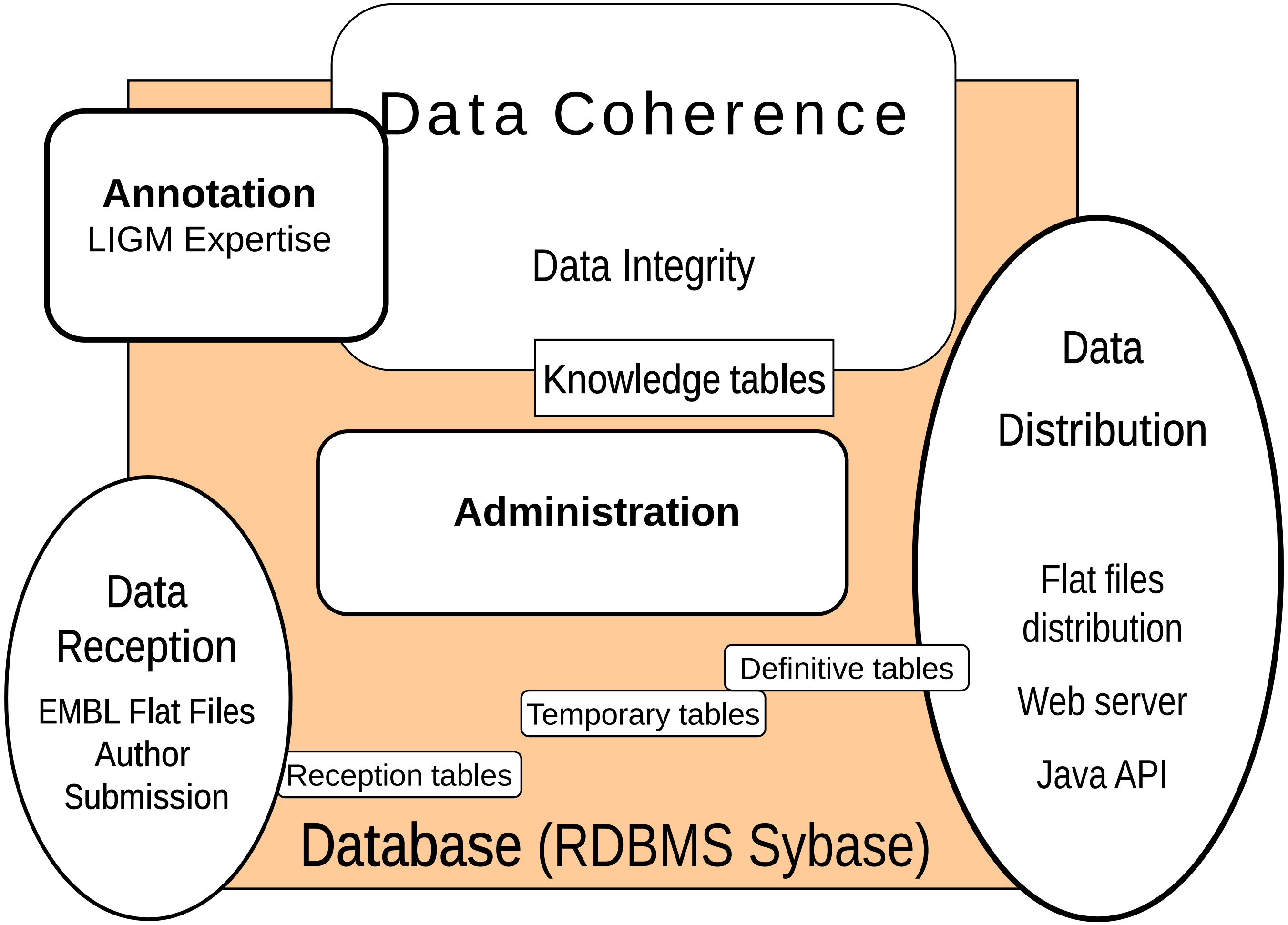
1- Coordination des flux des données
de différentes sources dans la base: comment gérer
l’entrée des séquences (700 arrivées par semaine) ,
des annotations et leur distribution?

2- Gestion des accès aux données: qui annote les séquences,
quelles informations peuvent être consultées, extraites
et modifiées, et par qui?

1- Coordination des flux de données

Information system®

http://www.imgt.org



Data Coherence

Data Integrity

Knowledge tables

Annotation
LIGM Expertise

Administration

Data Distribution

Flat files distribution

Web server

Java API

Definitive tables

Temporary tables

Reception tables

Data Reception

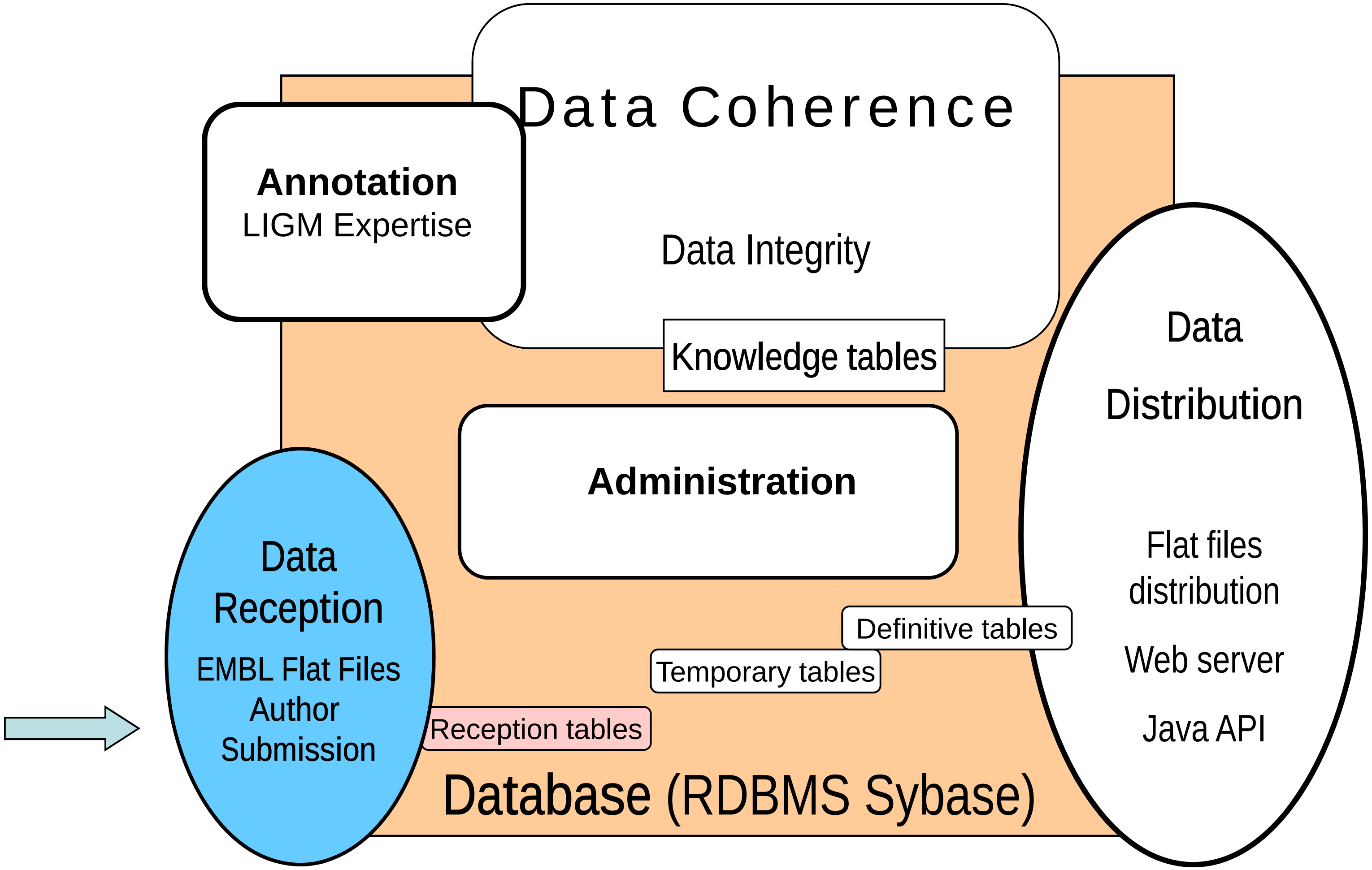
EMBL Flat Files

Author Submission

Database (RDBMS Sybase)

1- Coordination des flux de données

http://www.imgt.org



Data Coherence

Data Integrity

Knowledge tables

Annotation
LIGM Expertise

Administration

Data
Reception
EMBL Flat Files
Author
Submission

Reception tables

Temporary tables

Definitive tables

Data
Distribution

Flat files
distribution

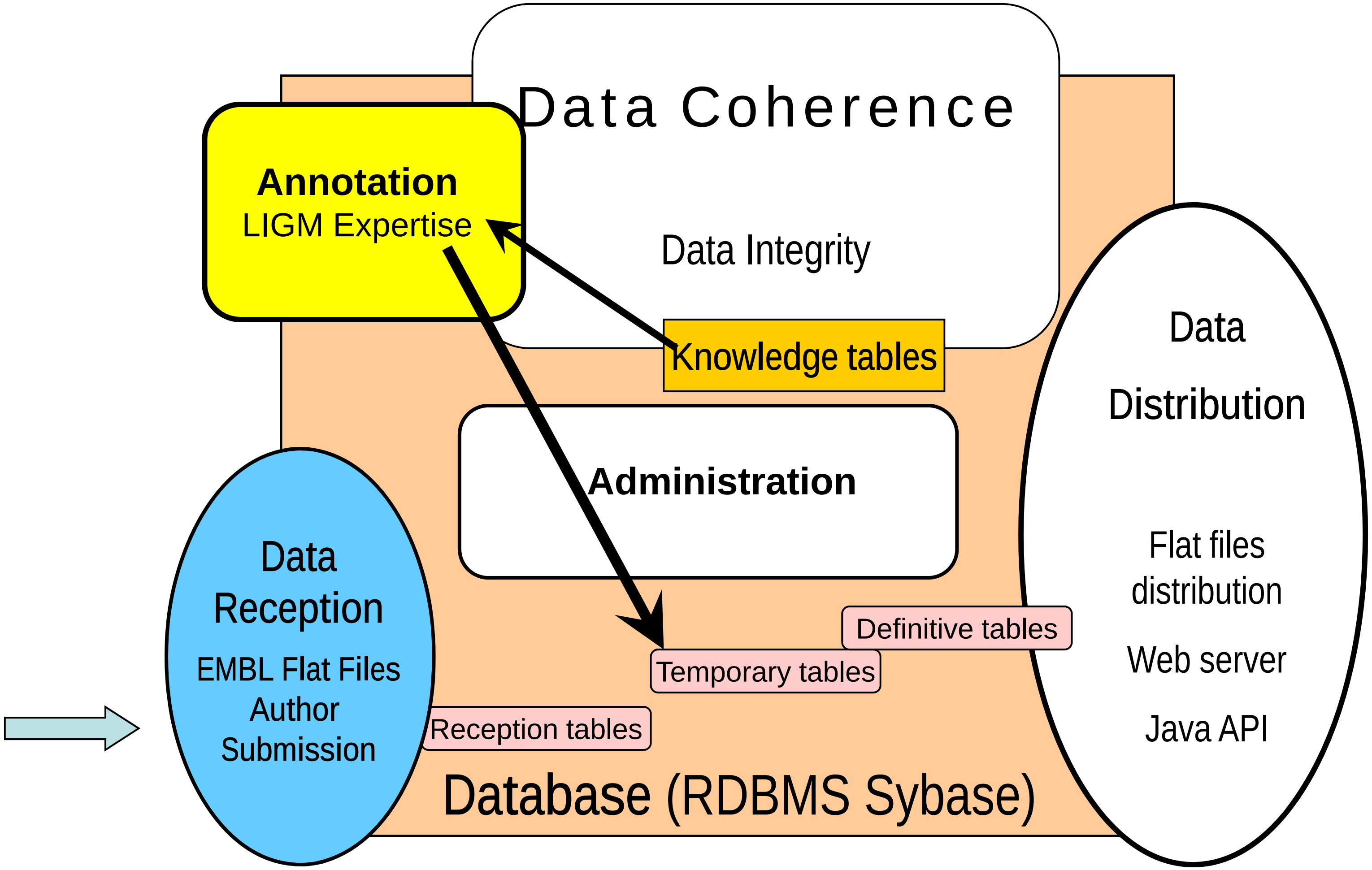
Web server

Java API

Database (RDBMS Sybase)

1- Coordination des flux de données

http://www.imgt.org



Data Coherence

Annotation
LIGM Expertise

Data Integrity

Knowledge tables

Administration

Data
Distribution

Flat files
distribution

Web server

Java API

Data
Reception

EMBL Flat Files
Author
Submission

Definitive tables

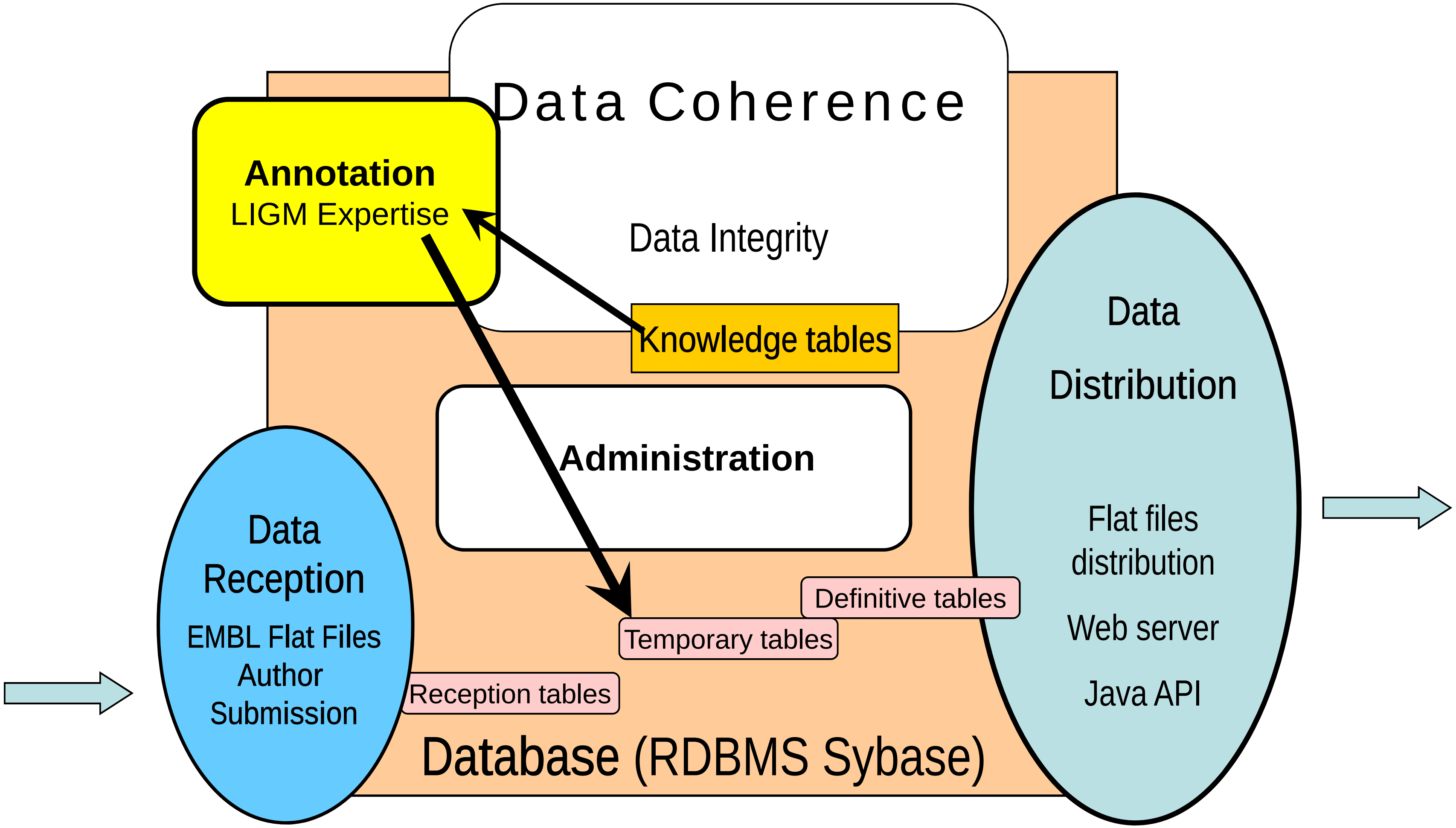
Temporary tables

Reception tables

Database (RDBMS Sybase)

1- Coordination des flux de données

http://www.imgt.org



Data Coherence

Annotation
LIGM Expertise

Data Integrity

Knowledge tables

Administration

Data
Distribution

Flat files
distribution

Web server

Java API

Data
Reception

EMBL Flat Files
Author
Submission

Definitive tables

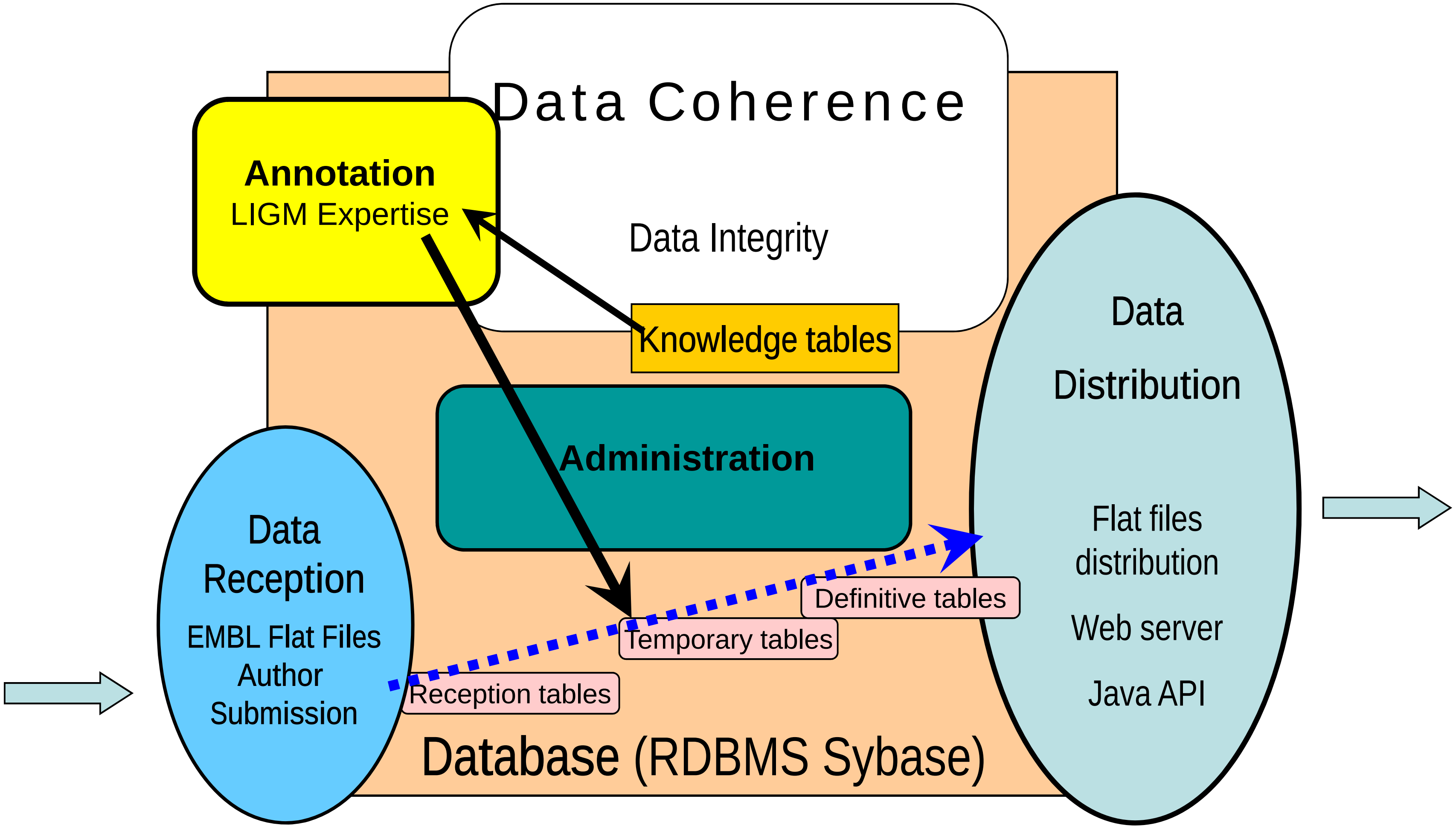
Temporary tables

Reception tables

Database (RDBMS Sybase)

1- Coordination des flux de données

http://www.imgt.org



Data Coherence

Data Integrity

Annotation
LIGM Expertise

Knowledge tables

Administration

Data
Distribution

Flat files
distribution

Web server

Java API

Data
Reception

EMBL Flat Files
Author
Submission

Reception tables

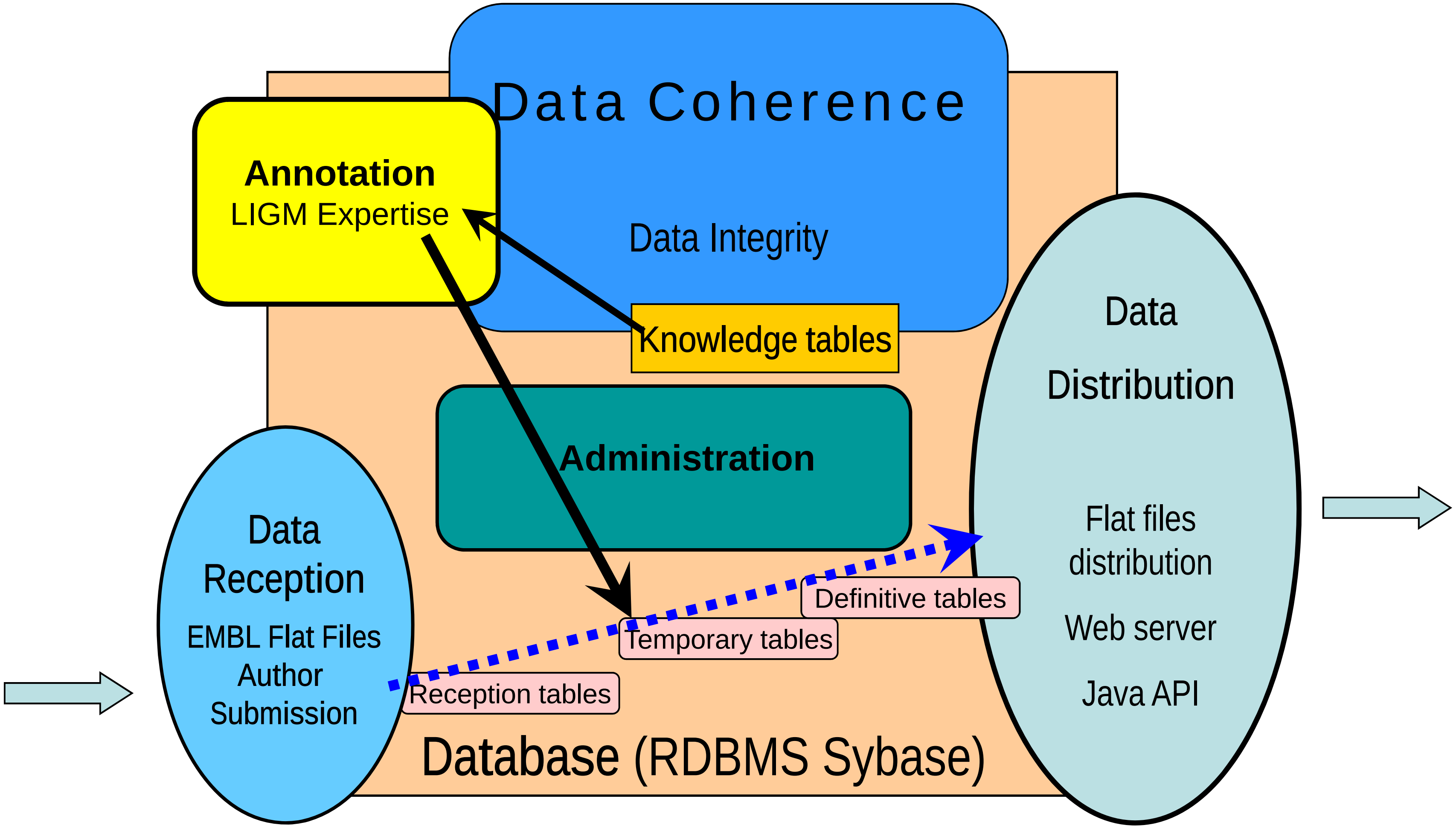
Temporary tables

Definitive tables

Database (RDBMS Sybase)

1- Coordination des flux de données

IMunoGeneTics Information system®
http://www.imgt.org



Data Coherence

Data Integrity

Annotation
LIGM Expertise

Knowledge tables

Administration

Data Reception
EMBL Flat Files
Author Submission

Reception tables

Temporary tables

Definitive tables

Database (RDBMS Sybase)

Data Distribution

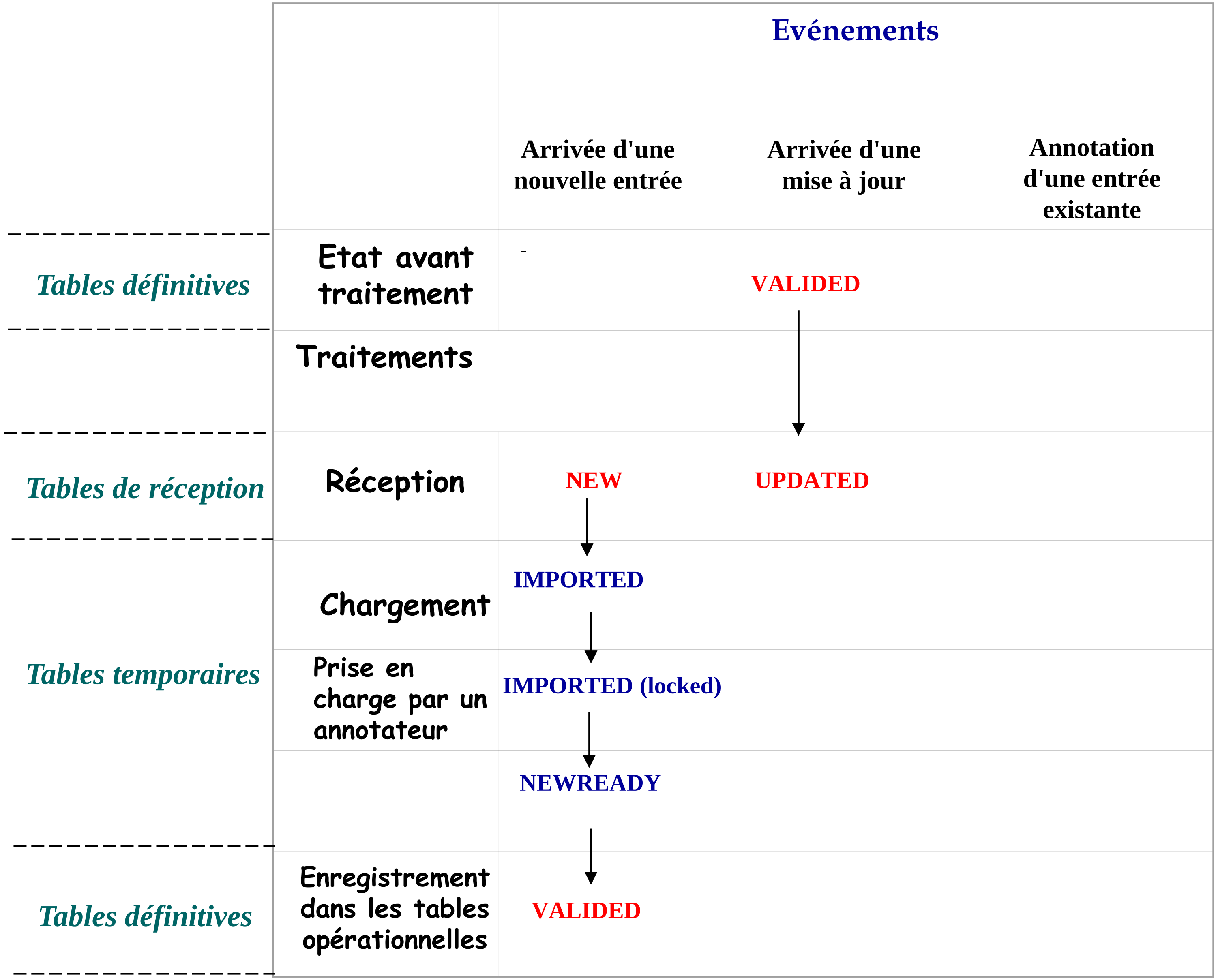
Flat files distribution

Web server

Java API

2 - Gestion accès aux données

http://www.imgt.org

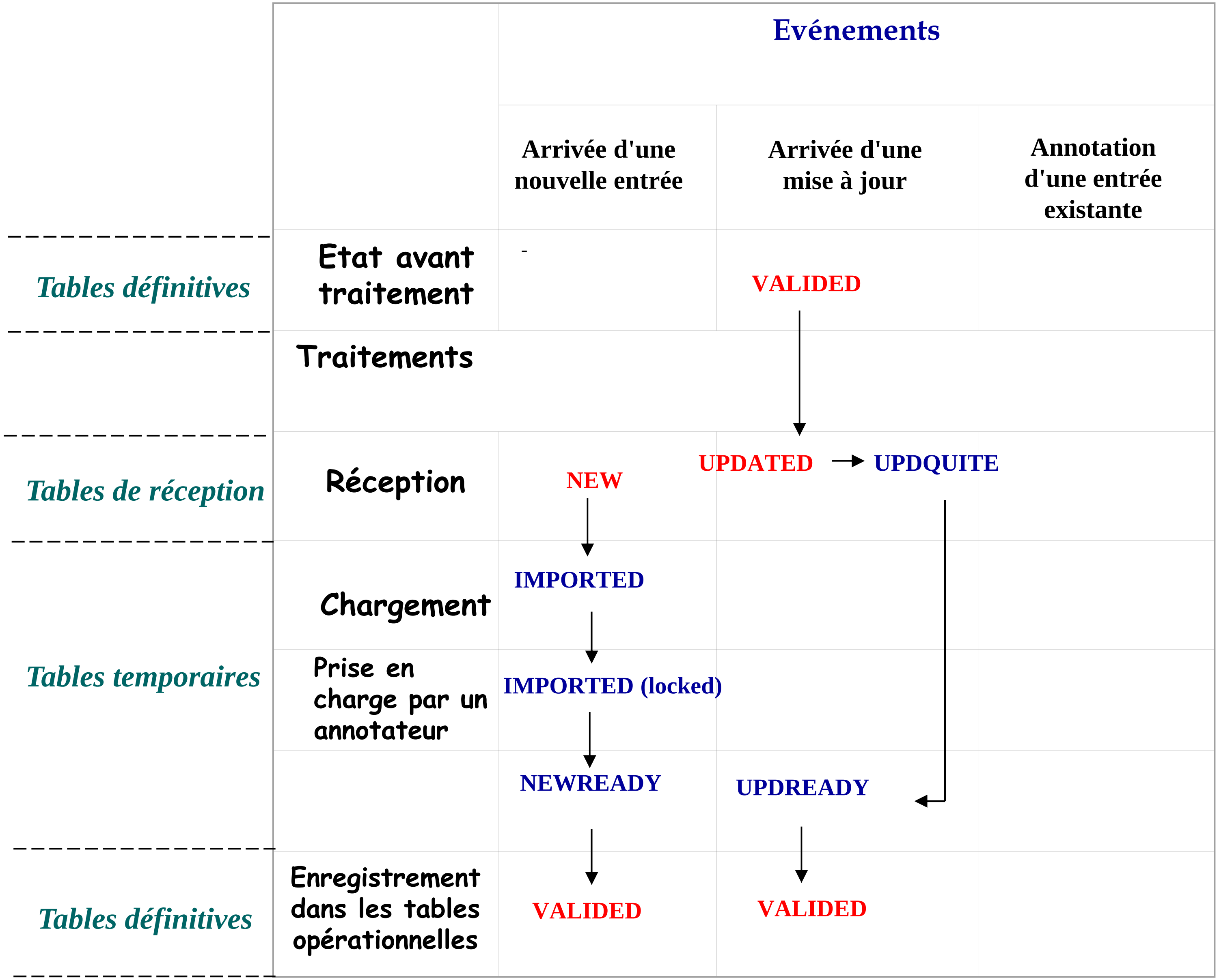


		Evénements		
		Arrivée d'une nouvelle entrée	Arrivée d'une mise à jour	Annotation d'une entrée existante
Tables définitives	**Etat avant traitement**	-	VALIDED	
	Traitements			
Tables de réception	**Réception**	NEW	UPDATED	
Tables temporaires	**Chargement**	IMPORTED		
	Prise en charge par un annotateur	IMPORTED (locked)		
		NEWREADY		
Tables définitives	**Enregistrement dans les tables opérationnelles**	VALIDED		

(locked) : indique que l'entrée est prise en charge par un annotateur

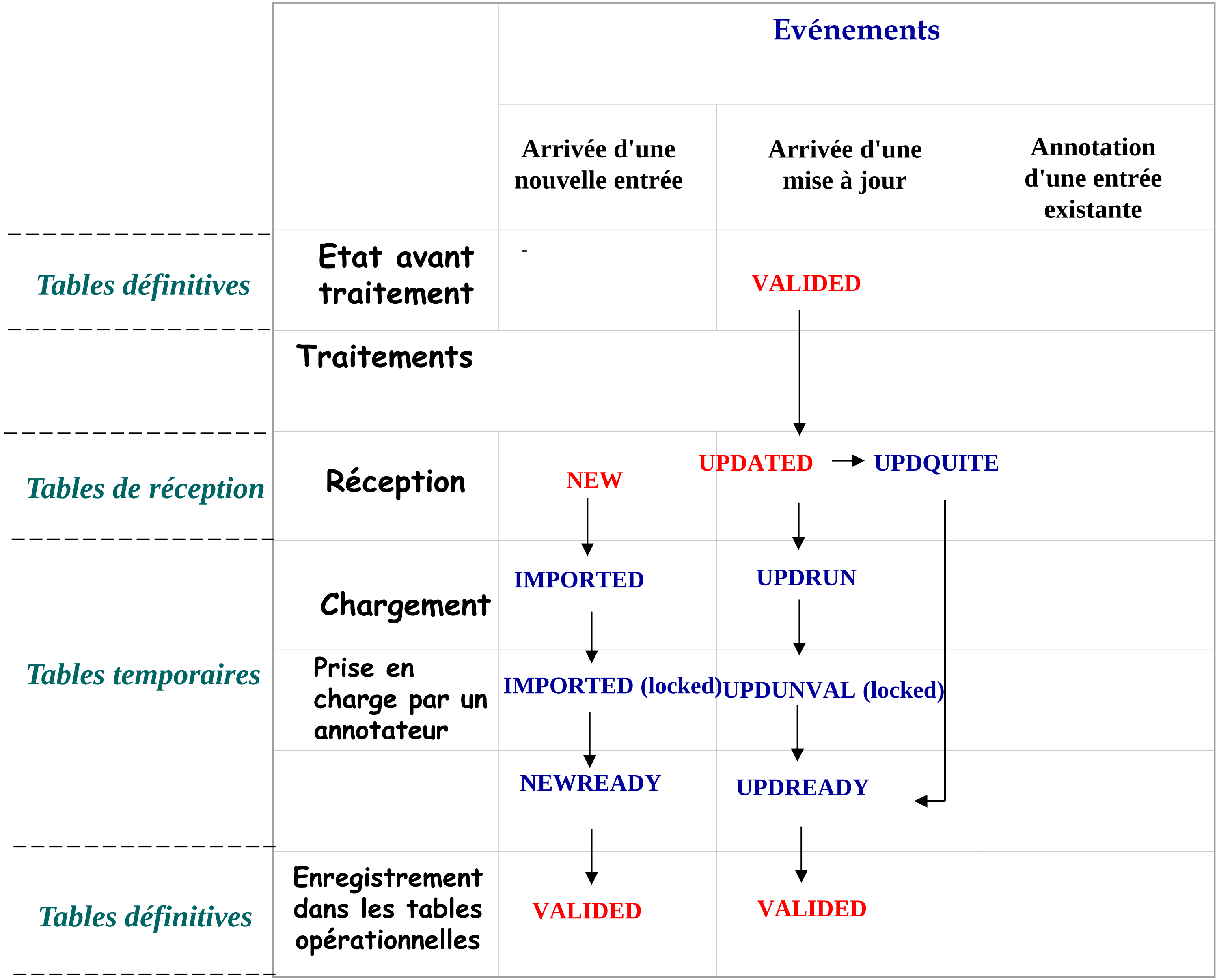
Information system®

http://www.imgt.org



		Evénements		
		Arrivée d'une nouvelle entrée	Arrivée d'une mise à jour	Annotation d'une entrée existante
Tables définitives	**Etat avant traitement**	-	VALIDED	
	Traitements			
Tables de réception	**Réception**	NEW	UPDATED → UPDQUITE	
Tables temporaires	**Chargement**	IMPORTED		
	Prise en charge par un annotateur	IMPORTED (locked)		
		NEWREADY	UPDREADY	
Tables définitives	**Enregistrement dans les tables opérationnelles**	VALIDED	VALIDED	

(locked) : indique que l'entrée est prise en charge par un annotateur

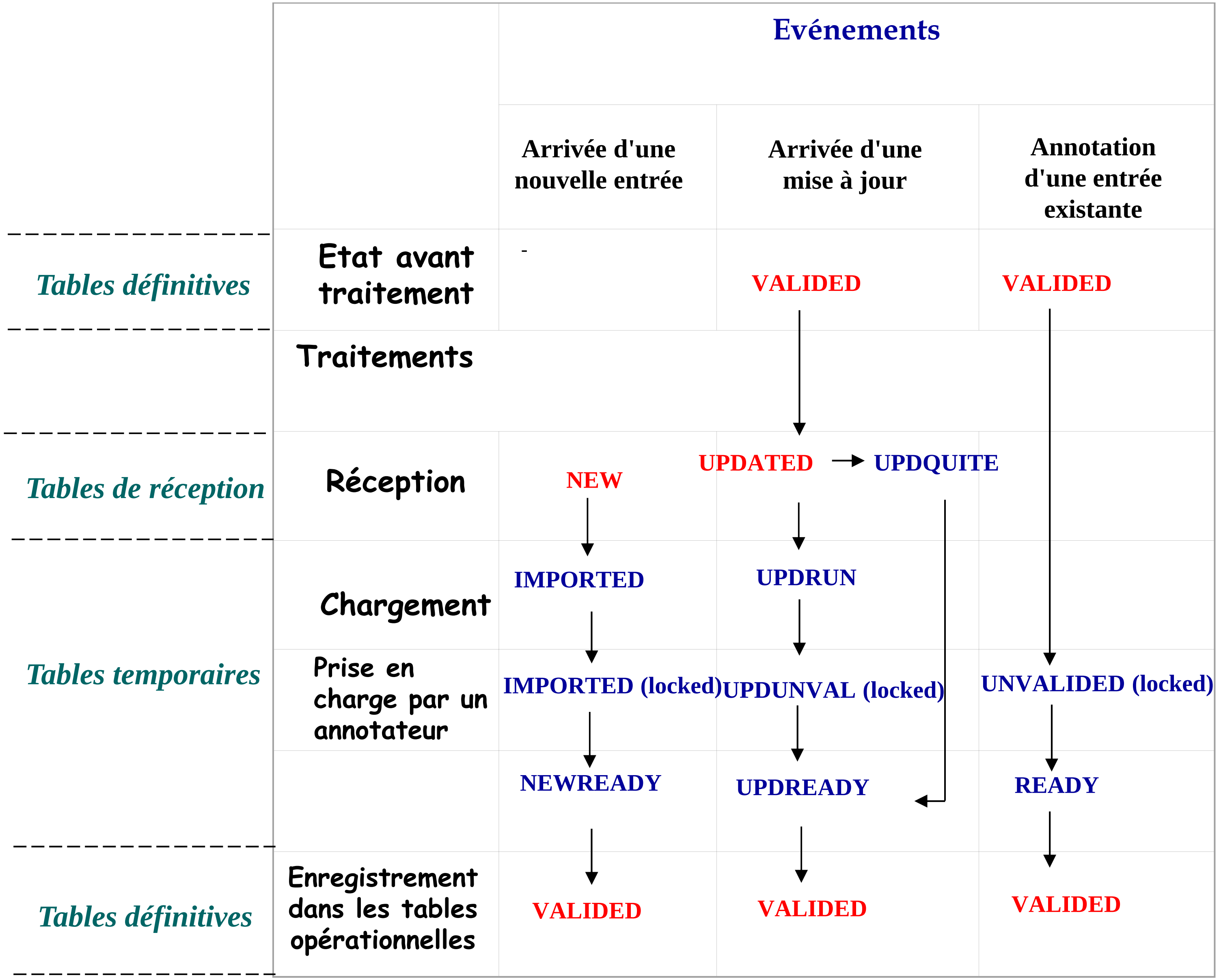


		Evénements		
		Arrivée d'une nouvelle entrée	Arrivée d'une mise à jour	Annotation d'une entrée existante
Tables définitives	**Etat avant traitement**	-	VALIDED	
	Traitements		↓	
Tables de réception	**Réception**	NEW	UPDATED → UPDQUITE	
Tables temporaires	**Chargement**	IMPORTED	UPDRUN	
	Prise en charge par un annotateur	IMPORTED (locked)	UPDUNVAL (locked)	
		NEWREADY	UPDREADY	
Tables définitives	**Enregistrement dans les tables opérationnelles**	VALIDED	VALIDED	

Information system®

http://www.imgt.org

(locked) : indique que l'entrée est prise en charge par un annotateur



		Evénements		
		Arrivée d'une nouvelle entrée	Arrivée d'une mise à jour	Annotation d'une entrée existante
Tables définitives	**Etat avant traitement**	-	VALIDED	VALIDED
	Traitements			
Tables de réception	**Réception**	NEW	UPDATED → UPDQUITE	
Tables temporaires	**Chargement**	IMPORTED	UPDRUN	
	Prise en charge par un annotateur	IMPORTED (locked)	UPDUNVAL (locked)	UNVALIDED (locked)
		NEWREADY	UPDREADY	READY
Tables définitives	**Enregistrement dans les tables opérationnelles**	VALIDED	VALIDED	VALIDED

http://www.imgt.org

(locked) : indique que l'entrée est prise en charge par un annotateur

Information et connaissance

- **Une Information** comprend des données primaires issues l’expérimentation (mesures, images, séquences) ainsi que les données secondaires qui comprennent aussi ce qu’il est nécessaire de connaître pour leur analyse.

  (résultats + matériels & méthodes)

- **La connaissance** inclut tout ce qu’il est nécessaire pour réaliser l’expertise de données réalisée par des spécialistes dans un domaine scientifique particulier.

Pour partager les informations,
Les données, les connaissances,…

Il faut utiliser/parler le même langage

⇒Nécessité d’utiliser le même vocabulaire standardisé, contrôlé (terminologie)

=> Nécessité de définir les dépendances entre les termes (ontologie)

Un exemple de question

Qu’est-ce qu’un gène?

