http://imgt.cines.fr

Systèmes d'information
-
Ontologies
Bases de connaissances

UE FMBS312 Bioinformatique (TC3 du Master de recherche 2)
Universités Montpellier 1 et Montpellier 2, ENSCM

Véronique Giudicelli
LIGM, IGH, UPR CNRS 1142
141 rue de la Cardonille
34296 Montpellier Cedex 5
tel: 04.99.61.99.65 fax: 04.99.61.99.01
e-mail: Veronique.Giudicelli@igh.cnrs.fr

29 septembre 2008

http://www.cbs.cnrs.fr/spip.php?rubrique134

login: CBS mot de passe: coursCBS

Centre de Biochimie Structurale - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.cbs.cnrs.fr/spip.php?rubrique174

Most Visited | Getting Started | Latest Headlines | tomcat | Outils linguistiques

CNRS-IGH WebMail | Centre de Biochimie Struct...

Centre de Biochimie Structurale Montpellier

Le CNRS | Accueil CNRS | Autres sites CNRS

Accès aux enseignements en ligne > Réservé aux étudiants avec mot de passe > Année 2008-2009 > Master Mention Biologie Santé > M2__FMBS312_Bioinformatique

Accueil CBS
Contacts /
Thèmes de recherche /
Plateforme R.I.O.
Publications
Pôles Techniques / Technical poles
A.B.C.I.S.
Formations
Partenariats /

M2__FMBS312_Bioinformatique

Accès validé pour **Etudiant Master** . Cliquez ici pour vous déconnecter

Enseignement : **planning** réalisé par : Stefano Trapani

Responsables :

- Marie-Paule Lefranc
- Stefano Trapani

PLANNING

lun 29 sept 2008

14:30—16:20 Bases de connaissances, systèmes d'information et ontologies. Véronique Giudicelli

16:40—18:30 Concepts de bases de données. Cyril Sarrauste de Menthière

mar 30 sept 2008

14:30—16:20 Analyse de séquences de protéines Stefano Trapani

16:40—18:30 Protéomiques et analyse de données expérimentales (MS, RMN). Alain Chavanieu

mer 1 oct 2008

Done

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://imgt.cines.fr)

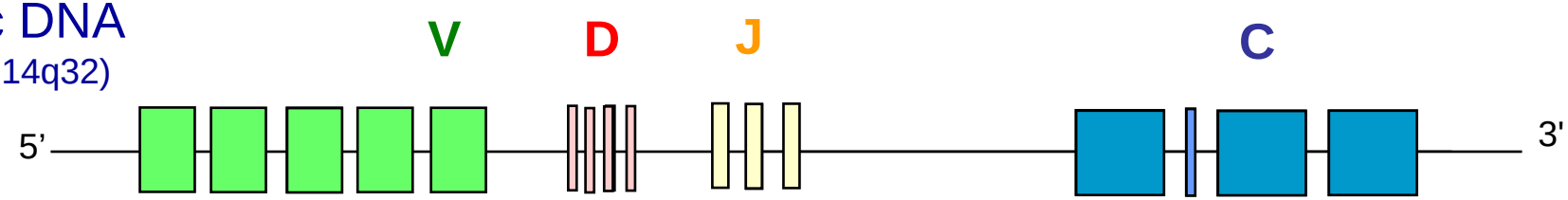
Source unique de connaissances en immunogénétique et immunoinformatique, au niveau international.

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.

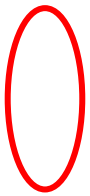
Le contexte biologique:
la synthèse des IG et des TR

http://imgt.cines.fr

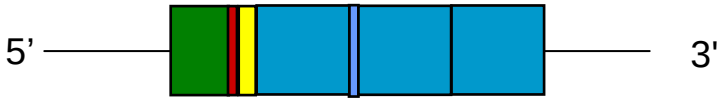
genomic DNA
(IGH Locus 14q32)