=> plusieurs définitions, évolution

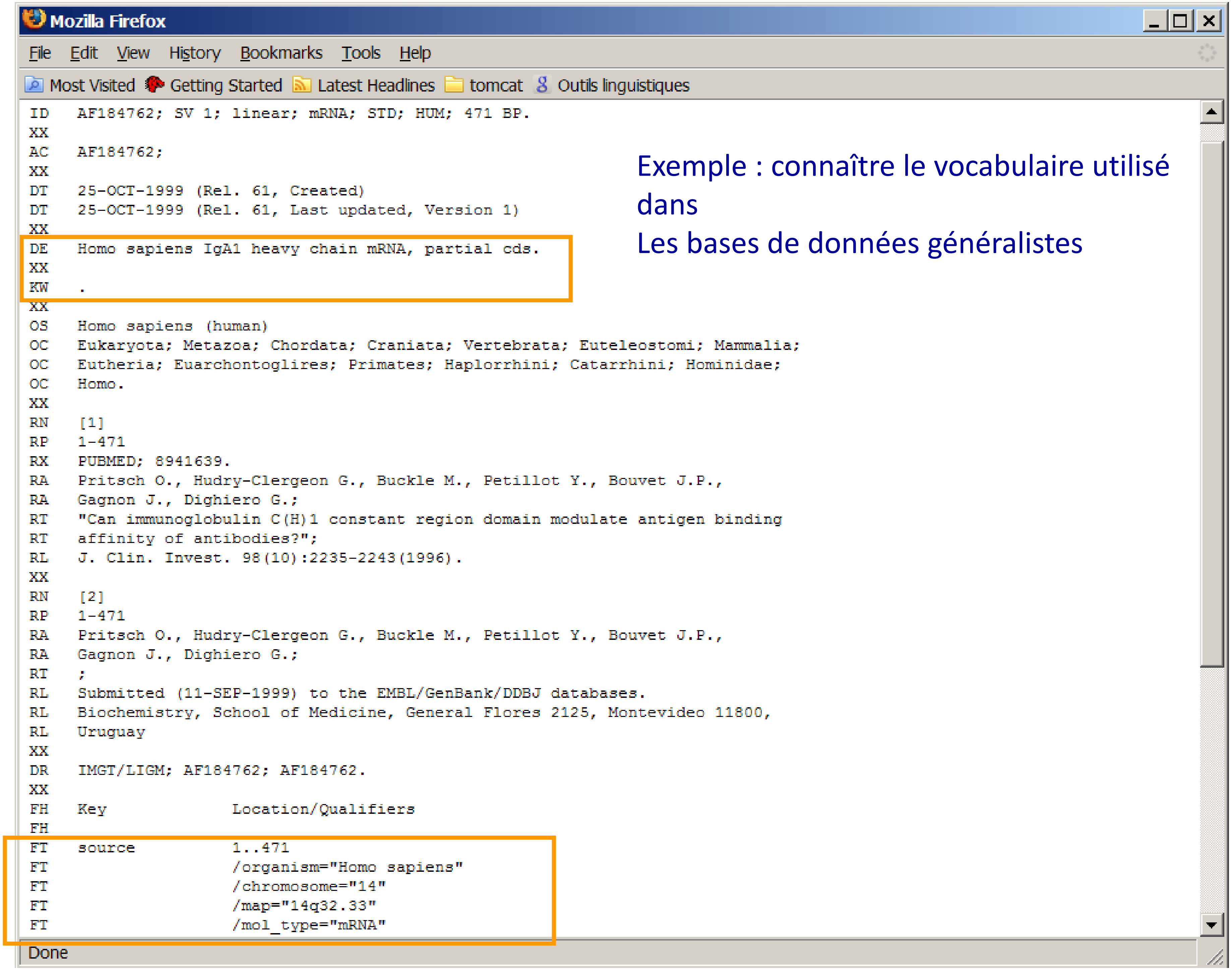
(**What is a gene, post-ENCODE? History and updated definition. Genome Res. 2007 Jun;17(6):669-81)**

Dans une base de données:

- établissement d’un dictionnaire des données
- mise en place de contraintes à travers un vocabulaire contrôlé

Recherches bibliographiques à travers des thesaurus de mots clés

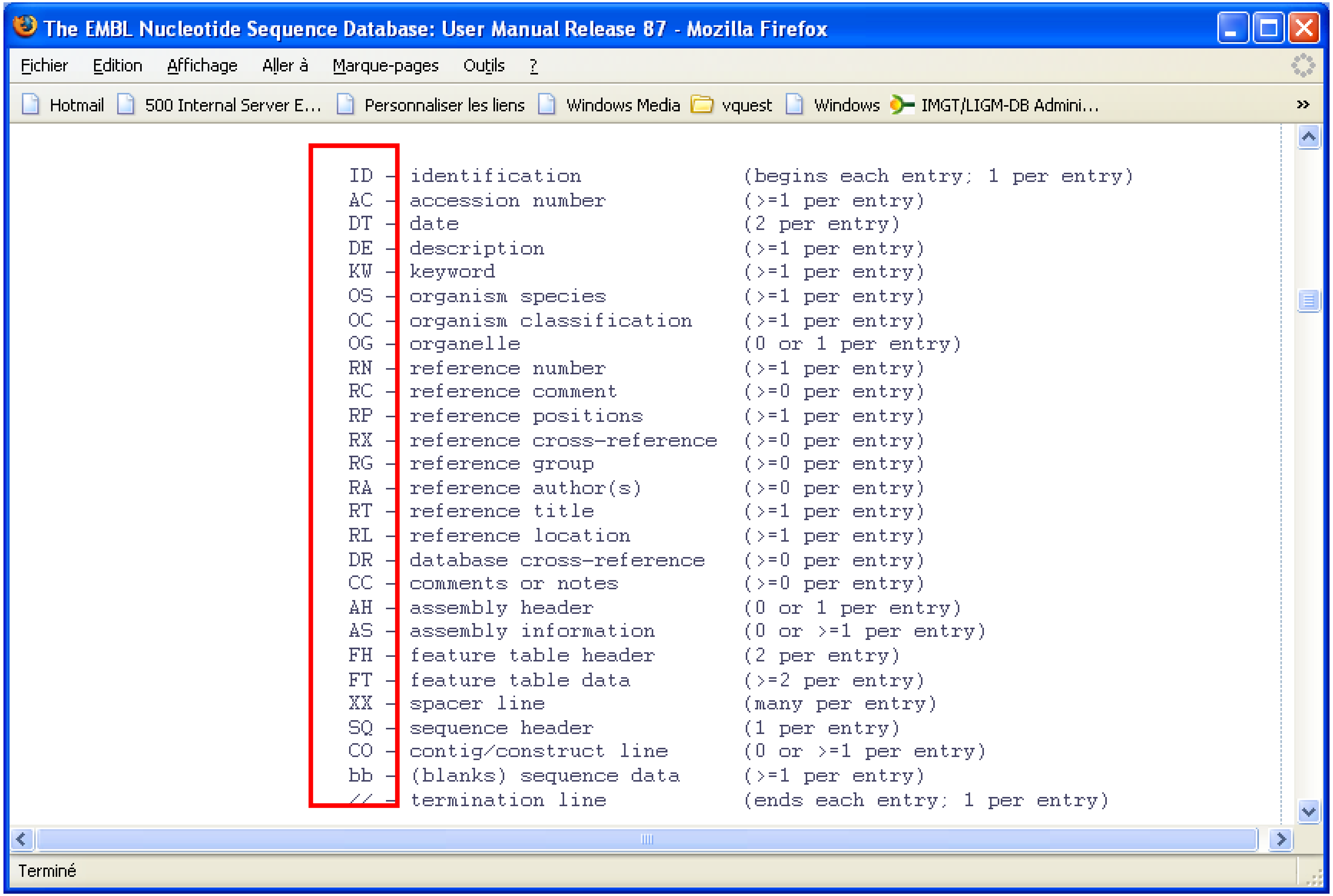
Exemple : connaître le vocabulaire utilisé dans
Les bases de données généralistes



```
ID   AF184762; SV 1; linear; mRNA; STD; HUM; 471 BP.
XX
AC   AF184762;
XX
DT   25-OCT-1999 (Rel. 61, Created)
DT   25-OCT-1999 (Rel. 61, Last updated, Version 1)
XX
DE   Homo sapiens IgA1 heavy chain mRNA, partial cds.
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   1-471
RX   PUBMED; 8941639.
RA   Pritsch O., Hudry-Clergeon G., Buckle M., Petillot Y., Bouvet J.P.,
RA   Gagnon J., Dighiero G.;
RT   "Can immunoglobulin C(H)1 constant region domain modulate antigen binding
RT   affinity of antibodies?";
RL   J. Clin. Invest. 98(10):2235-2243(1996).
XX
RN   [2]
RP   1-471
RA   Pritsch O., Hudry-Clergeon G., Buckle M., Petillot Y., Bouvet J.P.,
RA   Gagnon J., Dighiero G.;
RT   ;
RL   Submitted (11-SEP-1999) to the EMBL/GenBank/DDBJ databases.
RL   Biochemistry, School of Medicine, General Flores 2125, Montevideo 11800,
RL   Uruguay
XX
DR   IMGT/LIGM; AF184762; AF184762.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..471
FT                   /organism="Homo sapiens"
FT                   /chromosome="14"
FT                   /map="14q32.33"
FT                   /mol_type="mRNA"
```

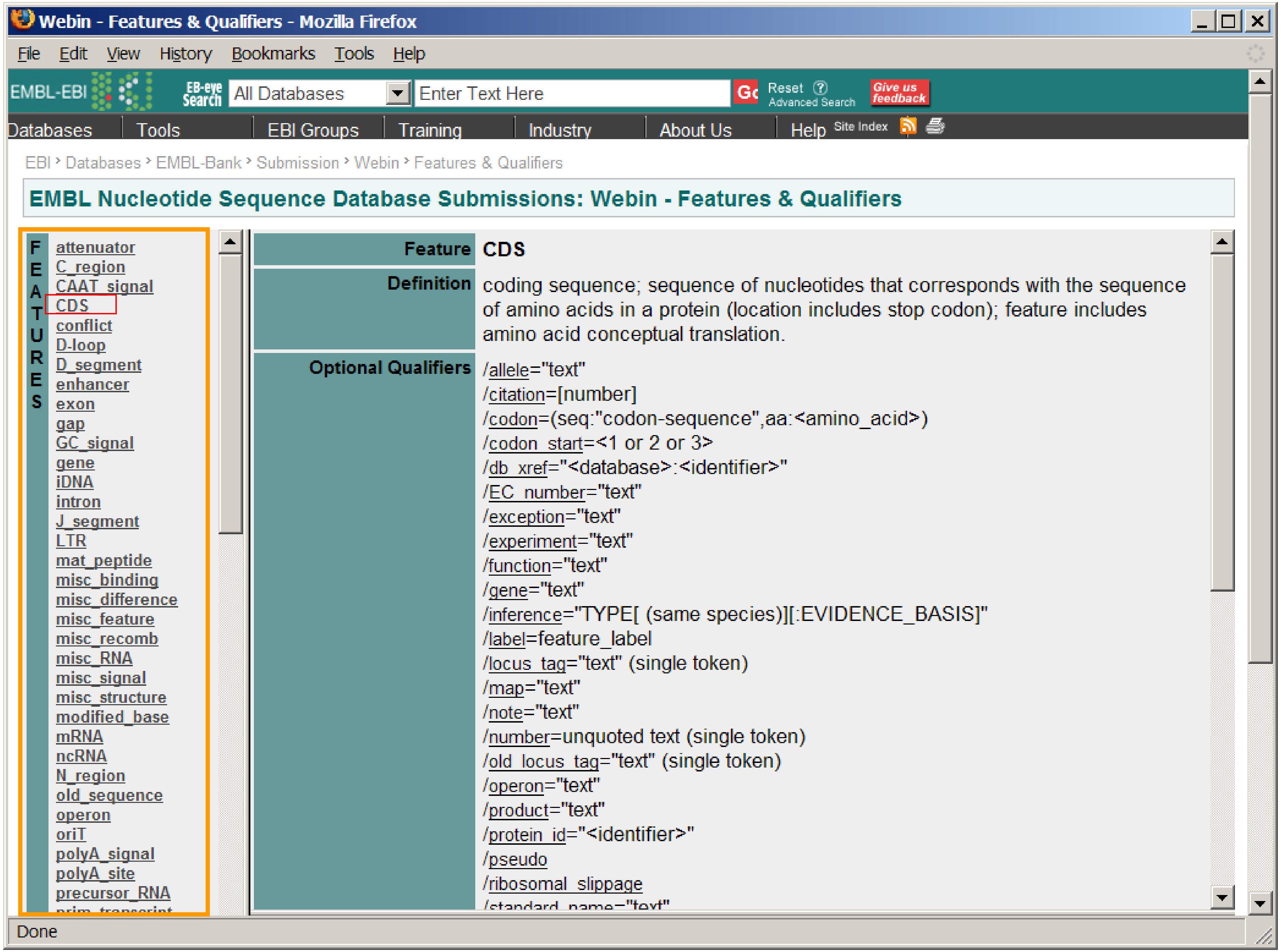
Le vocabulaire contrôlé des bases de données généralistes DDBJ/EMBL/GenBank

Des codes pour indiquer les types d'information



```
ID - identification               (begins each entry; 1 per entry)
AC - accession number             (>=1 per entry)
DT - date                         (2 per entry)
DE - description                  (>=1 per entry)
KW - keyword                      (>=1 per entry)
OS - organism species             (>=1 per entry)
OC - organism classification      (>=1 per entry)
OG - organelle                    (0 or 1 per entry)
RN - reference number             (>=1 per entry)
RC - reference comment            (>=0 per entry)
RP - reference positions          (>=1 per entry)
RX - reference cross-reference    (>=0 per entry)
RG - reference group              (>=0 per entry)
RA - reference author(s)          (>=0 per entry)
RT - reference title              (>=1 per entry)
RL - reference location           (>=1 per entry)
DR - database cross-reference     (>=0 per entry)
CC - comments or notes            (>=0 per entry)
AH - assembly header              (0 or 1 per entry)
AS - assembly information         (0 or >=1 per entry)
FH - feature table header         (2 per entry)
FT - feature table data           (>=2 per entry)
XX - spacer line                  (many per entry)
SQ - sequence header              (1 per entry)
CO - contig/construct line        (0 or >=1 per entry)
bb - (blanks) sequence data       (>=1 per entry)
// - termination line             (ends each entry; 1 per entry)
```

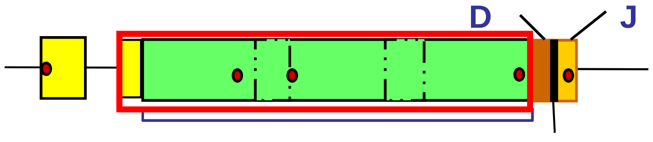
EMBL Feature labels (http://www.ebi.ac.uk/embl/WebFeat/index.html)



EMBL Nucleotide Sequence Database Submissions: Webin - Features & Qualifiers

Feature	**CDS**
Definition	coding sequence; sequence of nucleotides that corresponds with the sequence of amino acids in a protein (location includes stop codon); feature includes amino acid conceptual translation.
Optional Qualifiers	/allele="text" /citation=[number] /codon=(seq:"codon-sequence",aa:<amino_acid>) /codon_start=<1 or 2 or 3> /db_xref="<database>:<identifier>" /EC_number="text" /exception="text" /experiment="text" /function="text" /gene="text" /inference="TYPE[(same species)][:EVIDENCE_BASIS]" /label=feature_label /locus_tag="text" (single token) /map="text" /note="text" /number=unquoted text (single token) /old_locus_tag="text" (single token) /operon="text" /product="text" /protein_id="<identifier>" /pseudo /ribosomal_slippage /standard_name="text"

EMBL Feature labels



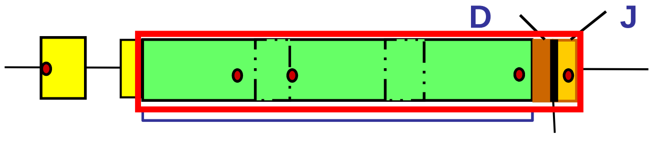
EMBL Nucleotide Sequence Database Submissions: Webin - Features & Qualifiers

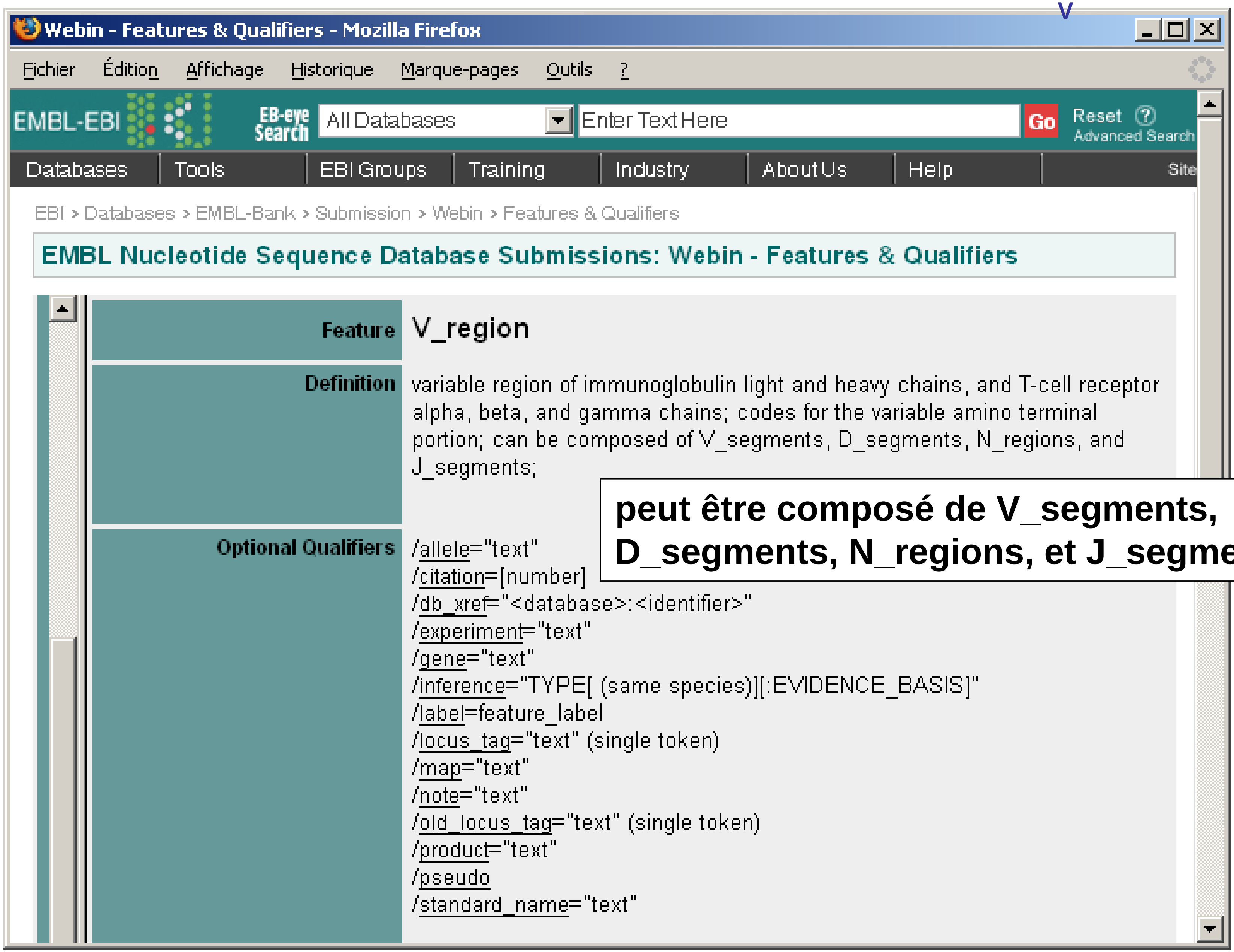
Feature	V_segment
Definition	variable segment of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; codes for most of the variable region (V_region) and the last few amino acids of the leader peptide;
Optional Qualifiers	/allele="text" /citation=[number] /db_xref="<databas /experiment="text" /gene="text" /inference="TYPE[(same species)][:EVIDENCE_BASIS]" /label=feature_label /locus_tag="text" (single token) /map="text" /note="text" /old_locus_tag="text" (single token) /product="text" /pseudo /standard_name="text"

code la majeur partie de la région variable (V_region) et les quelques acides aminés du peptide signal

?

EMBL Feature labels





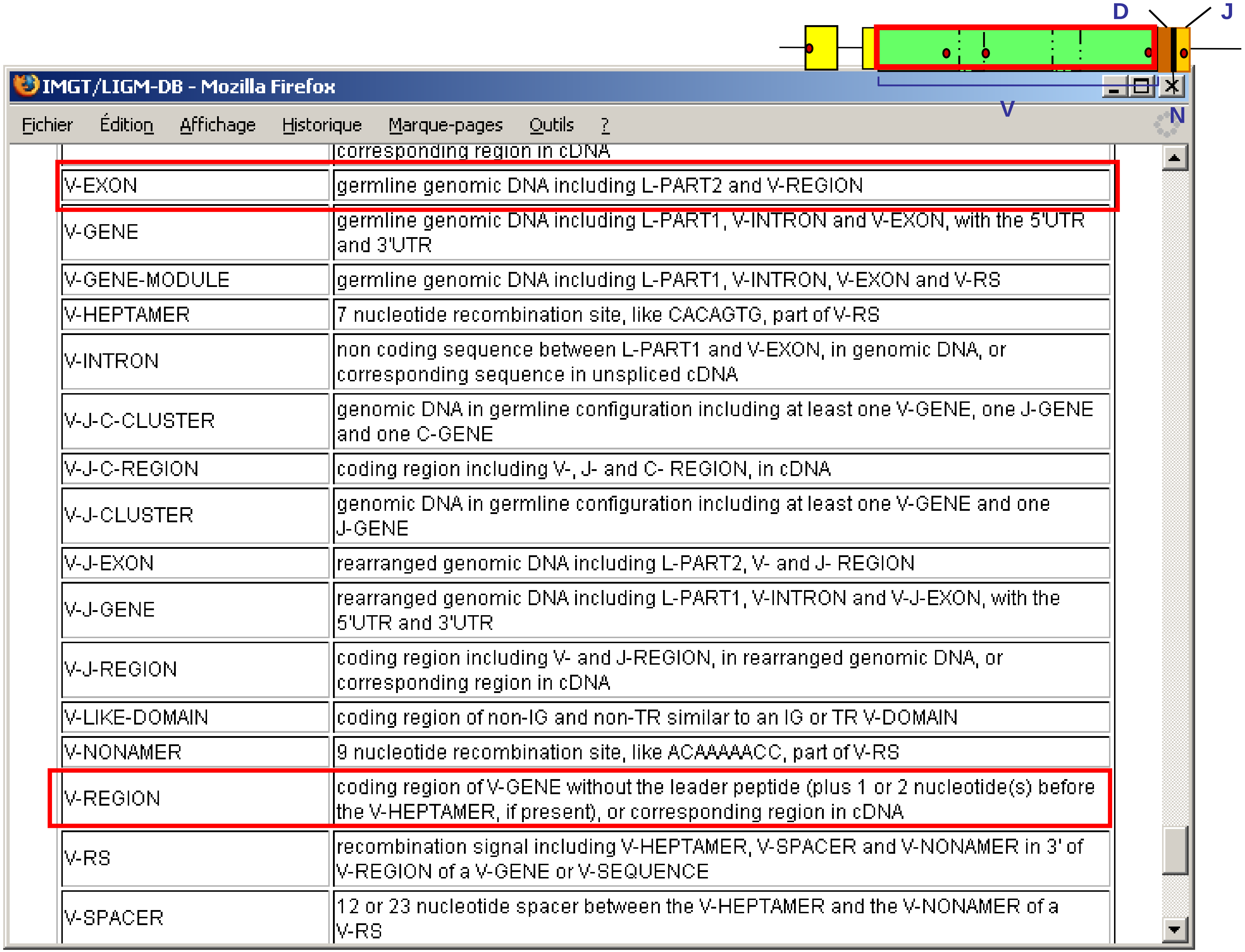
EMBL Nucleotide Sequence Database Submissions: Webin - Features & Qualifiers

Feature	V_region
Definition	variable region of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; codes for the variable amino terminal portion; can be composed of V_segments, D_segments, N_regions, and J_segments;
Optional Qualifiers	/allele="text" /citation=[number] /db_xref="<database>:<identifier>" /experiment="text" /gene="text" /inference="TYPE[(same species)][:EVIDENCE_BASIS]" /label=feature_label /locus_tag="text" (single token) /map="text" /note="text" /old_locus_tag="text" (single token) /product="text" /pseudo /standard_name="text"

peut être composé de V_segments, D_segments, N_regions, et J_segments

?

List and definition of IMGT standardized labels



	corresponding region in cDNA
V-EXON	germline genomic DNA including L-PART2 and V-REGION
V-GENE	germline genomic DNA including L-PART1, V-INTRON and V-EXON, with the 5'UTR and 3'UTR
V-GENE-MODULE	germline genomic DNA including L-PART1, V-INTRON, V-EXON and V-RS
V-HEPTAMER	7 nucleotide recombination site, like CACAGTG, part of V-RS
V-INTRON	non coding sequence between L-PART1 and V-EXON, in genomic DNA, or corresponding sequence in unspliced cDNA
V-J-C-CLUSTER	genomic DNA in germline configuration including at least one V-GENE, one J-GENE and one C-GENE
V-J-C-REGION	coding region including V-, J- and C- REGION, in cDNA
V-J-CLUSTER	genomic DNA in germline configuration including at least one V-GENE and one J-GENE
V-J-EXON	rearranged genomic DNA including L-PART2, V- and J- REGION
V-J-GENE	rearranged genomic DNA including L-PART1, V-INTRON and V-J-EXON, with the 5'UTR and 3'UTR
V-J-REGION	coding region including V- and J-REGION, in rearranged genomic DNA, or corresponding region in cDNA
V-LIKE-DOMAIN	coding region of non-IG and non-TR similar to an IG or TR V-DOMAIN
V-NONAMER	9 nucleotide recombination site, like ACAAAAACC, part of V-RS
V-REGION	coding region of V-GENE without the leader peptide (plus 1 or 2 nucleotide(s) before the V-HEPTAMER, if present), or corresponding region in cDNA
V-RS	recombination signal including V-HEPTAMER, V-SPACER and V-NONAMER in 3' of V-REGION of a V-GENE or V-SEQUENCE
V-SPACER	12 or 23 nucleotide spacer between the V-HEPTAMER and the V-NONAMER of a V-RS

OK

OK

Quelle est la signification des ces termes dans le contexte de la biologie?

Comment ces termes sont–ils organisés?

Comment sont-il reliés?

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

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 les relations entre les connaissances

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.**

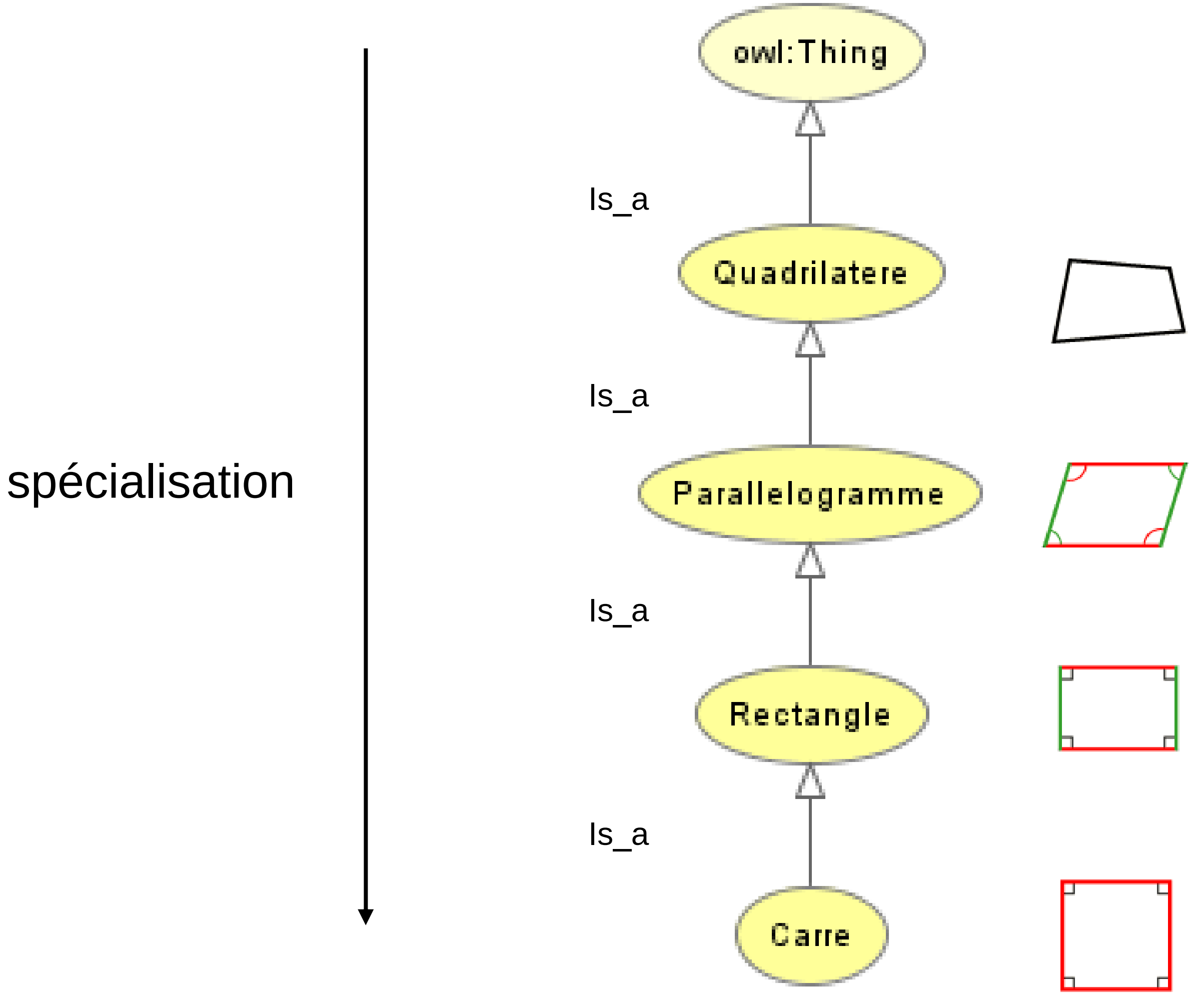
Bases de connaissances:

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.

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.



Def : polygone à 4 côtés

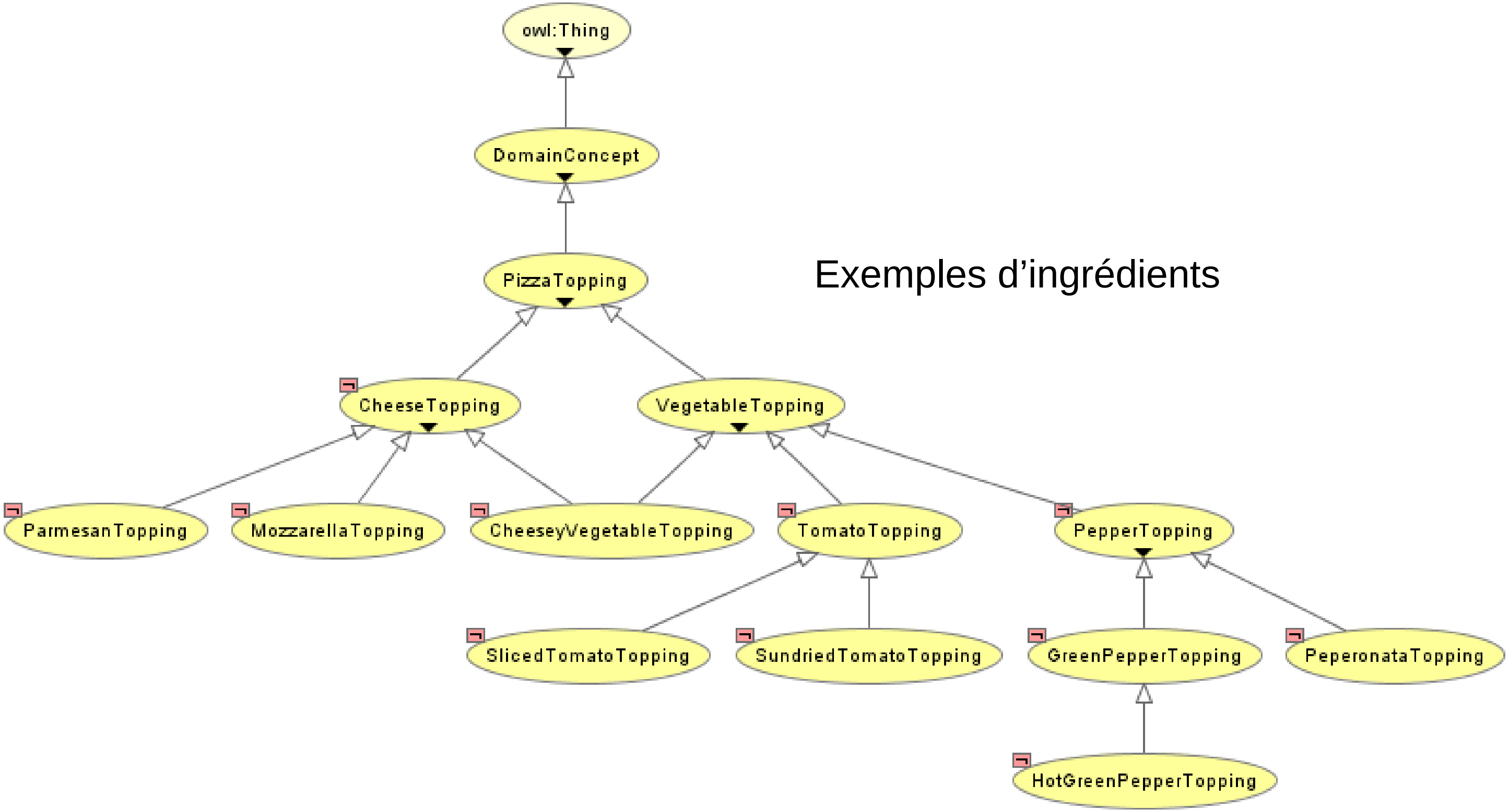
Def: quadrilatère (convexe) dont les côtés sont parallèles deux à deux

Def: Parallélogramme à angles droits

Def: rectangle dont les cotés sont égaux

Exemple de hiérarchie : ontologie des pizzas

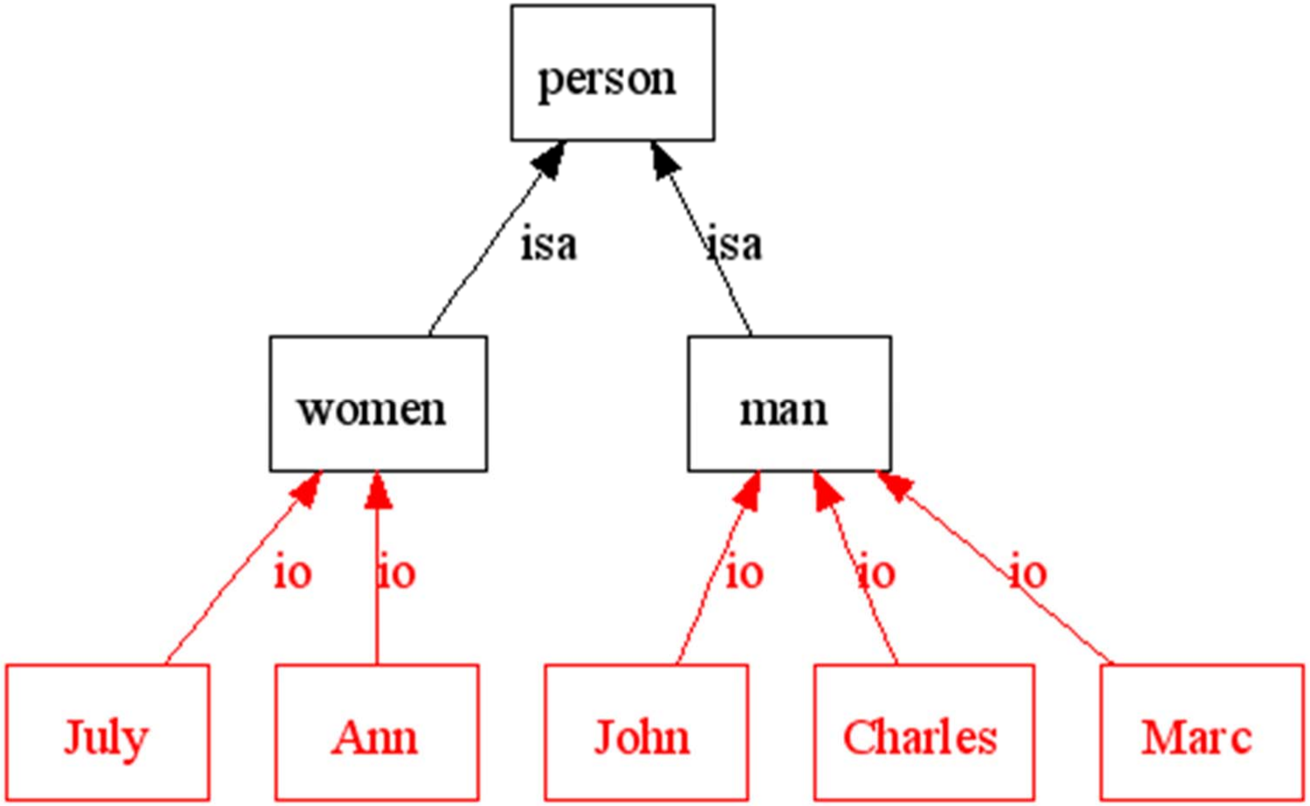
(exemple fournit avec l'éditeur d'ontologies Protégé http://protege.stanford.edu/)



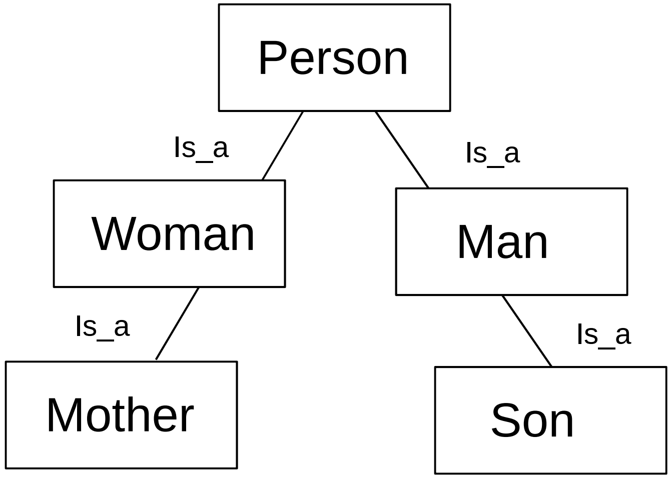
Exemples d'ingrédients

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?

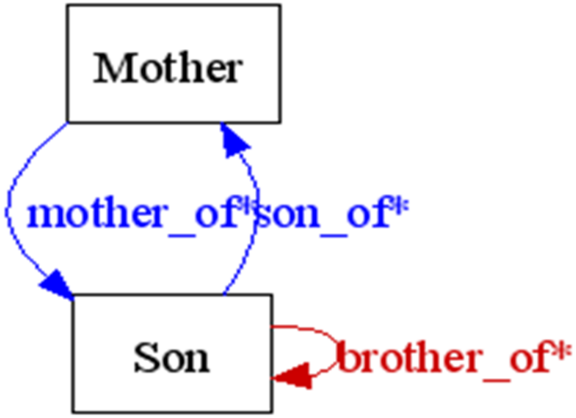
Propriétés/Relations
(OWL, Web Ontology Language)

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

- date de naissance
- numéro de sécurité sociale

…..

- Propriété d'objet : classe => classe

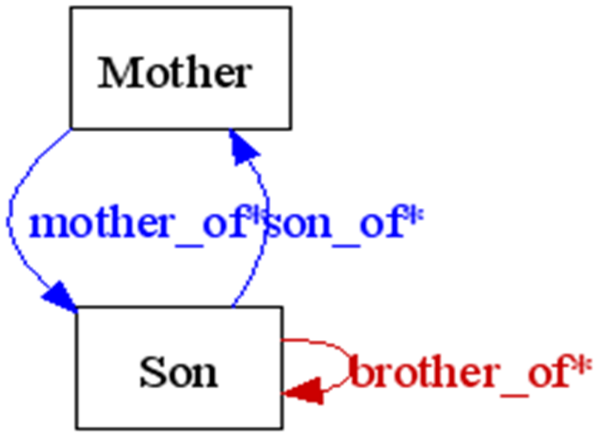


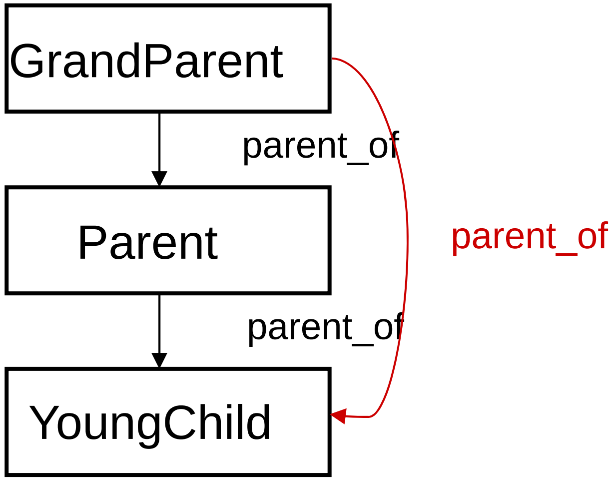
Définir et caractériser les Propriétés/Relations (OWL, Web Ontology Language)

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

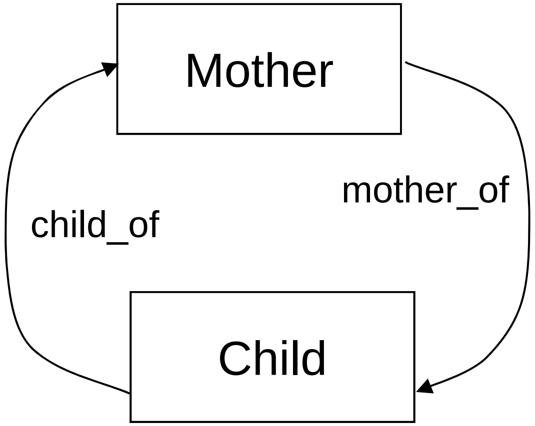




(familles de 3 générations avec de jeunes enfants)

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

Caractérisation des 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

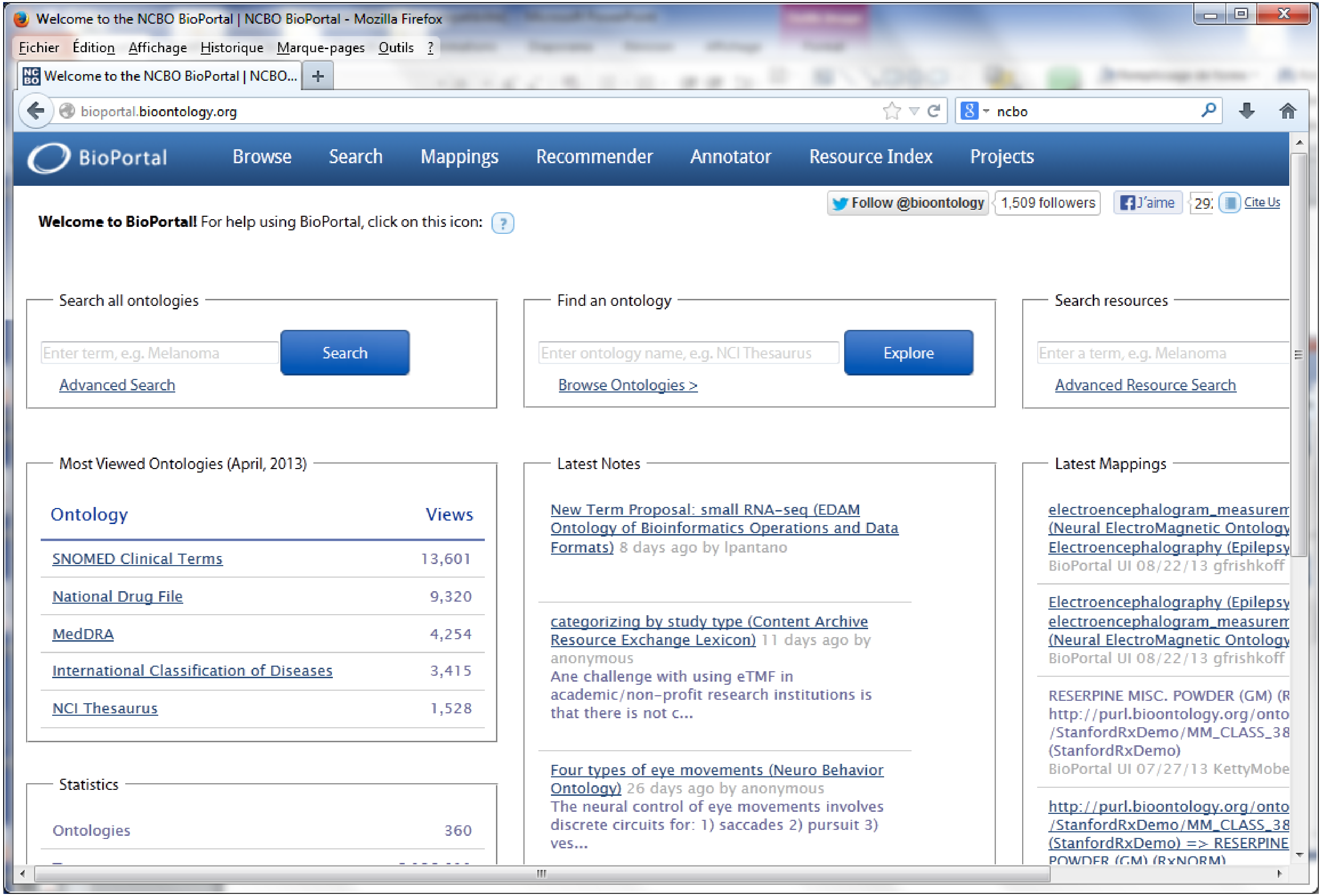
Les ontologies en Biologie

- 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/)



Welcome to BioPortal! For help using BioPortal, click on this icon: ?

Search all ontologies

Enter term, e.g. Melanoma [Search]

Advanced Search

Find an ontology

Enter ontology name, e.g. NCI Thesaurus [Explore]

Browse Ontologies >

Search resources

Enter a term, e.g. Melanoma

Advanced Resource Search

Most Viewed Ontologies (April, 2013)

Ontology	Views
SNOMED Clinical Terms	13,601
National Drug File	9,320
MedDRA	4,254
International Classification of Diseases	3,415
NCI Thesaurus	1,528

Statistics

Ontologies	360

Latest Notes

New Term Proposal: small RNA-seq (EDAM Ontology of Bioinformatics Operations and Data Formats) 8 days ago by lpantano

categorizing by study type (Content Archive Resource Exchange Lexicon) 11 days ago by anonymous
Ane challenge with using eTMF in academic/non-profit research institutions is that there is not c...

Four types of eye movements (Neuro Behavior Ontology) 26 days ago by anonymous
The neural control of eye movements involves discrete circuits for: 1) saccades 2) pursuit 3) ves...

Latest Mappings

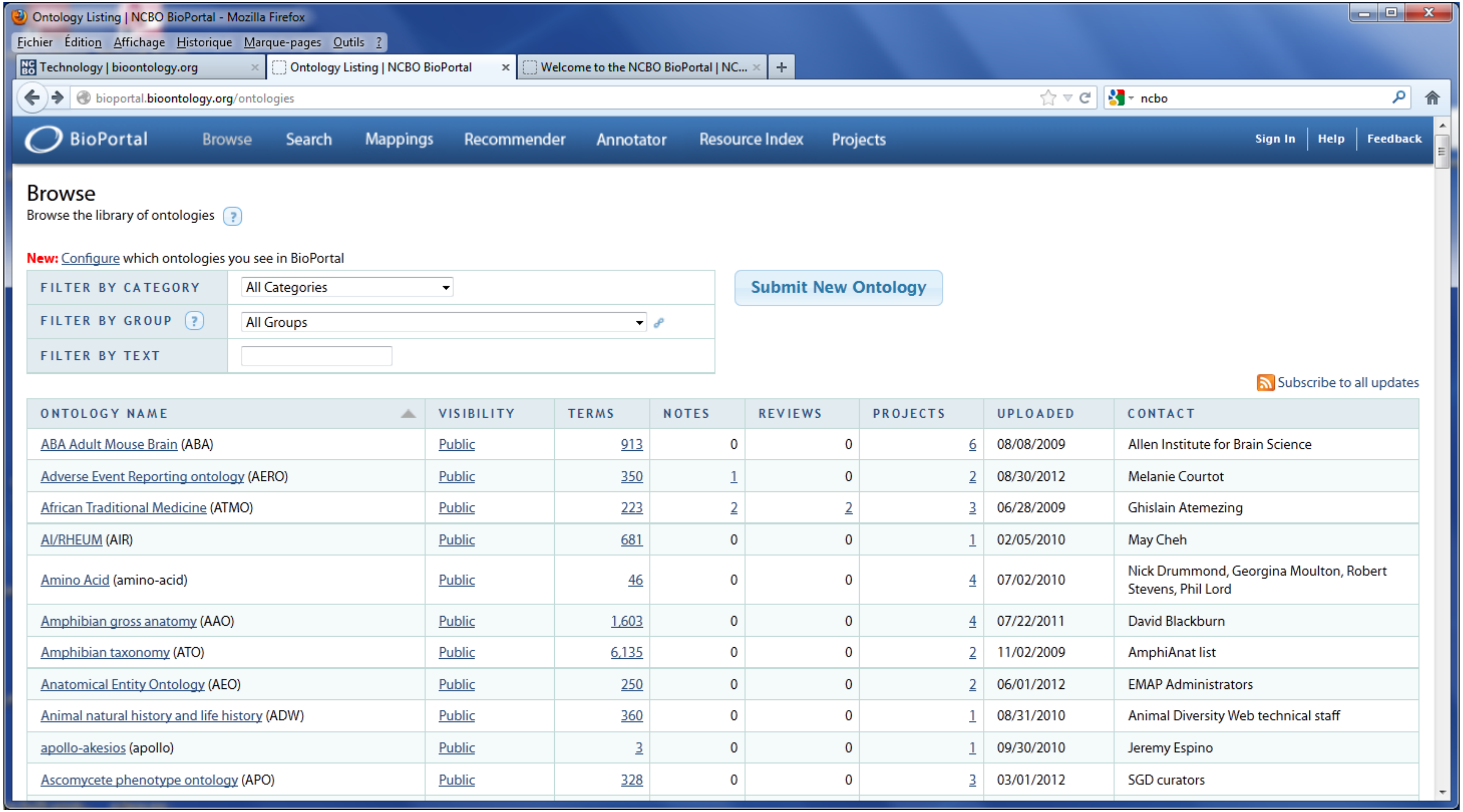
electroencephalogram_measurem
(Neural ElectroMagnetic Ontology
Electroencephalography (Epilepsy
BioPortal UI 08/22/13 gfrishkoff

Electroencephalography (Epilepsy
electroencephalogram_measurem
(Neural ElectroMagnetic Ontology
BioPortal UI 08/22/13 gfrishkoff

RESERPINE MISC. POWDER (GM) (R
http://purl.bioontology.org/onto
/StanfordRxDemo/MM_CLASS_38
(StanfordRxDemo)
BioPortal UI 07/27/13 KettyMobe

http://purl.bioontology.org/onto
/StanfordRxDemo/MM_CLASS_38
(StanfordRxDemo) => RESERPINE
POWDER (GM) (RXNORM)

NCBO Bioportal



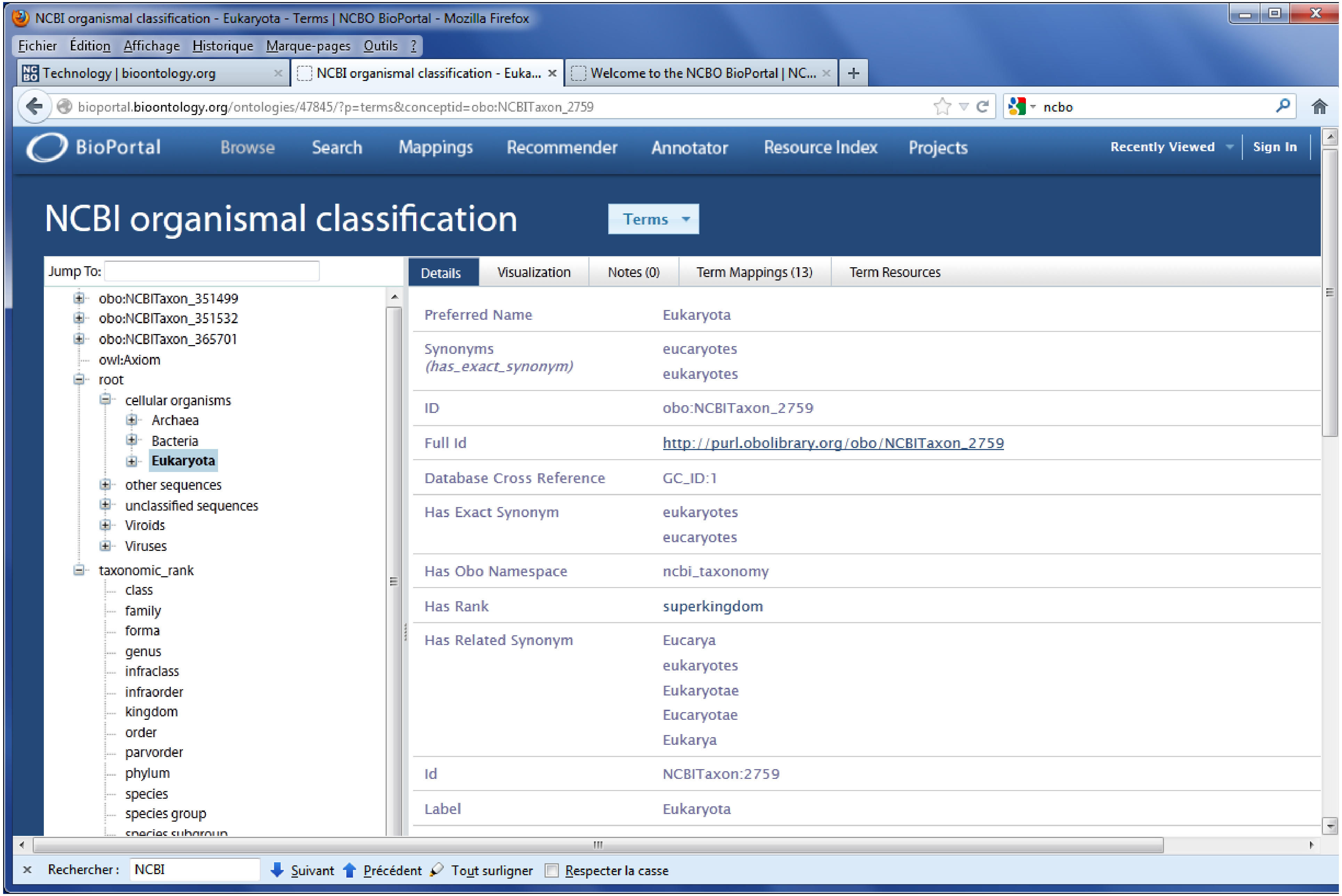
Browse

Browse the library of ontologies

New: Configure which ontologies you see in BioPortal

FILTER BY CATEGORY	All Categories
FILTER BY GROUP	All Groups
FILTER BY TEXT	

ONTOLOGY NAME	VISIBILITY	TERMS	NOTES	REVIEWS	PROJECTS	UPLOADED	CONTACT
ABA Adult Mouse Brain (ABA)	Public	913	0	0	6	08/08/2009	Allen Institute for Brain Science
Adverse Event Reporting ontology (AERO)	Public	350	1	0	2	08/30/2012	Melanie Courtot
African Traditional Medicine (ATMO)	Public	223	2	2	3	06/28/2009	Ghislain Atemezing
AI/RHEUM (AIR)	Public	681	0	0	1	02/05/2010	May Cheh
Amino Acid (amino-acid)	Public	46	0	0	4	07/02/2010	Nick Drummond, Georgina Moulton, Robert Stevens, Phil Lord
Amphibian gross anatomy (AAO)	Public	1,603	0	0	4	07/22/2011	David Blackburn
Amphibian taxonomy (ATO)	Public	6,135	0	0	2	11/02/2009	AmphiAnat list
Anatomical Entity Ontology (AEO)	Public	250	0	0	2	06/01/2012	EMAP Administrators
Animal natural history and life history (ADW)	Public	360	0	0	1	08/31/2010	Animal Diversity Web technical staff
apollo-akesios (apollo)	Public	3	0	0	1	09/30/2010	Jeremy Espino
Ascomycete phenotype ontology (APO)	Public	328	0	0	3	03/01/2012	SGD curators



bioportal.bioontology.org/ontologies/47845/?p=terms&conceptid=obo:NCBITaxon_2759

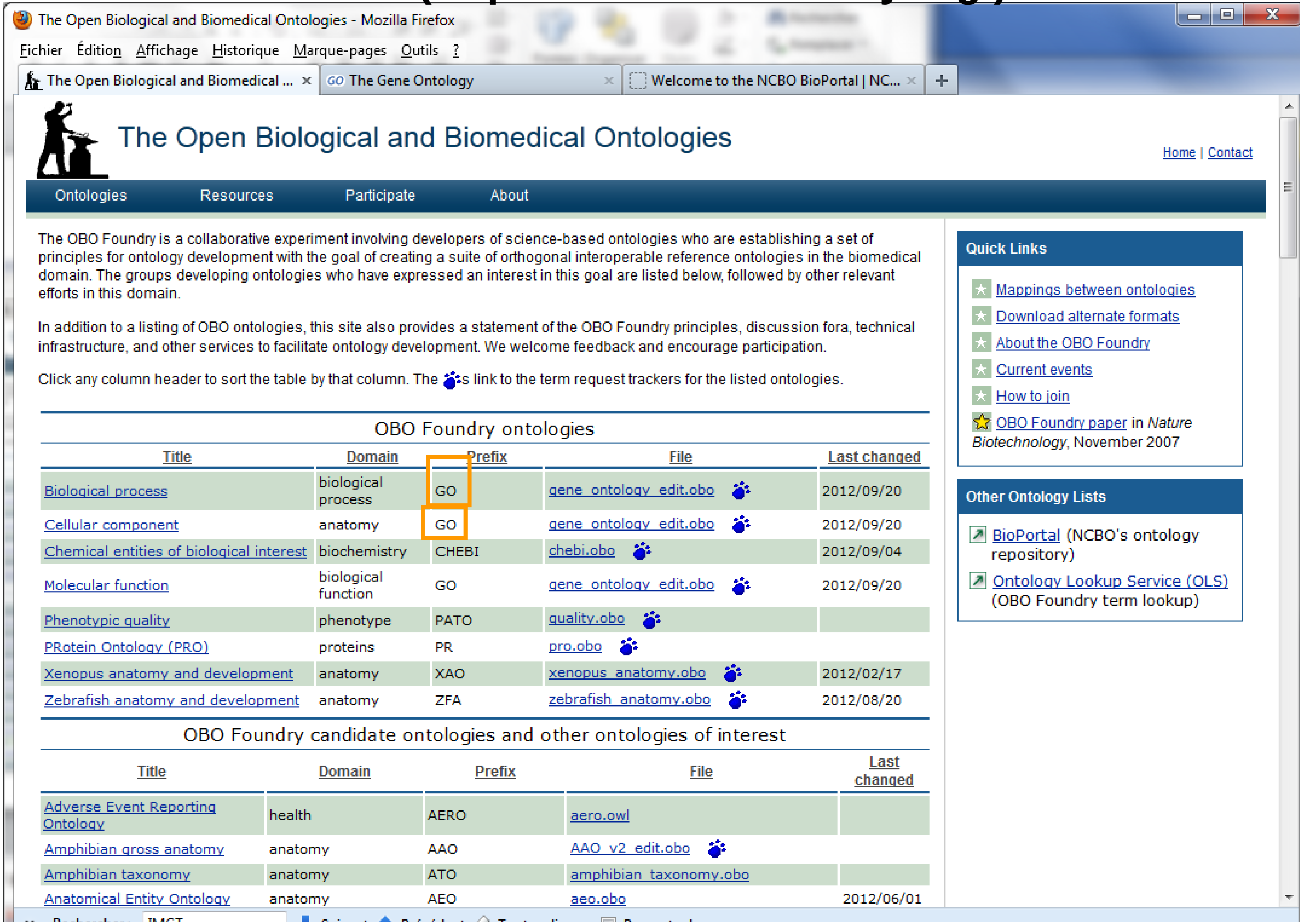
Rechercher : NCBI

NCBI organismal classification

- obo:NCBITaxon_351499
- obo:NCBITaxon_351532
- obo:NCBITaxon_365701
- owl:Axiom
- 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

Preferred Name	Eukaryota
Synonyms *(has_exact_synonym)*	eucaryotes eukaryotes
ID	obo:NCBITaxon_2759
Full Id	http://purl.obolibrary.org/obo/NCBITaxon_2759
Database Cross Reference	GC_ID:1
Has Exact Synonym	eukaryotes eucaryotes
Has Obo Namespace	ncbi_taxonomy
Has Rank	superkingdom
Has Related Synonym	Eucarya eukaryotes Eukaryotae Eucaryotae Eukarya
Id	NCBITaxon:2759
Label	Eukaryota

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



The Open Biological and Biomedical Ontologies

The OBO Foundry is a collaborative experiment involving developers of science-based ontologies who are establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain. The groups developing ontologies who have expressed an interest in this goal are listed below, followed by other relevant efforts in this domain.

In addition to a listing of OBO ontologies, this site also provides a statement of the OBO Foundry principles, discussion fora, technical infrastructure, and other services to facilitate ontology development. We welcome feedback and encourage participation.

Click any column header to sort the table by that column. The 🐾s link to the term request trackers for the listed ontologies.

OBO Foundry ontologies

Title	Domain	Prefix	File	Last changed
Biological process	biological process	GO	gene_ontology_edit.obo 🐾	2012/09/20
Cellular component	anatomy	GO	gene_ontology_edit.obo 🐾	2012/09/20
Chemical entities of biological interest	biochemistry	CHEBI	chebi.obo 🐾	2012/09/04
Molecular function	biological function	GO	gene_ontology_edit.obo 🐾	2012/09/20
Phenotypic quality	phenotype	PATO	quality.obo 🐾	
PRotein Ontology (PRO)	proteins	PR	pro.obo 🐾	
Xenopus anatomy and development	anatomy	XAO	xenopus_anatomy.obo 🐾	2012/02/17
Zebrafish anatomy and development	anatomy	ZFA	zebrafish_anatomy.obo 🐾	2012/08/20

OBO Foundry candidate ontologies and other ontologies of interest

Title	Domain	Prefix	File	Last changed
Adverse Event Reporting Ontology	health	AERO	aero.owl	
Amphibian gross anatomy	anatomy	AAO	AAO_v2_edit.obo 🐾	
Amphibian taxonomy	anatomy	ATO	amphibian_taxonomy.obo	
Anatomical Entity Ontology	anatomy	AEO	aeo.obo	2012/06/01

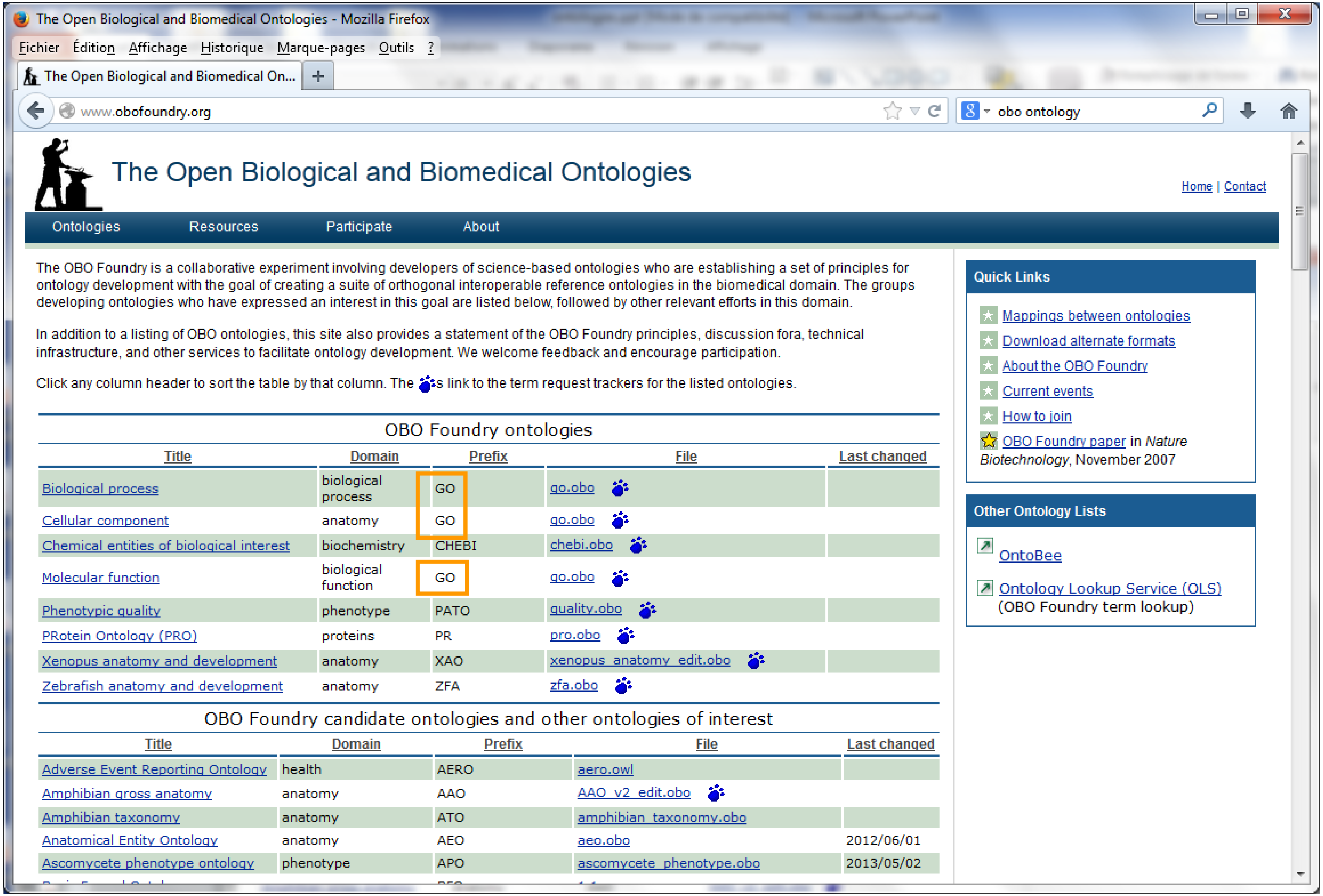
Quick Links

- Mappings between ontologies
- Download alternate formats
- About the OBO Foundry
- Current events
- How to join
- OBO Foundry paper in *Nature Biotechnology*, November 2007

Other Ontology Lists

- BioPortal (NCBO's ontology repository)
- Ontology Lookup Service (OLS) (OBO Foundry term lookup)

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



The Open Biological and Biomedical Ontologies

The OBO Foundry is a collaborative experiment involving developers of science-based ontologies who are establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain. The groups developing ontologies who have expressed an interest in this goal are listed below, followed by other relevant efforts in this domain.

In addition to a listing of OBO ontologies, this site also provides a statement of the OBO Foundry principles, discussion fora, technical infrastructure, and other services to facilitate ontology development. We welcome feedback and encourage participation.

Click any column header to sort the table by that column. The 🐾s link to the term request trackers for the listed ontologies.

OBO Foundry ontologies

Title	Domain	Prefix	File	Last changed
Biological process	biological process	GO	go.obo 🐾	
Cellular component	anatomy	GO	go.obo 🐾	
Chemical entities of biological interest	biochemistry	CHEBI	chebi.obo 🐾	
Molecular function	biological function	GO	go.obo 🐾	
Phenotypic quality	phenotype	PATO	quality.obo 🐾	
PRotein Ontology (PRO)	proteins	PR	pro.obo 🐾	
Xenopus anatomy and development	anatomy	XAO	xenopus_anatomy_edit.obo 🐾	
Zebrafish anatomy and development	anatomy	ZFA	zfa.obo 🐾	

OBO Foundry candidate ontologies and other ontologies of interest

Title	Domain	Prefix	File	Last changed
Adverse Event Reporting Ontology	health	AERO	aero.owl	
Amphibian gross anatomy	anatomy	AAO	AAO_v2_edit.obo 🐾	
Amphibian taxonomy	anatomy	ATO	amphibian_taxonomy.obo	
Anatomical Entity Ontology	anatomy	AEO	aeo.obo	2012/06/01
Ascomycete phenotype ontology	phenotype	APO	ascomycete_phenotype.obo	2013/05/02

Quick Links

- Mappings between ontologies
- Download alternate formats
- About the OBO Foundry
- Current events
- How to join
- OBO Foundry paper in *Nature Biotechnology*, November 2007

Other Ontology Lists

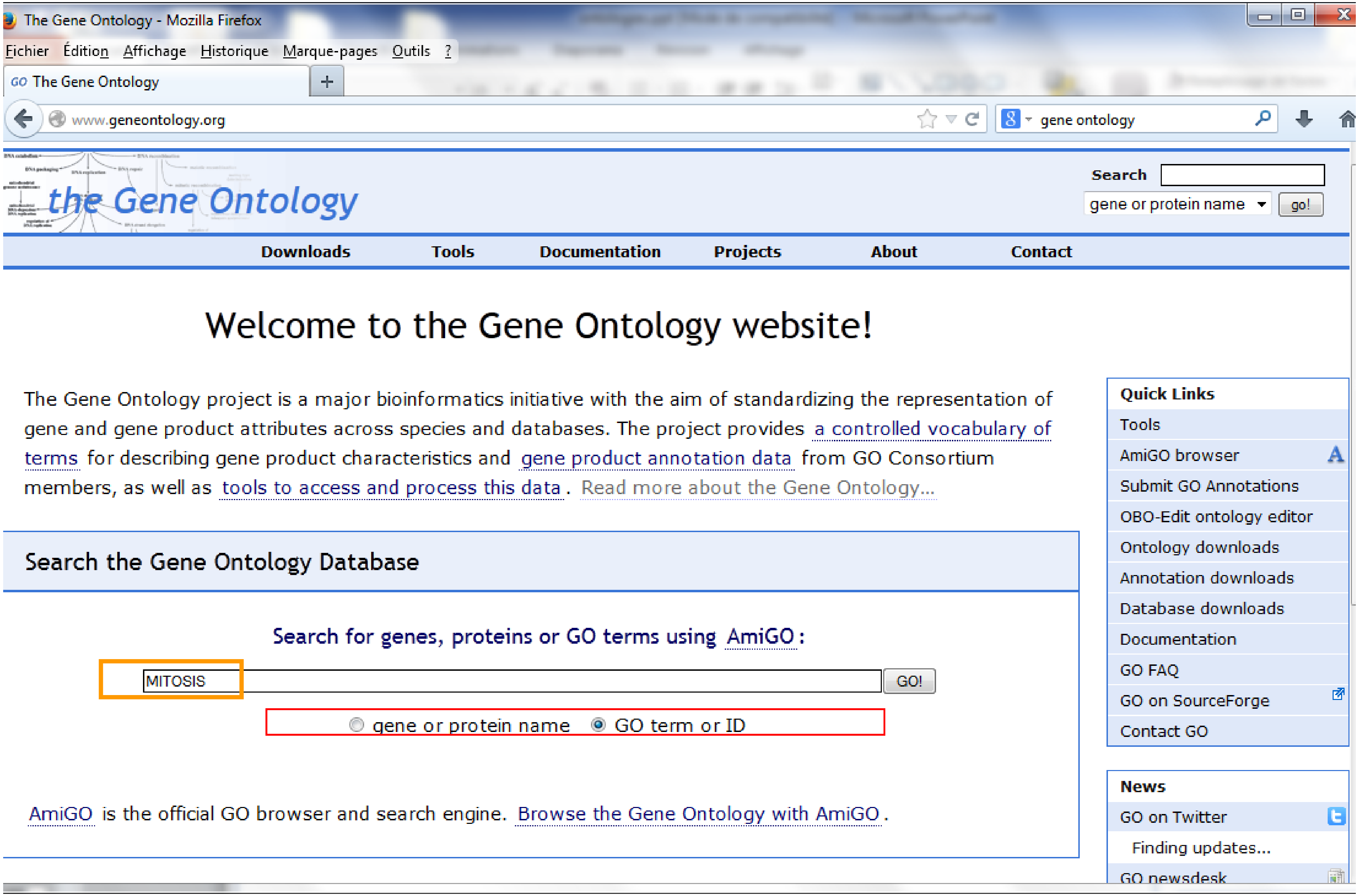
- OntoBee
- Ontology Lookup Service (OLS) (OBO Foundry term lookup)

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/)



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...

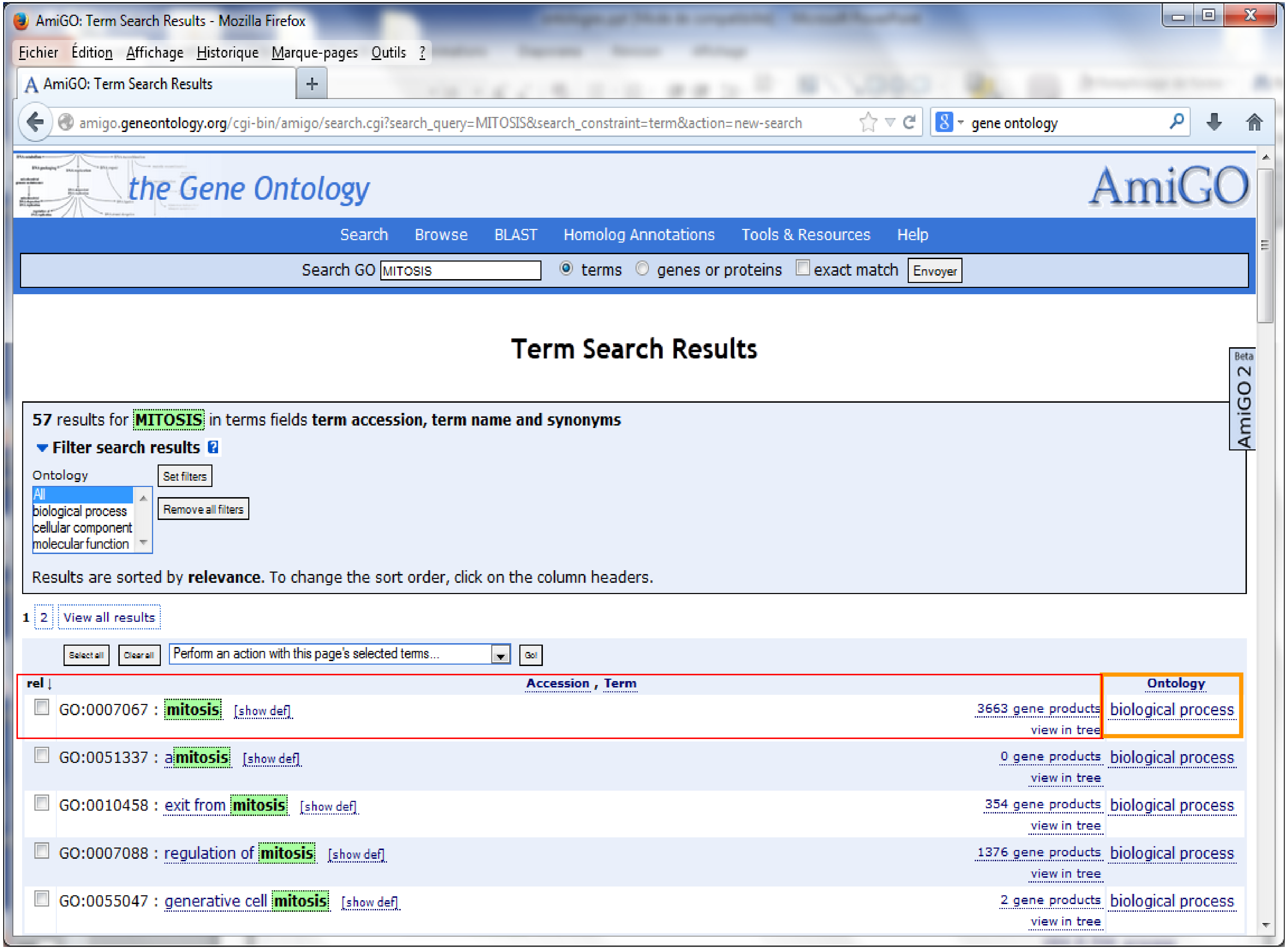
Search the Gene Ontology Database

Search for genes, proteins or GO terms using AmiGO:

MITOSIS GO!

○ gene or protein name ● GO term or ID

AmiGO is the official GO browser and search engine. Browse the Gene Ontology with AmiGO.



Search GO: MITOSIS

Term Search Results

57 results for **MITOSIS** in terms fields **term accession, term name and synonyms**

Filter search results

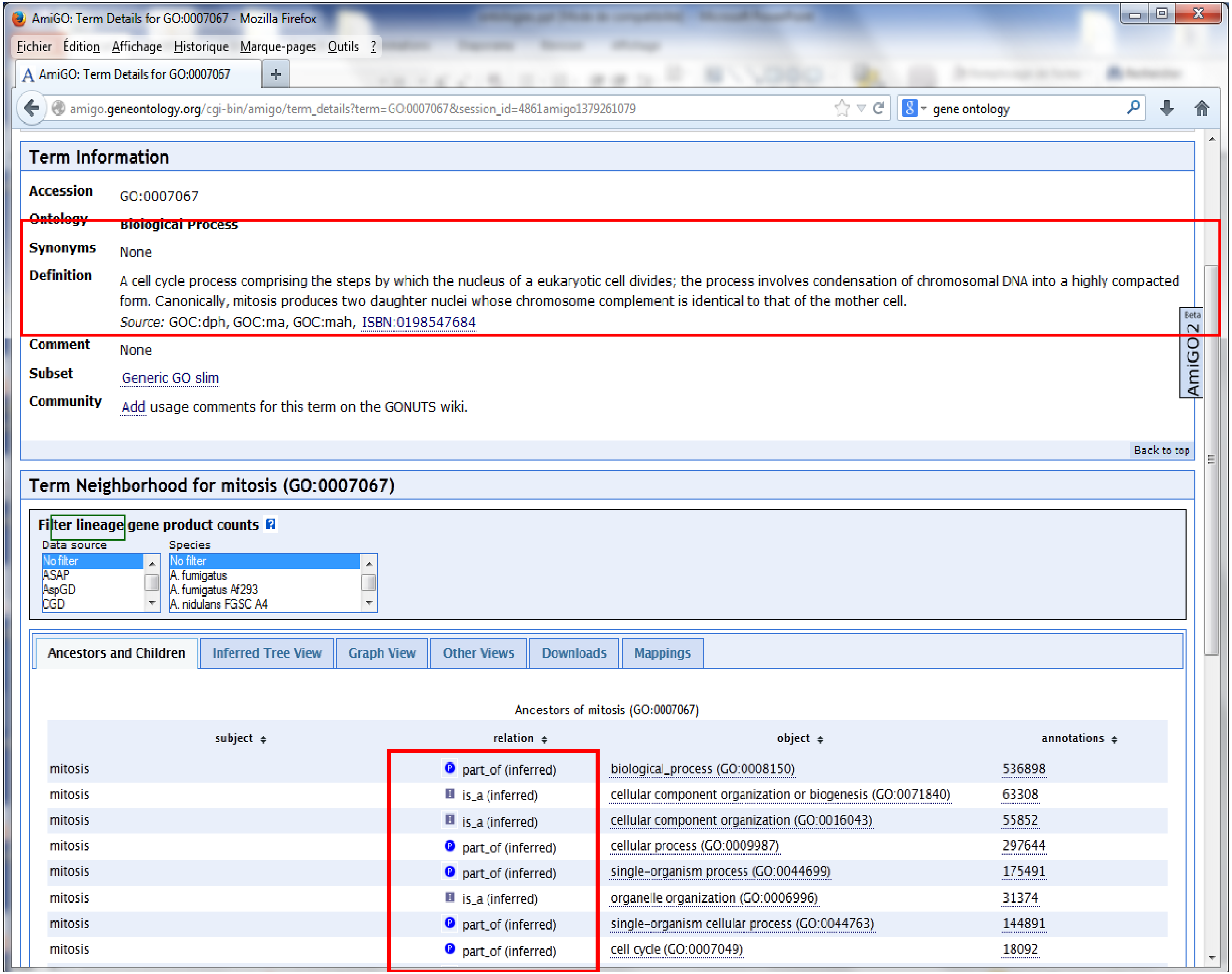
Ontology: All, biological process, cellular component, molecular function

Results are sorted by **relevance.** To change the sort order, click on the column headers.

1 2 View all results

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

rel ↓	Accession , Term		Ontology
☐	GO:0007067 : **mitosis** [show def]	3663 gene products view in tree	biological process
☐	GO:0051337 : a**mitosis** [show def]	0 gene products view in tree	biological process
☐	GO:0010458 : exit from **mitosis** [show def]	354 gene products view in tree	biological process
☐	GO:0007088 : regulation of **mitosis** [show def]	1376 gene products view in tree	biological process
☐	GO:0055047 : generative cell **mitosis** [show def]	2 gene products view in tree	biological process



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

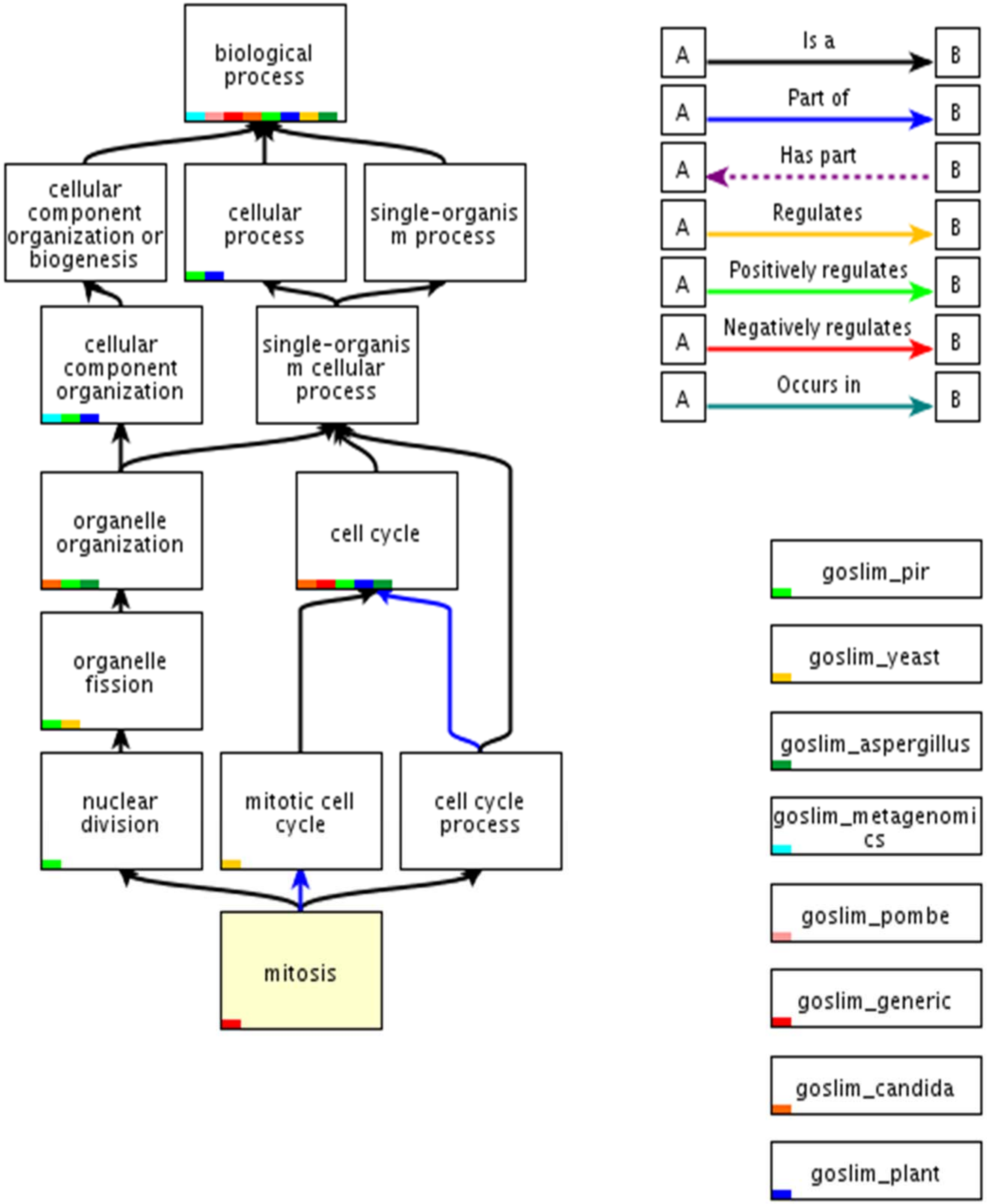
Subset Generic GO slim

Community Add usage comments for this term on the GONUTS wiki.

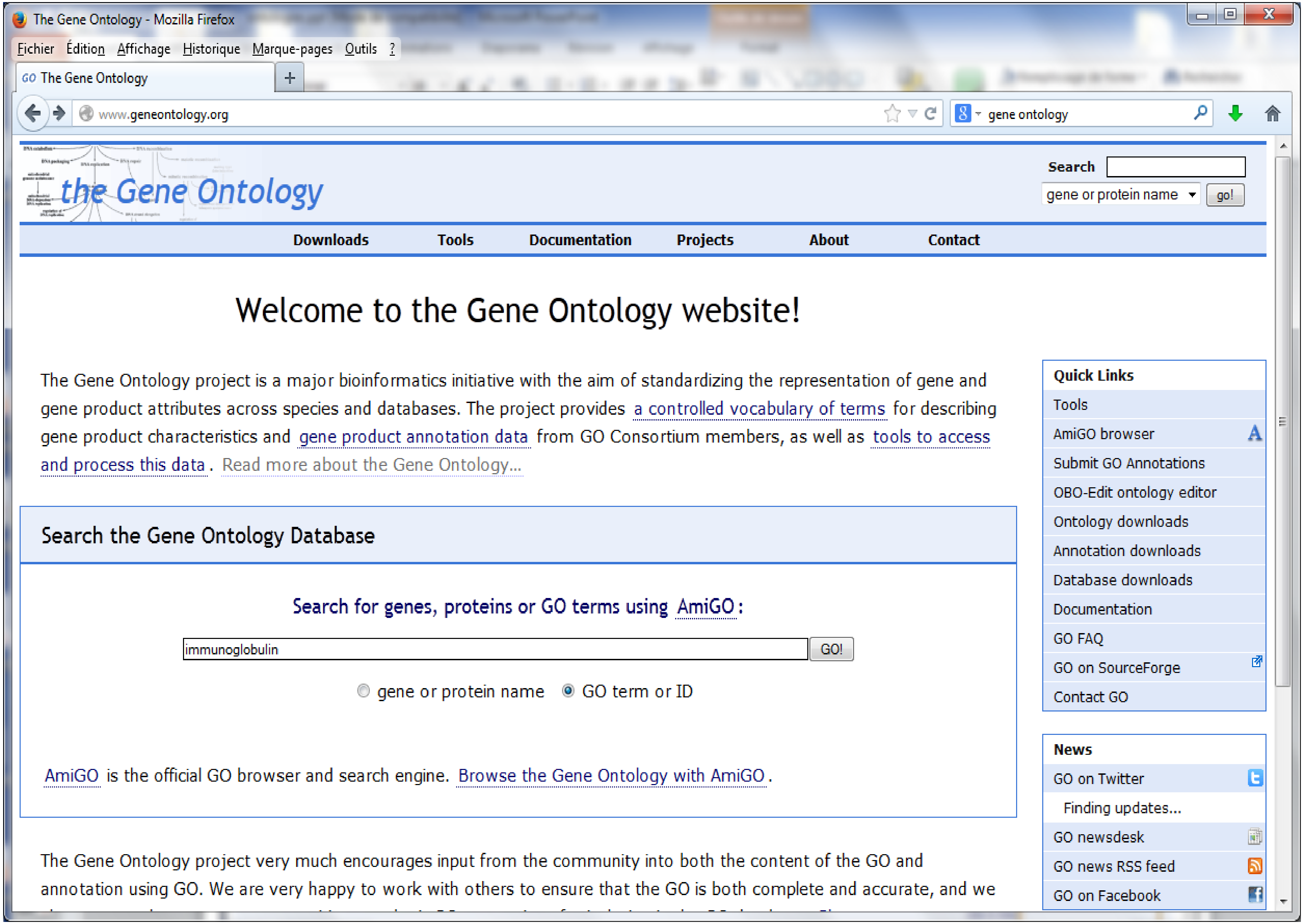
Term Neighborhood for mitosis (GO:0007067)

Ancestors of mitosis (GO:0007067)

subject	relation	object	annotations
mitosis	P part_of (inferred)	biological_process (GO:0008150)	536898
mitosis	I is_a (inferred)	cellular component organization or biogenesis (GO:0071840)	63308
mitosis	I is_a (inferred)	cellular component organization (GO:0016043)	55852
mitosis	P part_of (inferred)	cellular process (GO:0009987)	297644
mitosis	P part_of (inferred)	single-organism process (GO:0044699)	175491
mitosis	I is_a (inferred)	organelle organization (GO:0006996)	31374
mitosis	P part_of (inferred)	single-organism cellular process (GO:0044763)	144891
mitosis	P part_of (inferred)	cell cycle (GO:0007049)	18092



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



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

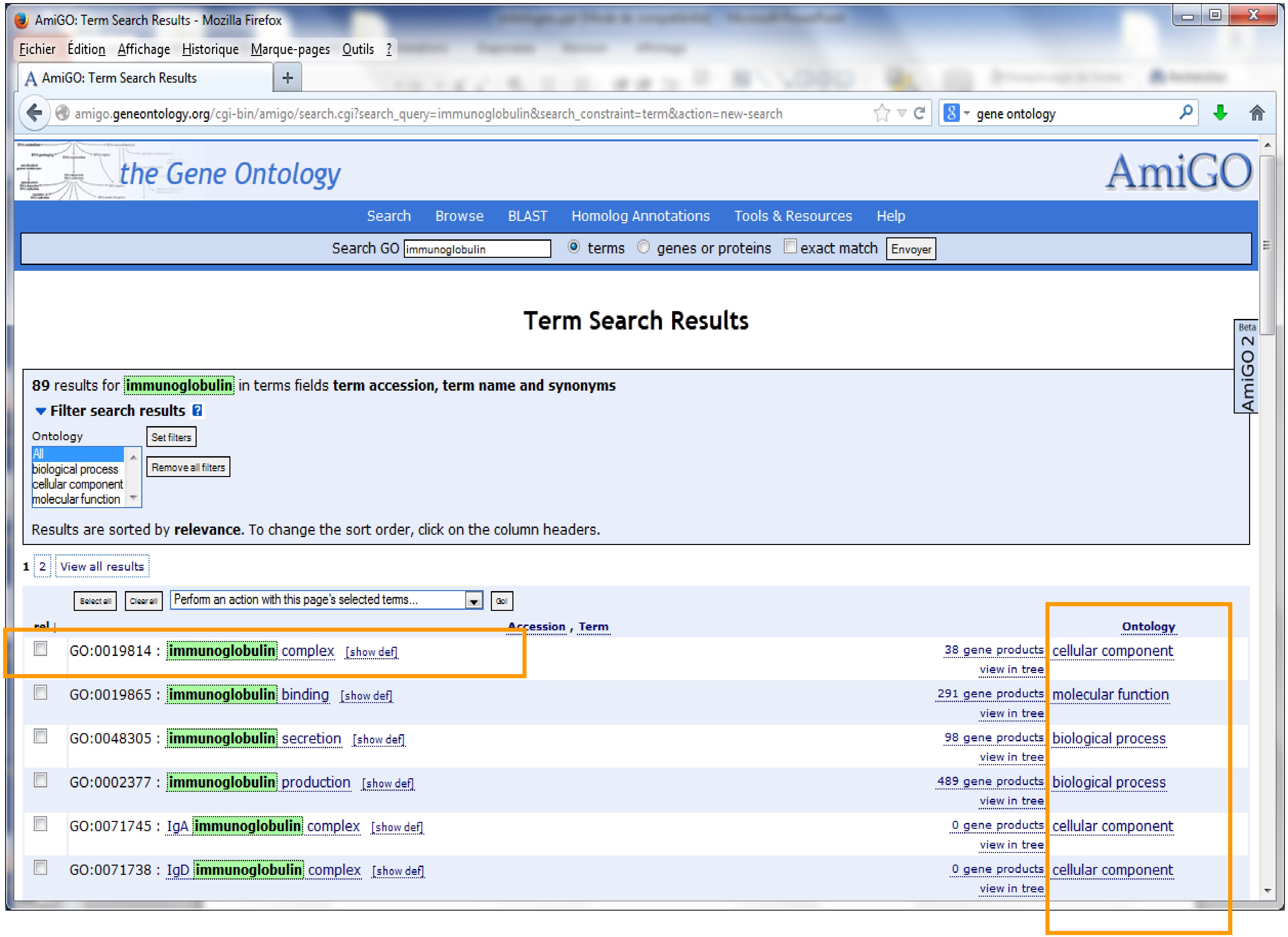
Search for genes, proteins or GO terms using AmiGO:

immunoglobulin GO!