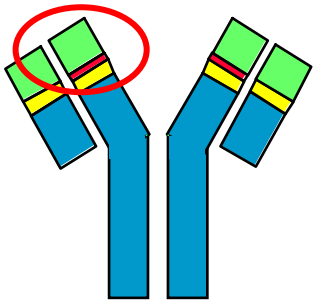
rearranged
DNA



mRNA



Potentiellement 2 x 10^{12}
differentes IG
Par individu



Diversité des junctions : exemple d'une chaîne lourde d'IG

http://imgt.cines.fr

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

tgtgcgaaaga tacg agcatattgtggtggtgactgctattcc gat acaactggttcgacccc**tgg**

2nd-CYS

JUNCTION

J-TRP

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

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

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, …

http://imgt.cines.fr

IMGT Home page - Mozilla Firefox

File Edit View History Bookmarks Tools Help

IMGT databases

- IMGT/LIGM-DB (IG and TR from 223 species) (LIGM, Montpellier, France) (**126 667 entries**)
- IMGT/MHC-DB (IMGT/MHC-HLA, -NHP, -DLA, -FLA) (ANRI, BPRC, hosted at EBI)
- IMGT/PRIMER-DB (IG and TR from 11 species) (LIGM, Montpellier, France) (**1 864 entries**)

- IMGT/GENE-DB (IG and TR genes from human, mouse and rat) (LIGM, Montpellier, France) (**1 911 genes, 2 909 alleles**)

- IMGT/3Dstructure-DB (IG and TR, MHC and RPI gene and allele identification and IMGT Colliers de Perles) (**1 461 entries**)

IMGT Web resources

- **IMGT Repertoire** (IG and TR, MHC and RPI)
- **IMGT Index** (FactsBook)
- **IMGT Scientific chart** (Sequence description, Numbering, Nomenclature, Representation rules)
- **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 Medical page, IMGT Veterinary page, IMGT Biotechnology page
- IMGT Posters and diaporama
- **The IMGT Immunoinformatics page**

IMGT tools

- IMGT/V-QUEST (sequence alignment software for IG, TR and HLA)
- IMGT/JunctionAnalysis (for human and mouse IG and TR)
- IMGT/Allele-Align
- IMGT/PhyloGene
- IMGT/DomainDisplay (Amino acid sequences)

- IMGT/LocusView, IMGT/GeneView, IMGT/GeneSearch, IMGT/CloneSearch (for human IGK, IGL, IGH, TRA/TRD, TRB, TRG, mouse TRA/TRD and human MHC)
- IMGT/GeneInfo (TIMC and ICH, Grenoble; LIGM, Montpellier)
- IMGT/GeneFrequency

- IMGT/DomainGapAlign
- IMGT/Collier-de-Perles
- IMGT/DomainSuperimpose
- IMGT/StructuralQuery

IMGT other accesses

- IMGT Other accesses (SRS, MRS)
- Compare your sequence against IMGT (BLAST, FASTA)
- IMGT/LIGM-DB Sequence submission
- IMGT flat file release information

Search

Google

WWW IMGT domain

Done

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

http://imgt.cines.fr

IMGT Home page - Mozilla Firefox

File Edit View History Bookmarks Tools Help

IMGT databases

- IMGT/LIGM-DB (IG and TR from 223 species) (LIGM, Montpellier, France) (**126 667 entries**)
- IMGT/MHC-DB (IMGT/MHC-HLA, -NHP, -DLA, -FLA) (ANRI, BPRC, hosted at EBI)
- IMGT/PRIMER-DB (IG and TR from 11 species) (LIGM, Montpellier, France) (**1 864 entries**)

- IMGT/GENE-DB (IG and TR genes from human, mouse and rat) (LIGM, Montpellier, France) (**1 911 genes, 2 909 alleles**)

- IMGT/3Dstructure-DB (IG and TR, MHC and RPI gene and allele identification and IMGT Colliers de Perles) (**1 461 entries**)

IMGT Web resources

- **IMGT Repertoire** (IG and TR, MHC and RPI)
- **IMGT Index** (FactsBook)
- **IMGT Scientific chart** (Sequence description, Numbering, Nomenclature, Representation rules)
- **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 Medical page, IMGT Veterinary page, IMGT Biotechnology page
- IMGT Posters and diaporama
- **The IMGT Immunoinformatics page**

IMGT tools

- IMGT/V-QUEST (sequence alignment software for IG, TR and HLA)
- IMGT/JunctionAnalysis (for human and mouse IG and TR)
- IMGT/Allele-Align
- IMGT/PhyloGene
- IMGT/DomainDisplay (Amino acid sequences)

- IMGT/LocusView, IMGT/GeneView, IMGT/GeneSearch, IMGT/CloneSearch (for human IGK, IGL, IGH, TRA/TRD, TRB, TRG, mouse TRA/TRD and human MHC)
- IMGT/GeneInfo (TIMC and ICH, Grenoble; LIGM, Montpellier)
- IMGT/GeneFrequency

- IMGT/DomainGapAlign
- IMGT/Collier-de-Perles
- IMGT/DomainSuperimpose
- IMGT/StructuralQuery