○ 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



Search GO immunoglobulin ● terms ○ genes or proteins ☐ exact match Envoyer

Term Search Results

89 results for **immunoglobulin** in terms fields **term accession, term name and synonyms**

Filter search results

Ontology: All, biological process, cellular component, molecular function

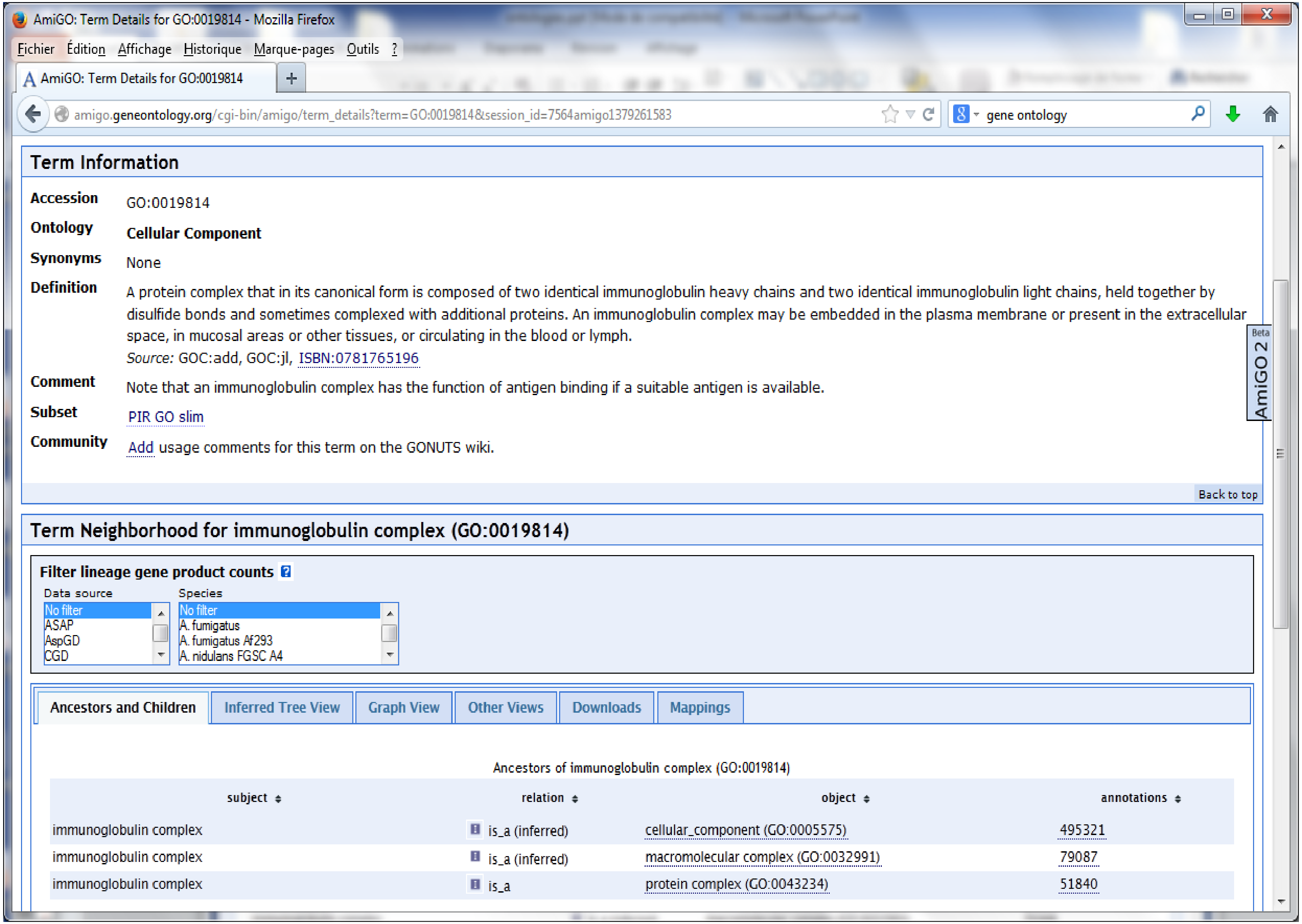
Set filters | Remove all filters

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:0019814 : **immunoglobulin** complex [show def]	38 gene products view in tree	cellular component
☐	GO:0019865 : **immunoglobulin** binding [show def]	291 gene products view in tree	molecular function
☐	GO:0048305 : **immunoglobulin** secretion [show def]	98 gene products view in tree	biological process
☐	GO:0002377 : **immunoglobulin** production [show def]	489 gene products view in tree	biological process
☐	GO:0071745 : IgA **immunoglobulin** complex [show def]	0 gene products view in tree	cellular component
☐	GO:0071738 : IgD **immunoglobulin** complex [show def]	0 gene products view in tree	cellular component



Term Information

Accession GO:0019814

Ontology **Cellular Component**

Synonyms None

Definition A protein complex that in its canonical form is composed of two identical immunoglobulin heavy chains and two identical immunoglobulin light chains, held together by disulfide bonds and sometimes complexed with additional proteins. An immunoglobulin complex may be embedded in the plasma membrane or present in the extracellular space, in mucosal areas or other tissues, or circulating in the blood or lymph.
Source: GOC:add, GOC:jl, ISBN:0781765196

Comment Note that an immunoglobulin complex has the function of antigen binding if a suitable antigen is available.

Subset PIR GO slim

Community Add usage comments for this term on the GONUTS wiki.

Term Neighborhood for immunoglobulin complex (GO:0019814)

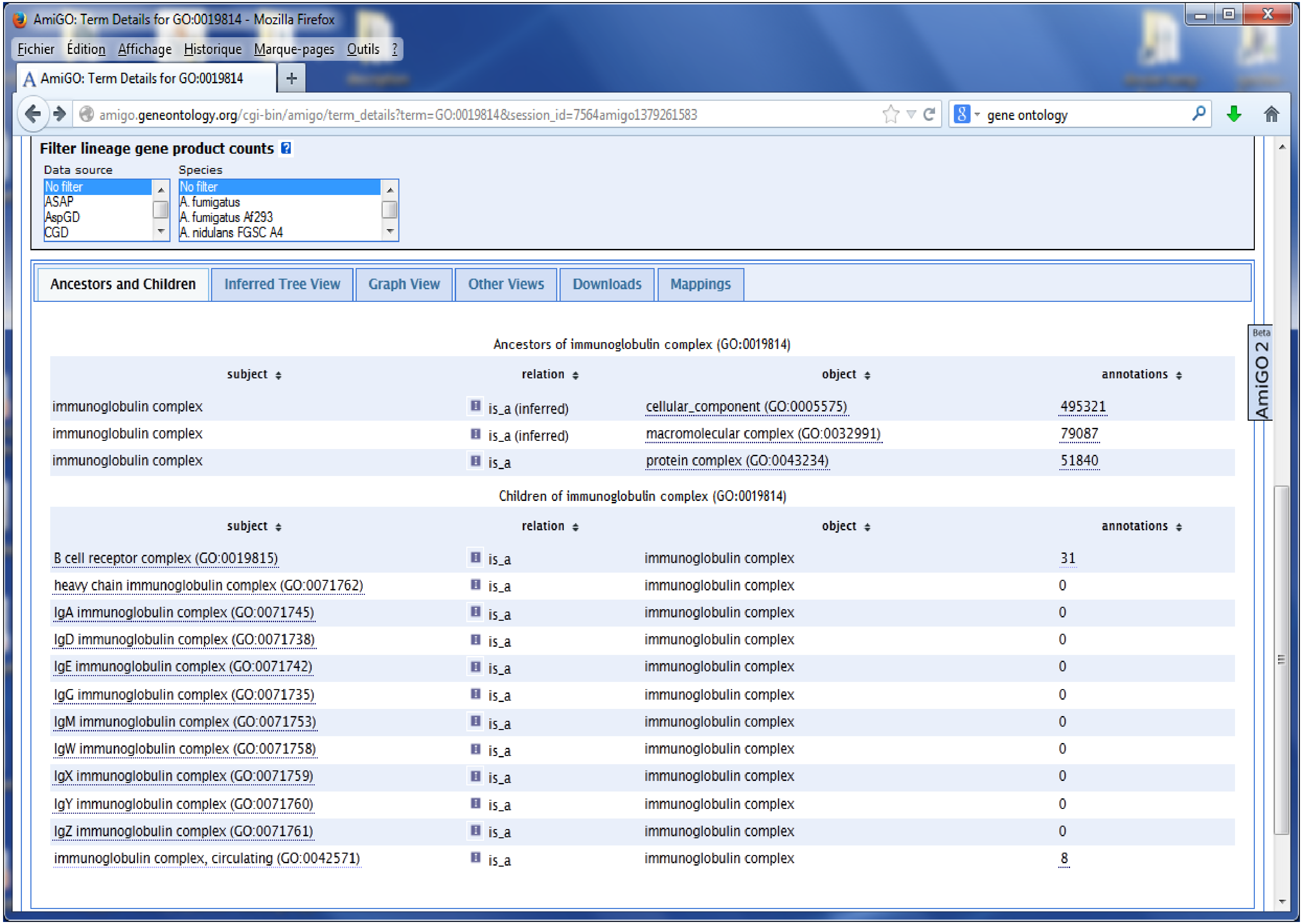
Filter lineage gene product counts

Data source: No filter, ASAP, AspGD, CGD

Species: No filter, A. fumigatus, A. fumigatus Af293, A. nidulans FGSC A4

Ancestors of immunoglobulin complex (GO:0019814)

subject	relation	object	annotations
immunoglobulin complex	is_a (inferred)	cellular_component (GO:0005575)	495321
immunoglobulin complex	is_a (inferred)	macromolecular complex (GO:0032991)	79087
immunoglobulin complex	is_a	protein complex (GO:0043234)	51840



Filter lineage gene product counts

Data source: No filter, ASAP, AspGD, CGD

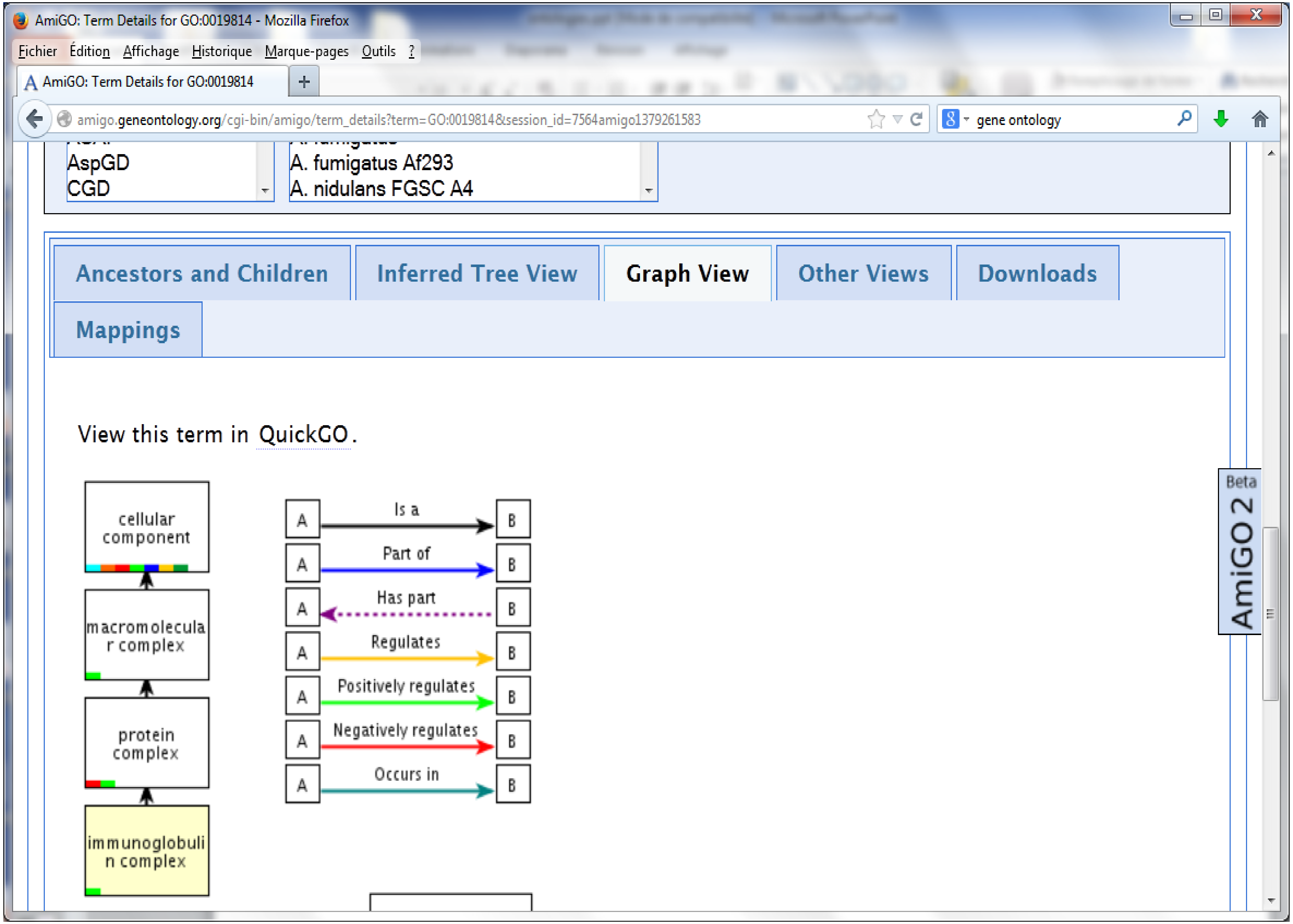
Species: No filter, A. fumigatus, A. fumigatus Af293, A. nidulans FGSC A4

Ancestors of immunoglobulin complex (GO:0019814)

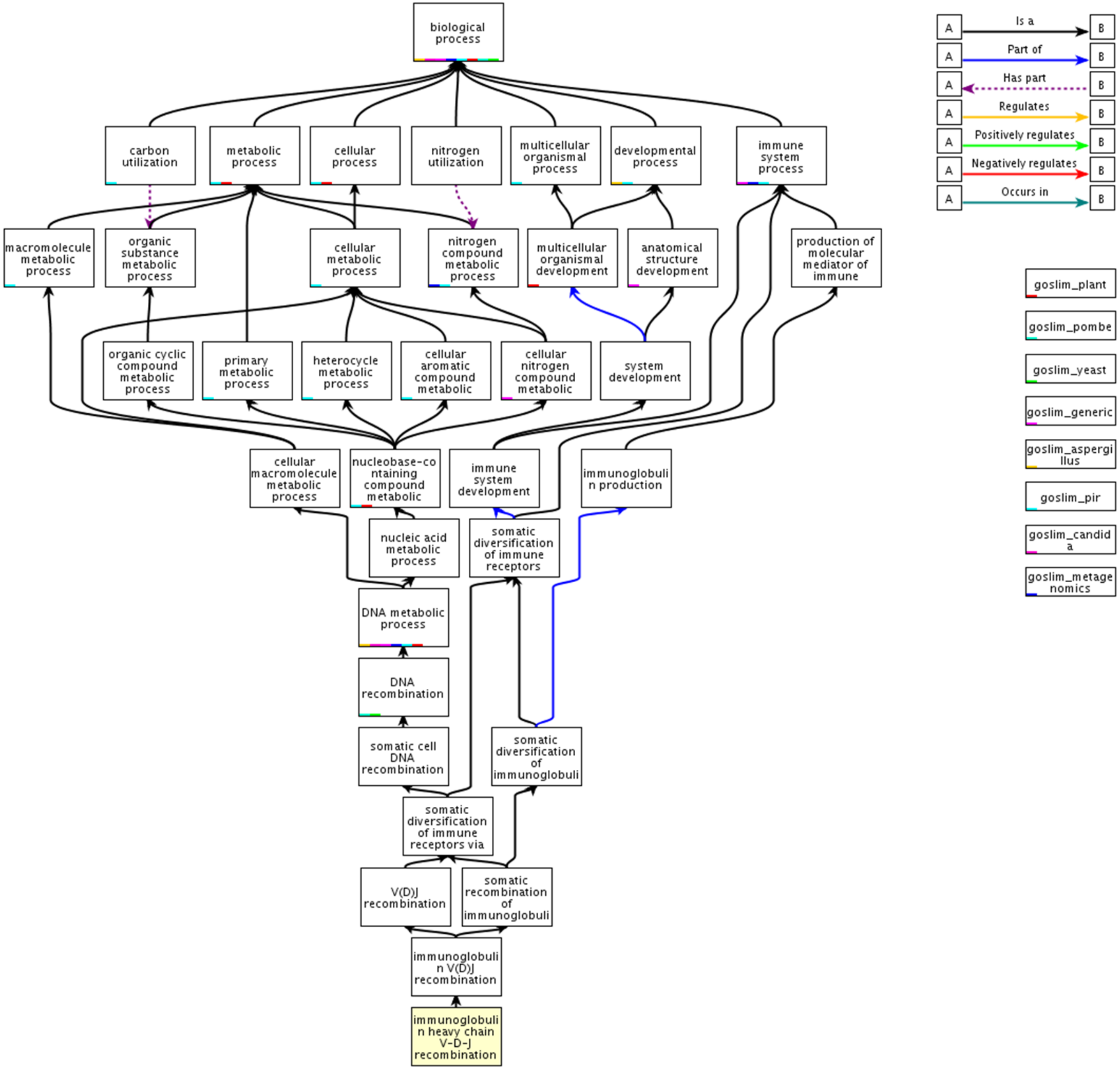
subject	relation	object	annotations
immunoglobulin complex	is_a (inferred)	cellular_component (GO:0005575)	495321
immunoglobulin complex	is_a (inferred)	macromolecular complex (GO:0032991)	79087
immunoglobulin complex	is_a	protein complex (GO:0043234)	51840

Children of immunoglobulin complex (GO:0019814)

subject	relation	object	annotations
B cell receptor complex (GO:0019815)	is_a	immunoglobulin complex	31
heavy chain immunoglobulin complex (GO:0071762)	is_a	immunoglobulin complex	0
IgA immunoglobulin complex (GO:0071745)	is_a	immunoglobulin complex	0
IgD immunoglobulin complex (GO:0071738)	is_a	immunoglobulin complex	0
IgE immunoglobulin complex (GO:0071742)	is_a	immunoglobulin complex	0
IgG immunoglobulin complex (GO:0071735)	is_a	immunoglobulin complex	0
IgM immunoglobulin complex (GO:0071753)	is_a	immunoglobulin complex	0
IgW immunoglobulin complex (GO:0071758)	is_a	immunoglobulin complex	0
IgX immunoglobulin complex (GO:0071759)	is_a	immunoglobulin complex	0
IgY immunoglobulin complex (GO:0071760)	is_a	immunoglobulin complex	0
IgZ immunoglobulin complex (GO:0071761)	is_a	immunoglobulin complex	0
immunoglobulin complex, circulating (GO:0042571)	is_a	immunoglobulin complex	8



View this term in QuickGO.



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

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 les séquences d’IG (.. et des TR): identification, classification des gènes, description, numérotation des acides aminés, obtention ?**

Sequence Ontology

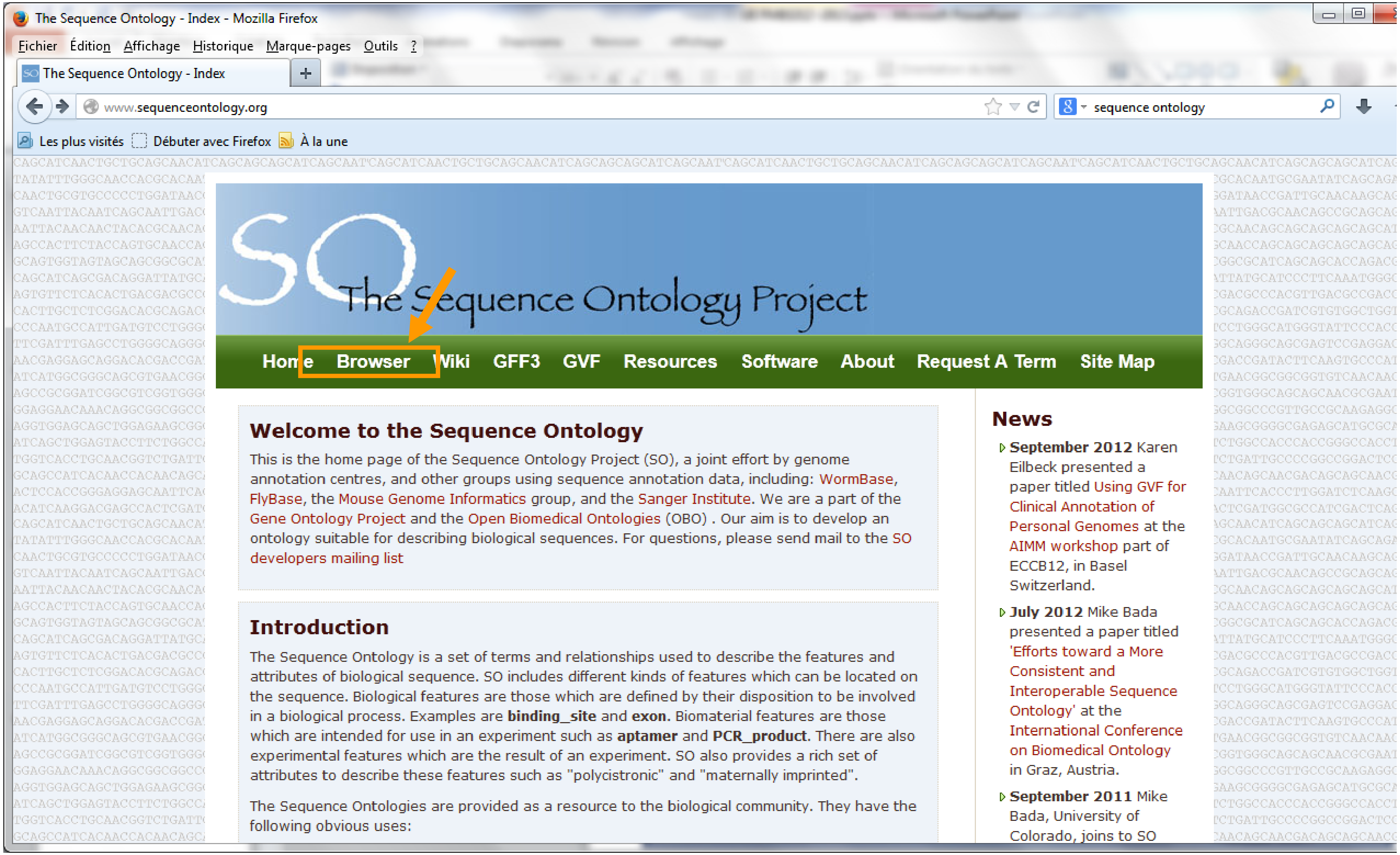
(http://www.sequenceontology.org/)

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

- **proposer une représentation structurée des annotations**

- **vocabulaire pour la description des mutations**

Sequence Ontology: une ontologie pour décrire les séquences



Welcome to the Sequence Ontology

This is the home page of the Sequence Ontology Project (SO), a joint effort by genome annotation centres, and other groups using sequence annotation data, including: WormBase, FlyBase, the Mouse Genome Informatics group, and the Sanger Institute. We are a part of the Gene Ontology Project and the Open Biomedical Ontologies (OBO) . Our aim is to develop an ontology suitable for describing biological sequences. For questions, please send mail to the SO developers mailing list

Introduction

The Sequence Ontology is a set of terms and relationships used to describe the features and attributes of biological sequence. SO includes different kinds of features which can be located on the sequence. Biological features are those which are defined by their disposition to be involved in a biological process. Examples are **binding_site** and **exon.** Biomaterial features are those which are intended for use in an experiment such as **aptamer** and **PCR_product**. There are also experimental features which are the result of an experiment. SO also provides a rich set of attributes to describe these features such as "polycistronic" and "maternally imprinted".

The Sequence Ontologies are provided as a resource to the biological community. They have the following obvious uses:

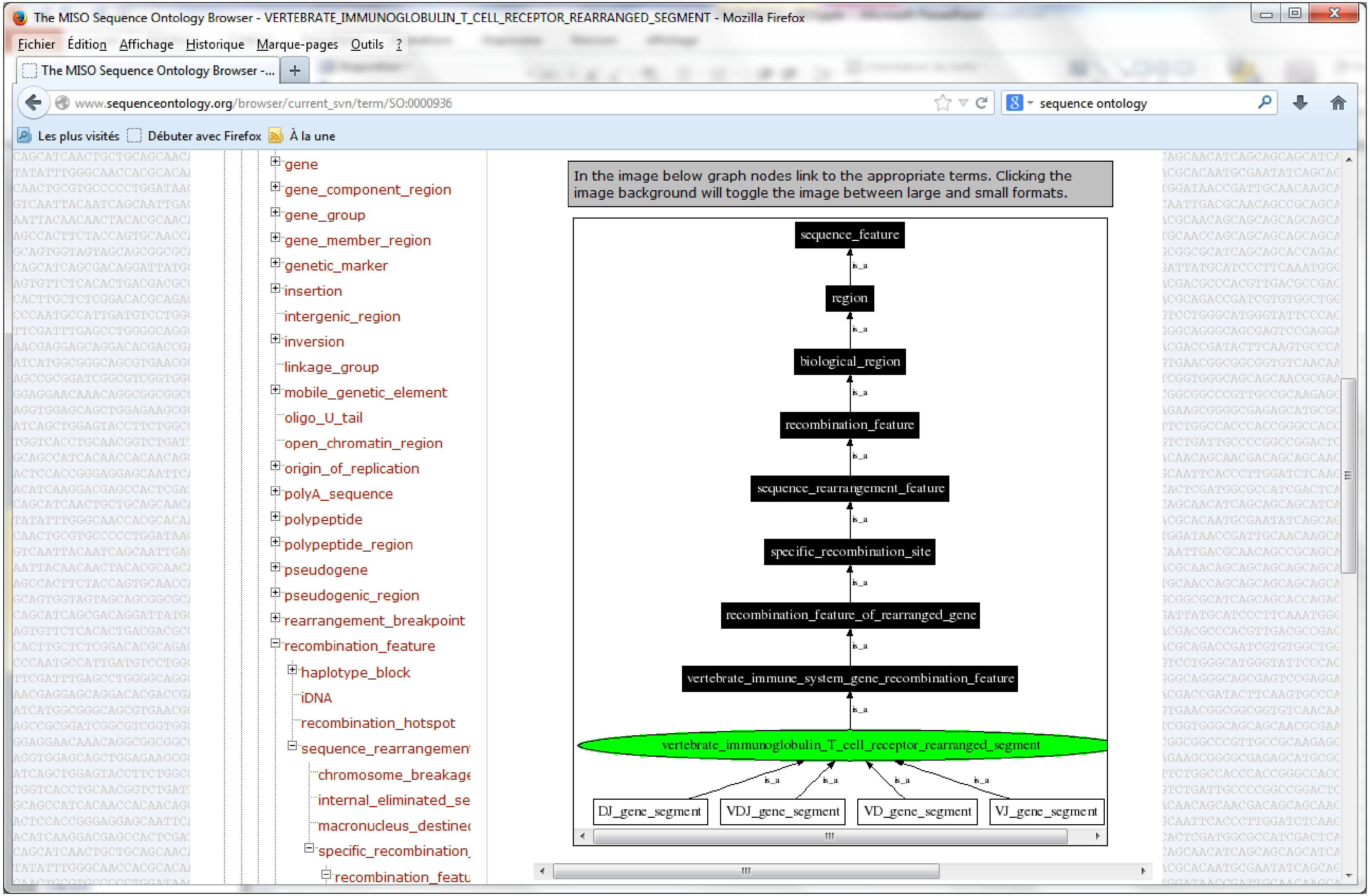
News

- **September 2012** Karen Eilbeck presented a paper titled Using GVF for Clinical Annotation of Personal Genomes at the AIMM workshop part of ECCB12, in Basel Switzerland.
- **July 2012** Mike Bada presented a paper titled 'Efforts toward a More Consistent and Interoperable Sequence Ontology' at the International Conference on Biomedical Ontology in Graz, Austria.
- **September 2011** Mike Bada, University of Colorado, joins to SO

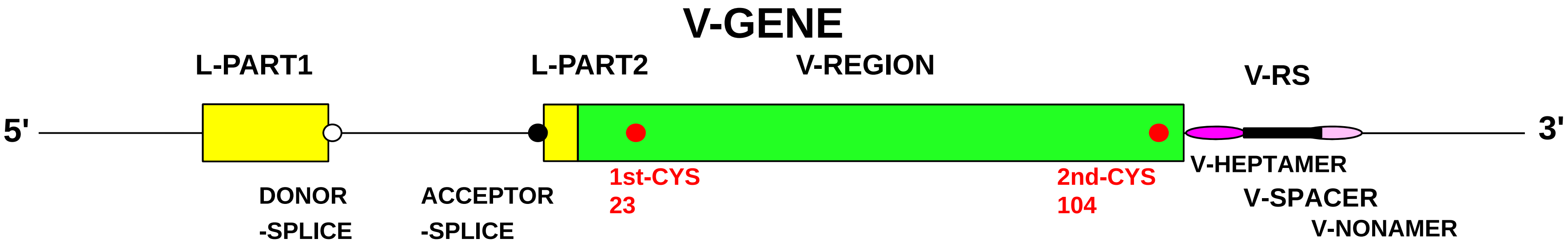
Welcome to MISO, the Sequence Ontology Browser

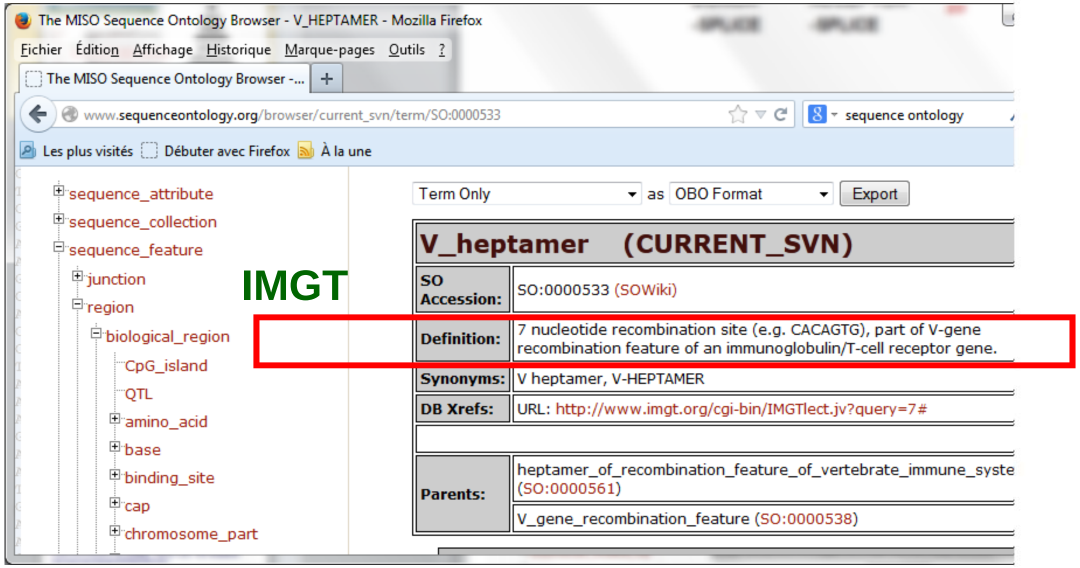
- Search for a SO term by entering a SO term name or synonym in the query box above;
- Explore the structure of SO and browse for SO terms using the expandable, cascading tree on the left;
- Go to the detail page for a term where you can:
  - Get details about a term, it's definition and relationships;
  - 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



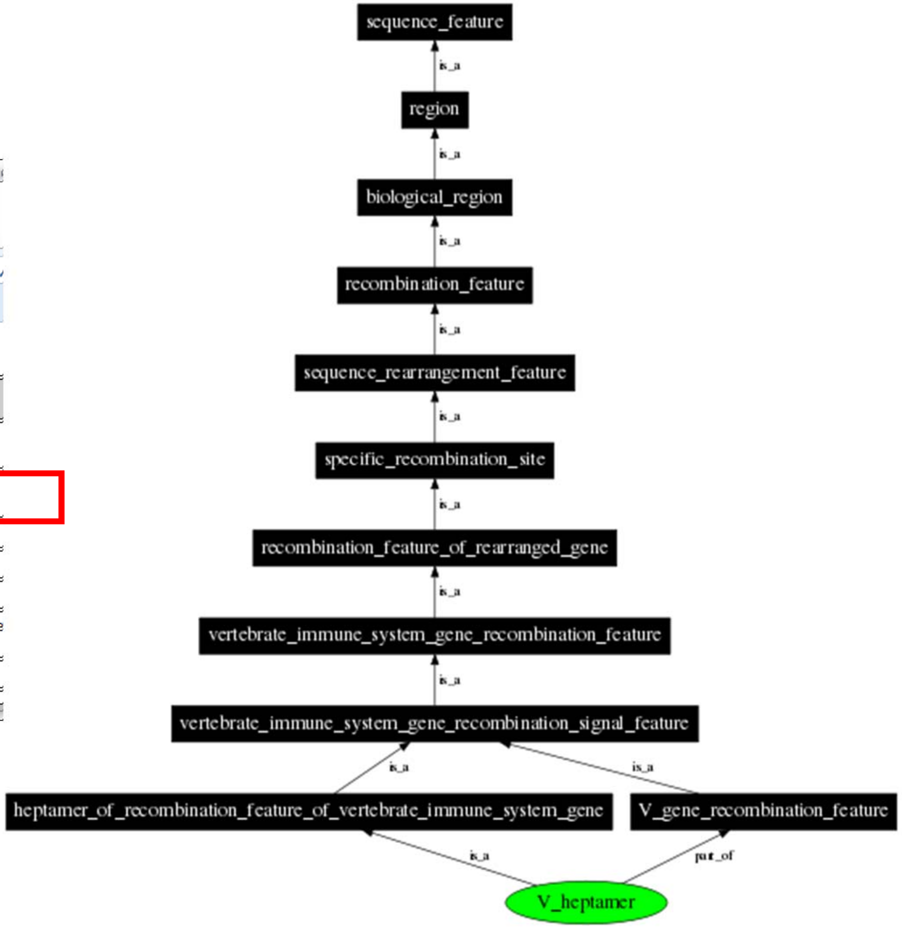
In the image below graph nodes link to the appropriate terms. Clicking the image background will toggle the image between large and small formats.





V_heptamer (CURRENT_SVN)

SO Accession:	SO:0000533 (SOWiki)
Definition:	7 nucleotide recombination site (e.g. CACAGTG), part of V-gene recombination feature of an immunoglobulin/T-cell receptor gene.
Synonyms:	V heptamer, V-HEPTAMER
DB Xrefs:	URL: http://www.imgt.org/cgi-bin/IMGTlect.jv?query=7#
Parents:	heptamer_of_recombination_feature_of_vertebrate_immune_syste (SO:0000561) V_gene_recombination_feature (SO:0000538)



Dans le domaine de l’immunogénétique

Limites de SO:

- Il est nécessaire de prendre en compte d’autres informations:
type de gènes, type de chaînes, …

- Il faut des relations plus précises que « part_of » entre les motifs

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

=> nécessité d’une ontologie spécifique

IMGT-ONTOLOGY

http://www.imgt.org

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

IMGT-ONTOLOGY repose sur 7 axiomes:

- **IDENTIFICATION**
- **CLASSIFICATION**
- **DESCRIPTION**
- **NUMEROTATION**
- **LOCALIZATION**
- **ORIENTATION**
- **OBTENTION**

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

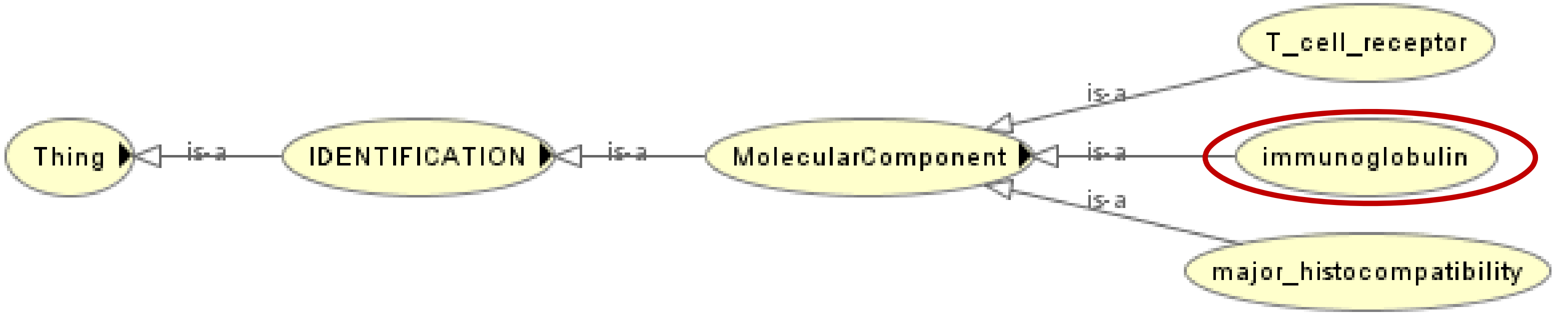
Version v1.0.1 sur le site BioPortal

(http://bioportal.bioontology.org/ontologies/1491)

1 – Concepts d'identification d'IMGT-ONTOLOGY: keywords

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

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

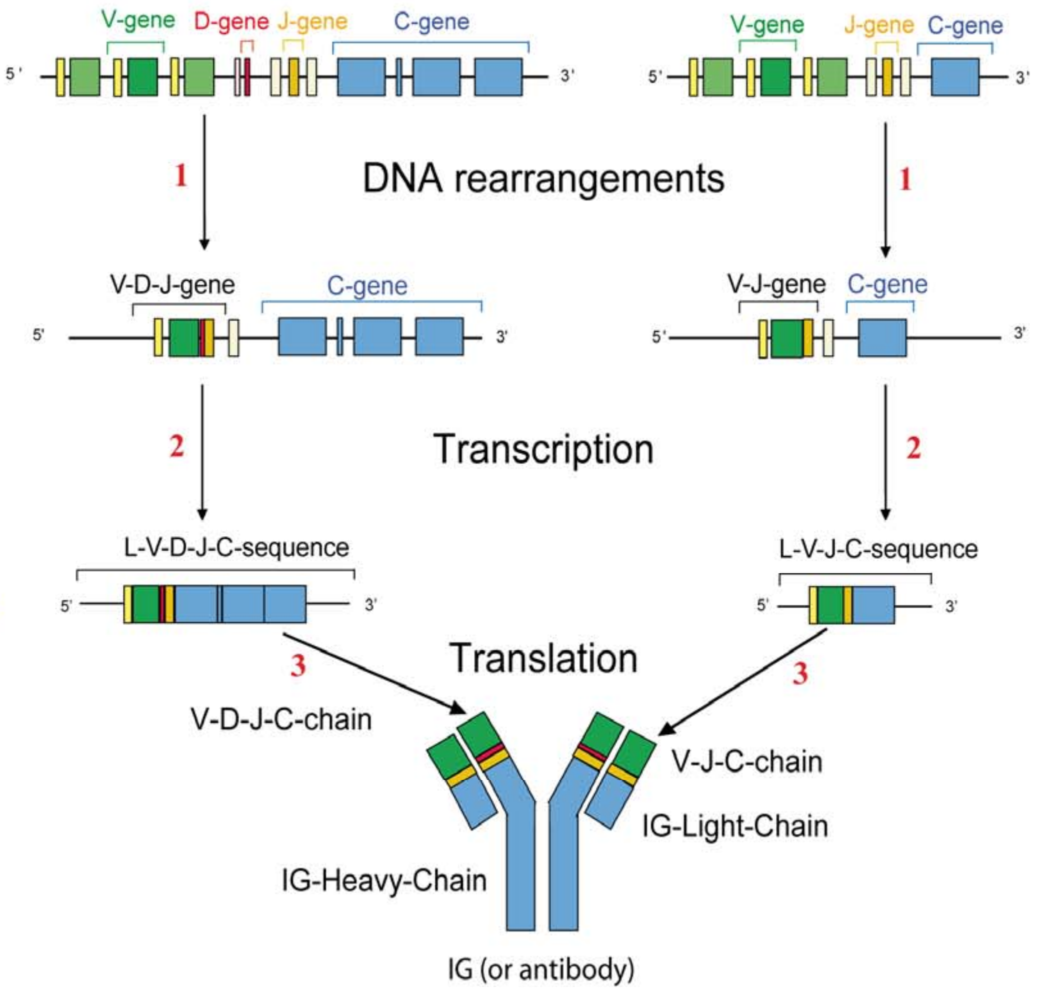


Synthèse des immunoglobulines

http://www.imgt.org

Chaînes lourde (IGH)

Chaînes légère (IGK/IGL)