IMGT other accesses

- IMGT Other accesses (SRS, MRS)
- Compare your sequence against IMGT (BLAST, FASTA)
- IMGT/LIGM-DB Sequence submission
- IMGT flat file release information

Search

Google

WWW IMGT domain

Done

Les objectifs de IMGT/LIGM-DB

http://imgt.cines.fr

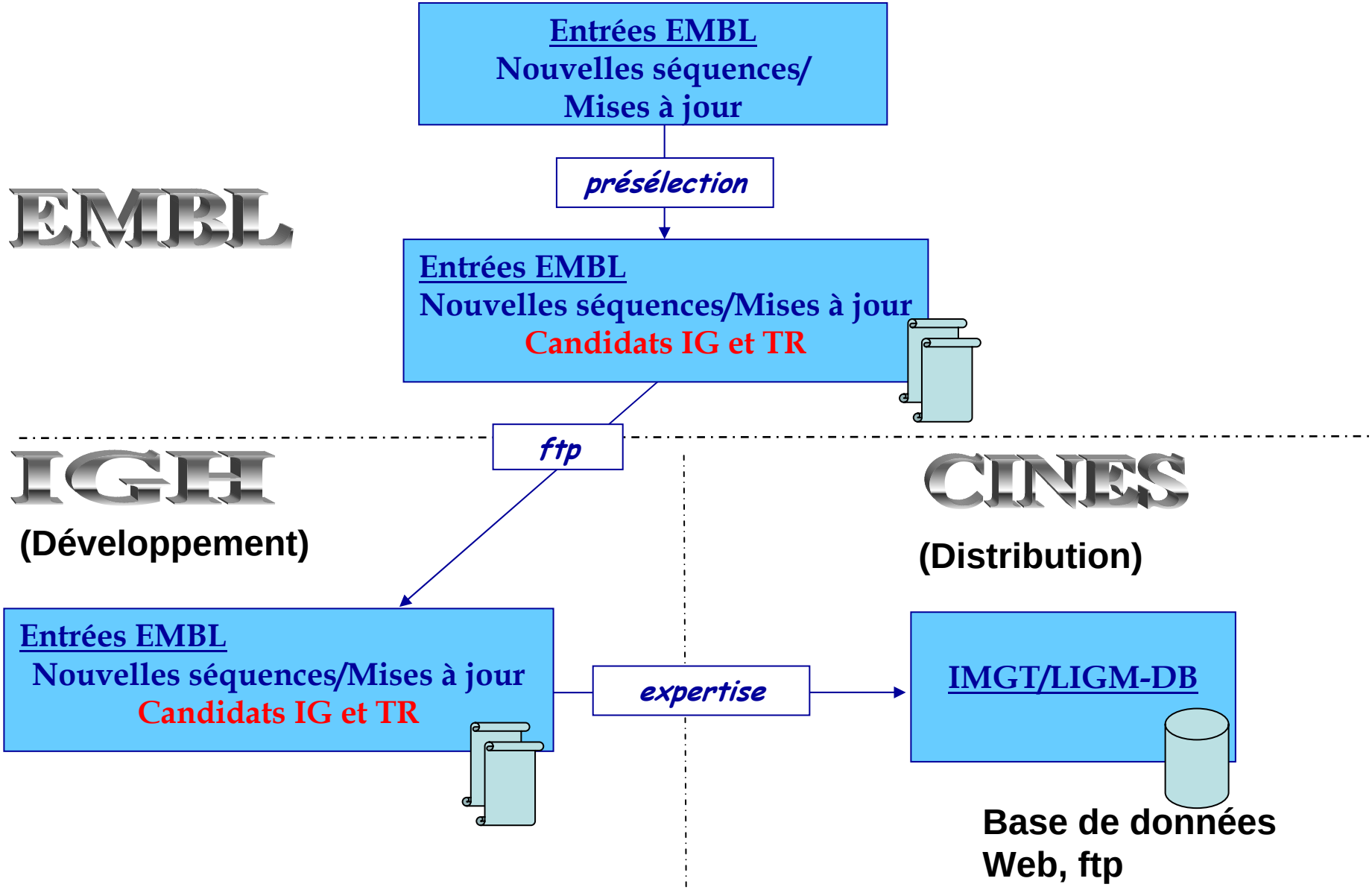
- Contenir toutes les séquences IG et TR des bases EMBL/GenBank/DDBJ
  + de 126.600 séquences de 223 espèces