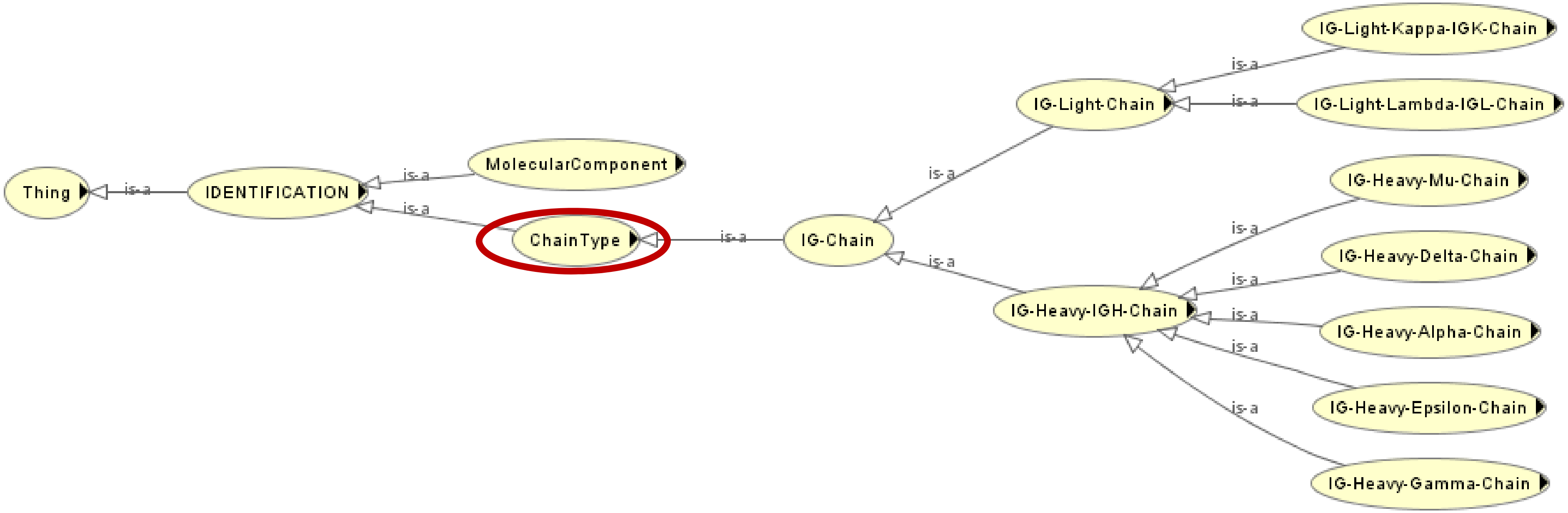


	"MoleculeType" concept	**"ConfigurationType" concept**
Genome	gDNA	germline **(undefined for C)**
	gDNA	rearranged **(undefined for C)**
Transcriptome	mRNA (or *in vitro* cDNA)	rearranged
Proteome	protein	rearranged

1 – Concepts d'identification d'IMGT-ONTOLOGY: keywords

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

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

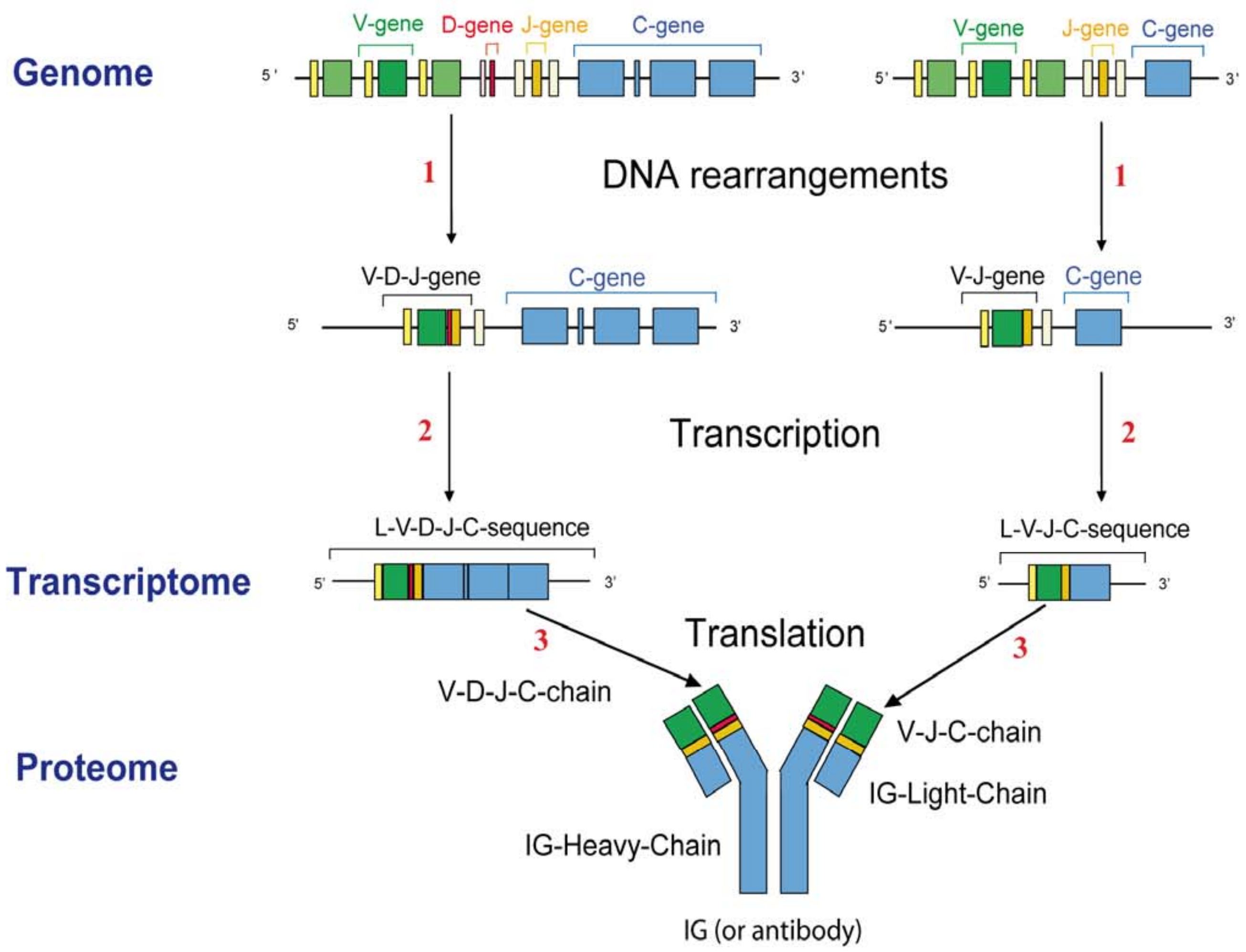


Synthèse des immunoglobulines

http://www.imgt.org

Chaînes lourde
(IGH)

Chaînes légère
(IGK/IGL)

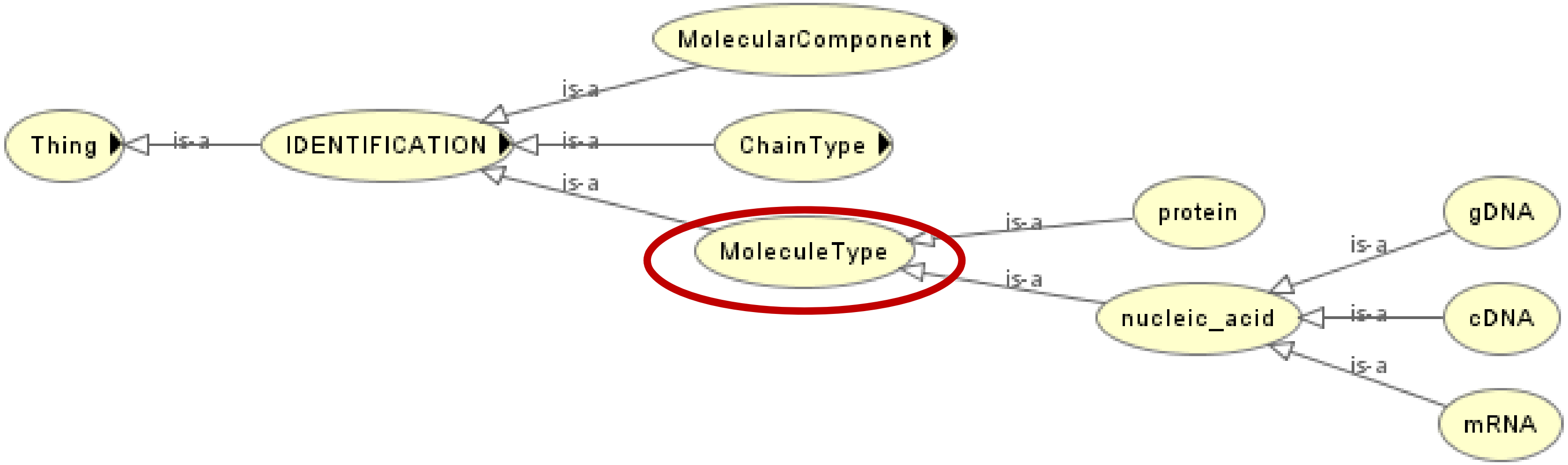


"MoleculeType" concept	**"ConfigurationType" concept**
gDNA	germline **(undefined for C)**
gDNA	rearranged **(undefined for C)**
mRNA (or *in vitro* cDNA)	rearranged
protein	rearranged

1 – Concepts d'identification d'IMGT-ONTOLOGY: keywords

Les types de molécules (MoleculeType) :

- gDNA
- mRNA
- cDNA
- protein

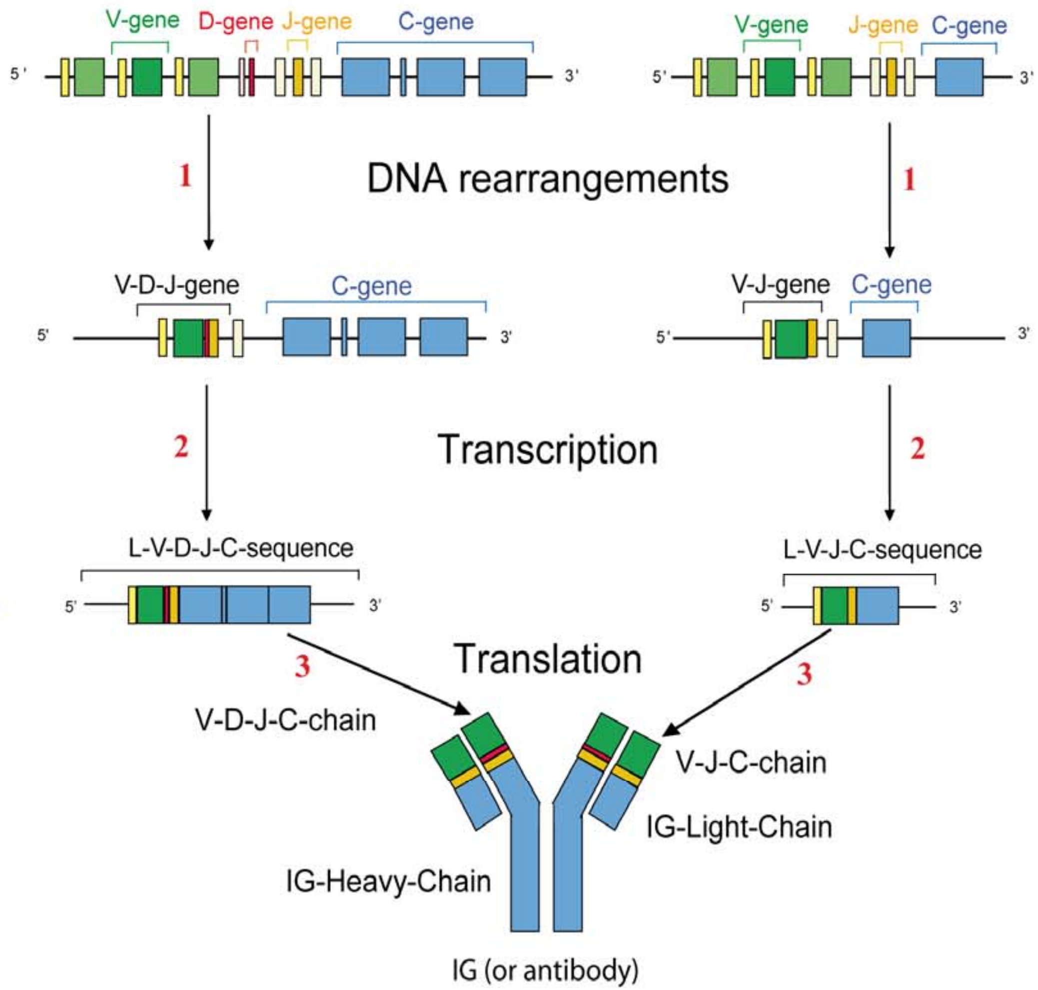


Synthèse des immunoglobulines

http://www.imgt.org

Chaînes lourde
(IGH)

Chaînes légère
(IGK/IGL)

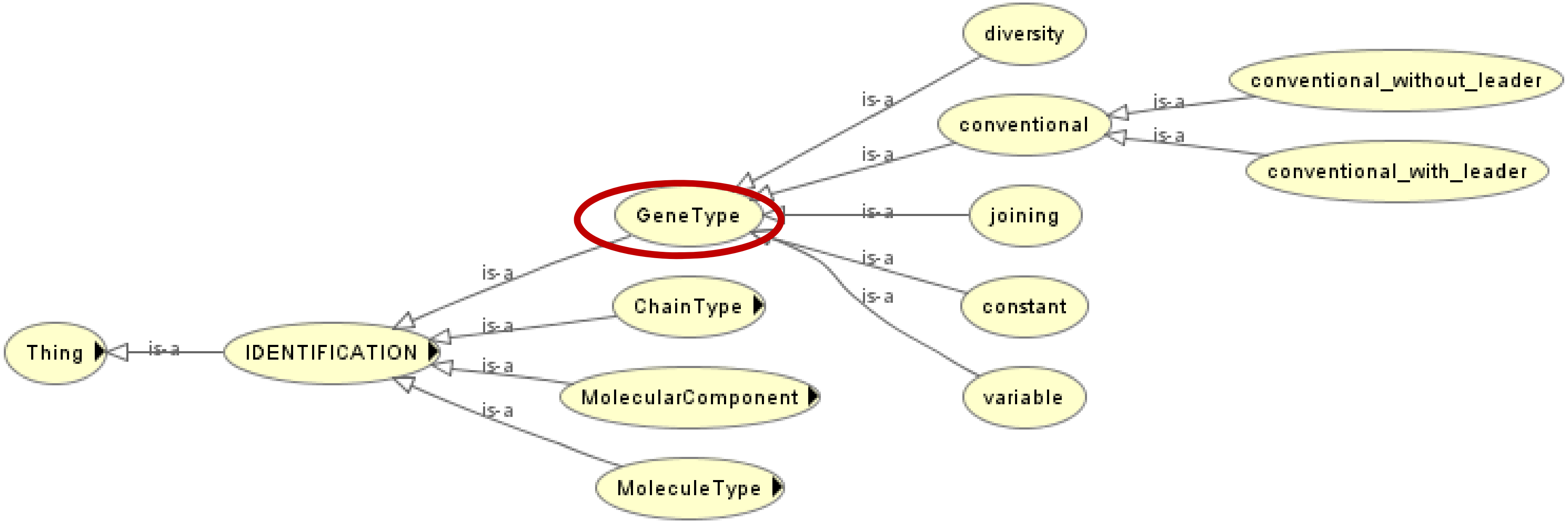


	"MoleculeType" concept	**"ConfigurationType" concept**
Genome	gDNA	germline **(undefined for C)**
	gDNA	rearranged **(undefined for C)**
Transcriptome	mRNA (or *in vitro* cDNA)	rearranged
Proteome	protein	rearranged

1 – Concepts d’identification d’IMGT-ONTOLOGY: keywords

Les types de gène (GeneType) :

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

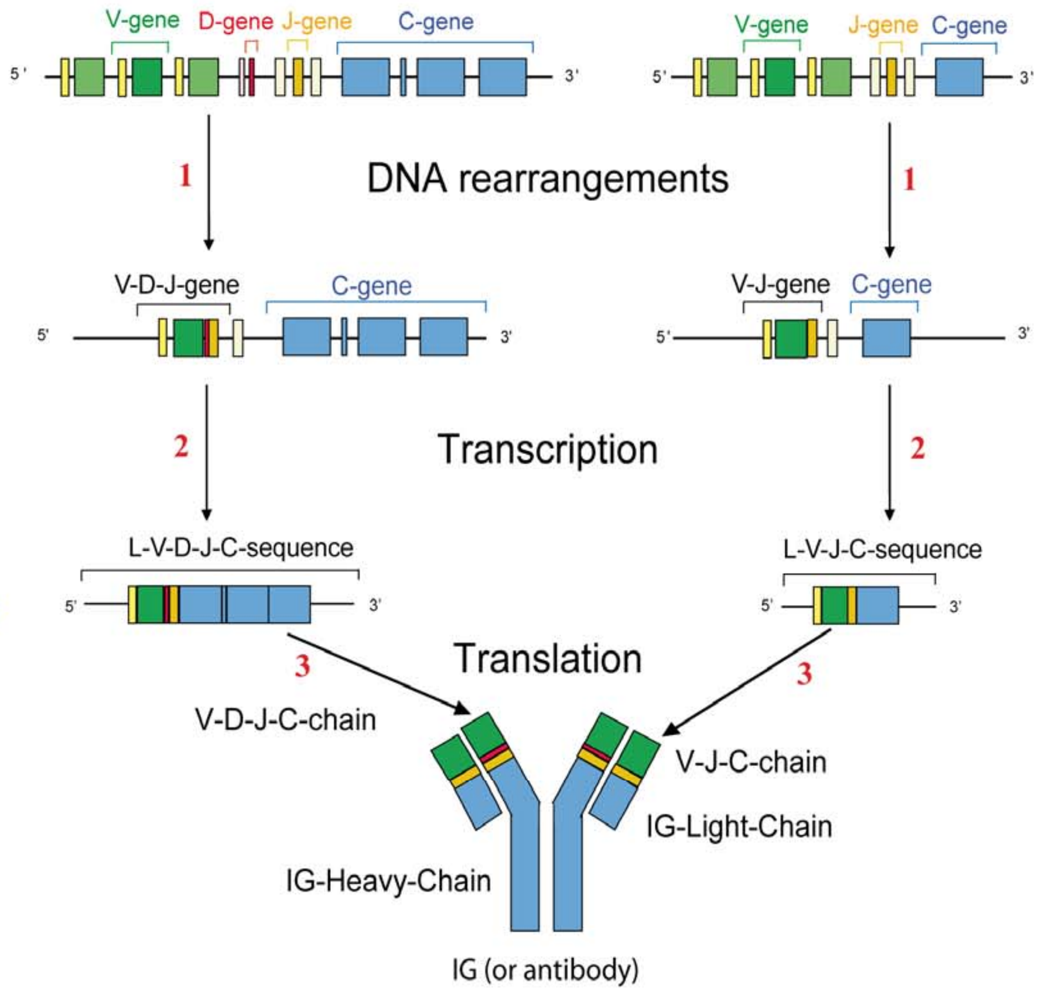


Synthèse des immunoglobulines

IMGT
http://www.imgt.org

Chaînes lourde
(IGH)

Chaînes légère
(IGK/IGL)

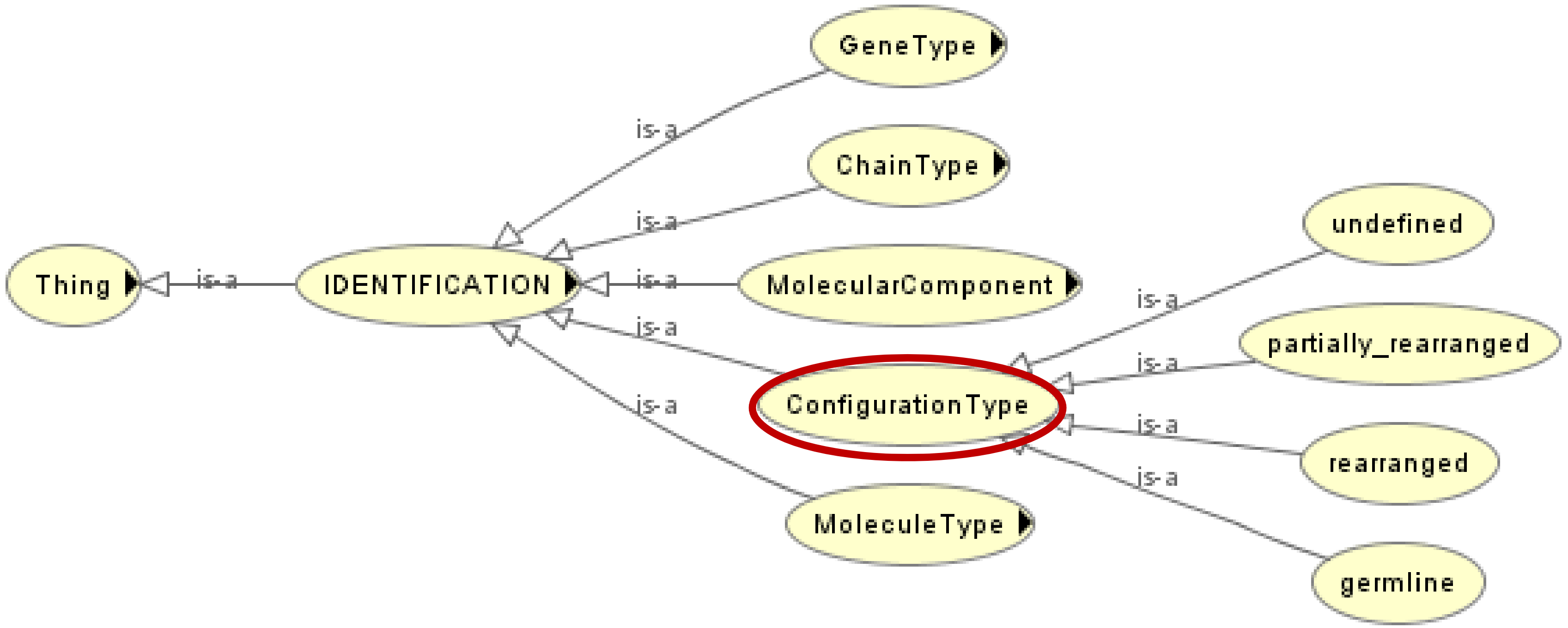


	"MoleculeType" concept	**"ConfigurationType" concept**
Genome	gDNA	germline **(undefined for C)**
	gDNA	rearranged **(undefined for C)**
Transcriptome	mRNA (or *in vitro* cDNA)	rearranged
Proteome	protein	rearranged

1 – Concepts d'identification d'IMGT-ONTOLOGY: keywords

Les types de configuration (ConfigurationType) :

- germline
- rearranged
- partially-rearranged
- undefined

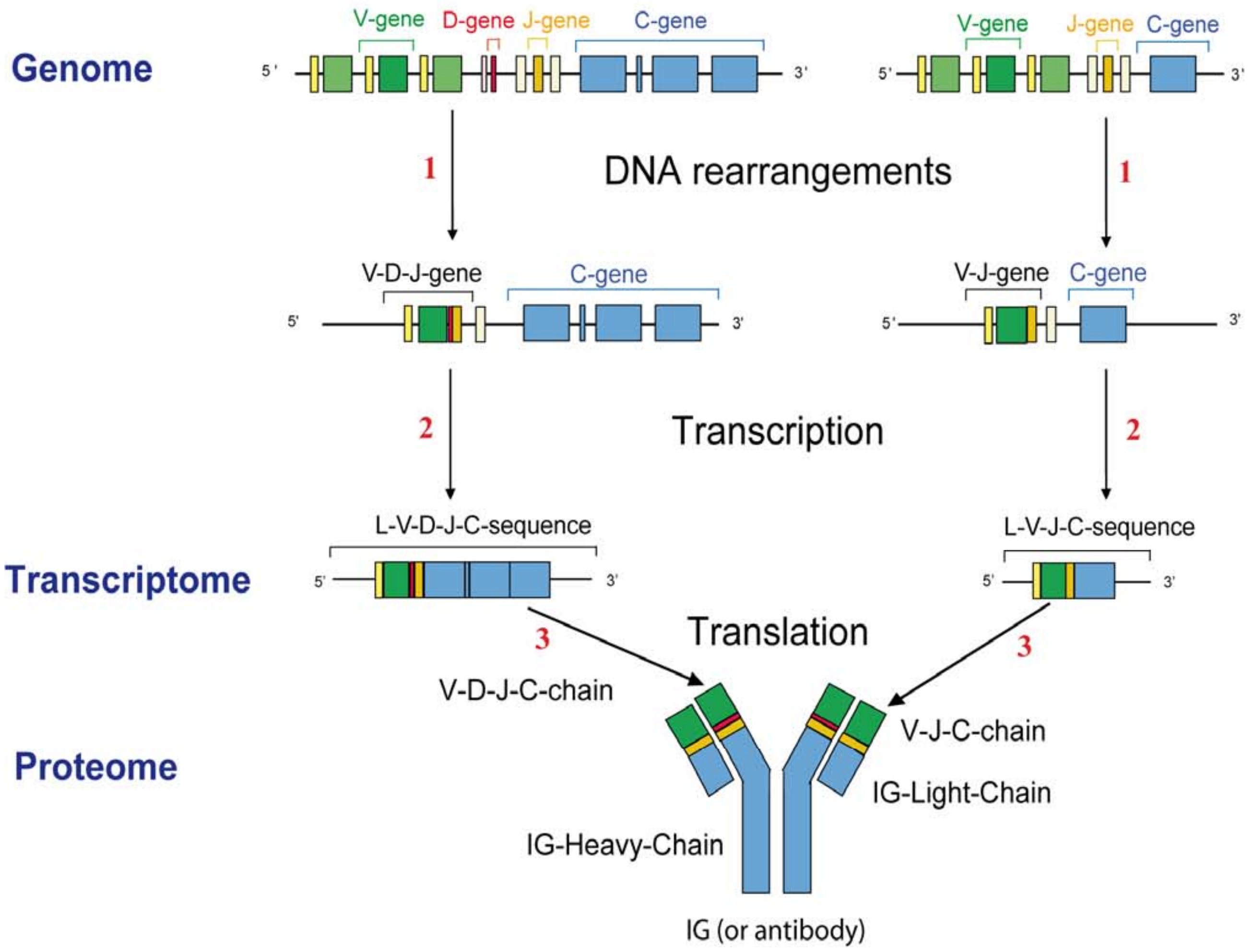


Synthèse des immunoglobulines

ImMunoGeneTics Information system®
http://www.imgt.org

Chaînes lourde (IGH)

Chaînes légère (IGK/IGL)

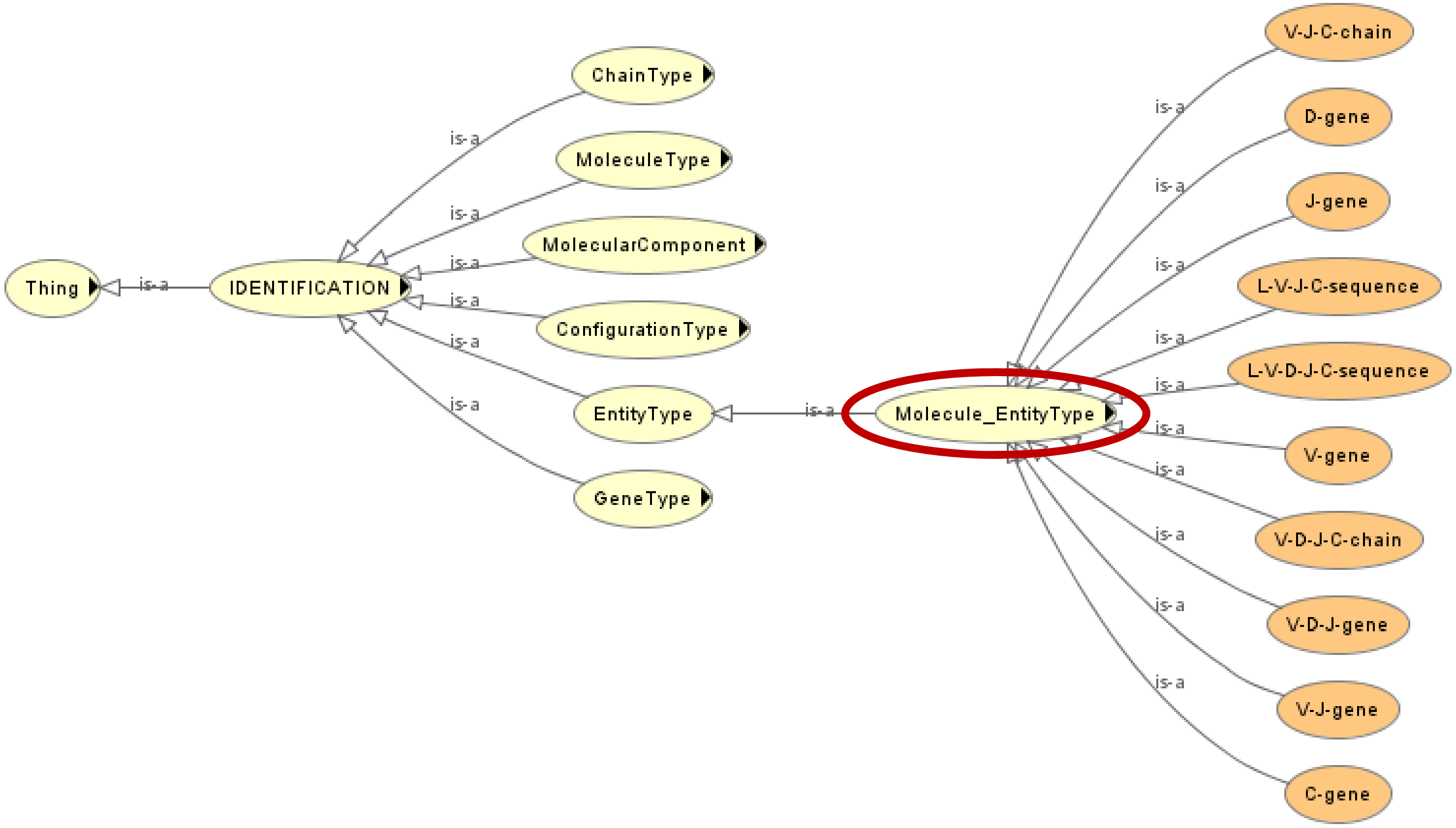


	"MoleculeType" concept	"ConfigurationType" concept
Genome	gDNA	germline **(undefined for C)**
	gDNA	rearranged **(undefined for C)**
Transcriptome	mRNA (or *in vitro* cDNA)	rearranged
Proteome	protein	rearranged

1 – Concepts d'identification d'IMGT-ONTOLOGY: keywords

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

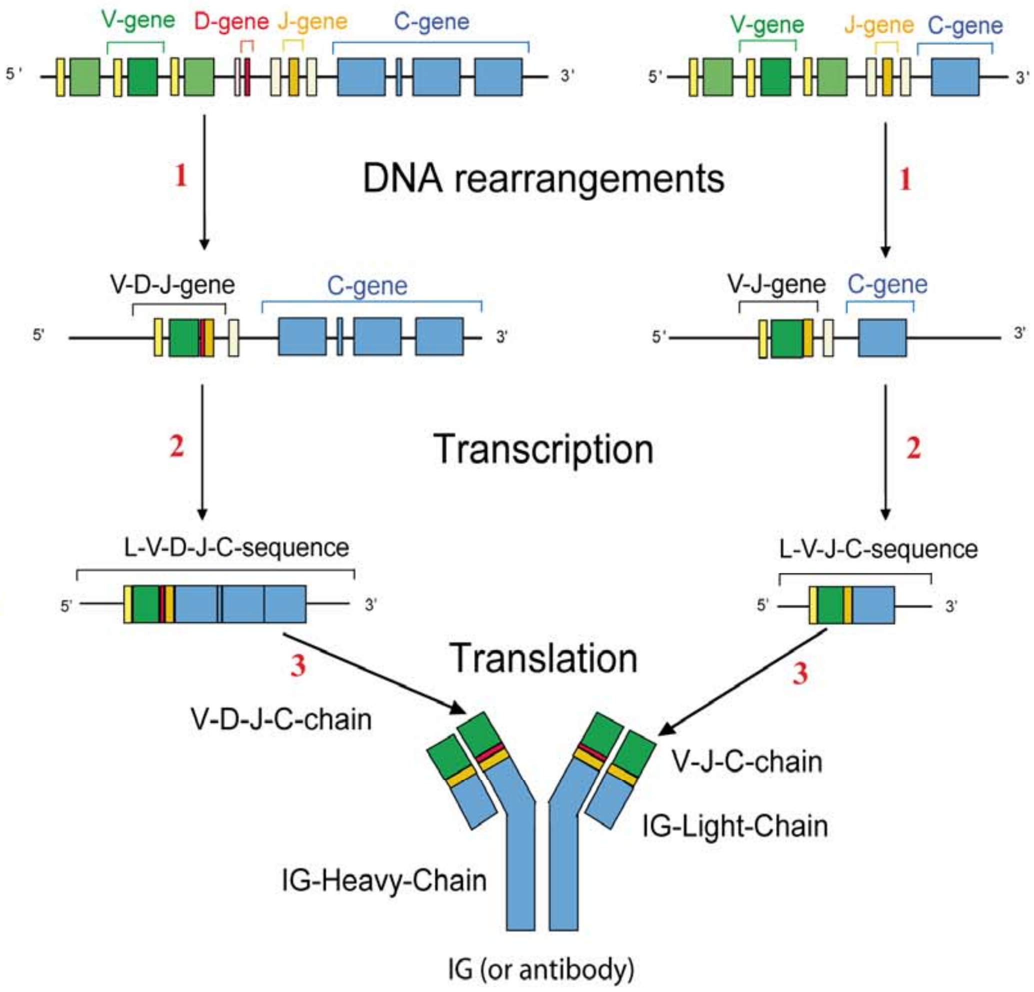


Synthèse des immunoglobulines

http://www.imgt.org

Chaînes lourde (IGH)

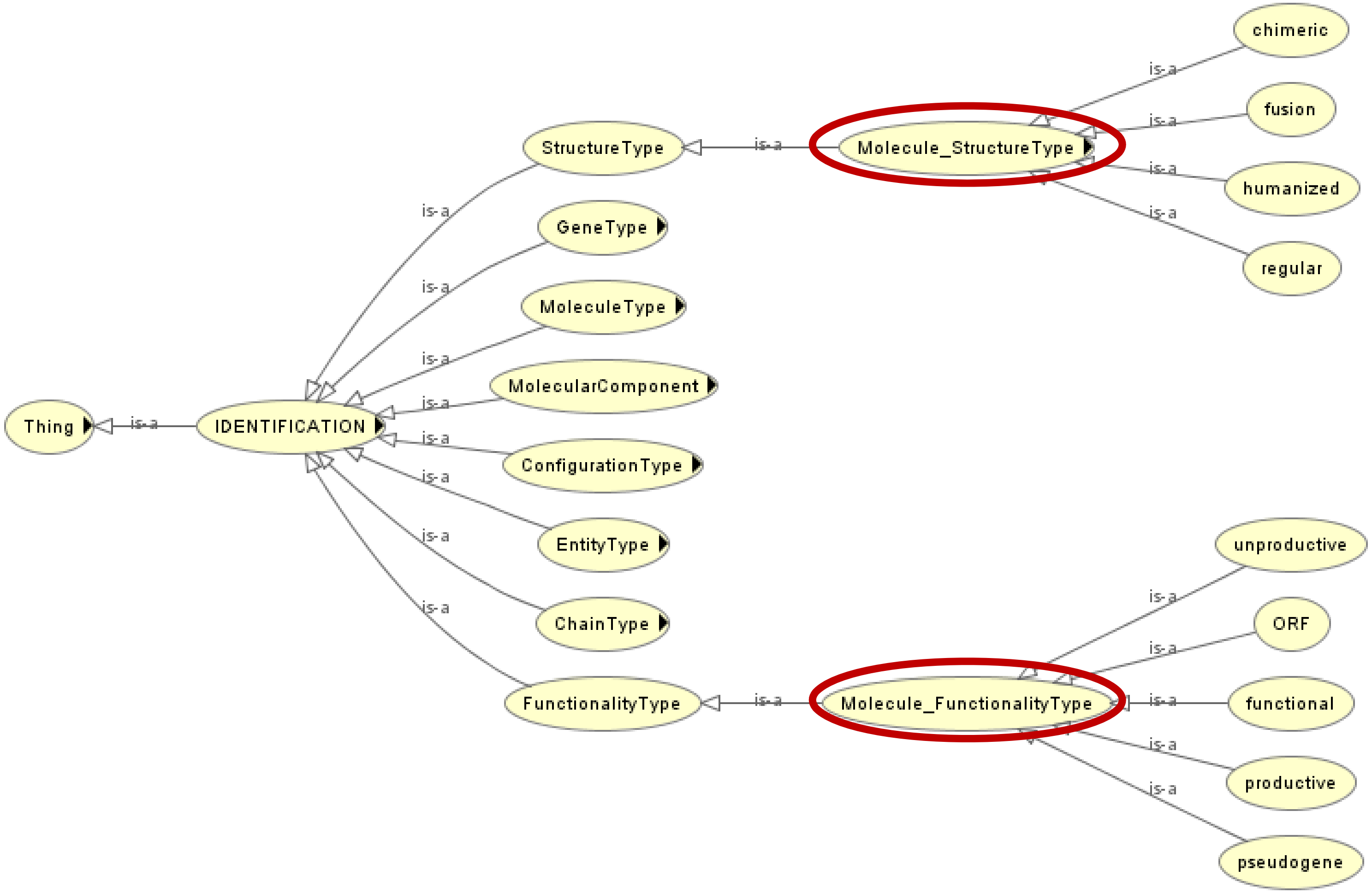
Chaînes légère (IGK/IGL)



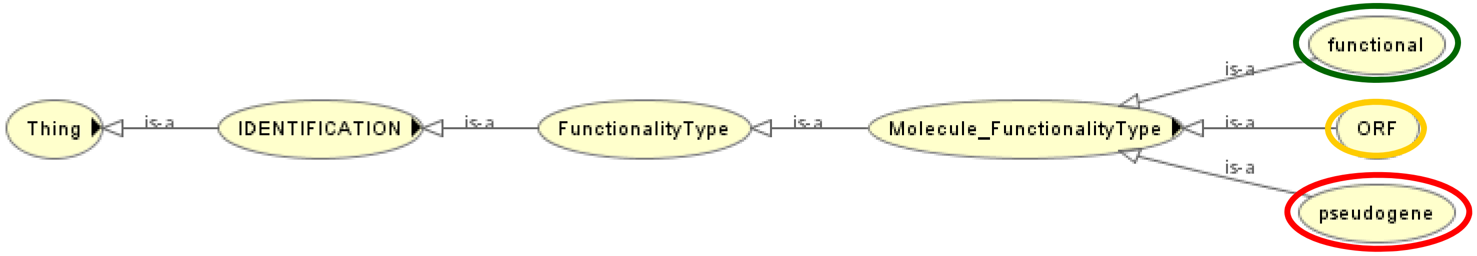
	"MoleculeType" concept	"ConfigurationType" concept
Genome	gDNA	germline **(undefined for C)**
	gDNA	rearranged **(undefined for C)**
Transcriptome	mRNA (or *in vitro* cDNA)	rearranged
Proteome	protein	rearranged

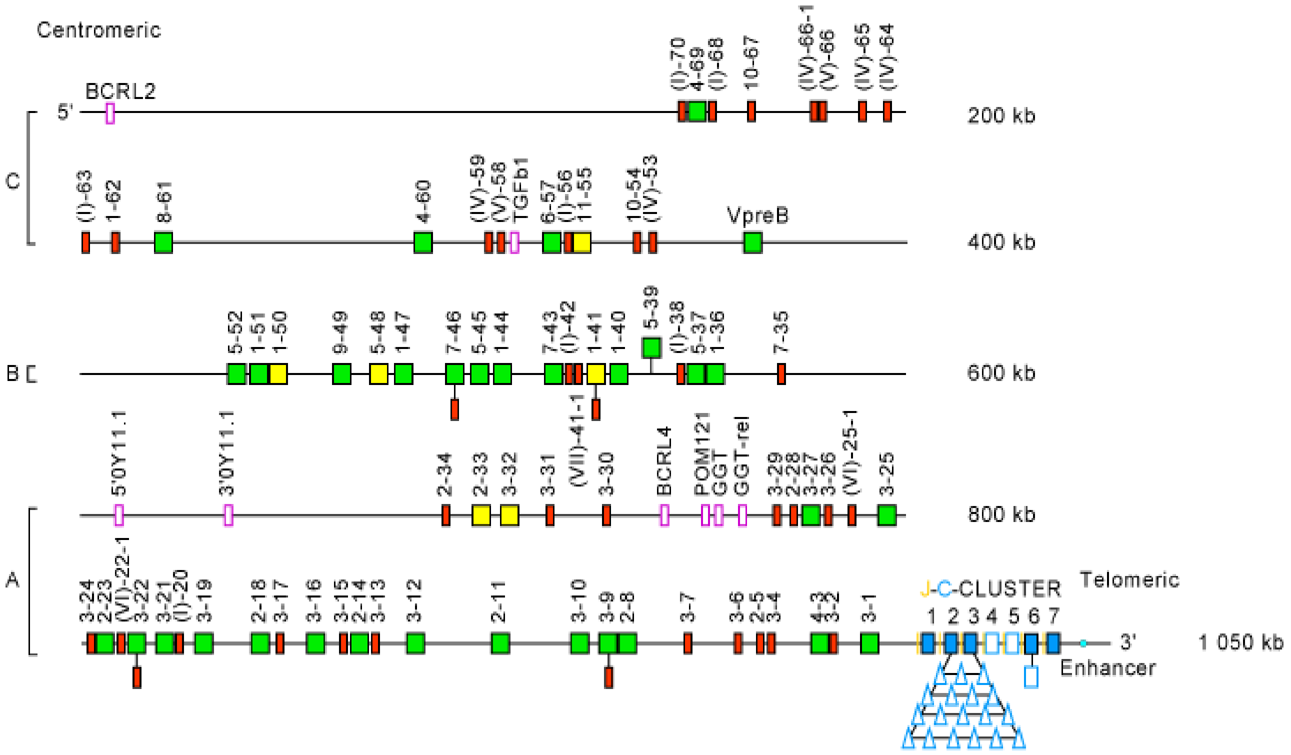
1 – Concepts d’identification d’IMGT-ONTOLOGY: keywords

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



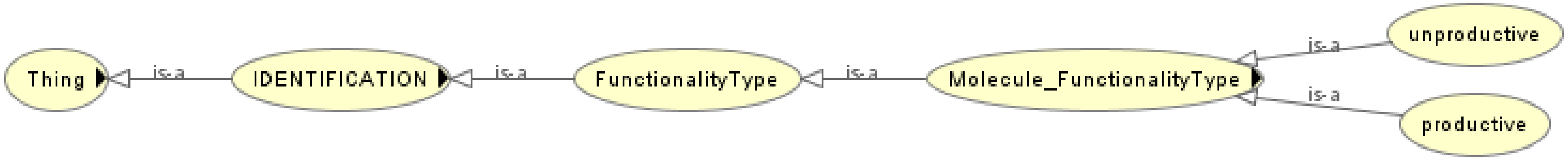
Les types de fonctionnalité pour les entités moléculaires en configuration germline ou undefined

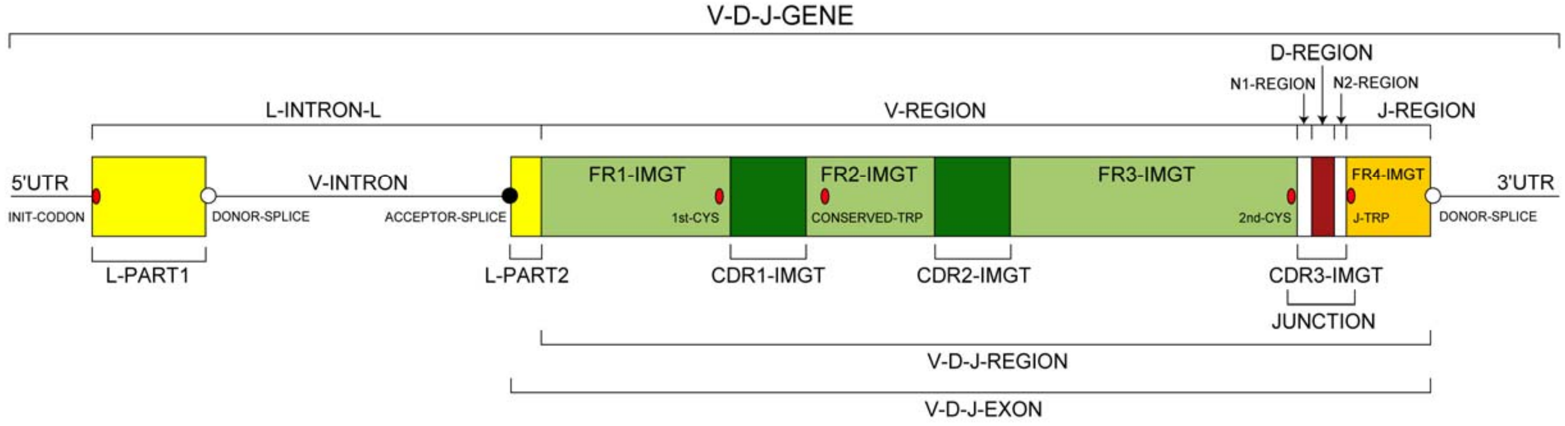




Human IGL locus (22q11.2)

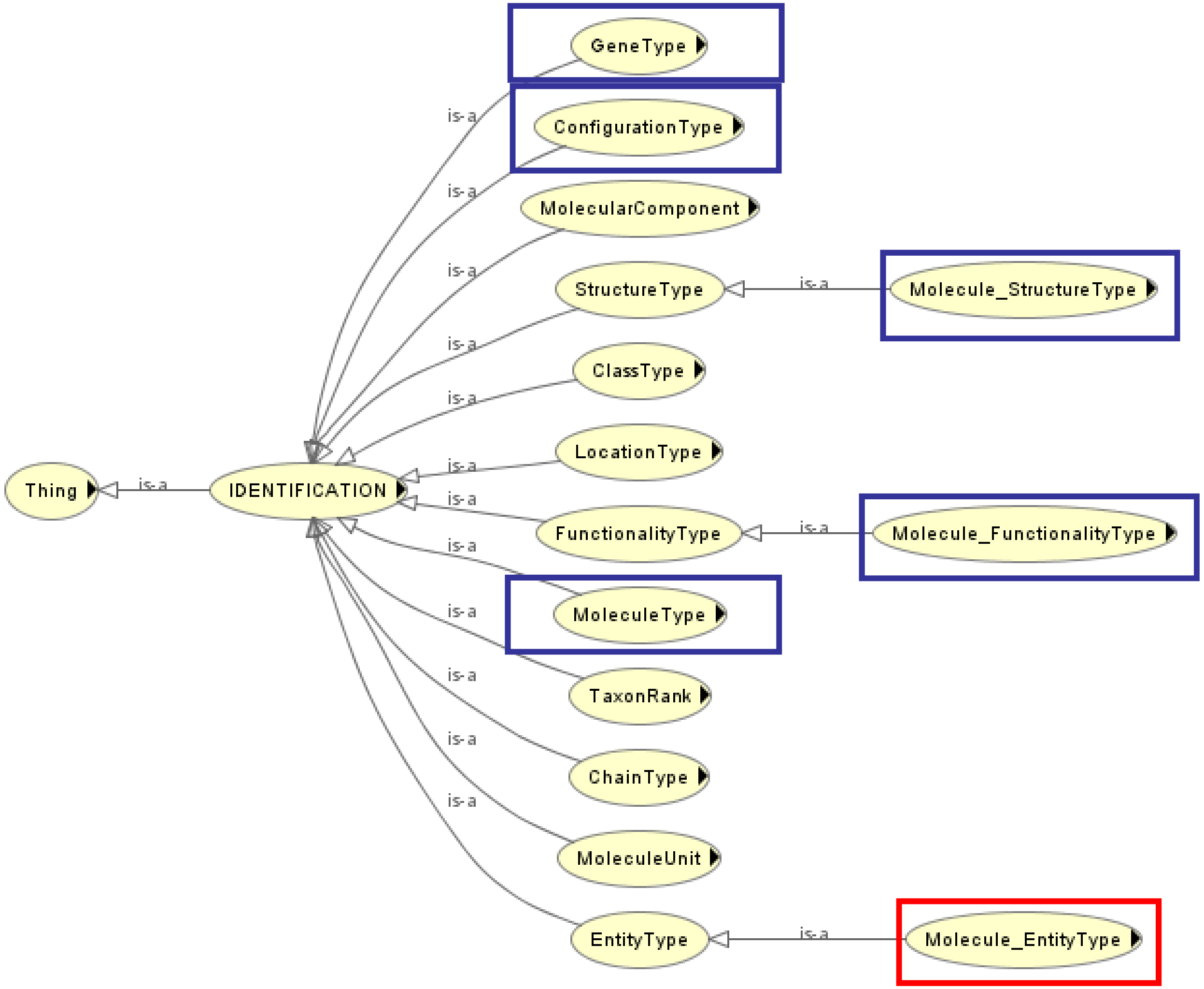
Les types de fonctionnalité pour les entités moléculaires en configuration réarrangées





Concepts d'Indentification

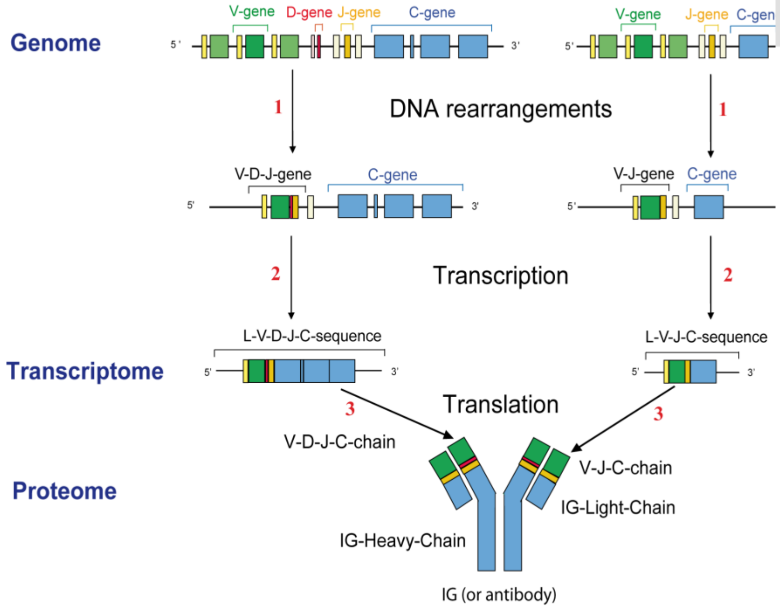
http://www.imgt.org

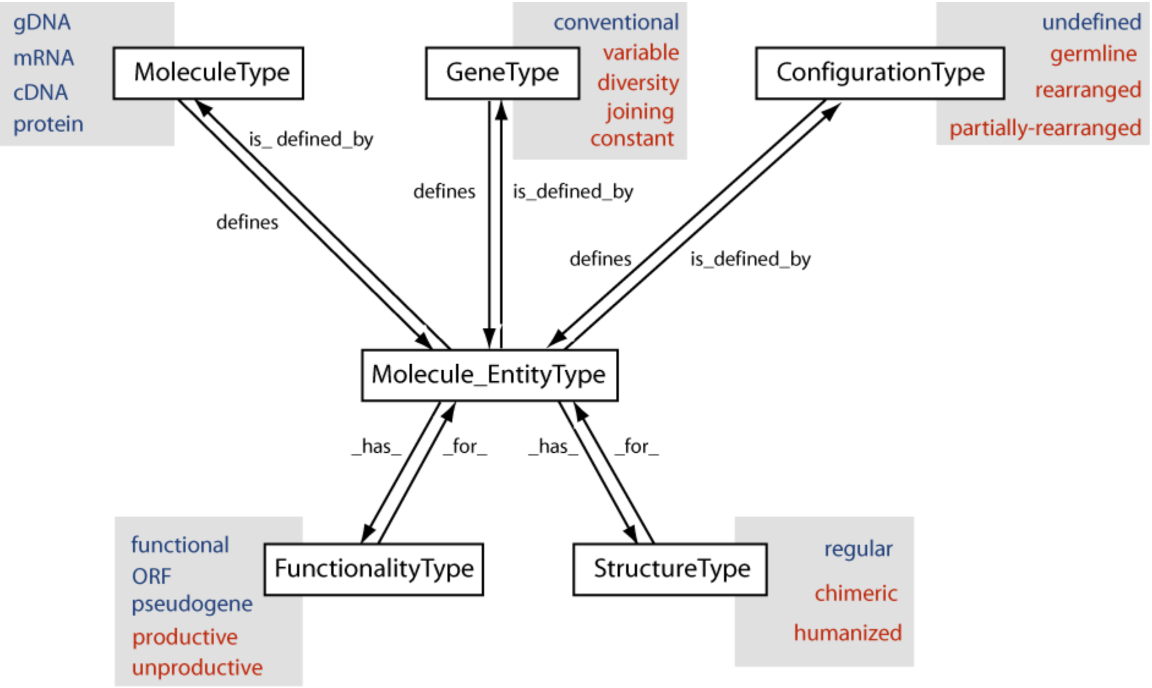


Quelles sont relations/contraintes sémantiques qui relient les concepts entre eux?

Relationships between Molecule_EntityType and the other IDENTIFICATION concepts

http://www.imgt.org





gDNA
mRNA
cDNA
protein

MoleculeType

GeneType

conventional
variable
diversity
joining
constant

ConfigurationType

undefined
germline
rearranged
partially-rearranged

is_ defined_by
defines
defines
is_defined_by
defines
is_defined_by

Molecule_EntityType

has
for
has
for

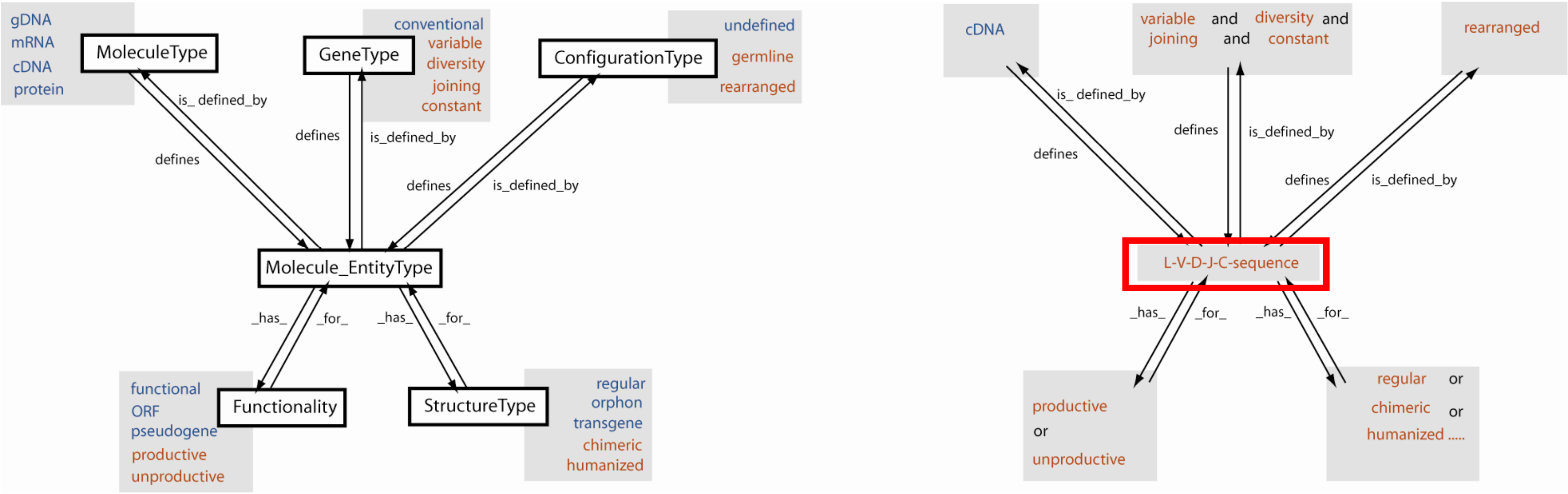
functional
ORF
pseudogene
productive
unproductive

FunctionalityType

StructureType

regular
chimeric
humanized

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



gDNA
mRNA
cDNA
protein

MoleculeType

is_ defined_by

defines

GeneType

conventional
variable
diversity
joining
constant

defines

is_defined_by

ConfigurationType

undefined
germline
rearranged

defines

is_defined_by

Molecule_EntityType

has _for_ _has_ _for_

Functionality

functional
ORF
pseudogene
productive
unproductive

StructureType

regular
orphon
transgene
chimeric
humanized

cDNA

variable and diversity and
joining and constant

rearranged

is_ defined_by

defines

defines

is_defined_by

defines

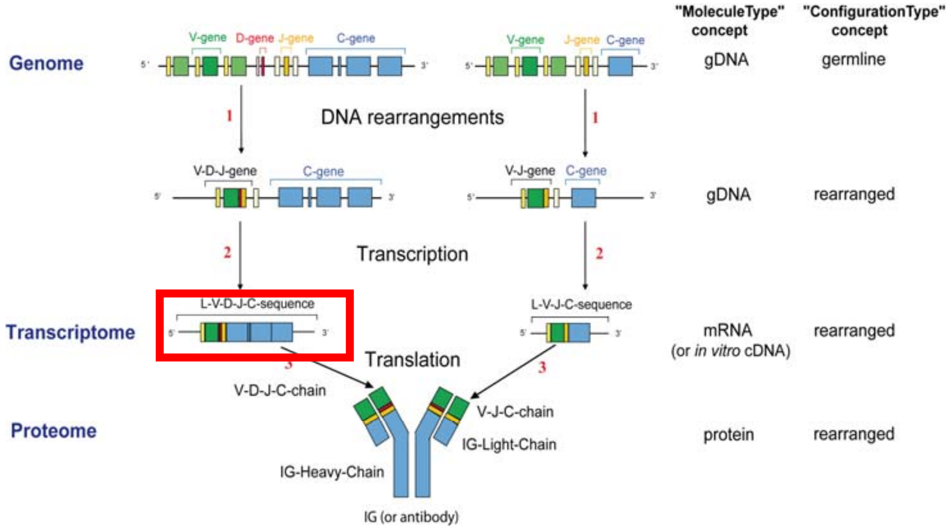
is_defined_by

L-V-D-J-C-sequence

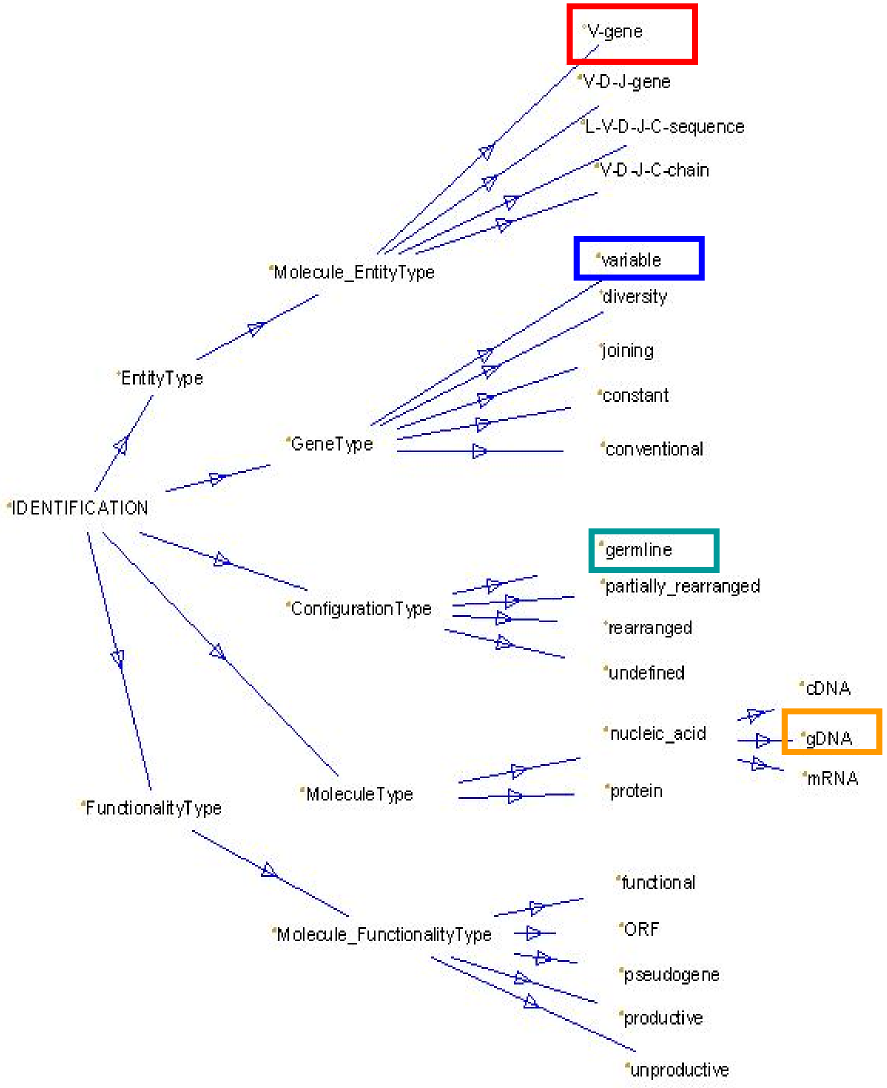
has _for_ _has_ _for_

productive
or
unproductive

regular or
chimeric or
humanized .....



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



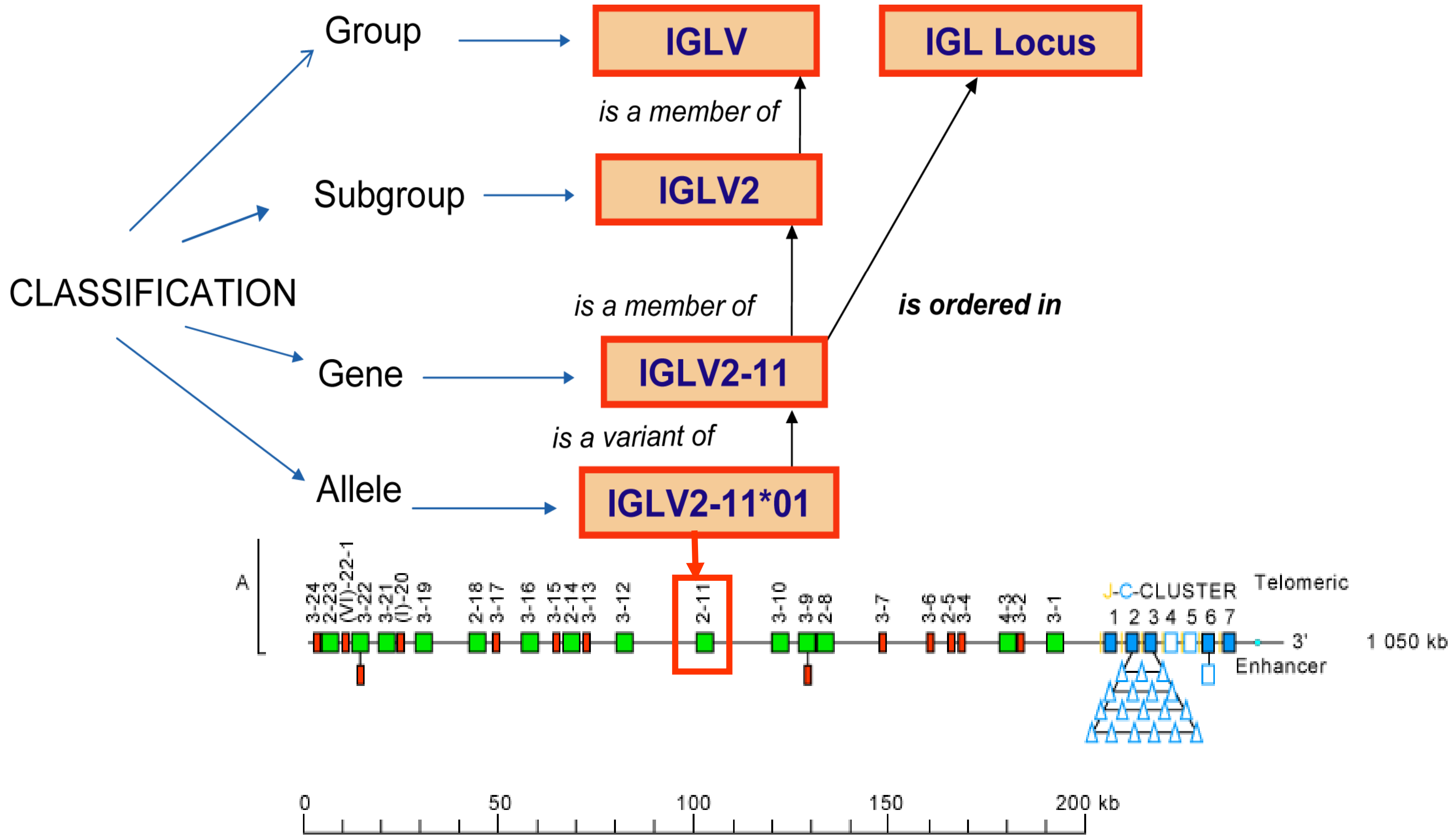
IMGT/LIGM-DB flatfile

```
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
```

2 – Concepts de Classification d’IMGT-ONTOLOGY: Principes de nomenclature des gènes IG et TR

http://www.imgt.org

Human IGL locus (22q11.2)



CLASSIFICATION

Group → **IGLV**

is a member of

Subgroup → **IGLV2**

is a member of

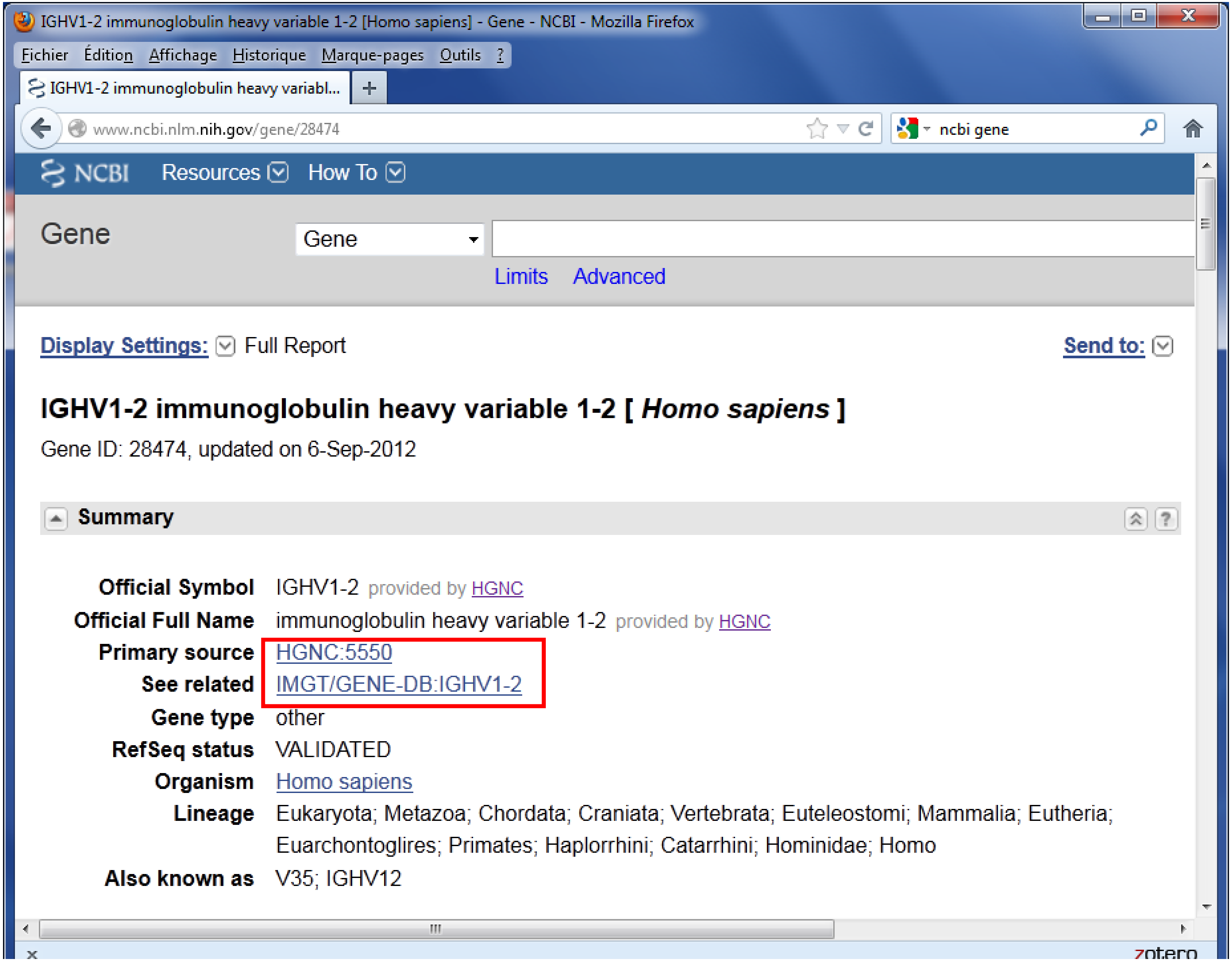
Gene → **IGLV2-11**

is a variant of

Allele → **IGLV2-11*01**

is ordered in

IGL Locus



www.ncbi.nlm.nih.gov/gene/28474

ncbi gene

Display Settings: Full Report

IGHV1-2 immunoglobulin heavy variable 1-2 [*Homo sapiens*]

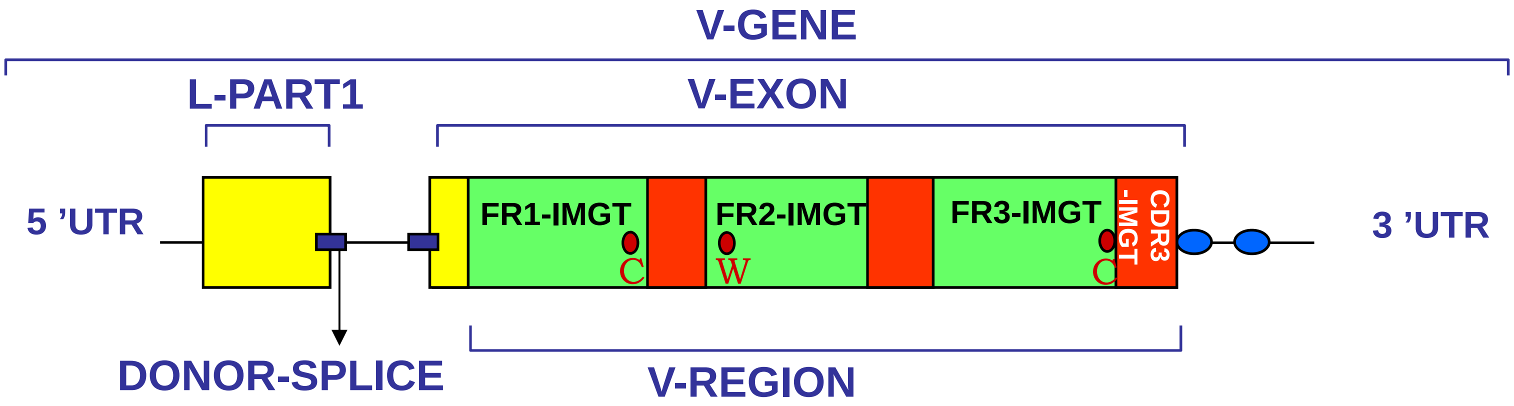
Gene ID: 28474, updated on 6-Sep-2012

Summary

Official Symbol	IGHV1-2 provided by HGNC
Official Full Name	immunoglobulin heavy variable 1-2 provided by HGNC
Primary source	HGNC:5550
See related	IMGT/GENE-DB:IGHV1-2
Gene type	other
RefSeq status	VALIDATED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as	V35; IGHV12

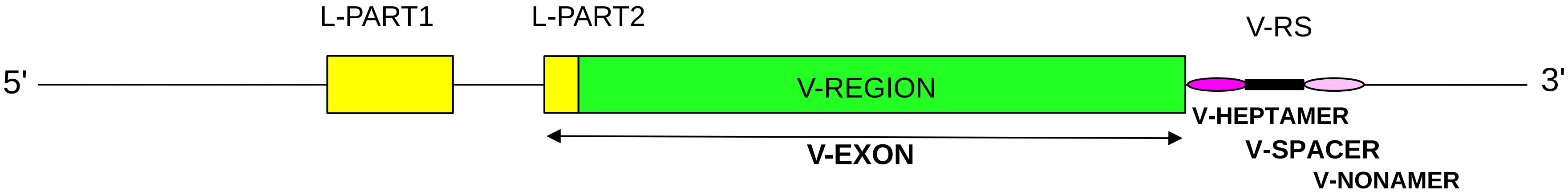
3 – Concepts de Description d’IMGT-ONTOLOGY: Labels et features

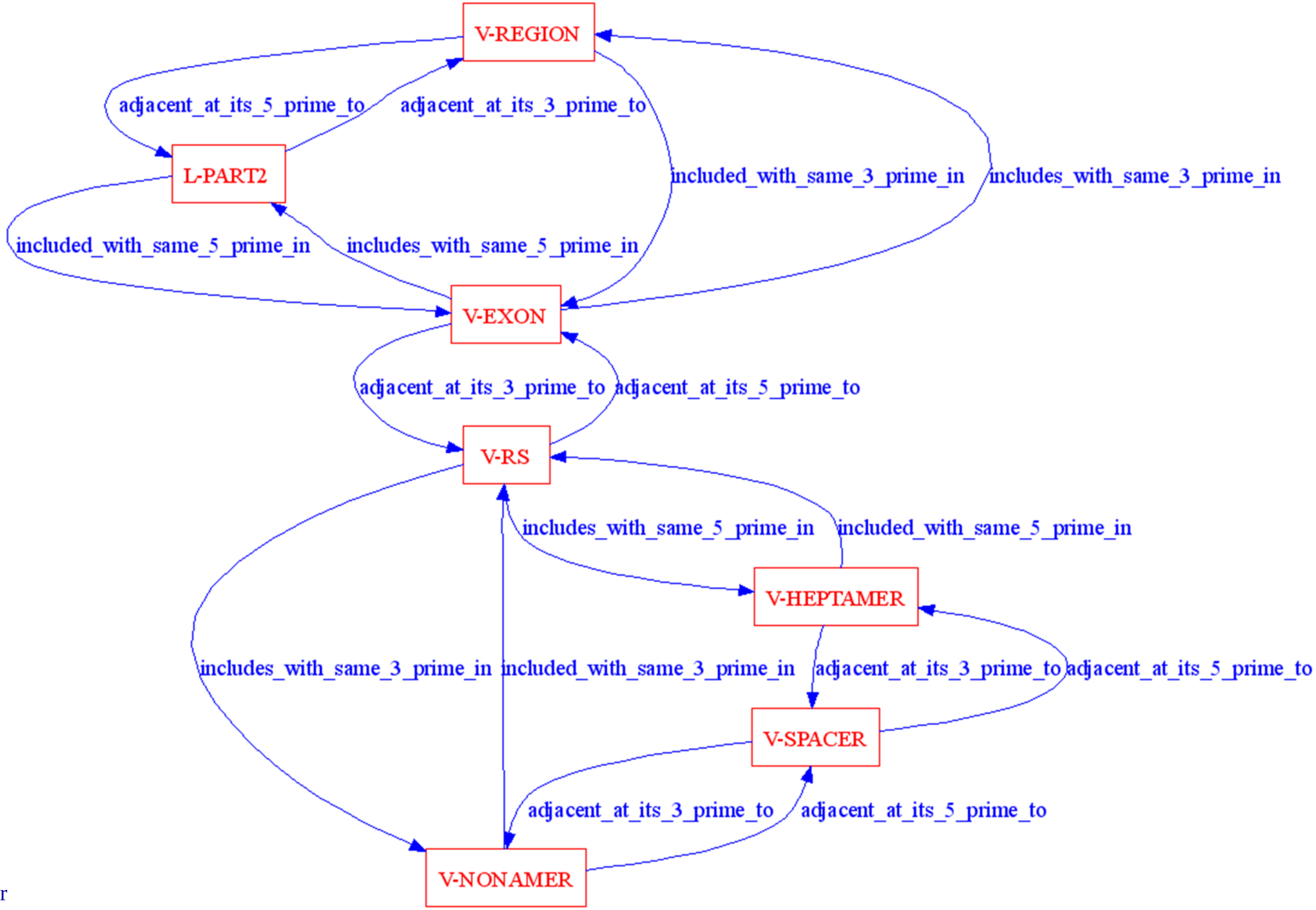
http://imgt.cines.fr
http://www.imgt.org



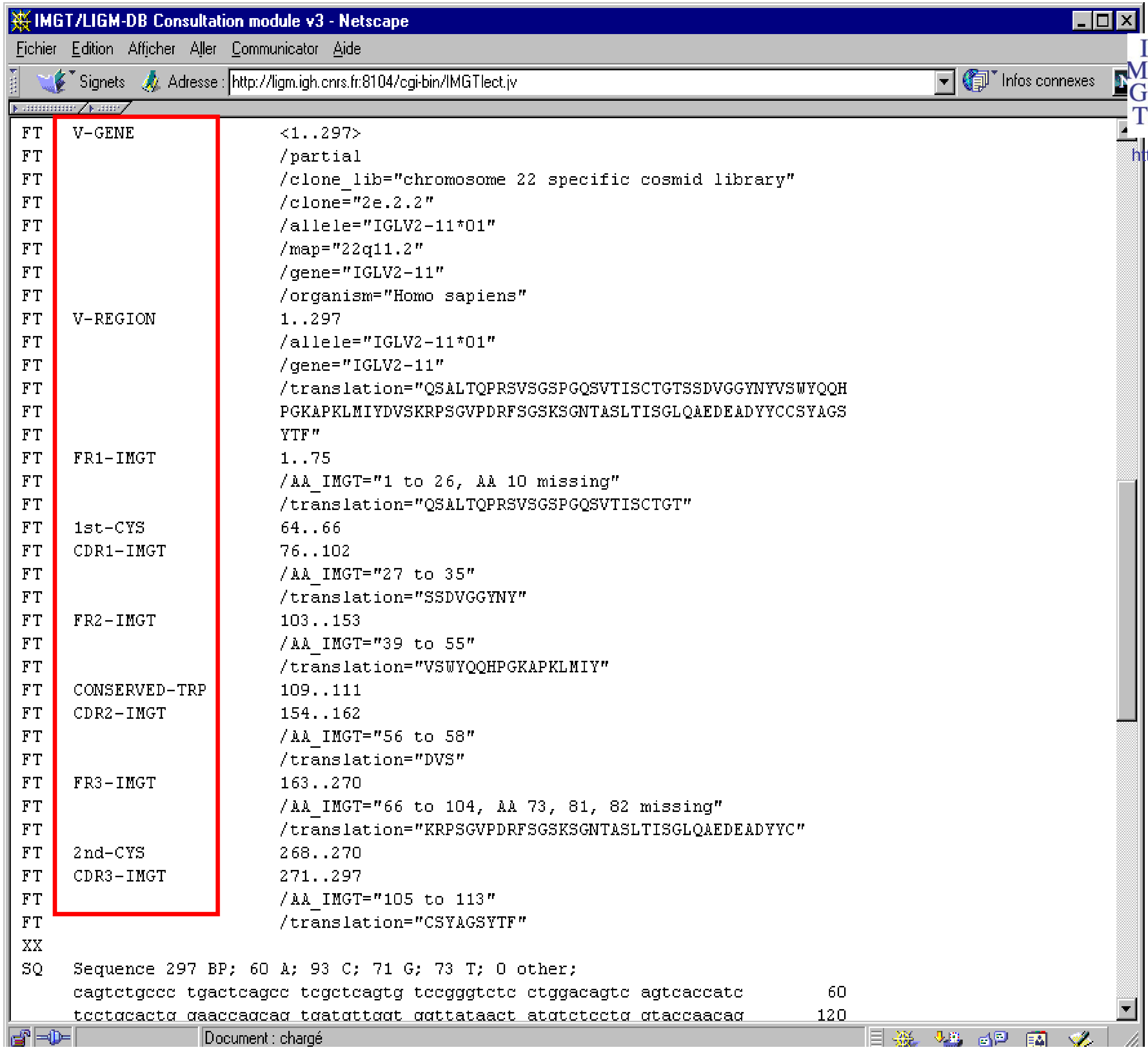
Label 1	Label	Label relations
V-GENE	V-EXON	
FR3-IMGT	CDR3-IMGT	
L-PART1	DONOR-SPLICE	
V-REGION	FR1-IMGT	
V-REGION	CDR3-IMGT	

V-GENE





http://imgt.cines.fr

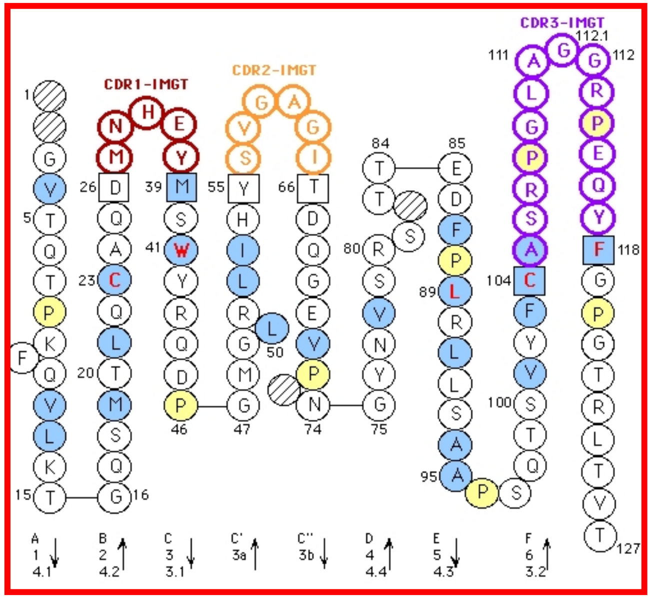


```
FT   V-GENE              <1..297>
FT                       /partial
FT                       /clone_lib="chromosome 22 specific cosmid library"
FT                       /clone="2e.2.2"
FT                       /allele="IGLV2-11*01"
FT                       /map="22q11.2"
FT                       /gene="IGLV2-11"
FT                       /organism="Homo sapiens"
FT   V-REGION            1..297
FT                       /allele="IGLV2-11*01"
FT                       /gene="IGLV2-11"
FT                       /translation="QSALTQPRSVSGSPGQSVTISCTGTSSDVGGYNYVSWYQQH
FT                       PGKAPKLMIYDVSKRPSGVPDRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGS
FT                       YTF"
FT   FR1-IMGT            1..75
FT                       /AA_IMGT="1 to 26, AA 10 missing"
FT                       /translation="QSALTQPRSVSGSPGQSVTISCTGT"
FT   1st-CYS             64..66
FT   CDR1-IMGT           76..102
FT                       /AA_IMGT="27 to 35"
FT                       /translation="SSDVGGYNY"
FT   FR2-IMGT            103..153
FT                       /AA_IMGT="39 to 55"
FT                       /translation="VSWYQQHPGKAPKLMIY"
FT   CONSERVED-TRP       109..111
FT   CDR2-IMGT           154..162
FT                       /AA_IMGT="56 to 58"
FT                       /translation="DVS"
FT   FR3-IMGT            163..270
FT                       /AA_IMGT="66 to 104, AA 73, 81, 82 missing"
FT                       /translation="KRPSGVPDRFSGSKSGNTASLTISGLQAEDEADYYC"
FT   2nd-CYS             268..270
FT   CDR3-IMGT           271..297
FT                       /AA_IMGT="105 to 113"
FT                       /translation="CSYAGSYTF"
XX
SQ   Sequence 297 BP; 60 A; 93 C; 71 G; 73 T; 0 other;
     cagtctgccc tgactcagcc tcgctcagtg tccgggtctc ctggacagtc agtcaccatc        60
     tcctgcactg gaaccagcag tgatgttggt ggttataact atgtctcctg gtaccaacag       120
```

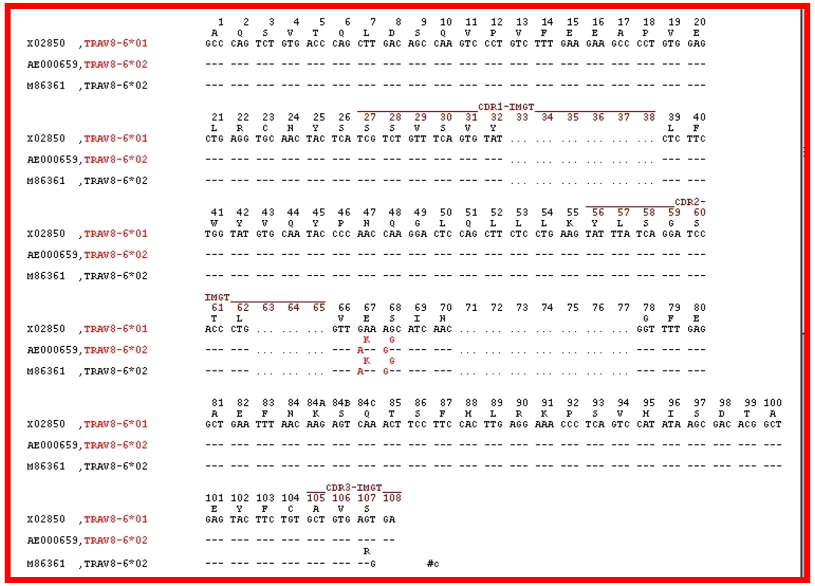
http://www.imgt.org

4 – Concepts de Numérotation d’IMGT-ONTOLOGY: Numérotation standardisée des domaines des IG et TR

Collier De Perles

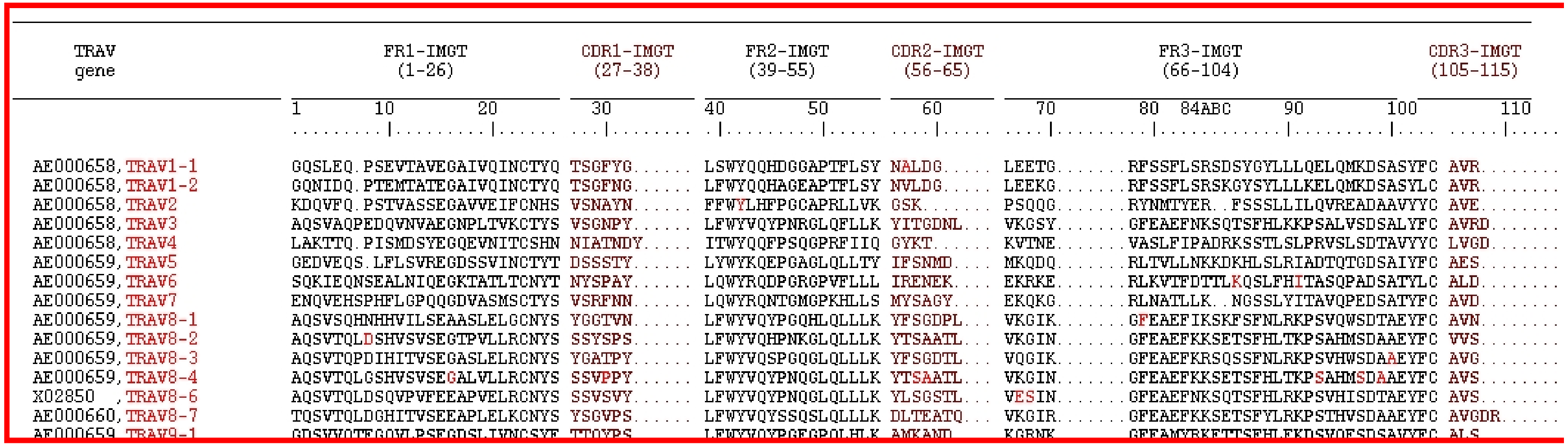


Alignment of alleles



```
                       1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
                       A   Q   S   V   T   Q   L   D   S   Q   V   P   V   F   E   E   A   P   V   E
X02850  ,TRAV8-6*01   GCC CAG TCT GTG ACC CAG CTT GAC AGC CAA GTC CCT GTC TTT GAA GAA GCC CCT GTG GAG

AE000659,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

M86361  ,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                              ______________CDR1-IMGT______________
                      21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
                       L   R   C   N   Y   S   S   S   V   S   V   Y                           L   F
X02850  ,TRAV8-6*01   CTG AGG TGC AAC TAC TCA TCG TCT GTT TCA GTG TAT ... ... ... ... ... ... CTC TTC

AE000659,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- ... ... ... ... ... ... --- ---

M86361  ,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- ... ... ... ... ... ... --- ---

                                                                                  _________CDR2-
                      41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
                       W   Y   V   Q   Y   P   N   Q   G   L   Q   L   L   L   K   Y   L   S   G   S
X02850  ,TRAV8-6*01   TGG TAT GTG CAA TAC CCC AAC CAA GGA CTC CAG CTT CTC CTG AAG TAT TTA TCA GGA TCC

AE000659,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

M86361  ,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                      IMGT_______________
                      61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
                       T   L               V   E   S   I   N                               G   F   E
X02850  ,TRAV8-6*01   ACC CTG ... ... ... GTT GAA AGC ATC AAC ... ... ... ... ... ... ... GGT TTT GAG
                                               K   G
AE000659,TRAV8-6*02   --- --- ... ... ... --- A-- G-- --- --- ... ... ... ... ... ... ... --- --- ---
                                               K   G
M86361  ,TRAV8-6*02   --- --- ... ... ... --- A-- G-- --- --- ... ... ... ... ... ... ... --- --- ---

                      81  82  83  84  84A 84B 84C 85  86  87  88  89  90  91  92  93  94  95  96  97  98  99  100
                       A   E   F   N   K   S   Q   T   S   F   H   L   R   K   P   S   V   H   I   S   D   T   A
X02850  ,TRAV8-6*01   GCT GAA TTT AAC AAG AGT CAA ACT TCC TTC CAC TTG AGG AAA CCC TCA GTC CAT ATA AGC GAC ACG GCT

AE000659,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

M86361  ,TRAV8-6*02   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                          ___CDR3-IMGT___
                      101 102 103 104 105 106 107 108
                       E   Y   F   C   A   V   S
X02850  ,TRAV8-6*01   GAG TAC TTC TGT GCT GTG AGT GA

AE000659,TRAV8-6*02   --- --- --- --- --- --- --- --
                                               R
M86361  ,TRAV8-6*02   --- --- --- --- --- --- --G         #c
```