- 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

http://imgt.cines.fr



A l'arrivée dans IMGT®

http://imgt.cines.fr

IMGT/LIGM-DB - Mozilla Firefox

File Edit View History Bookmarks Tools Help

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.
```

Done

Après expertise dans IMGT®

http://imgt.cines.fr

IMGT/LIGM-DB - Mozilla Firefox

File Edit View History Bookmarks Tools Help

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
```

Done

A l'arrivée dans IMGT®

http://imgt.cines.fr

IMGT/LIGM-DB - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

```
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://imgt.cines.fr

IMGT/LIGM-DB - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

```
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"
```

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'annotation des données, telle qu'elle peut être réalisée par des experts dans un domaine particulier.

=> vocabulaire standardisé, contrôlé,
règles d'annotation, dépendance entre les termes

Un exemple de question

Qu'est-ce qu'un gène?

=> plusieurs définitions

Dans une base de données:

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

Recherches bibliographique à travers des thesaurus de mots clés

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

http://imgt.cines.fr

```
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"
```

Done

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

Des codes pour indiquer les types d'information

The EMBL Nucleotide Sequence Database: User Manual Release 87 - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

Hotmail | 500 Internal Server E... | Personnaliser les liens | Windows Media | vquest | Windows | IMGT/LIGM-DB Admini...