Protein Display



```
    TRAV                    FR1-IMGT           CDR1-IMGT       FR2-IMGT      CDR2-IMGT                   FR3-IMGT                  CDR3-IMGT
    gene                     (1-26)             (27-38)        (39-55)        (56-65)                    (66-104)                  (105-115)

                   1        10        20         30         40        50         60         70        80   84ABC   90       100        110
                   .........|.........|...... ...|........ .|.........|..... ....|..... ....|.........|............|.........|.... .....|.....

AE000658,TRAV1-1   GQSLEQ.PSEVTAVEGAIVQINCTYQ TSGFYG...... LSWYQQHDGGAPTFLSY NALDG..... LEETG.......RFSSFLSRSDSYGYLLLQELQMKDSASYFC AVR........
AE000658,TRAV1-2   GQNIDQ.PTEMTATEGAIVQINCTYQ TSGFNG...... LFWYQQHAGEAPTFLSY NVLDG..... LEEKG.......RFSSFLSRSKGYSYLLLKELQMKDSASYLC AVR........
AE000658,TRAV2     KDQVFQ.PSTVASSEGAVVEIFCNHS VSNAYN...... FFWYLHFPGCAPRLLVK GSK....... PSQQG.......RYNMTYER..FSSSLLILQVREADAAVYYC AVE........
AE000658,TRAV3     AQSVAQPEDQVNVAEGNPLTVKCTYS VSGNPY...... LFWYVQYPNRGLQFLLK YITGDNL... VKGSY.......GFEAEFNKSQTSFHLKKPSALVSDSALYFC AVRD.......
AE000658,TRAV4     LAKTTQ.PISMDSYEGQEVNITCSHN NIATNDY..... ITWYQQFPSQGPRFIIQ GYKT...... KVTNE.......VASLFIPADRKSSTLSLPRVSLSDTAVYYC LVGD.......
AE000659,TRAV5     GEDVEQS.LFLSVREGDSSVINCTYT DSSSTY...... LYWYKQEPGAGLQLLTY IFSNMD.... MKQDQ.......RLTVLLNKKDKHLSLRIADTQTGDSAIYFC AES........
AE000659,TRAV6     SQKIEQNSEALNIQEGKTATLTCNYT NYSPAY...... LQWYRQDPGRGPVFLLL IRENEK.... EKRKE.......RLKVTFDTTLKQSLFHITASQPADSATYLC ALD........
AE000659,TRAV7     ENQVEHSPHFLGPQQGDVASMSCTYS VSRFNN...... LQWYRQNTGMGPKHLLS MYSAGY.... EKQKG.......RLNATLLK..NGSSLYITAVQPEDSATYFC AVD........
AE000659,TRAV8-1   AQSVSQHNHHVILSEAASLELGCNYS YGGTVN...... LFWYVQYPGQHLQLLLK YFSGDPL... VKGIK.......GFEAEFIKSKFSFNLRKPSVQWSDTAEYFC AVN........
AE000659,TRAV8-2   AQSVTQLDSHVSVSEGTPVLLRCNYS SSYSPS...... LFWYVQHPNKGLQLLLK YTSAATL... VKGIN.......GFEAEFKKSETSFHLTKPSAHMSDAAEYFC VVS........
AE000659,TRAV8-3   AQSVTQPDIHITVSEGASLELRCNYS YGATPY...... LFWYVQSPGQGLQLLLK YFSGDTL... VQGIK.......GFEAEFKRSQSSFNLRKPSVHWSDAAEYFC AVG........
AE000659,TRAV8-4   AQSVTQLGSHVSVSEGALVLLRCNYS SSVPPY...... LFWYVQYPNQGLQLLLK YTSAATL... VKGIN.......GFEAEFKKSETSFHLTKPSAHMSDAAEYFC AVS........
X02850  ,TRAV8-6   AQSVTQLDSQVPVFEEAPVELRCNYS SSVSVY...... LFWYVQYPNQGLQLLLK YLSGSTL... VESIN.......GFEAEFNKSQTSFHLRKPSVHISDTAEYFC AVS........
AE000660,TRAV8-7   TQSVTQLDGHITVSEEAPLELKCNYS YSGVPS...... LFWYVQYSSQSLQLLLK DLTEATQ... VKGIR.......GFEAEFKKSETSFYLRKPSTHVSDAAEYFC AVGDR......
AE000659,TRAV9-1   GDSVVQTEGQVLPSEGDSLIVNCSYE TTQYPS       LFWYVQYPGEGPQLHLK AMKAND     KGRNK       GFEAMYRKETTSFHLEKDSVQESDSAVYFC ALS
```

Antibody humanization and engineering

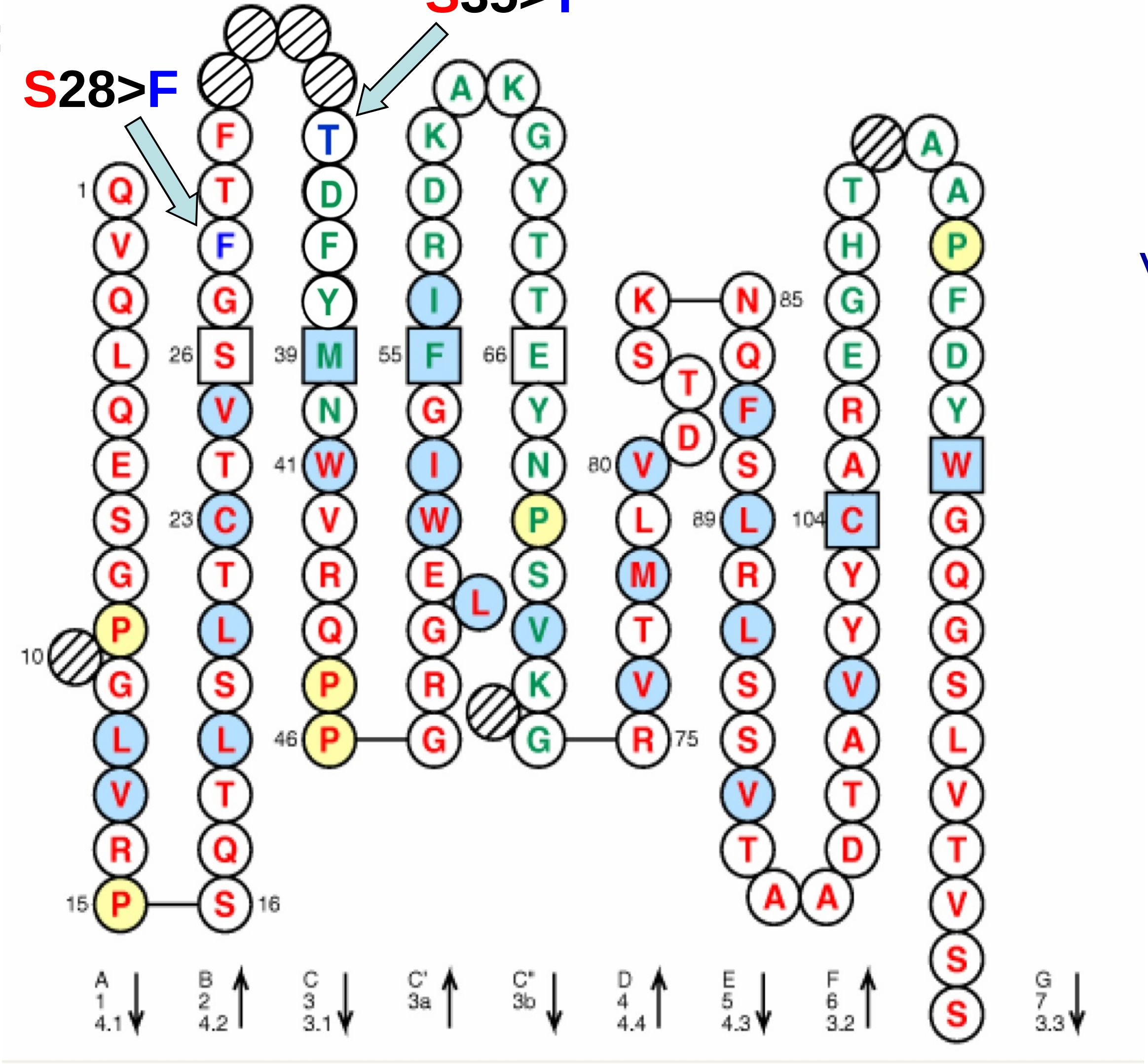
Alemtuzumab (CAMPATH®)

http://www.imgt.org

2 mutations:

S28>F

S35>T



VH domain
[8.10.12]

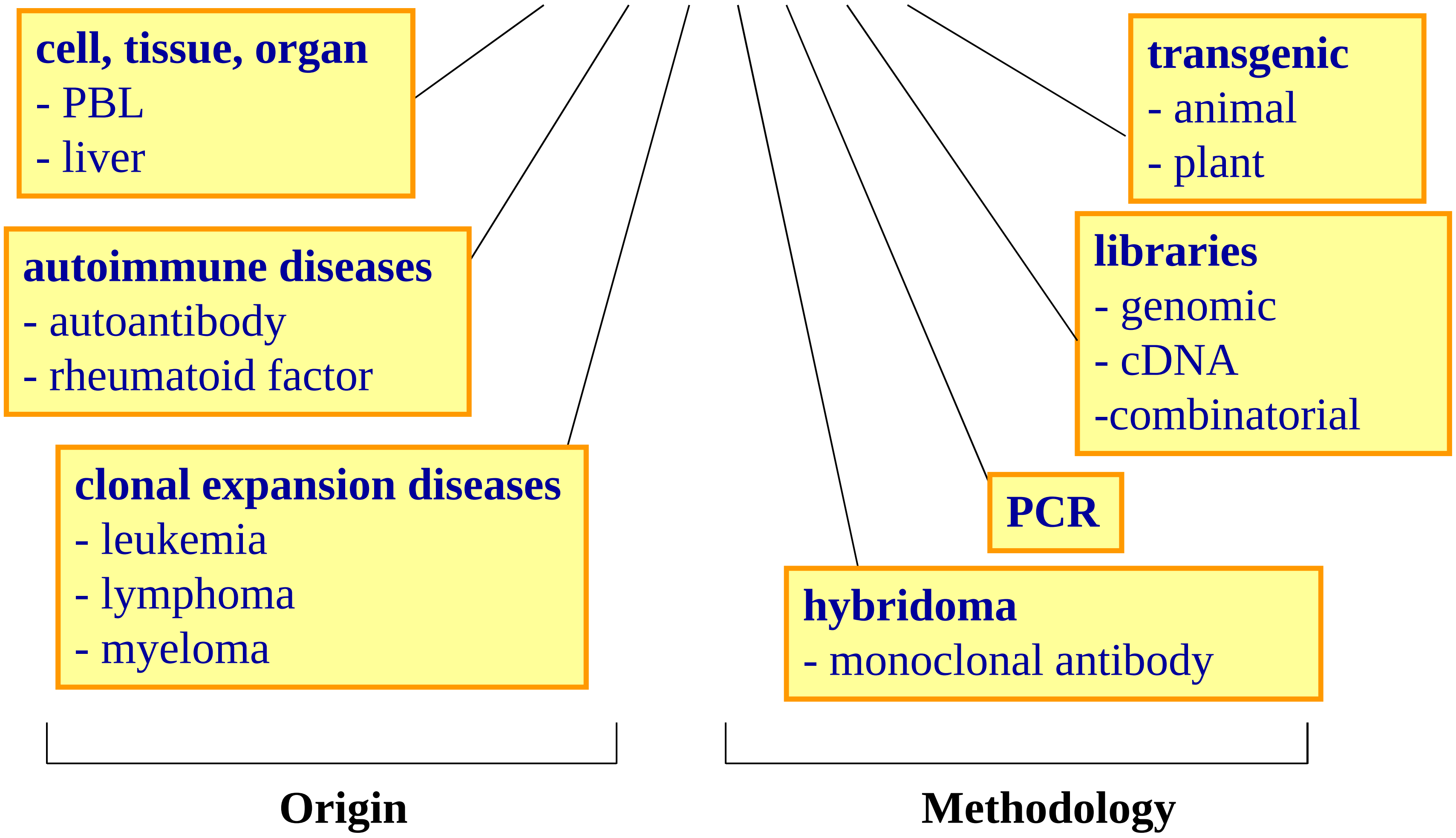
🟥 human

🟩 rat

IMGT® http://www.imgt.org

http://www.imgt.org

"OBTENTION"



cell, tissue, organ
- PBL
- liver

autoimmune diseases
- autoantibody
- rheumatoid factor

clonal expansion diseases
- leukemia
- lymphoma
- myeloma

transgenic
- animal
- plant

libraries
- genomic
- cDNA
-combinatorial

PCR

hybridoma
- monoclonal antibody

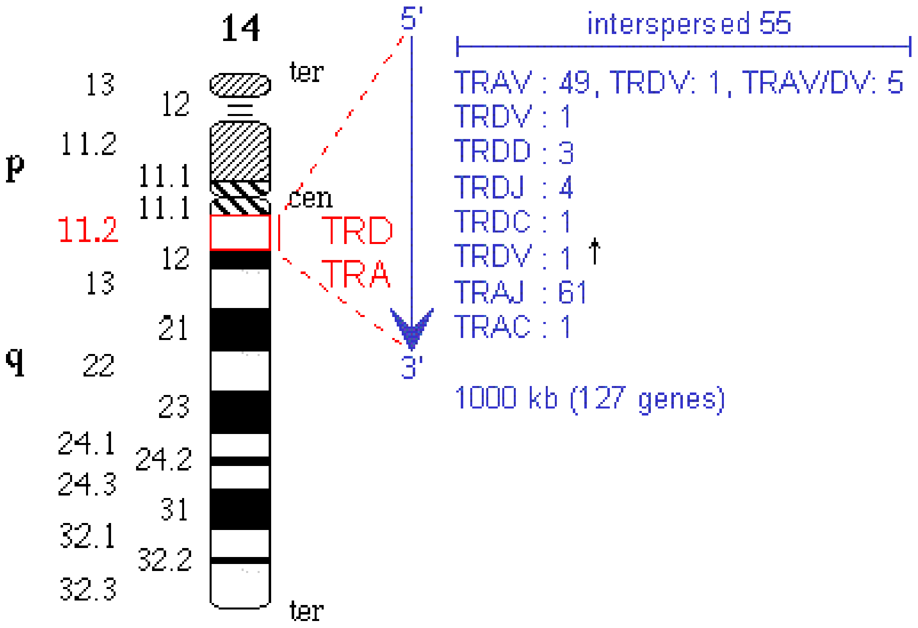
Origin

Methodology

The "OBTENTION" concepts specify the origin and methodology

"ORIENTATION"

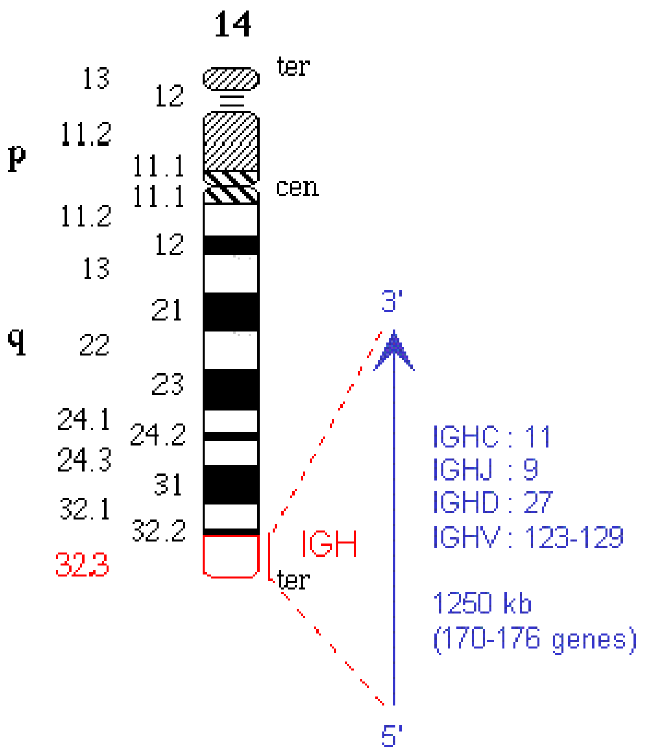
Locus TRA/TRD (human)
14q11.2
Orientation of the locus on the chromosome: FORWARD (FWD)



TRAV : 49, TRDV: 1, TRAV/DV: 5
TRDV : 1
TRDD : 3
TRDJ : 4
TRDC : 1
TRDV : 1 ↑
TRAJ : 61
TRAC : 1

1000 kb (127 genes)

Locus IGH (human)
14q32.33
Orientation of the locus on the chromosome: REVERSE (REV)

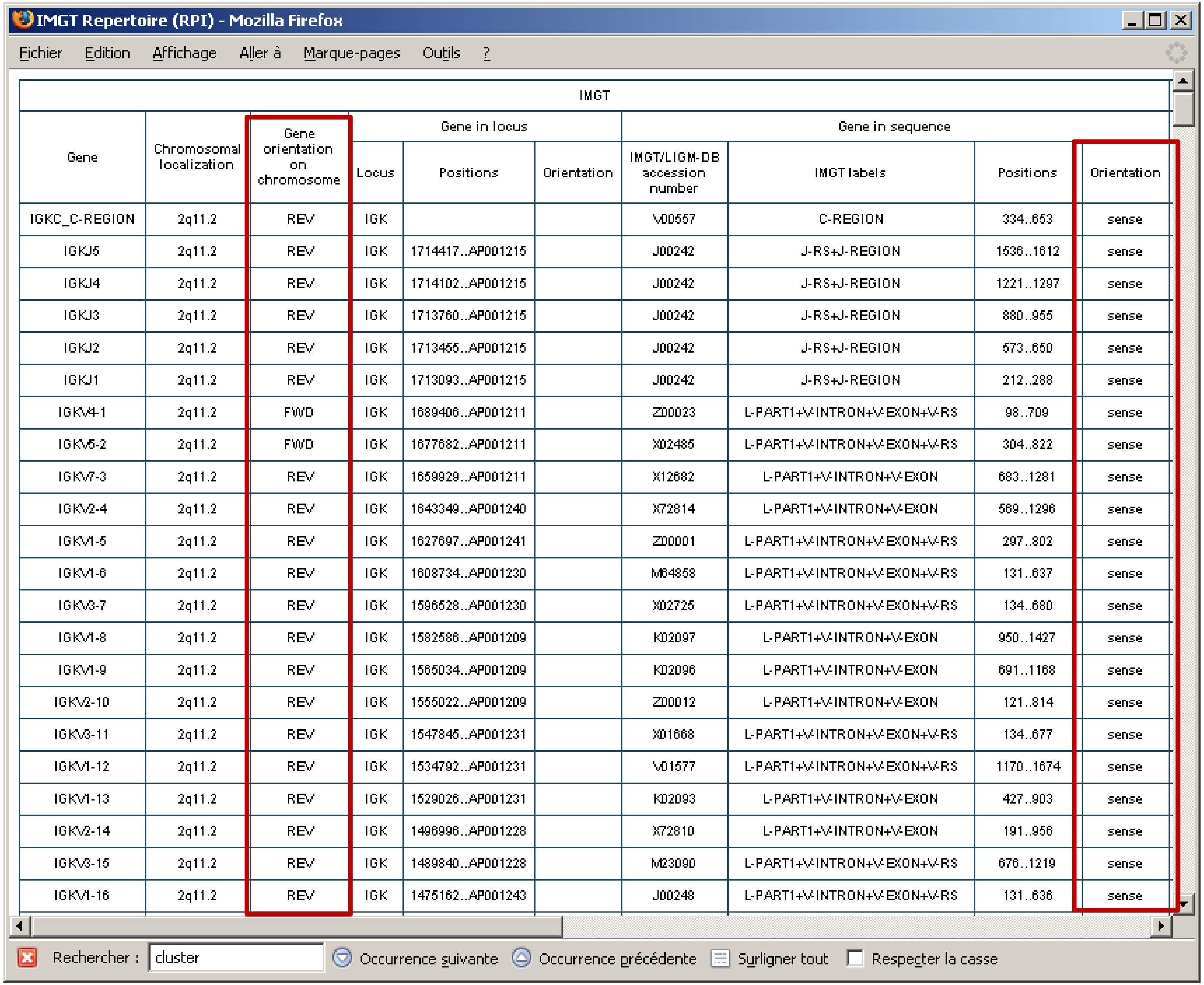


IGHC : 11
IGHJ : 9
IGHD : 27
IGHV : 123-129

1250 kb
(170-176 genes)

"ORIENTATION"

http://www.imgt.org



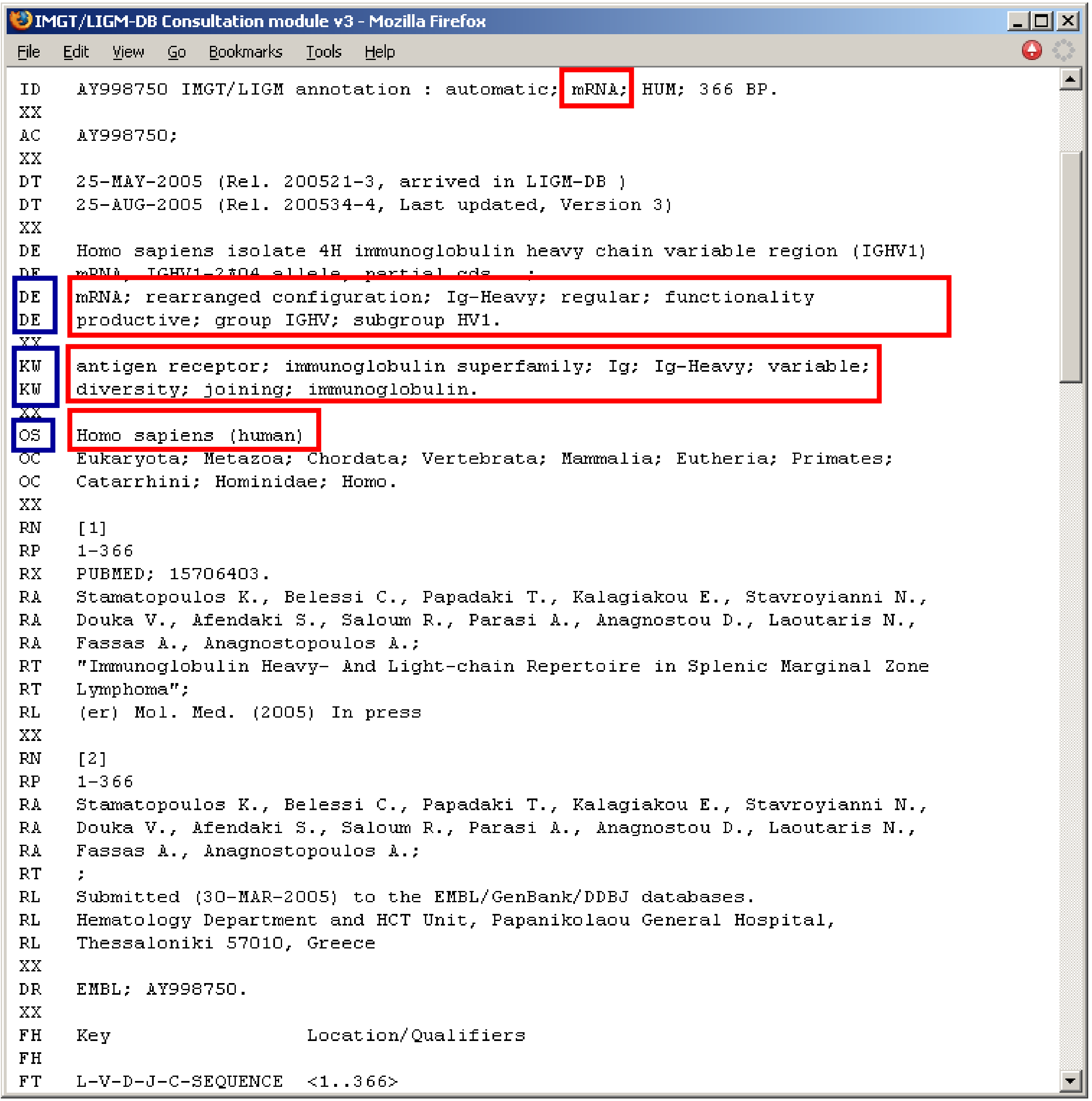
Rechercher : cluster

IMGT										
Gene	Chromosomal localization	Gene orientation on chromosome	Gene in locus			Gene in sequence				
			Locus	Positions	Orientation	IMGT/LIGM-DB accession number	IMGT labels	Positions	Orientation
IGKC_C-REGION	2q11.2	REV	IGK			V00557	C-REGION	334..653	sense
IGKJ5	2q11.2	REV	IGK	1714417..AP001215		J00242	J-RS+J-REGION	1536..1612	sense
IGKJ4	2q11.2	REV	IGK	1714102..AP001215		J00242	J-RS+J-REGION	1221..1297	sense
IGKJ3	2q11.2	REV	IGK	1713760..AP001215		J00242	J-RS+J-REGION	880..955	sense
IGKJ2	2q11.2	REV	IGK	1713455..AP001215		J00242	J-RS+J-REGION	573..650	sense
IGKJ1	2q11.2	REV	IGK	1713093..AP001215		J00242	J-RS+J-REGION	212..288	sense
IGKV4-1	2q11.2	FWD	IGK	1689406..AP001211		Z00023	L-PART1+V-INTRON+V-EXON+V-RS	98..709	sense
IGKV5-2	2q11.2	FWD	IGK	1677682..AP001211		X02485	L-PART1+V-INTRON+V-EXON+V-RS	304..822	sense
IGKV7-3	2q11.2	REV	IGK	1659929..AP001211		X12682	L-PART1+V-INTRON+V-EXON	683..1281	sense
IGKV2-4	2q11.2	REV	IGK	1643349..AP001240		X72814	L-PART1+V-INTRON+V-EXON	569..1296	sense
IGKV1-5	2q11.2	REV	IGK	1627697..AP001241		Z00001	L-PART1+V-INTRON+V-EXON+V-RS	297..802	sense
IGKV1-6	2q11.2	REV	IGK	1608734..AP001230		M64858	L-PART1+V-INTRON+V-EXON+V-RS	131..637	sense
IGKV3-7	2q11.2	REV	IGK	1596528..AP001230		X02725	L-PART1+V-INTRON+V-EXON+V-RS	134..680	sense
IGKV1-8	2q11.2	REV	IGK	1582586..AP001209		K02097	L-PART1+V-INTRON+V-EXON	950..1427	sense
IGKV1-9	2q11.2	REV	IGK	1565034..AP001209		K02096	L-PART1+V-INTRON+V-EXON	691..1168	sense
IGKV2-10	2q11.2	REV	IGK	1555022..AP001209		Z00012	L-PART1+V-INTRON+V-EXON	121..814	sense
IGKV3-11	2q11.2	REV	IGK	1547845..AP001231		X01668	L-PART1+V-INTRON+V-EXON+V-RS	134..677	sense
IGKV1-12	2q11.2	REV	IGK	1534792..AP001231		V01577	L-PART1+V-INTRON+V-EXON+V-RS	1170..1674	sense
IGKV1-13	2q11.2	REV	IGK	1529026..AP001231		K02093	L-PART1+V-INTRON+V-EXON	427..903	sense
IGKV2-14	2q11.2	REV	IGK	1496996..AP001228		X72810	L-PART1+V-INTRON+V-EXON	191..956	sense
IGKV3-15	2q11.2	REV	IGK	1489840..AP001228		M23090	L-PART1+V-INTRON+V-EXON+V-RS	676..1219	sense
IGKV1-16	2q11.2	REV	IGK	1475162..AP001243		J00248	L-PART1+V-INTRON+V-EXON+V-RS	131..636	sense

Annotation and IMGT-ONTOLOGY concepts

http://www.imgt.org

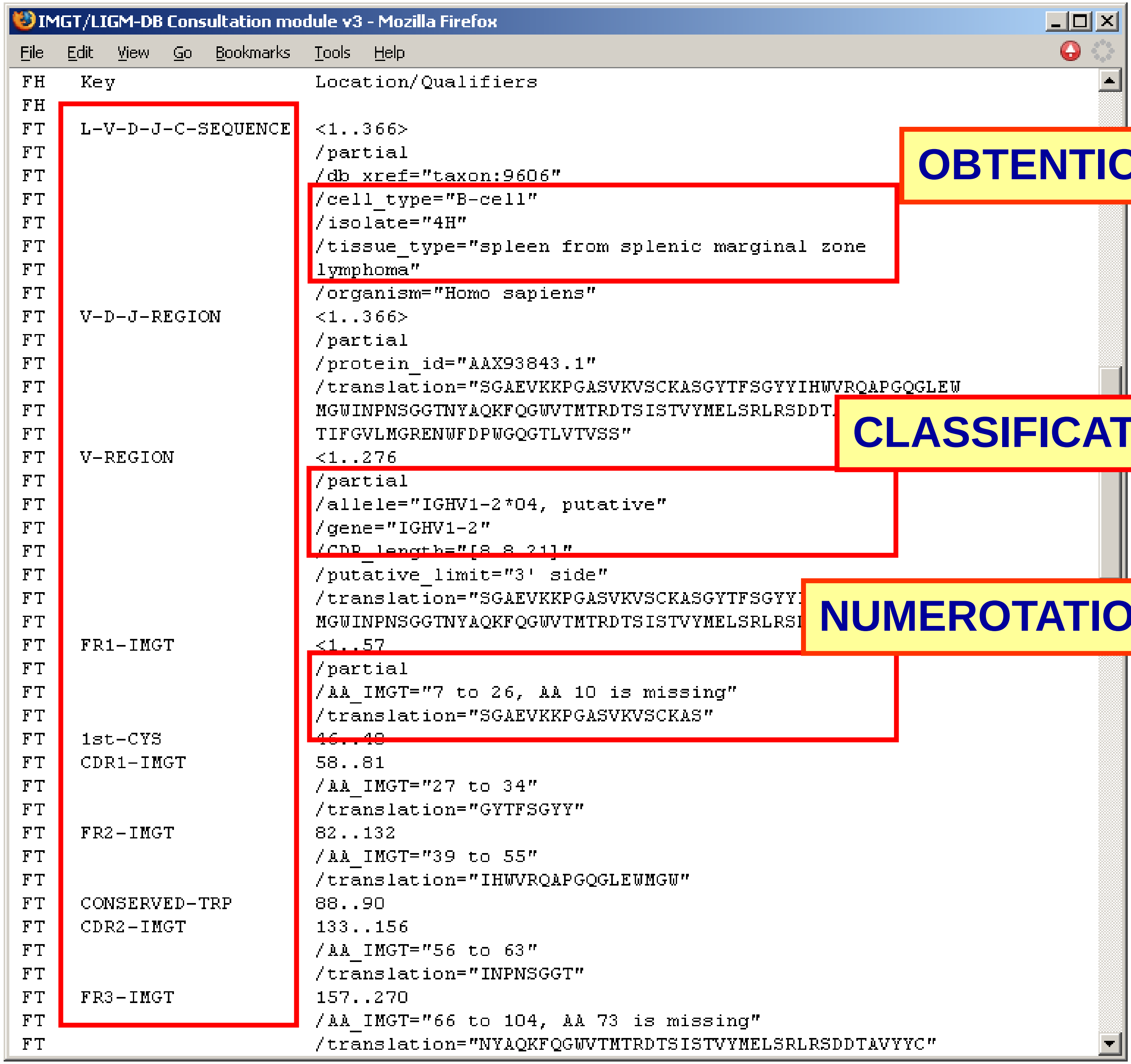
IDENTIFICATION



```
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>
```

Annotation and IMGT-ONTOLOGY concepts

ImMunoGeneTics Information system®
http://www.imgt.org



```
FH   Key                   Location/Qualifiers
FH
FT   L-V-D-J-C-SEQUENCE    <1..366>
FT                         /partial
FT                         /db_xref="taxon:9606"
FT                         /cell_type="B-cell"
FT                         /isolate="4H"
FT                         /tissue_type="spleen from splenic marginal zone
FT                         lymphoma"
FT                         /organism="Homo sapiens"
FT   V-D-J-REGION          <1..366>
FT                         /partial
FT                         /protein_id="AAX93843.1"
FT                         /translation="SGAEVKKPGASVKVSCKASGYTFSGYYIHWVRQAPGQGLEW
FT                         MGWINPNSGGTNYAQKFQGWVTMTRDTSISTVYMELSRLRSDDT
FT                         TIFGVLMGRENWFDPWGQGTLVTVSS"
FT   V-REGION              <1..276
FT                         /partial
FT                         /allele="IGHV1-2*04, putative"
FT                         /gene="IGHV1-2"
FT                         /CDR_length="[8.8.21]"
FT                         /putative_limit="3' side"
FT                         /translation="SGAEVKKPGASVKVSCKASGYTFSGYY
FT                         MGWINPNSGGTNYAQKFQGWVTMTRDTSISTVYMELSRLRS
FT   FR1-IMGT              <1..57
FT                         /partial
FT                         /AA_IMGT="7 to 26, AA 10 is missing"
FT                         /translation="SGAEVKKPGASVKVSCKAS"
FT   1st-CYS               46..48
FT   CDR1-IMGT             58..81
FT                         /AA_IMGT="27 to 34"
FT                         /translation="GYTFSGYY"
FT   FR2-IMGT              82..132
FT                         /AA_IMGT="39 to 55"
FT                         /translation="IHWVRQAPGQGLEWMGW"
FT   CONSERVED-TRP         88..90
FT   CDR2-IMGT             133..156
FT                         /AA_IMGT="56 to 63"
FT                         /translation="INPNSGGT"
FT   FR3-IMGT              157..270
FT                         /AA_IMGT="66 to 104, AA 73 is missing"
FT                         /translation="NYAQKFQGWVTMTRDTSISTVYMELSRLRSDDTAVYYC"
```

DESCRIPTION

OBTENTION

CLASSIFICATION

NUMEROTATION

Les atouts de IMGT-ONTOLOGY

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

Partage des connaissances

Partager/relier les différentes ontologies

⇒ ontologies généralistes/spécialisées

- ontologies développées pour la biologie « en général »
 partager les mêmes concepts
 quelque soit le domaine (ex: Gene Ontologie)

- ontologies spécialisées dans un domaine (ex: IMGT-ONTOLOGY dans le domaine de l’immunogénétique)

⇒ ontologies d’autres domaines scientifiques

    ⇒ physique
    ⇒ chimie
    ⇒ médecine

Une ontologie correspond à une représentation des connaissances

La subjectivité est liée aux nécessités du domaine d’application. Il faut accepter que les connaissances soient représentées sous différents points de vue, par exemple:

- GO: fonction moléculaire, composant cellulaire, processus biologique
- IMGT-ONTOLOGY: 7 axiomes

Il existe différentes terminologies, différentes conventions, différents niveaux de granularité

Les langages utilisés pour la formalisation des ontologies n’ont pas le même niveau d’expressivité

=> D’où la nécessité de comparer et de mettre en évidence les ressemblances

Chaque ontologie a sa propre organisation

Chaque concept :

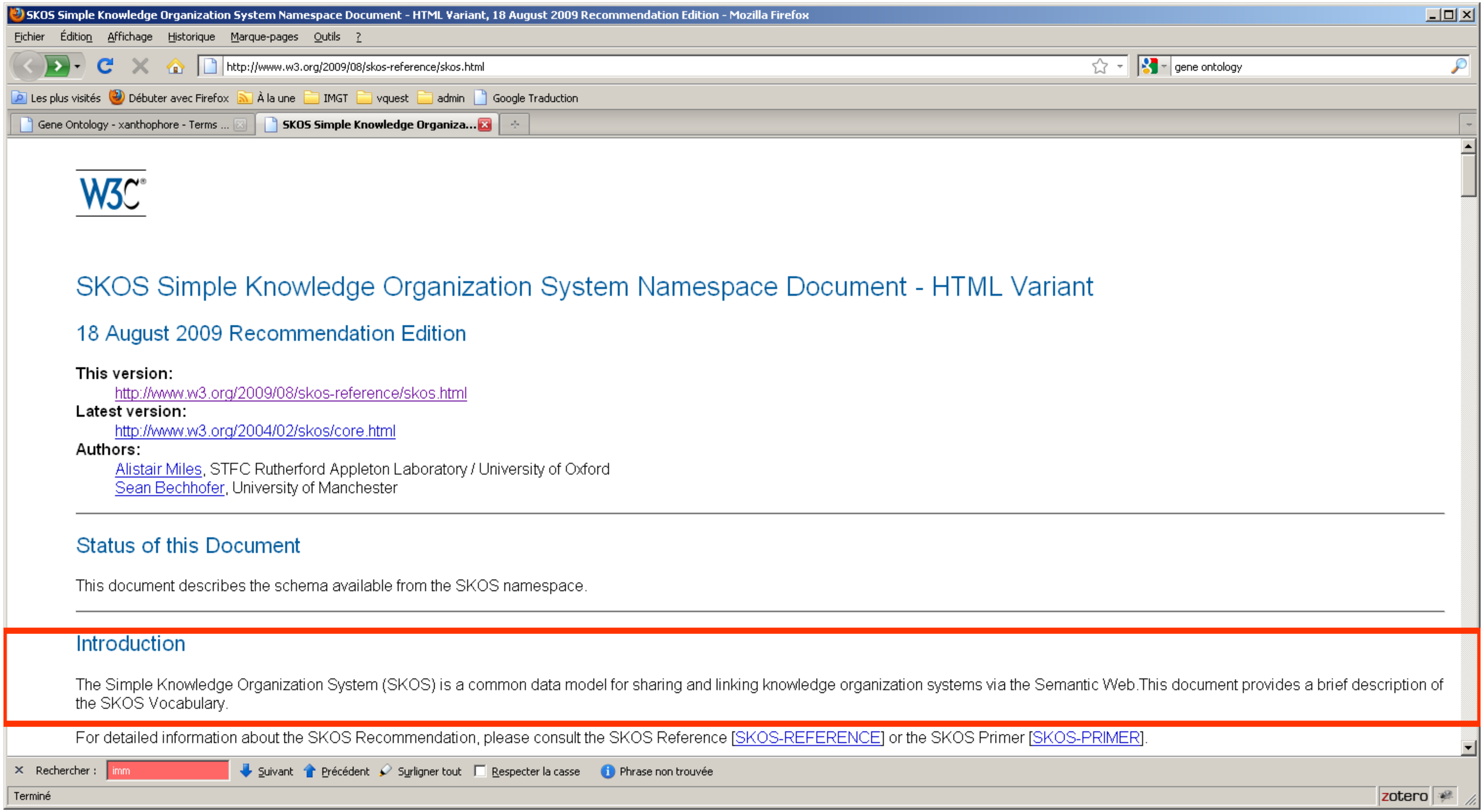
- a une position donnée dans le graphe
- est désigné par un terme
- possède une ensemble de propriétés
 (dont les relations avec les autres concepts)

Comment relier les ontologies?

Le mapping

- Correspondance entre les concepts/classes/entités de deux ontologies.

- Le mapping exprime une « similarité » sémantique



W3C

SKOS Simple Knowledge Organization System Namespace Document - HTML Variant

18 August 2009 Recommendation Edition

This version:
http://www.w3.org/2009/08/skos-reference/skos.html
Latest version:
http://www.w3.org/2004/02/skos/core.html
Authors:
Alistair Miles, STFC Rutherford Appleton Laboratory / University of Oxford
Sean Bechhofer, University of Manchester

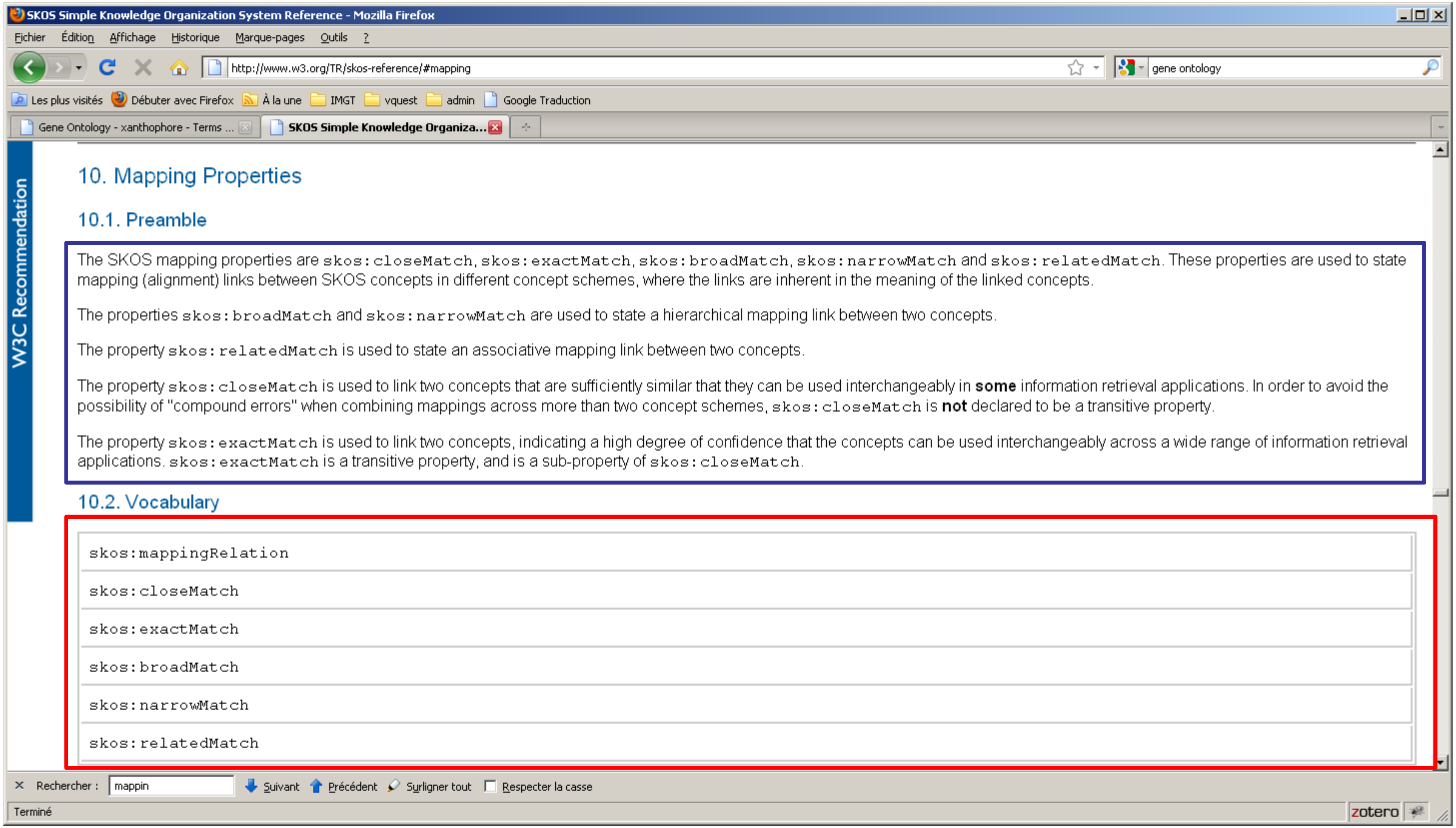
Status of this Document

This document describes the schema available from the SKOS namespace.

Introduction

The Simple Knowledge Organization System (SKOS) is a common data model for sharing and linking knowledge organization systems via the Semantic Web.This document provides a brief description of the SKOS Vocabulary.

For detailed information about the SKOS Recommendation, please consult the SKOS Reference [SKOS-REFERENCE] or the SKOS Primer [SKOS-PRIMER].



10. Mapping Properties

10.1. Preamble

The SKOS mapping properties are `skos:closeMatch`, `skos:exactMatch`, `skos:broadMatch`, `skos:narrowMatch` and `skos:relatedMatch`. These properties are used to state mapping (alignment) links between SKOS concepts in different concept schemes, where the links are inherent in the meaning of the linked concepts.

The properties `skos:broadMatch` and `skos:narrowMatch` are used to state a hierarchical mapping link between two concepts.

The property `skos:relatedMatch` is used to state an associative mapping link between two concepts.

The property `skos:closeMatch` is used to link two concepts that are sufficiently similar that they can be used interchangeably in **some** information retrieval applications. In order to avoid the possibility of "compound errors" when combining mappings across more than two concept schemes, `skos:closeMatch` is **not** declared to be a transitive property.

The property `skos:exactMatch` is used to link two concepts, indicating a high degree of confidence that the concepts can be used interchangeably across a wide range of information retrieval applications. `skos:exactMatch` is a transitive property, and is a sub-property of `skos:closeMatch`.

10.2. Vocabulary

`skos:mappingRelation`

`skos:closeMatch`

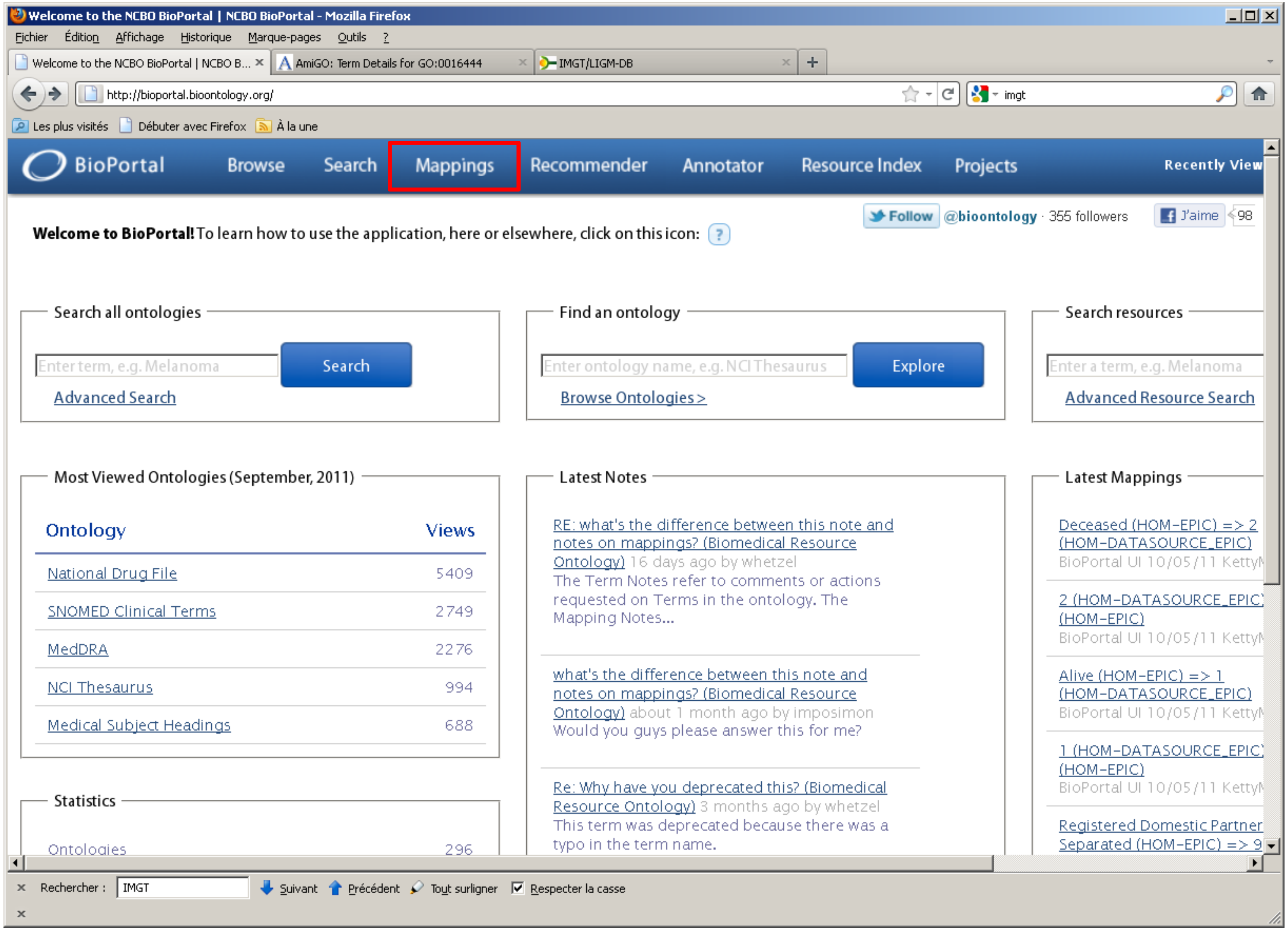
`skos:exactMatch`

`skos:broadMatch`

`skos:narrowMatch`

`skos:relatedMatch`

Mappings sur le site NCBO Bioportal



Welcome to BioPortal! To learn how to use the application, here or elsewhere, click on this icon: ?

Search all ontologies

Enter term, e.g. Melanoma

Advanced Search

Find an ontology

Enter ontology name, e.g. NCI Thesaurus

Browse Ontologies >

Search resources

Enter a term, e.g. Melanoma

Advanced Resource Search

Most Viewed Ontologies (September, 2011)

Ontology	Views
National Drug File	5409
SNOMED Clinical Terms	2749
MedDRA	2276
NCI Thesaurus	994
Medical Subject Headings	688

Statistics

Ontologies	296

Latest Notes

RE: what's the difference between this note and notes on mappings? (Biomedical Resource Ontology) 16 days ago by whetzel
The Term Notes refer to comments or actions requested on Terms in the ontology. The Mapping Notes...

what's the difference between this note and notes on mappings? (Biomedical Resource Ontology) about 1 month ago by imposimon
Would you guys please answer this for me?

Re: Why have you deprecated this? (Biomedical Resource Ontology) 3 months ago by whetzel
This term was deprecated because there was a typo in the term name.

Latest Mappings

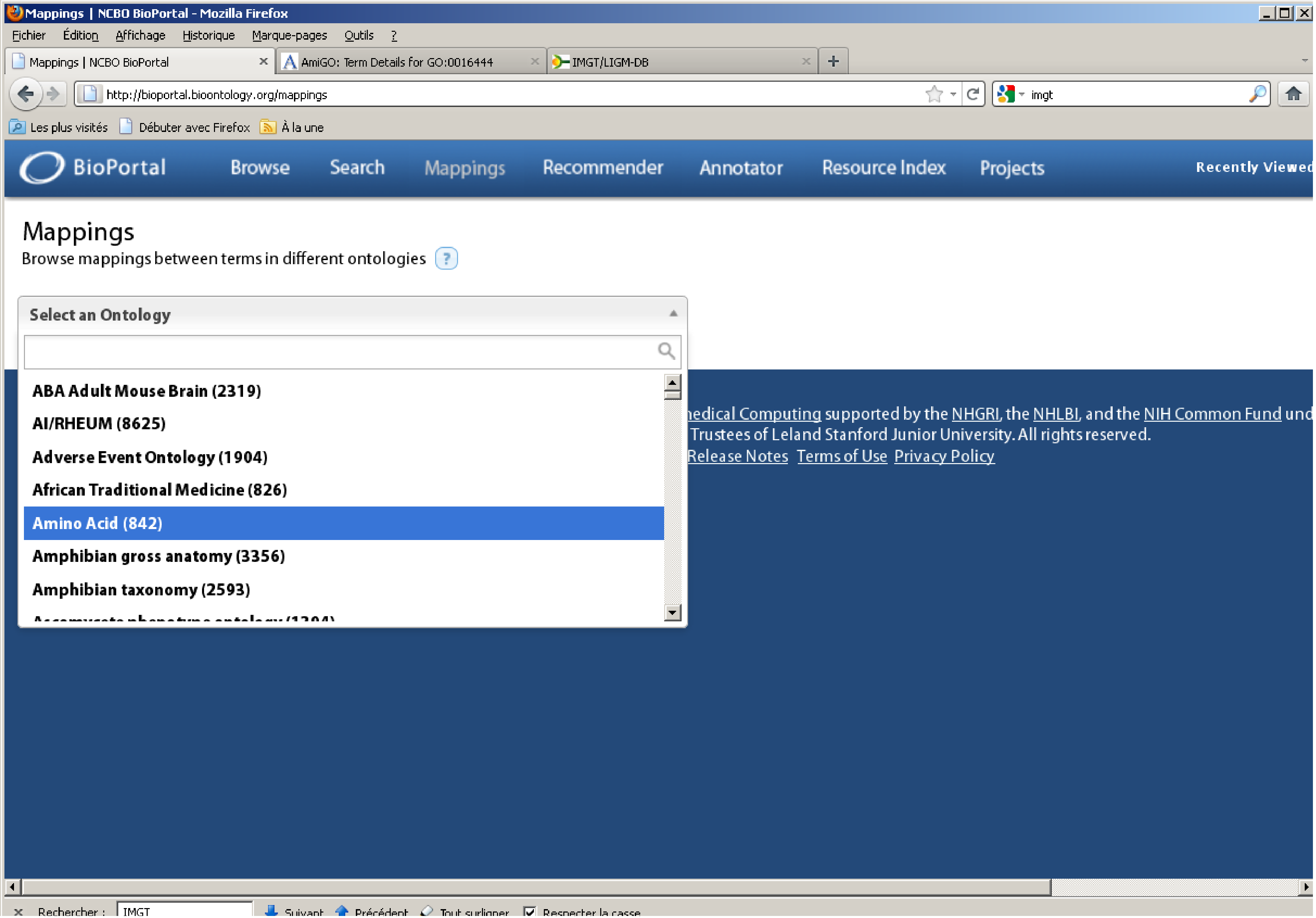
Deceased (HOM-EPIC) => 2 (HOM-DATASOURCE_EPIC)
BioPortal UI 10/05/11 Ketty

2 (HOM-DATASOURCE_EPIC) (HOM-EPIC)
BioPortal UI 10/05/11 Ketty

Alive (HOM-EPIC) => 1 (HOM-DATASOURCE_EPIC)
BioPortal UI 10/05/11 Ketty

1 (HOM-DATASOURCE_EPIC) (HOM-EPIC)
BioPortal UI 10/05/11 Ketty

Registered Domestic Partner Separated (HOM-EPIC) => 9



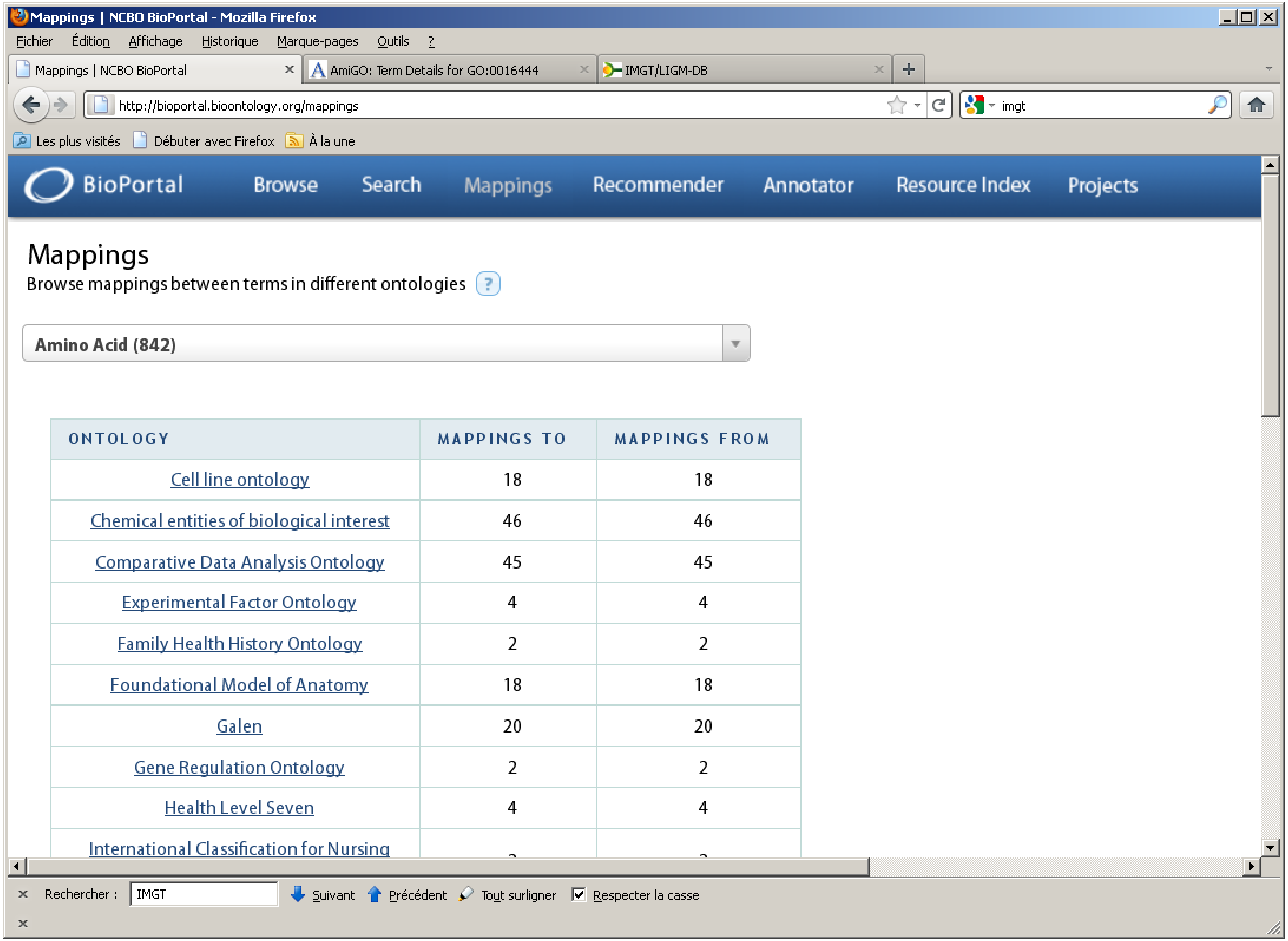
Mappings

Browse mappings between terms in different ontologies

Select an Ontology

- ABA Adult Mouse Brain (2319)
- AI/RHEUM (8625)
- Adverse Event Ontology (1904)
- African Traditional Medicine (826)
- Amino Acid (842)
- Amphibian gross anatomy (3356)
- Amphibian taxonomy (2593)

edical Computing supported by the NHGRI, the NHLBI, and the NIH Common Fund und
Trustees of Leland Stanford Junior University. All rights reserved.
Release Notes Terms of Use Privacy Policy

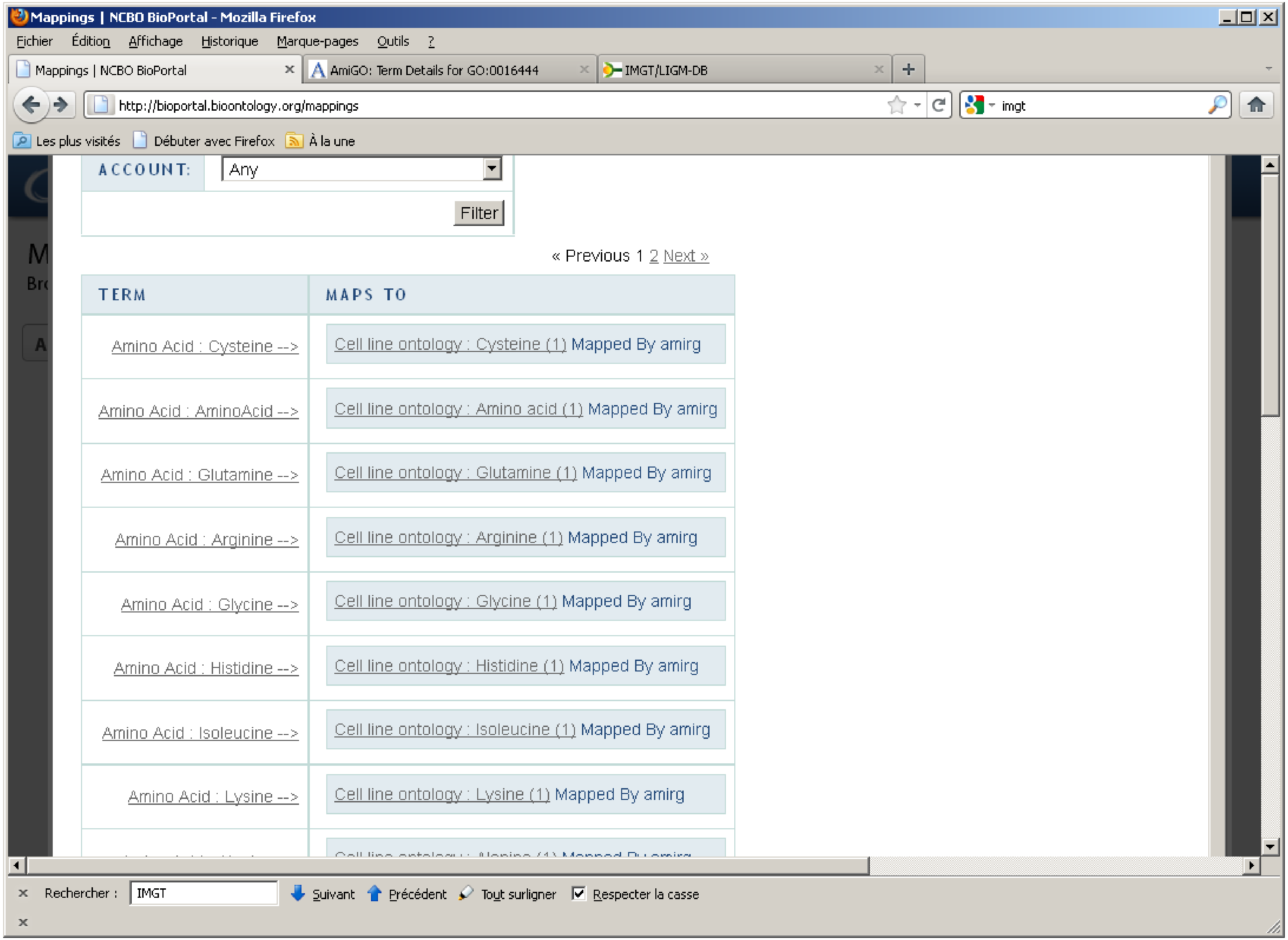


Mappings

Browse mappings between terms in different ontologies

Amino Acid (842)

ONTOLOGY	MAPPINGS TO	MAPPINGS FROM
Cell line ontology	18	18
Chemical entities of biological interest	46	46
Comparative Data Analysis Ontology	45	45
Experimental Factor Ontology	4	4
Family Health History Ontology	2	2
Foundational Model of Anatomy	18	18
Galen	20	20
Gene Regulation Ontology	2	2
Health Level Seven	4	4
International Classification for Nursing		



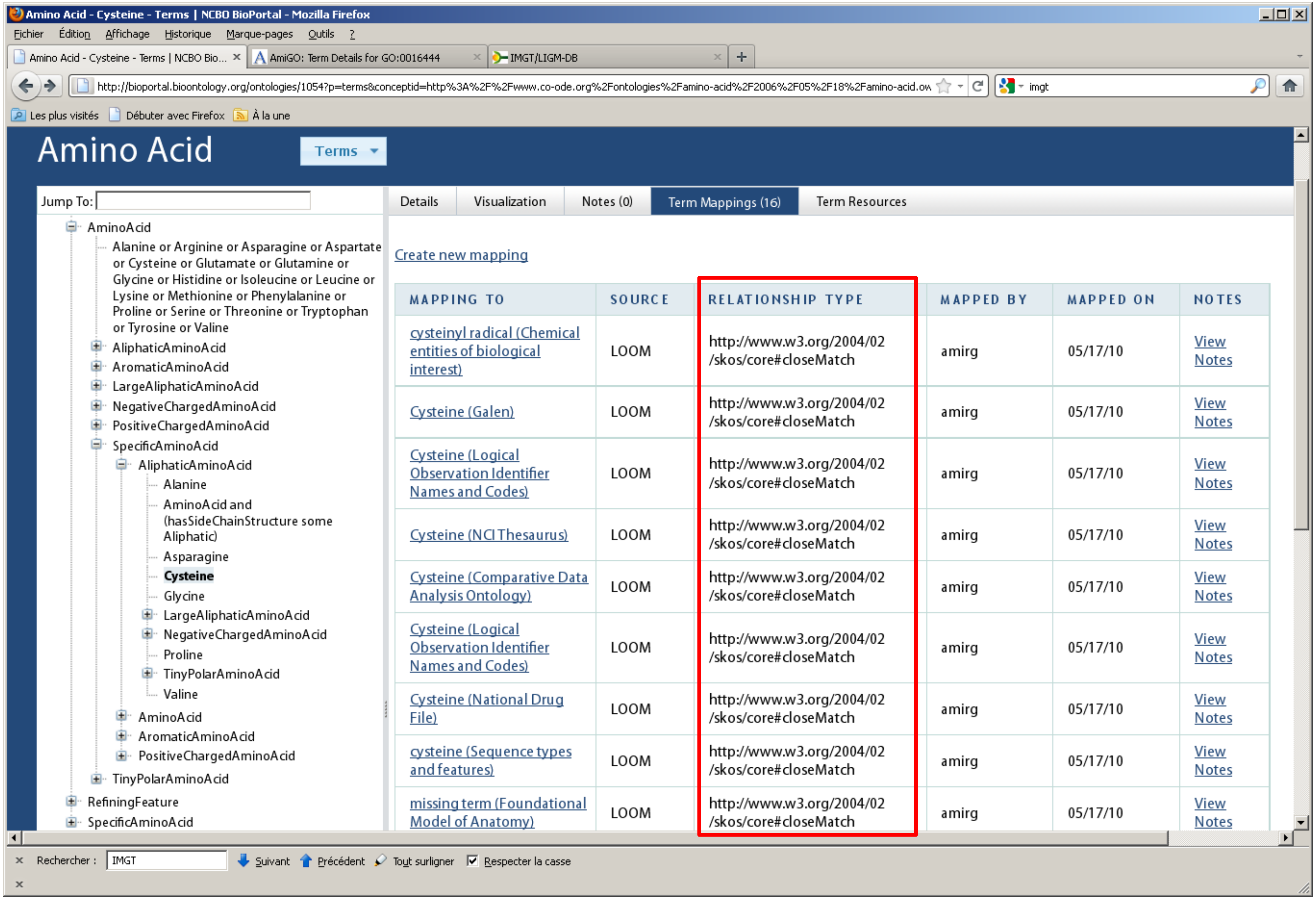
http://bioportal.bioontology.org/mappings

imgt

ACCOUNT: Any

TERM	MAPS TO
Amino Acid : Cysteine -->	Cell line ontology : Cysteine (1) Mapped By amirg
Amino Acid : AminoAcid -->	Cell line ontology : Amino acid (1) Mapped By amirg
Amino Acid : Glutamine -->	Cell line ontology : Glutamine (1) Mapped By amirg
Amino Acid : Arginine -->	Cell line ontology : Arginine (1) Mapped By amirg
Amino Acid : Glycine -->	Cell line ontology : Glycine (1) Mapped By amirg
Amino Acid : Histidine -->	Cell line ontology : Histidine (1) Mapped By amirg
Amino Acid : Isoleucine -->	Cell line ontology : Isoleucine (1) Mapped By amirg
Amino Acid : Lysine -->	Cell line ontology : Lysine (1) Mapped By amirg

Rechercher : IMGT ☑



MAPPING TO	SOURCE	RELATIONSHIP TYPE	MAPPED BY	MAPPED ON	NOTES
cysteinyl radical (Chemical entities of biological interest)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
Cysteine (Galen)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
Cysteine (Logical Observation Identifier Names and Codes)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
Cysteine (NCI Thesaurus)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
Cysteine (Comparative Data Analysis Ontology)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
Cysteine (Logical Observation Identifier Names and Codes)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
Cysteine (National Drug File)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
cysteine (Sequence types and features)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes
missing term (Foundational Model of Anatomy)	LOOM	http://www.w3.org/2004/02/skos/core#closeMatch	amirg	05/17/10	View Notes

Alignement d’ontologies:

L’alignement d'ontologies permet d’établir des liens de correspondance entre deux ontologies originales.

Les ontologies destinées à être alignées sont des ontologi
Distinctes.

L’alignement si les ontologies concernées deviennent « homogènes » et ceci tout en restant séparées

=>L’automatisation de ces tâches est encore du domaine de la recherche

Fusion d’ontologies :

La fusion d'ontologies permet de créer une seule ontologie qui regroupe les connaissances des deux (ou plus) ontologies fusionnées.

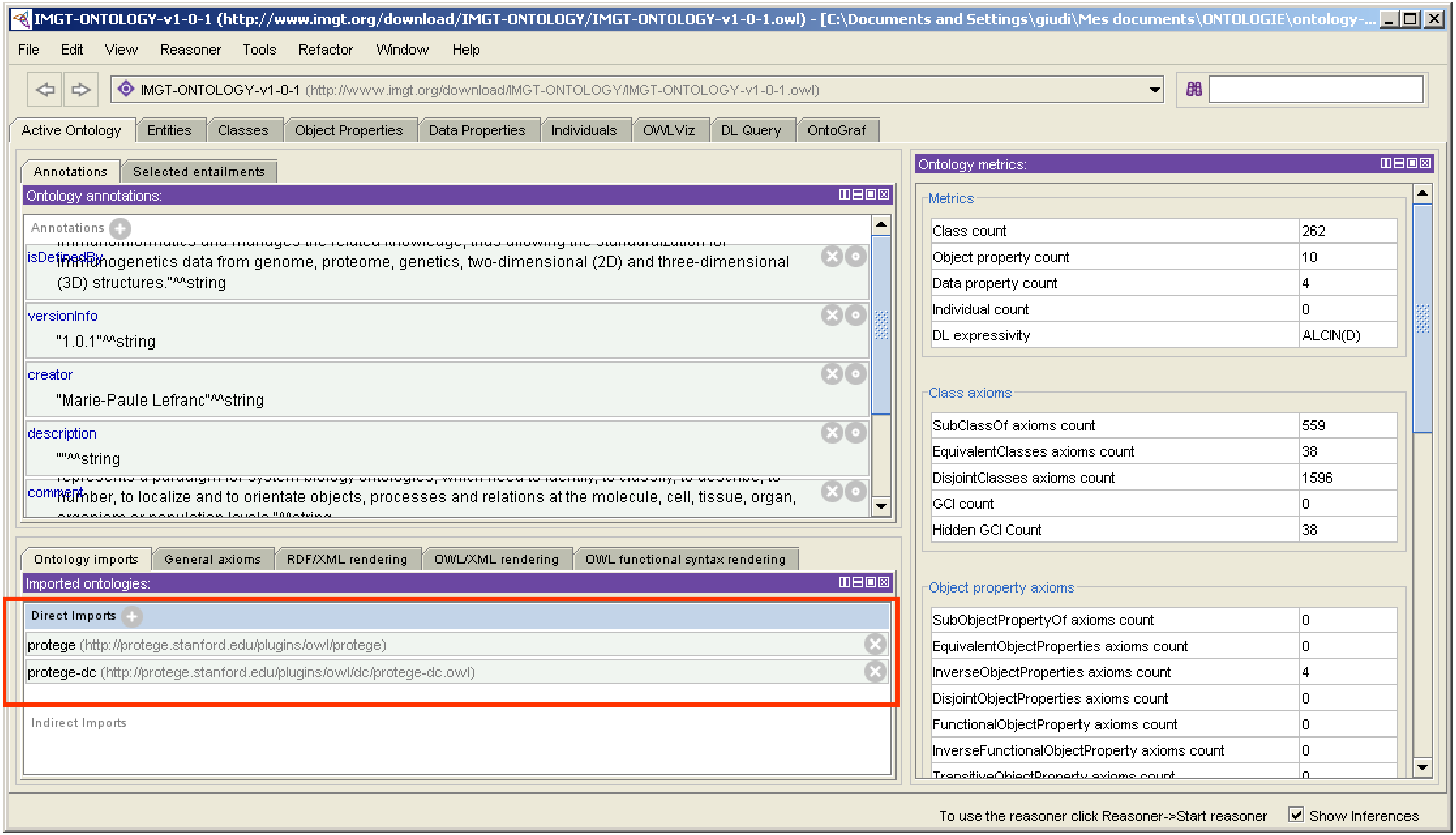
Ces ontologies appartiennent en général au même domaine d’application.

Intégration d’ontologies :

L'intégration d'ontologies permet de construire une nouvelle ontologie tout en utilisant des connaissances formalisées dans l’ontologie intégrée.

La nouvelle ontologie n'est pas destinée à remplacer l’ontologie intégrée.

IMGT-ONTOLOGY intègre les ontologies « protegé », « protege-dc »



isDefinedBy
immunogenetics data from genome, proteome, genetics, two-dimensional (2D) and three-dimensional (3D) structures."^^string

versionInfo
"1.0.1"^^string

creator
"Marie-Paule Lefranc"^^string

description
""^^string

comment
number, to localize and to orientate objects, processes and relations at the molecule, cell, tissue, organ,

Direct Imports

protege (http://protege.stanford.edu/plugins/owl/protege)

protege-dc (http://protege.stanford.edu/plugins/owl/dc/protege-dc.owl)

Indirect Imports

Metrics

Class count	262
Object property count	10
Data property count	4
Individual count	0
DL expressivity	ALCIN(D)

Class axioms

SubClassOf axioms count	559
EquivalentClasses axioms count	38
DisjointClasses axioms count	1596
GCI count	0
Hidden GCI Count	38

Object property axioms

SubObjectPropertyOf axioms count	0
EquivalentObjectProperties axioms count	0
InverseObjectProperties axioms count	4
DisjointObjectProperties axioms count	0
FunctionalObjectProperty axioms count	0
InverseFunctionalObjectProperty axioms count	0

To use the reasoner click Reasoner->Start reasoner ☑ Show Inferences

A quoi servent le mapping/alignement/fusion/intégration?

- Facilite l’extraction de données dans des bases hétérogène

- Correspondance entre les termes utilisés dans différents domaines pour l’expertise de données

- Pour le web sémantique

- Collaboration entre logiciels

Exemples :

Partage des données entre diverses bases de données

Relier des données entre les différentes bases

Quelle est la signification des liens?

Dans les fichiers EMBL:

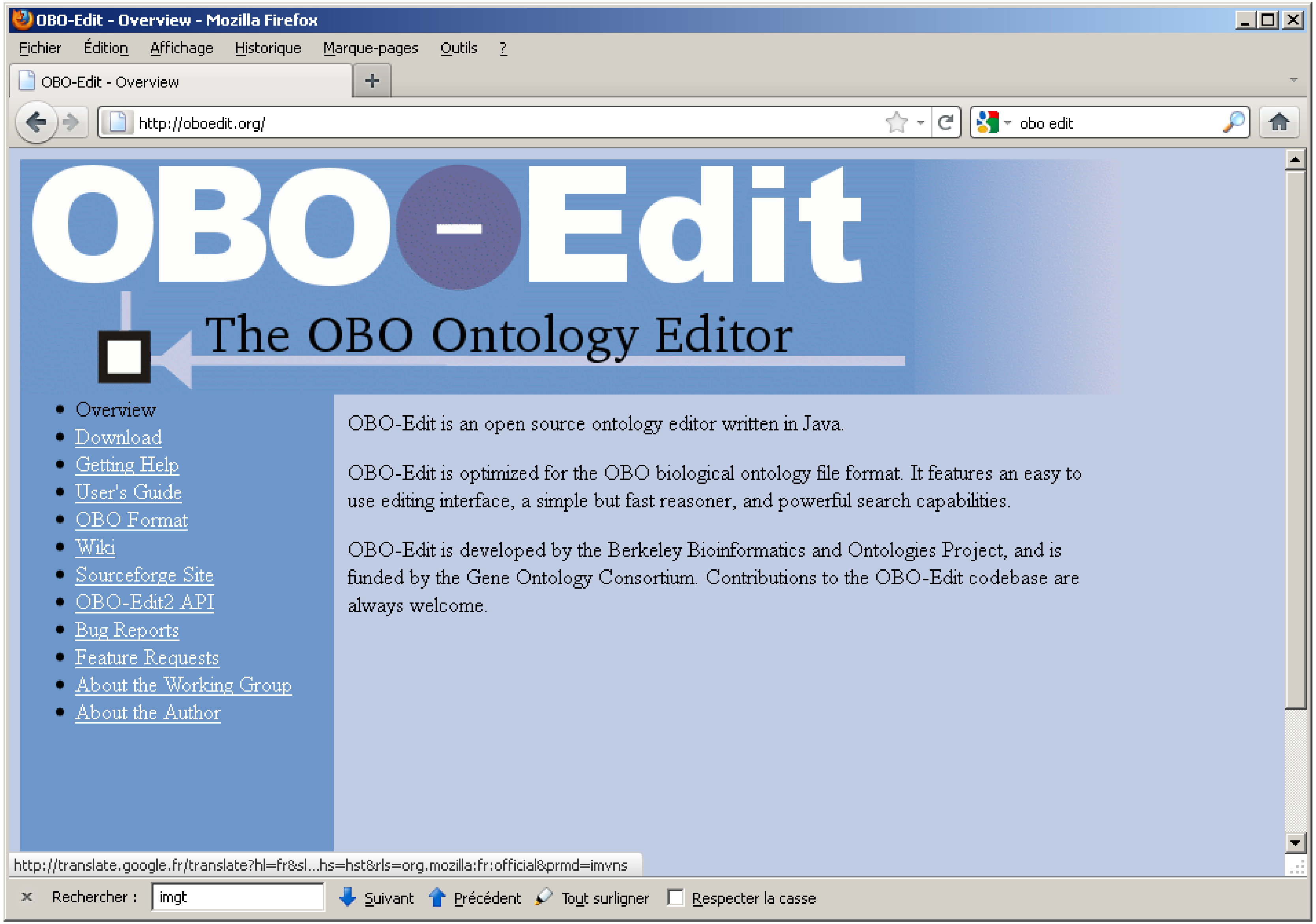
-**Les numéros d’accès reliés (intra EMBL): ligne AC**
  - **pour rappeler les anciens numéros d’accès**
  - **pour indiquer des séquences identiques, incluses l’une dans l’autre,….**

-**La liaison avec les données dans d’autres bases: ligne DR**
  - **identité**
  - **protéine correspondante**
  - **gènes impliqués ….**

⇒ **Compréhension implicite (ou presque) par l’être humain**
⇒ **qu’en est-il des machines?**

Exemple d'éditeur d'ontologie: OBO-Edit (GO, SO, …)

http://oboedit.org/



- Overview
- Download
- Getting Help
- User's Guide
- OBO Format
- Wiki
- Sourceforge Site
- OBO-Edit2 API
- Bug Reports
- Feature Requests
- About the Working Group
- About the Author

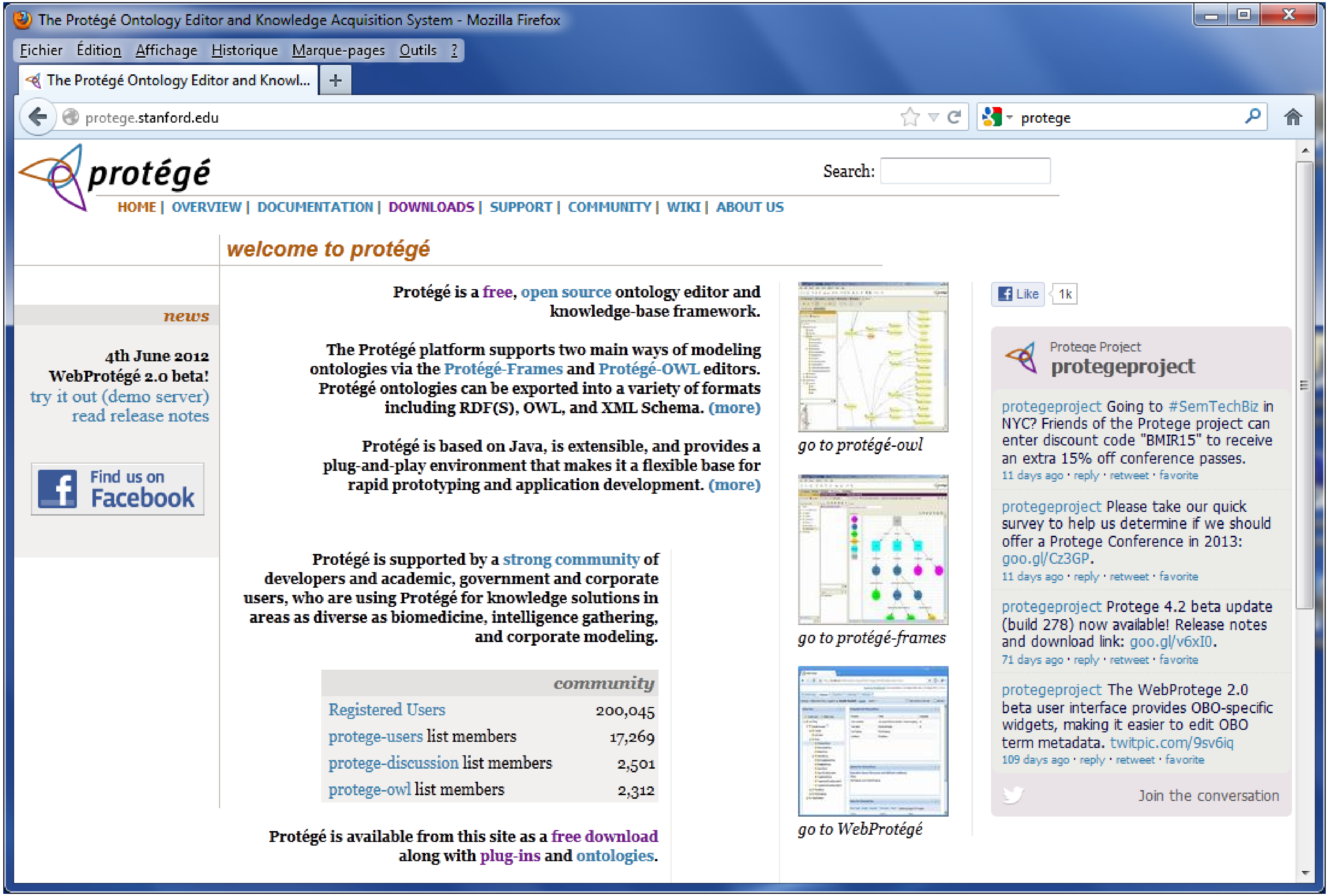
OBO-Edit is an open source ontology editor written in Java.

OBO-Edit is optimized for the OBO biological ontology file format. It features an easy to use editing interface, a simple but fast reasoner, and powerful search capabilities.

OBO-Edit is developed by the Berkeley Bioinformatics and Ontologies Project, and is funded by the Gene Ontology Consortium. Contributions to the OBO-Edit codebase are always welcome.

Exemple d’éditeur d’ontologie: Protégé

http://protege.stanford.edu/



welcome to protégé

news

4th June 2012
WebProtégé 2.0 beta!
try it out (demo server)
read release notes

Protégé is a free, open source ontology editor and knowledge-base framework.

The Protégé platform supports two main ways of modeling ontologies via the Protégé-Frames and Protégé-OWL editors. Protégé ontologies can be exported into a variety of formats including RDF(S), OWL, and XML Schema. (more)

Protégé is based on Java, is extensible, and provides a plug-and-play environment that makes it a flexible base for rapid prototyping and application development. (more)

Protégé is supported by a strong community of developers and academic, government and corporate users, who are using Protégé for knowledge solutions in areas as diverse as biomedicine, intelligence gathering, and corporate modeling.

community	
Registered Users	200,045
protege-users list members	17,269
protege-discussion list members	2,501
protege-owl list members	2,312

Protégé is available from this site as a free download along with plug-ins and ontologies.

go to protégé-owl

go to protégé-frames

go to WebProtégé

Protege Project
protegeproject

protegeproject Going to #SemTechBiz in NYC? Friends of the Protege project can enter discount code "BMIR15" to receive an extra 15% off conference passes.
11 days ago · reply · retweet · favorite

protegeproject Please take our quick survey to help us determine if we should offer a Protege Conference in 2013: goo.gl/Cz3GP.
11 days ago · reply · retweet · favorite

protegeproject Protege 4.2 beta update (build 278) now available! Release notes and download link: goo.gl/v6xI0.
71 days ago · reply · retweet · favorite

protegeproject The WebProtege 2.0 beta user interface provides OBO-specific widgets, making it easier to edit OBO term metadata. twitpic.com/9sv6iq
109 days ago · reply · retweet · favorite

Join the conversation

Rôles des responsables de SI

Assurer la cohérence d'un ensemble d'informations qui comprend des données (dans une base de données...), des applications, interfaces pour un ensemble d'utilisateurs

On demande en plus:

- la disponibilité de l'information le plus rapidement possible
- simultanément sous différents formats
- à travers diverses interfaces
- "tout le monde" veut offrir des services

Les utilisateurs ont des demandes difficiles à anticiper,
qui évoluent rapidement

Complexité des SI

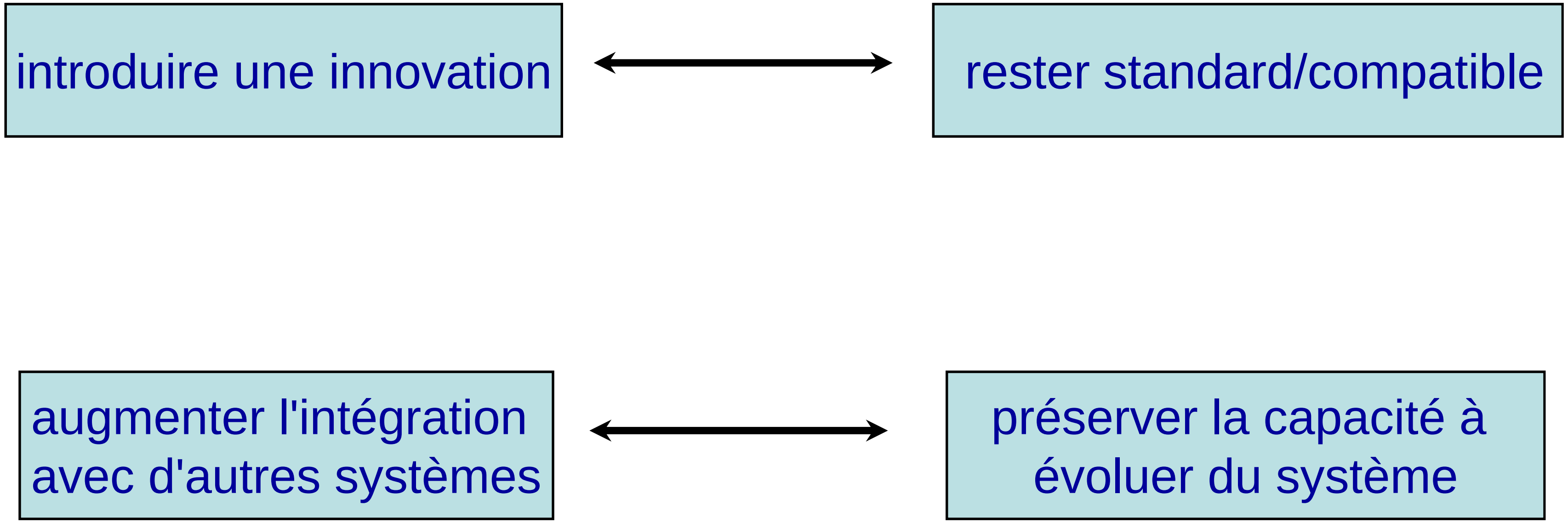
- Les systèmes d'information n'ont jamais été conçus pour durer si longtemps ou pour être résistants au temps qui passe (bug de l'an 2000)
- Il faudrait adapter certaines méthodes, standards, protocoles dépassés/oubliés.
- Les versions des outils, standards, changent plusieurs fois au cours d'un même projet.
- “La toute dernière release, (ou la version beta)" des outils disponibles comprend souvent de nombreux bugs.

Quand c'est enfin mature, c'est déjà obsolète !

La durée de vie d'un SI dépend de

- Portabilité: il existe plusieurs plate-formes alternatives qui peuvent faire tourner le système
- Evolutivité: un système est évolutif, on peut:
  - ajouter des fonctionnalités
  - augmenter ses capacités maximum
  - l'adapter à des besoins différents/nouveaux
  - corriger des dysfonctionnements
- Réutilisation: Un système est réutilisable s'il peut être intégré (en totalité ou en partie) dans un nouveau projet.
- Standardisation: **ensemble de propriétés décrites dans une norme, supportée ou certifiée par un organisme officiel.**

L’équilibre des SI



introduire une innovation

rester standard/compatible

augmenter l'intégration avec d'autres systèmes

préserver la capacité à évoluer du système

Im
Muno
Gene
Tics

Information
system®

http://www.imgt.org