```
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)
```

Terminé

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

Webin - Features & Qualifiers - Mozilla Firefox

File Edit View History Bookmarks Tools Help

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

EBI > Databases > EMBL-Bank > Submission > Webin > Features & Qualifiers

EMBL Nucleotide Sequence Database Submissions: Webin - Features & Qualifiers

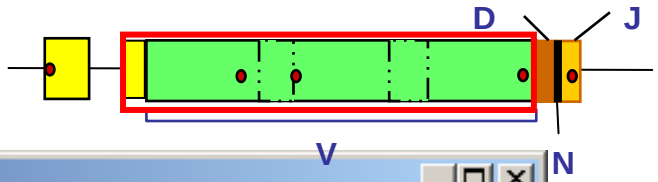
FEATURES

- attenuator
- C_region
- CAAT_signal
- CDS
- conflict
- D-loop
- D_segment
- enhancer
- exon
- gap
- GC_signal
- gene
- iDNA
- intron
- J_segment
- LTR
- mat_peptide
- misc_binding
- misc_difference
- misc_feature
- misc_recomb
- misc_RNA
- misc_signal
- misc_structure
- modified_base
- mRNA
- ncRNA
- N_region
- old_sequence
- operon
- oriT
- polyA_signal
- polyA_site
- precursor_RNA
- prim_transcript

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"

Done

EMBL Feature labels



Webin - Features & Qualifiers - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search

Databases Tools EBI Groups Training Industry About Us Help Site

EBI > Databases > EMBL-Bank > Submission > Webin > Features & Qualifiers

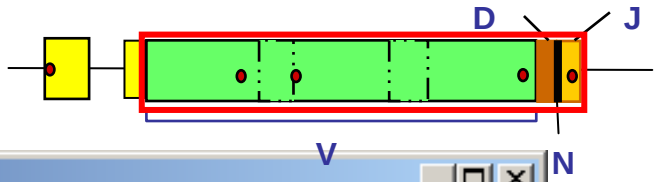
EMBL Nucleotide Sequence Database Submissions: Webin - Features & Qualifiers

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



Webin - Features & Qualifiers - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search

Databases | Tools | EBI Groups | Training | Industry | About Us | Help | Site

EBI > Databases > EMBL-Bank > Submission > Webin > Features & Qualifiers

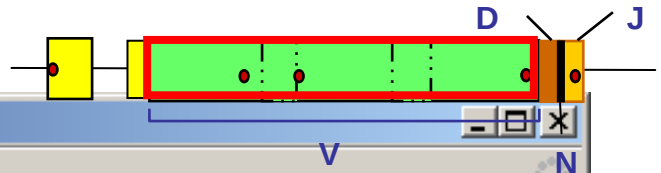
EMBL Nucleotide Sequence Database Submissions: Webin - Features & Qualifiers

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



IMGT/LIGM-DB - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

	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

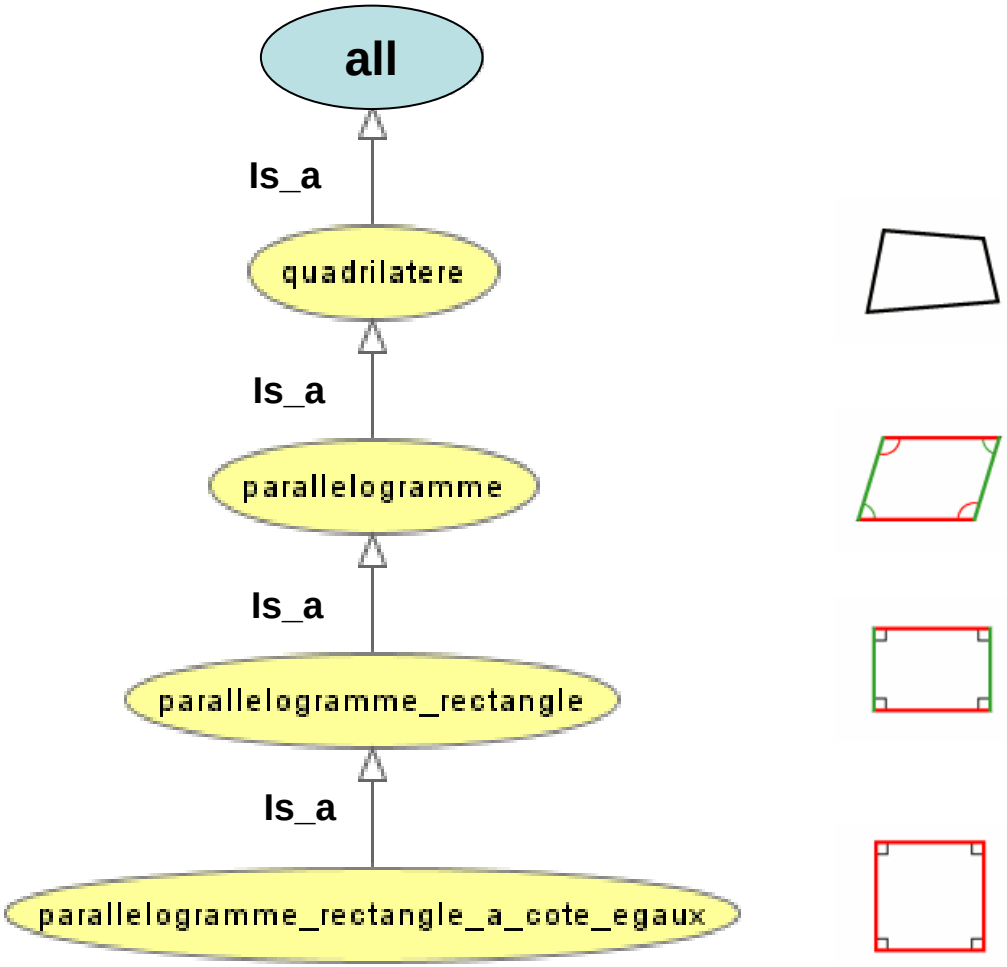
Ontologies

Au sein d'une ontologie, les termes sont regroupés sous forme de **concepts** (ou classes) sémantiques.

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**.

Exemple simple: les formes géométriques
représentation hiérarchique de concepts avec la relation « is_a »



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.

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.

OBO (Open Biological Ontologies)
recense les ontologies en biologie.
(http://obo.sourceforge.net/)

The Open Biomedical Ontologies - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.obofoundry.org/

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

The Open Biomedical Ontologies

Home | Contact

Ontologies Resources Participate About

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 candidate ontologies

Title	Domain	Prefix	File	Last changed
Amphibian gross anatomy	anatomy	AAO	amphibian_anatomy.obo	2008/06/19
Biological process	biological process	GO	gene_ontology_edit.obo 🐾	2008/09/27
C. elegans development	anatomy	WBls	worm_development.obo	2008/01/31
C. elegans gross anatomy	anatomy	WBbt	WBbt.obo 🐾	2008/07/09
C. elegans phenotype	phenotype	WBPhenotype	worm_phenotype.obo	2008/09/25
Cell type	anatomy	CL	cell.obo 🐾	2008/08/12
Cellular component	anatomy	GO	gene_ontology_edit.obo 🐾	2008/09/27
Cereal plant trait	phenotype	TO	plant_trait.obo 🐾	2008/04/05
Chemical entities of biological interest	biochemistry	CHEBI	chebi.obo 🐾	2008/09/25
Common Anatomy Reference Ontology	anatomy	CARO	caro.obo 🐾	2007/06/17
Dictyostelium discoideum anatomy	anatomy	DDANAT	dictyostelium_anatomy.obo 🐾	2008/05/29
Drosophila development	anatomy	FBdv	fly_development.obo 🐾	2007/03/20
Drosophila gross anatomy	anatomy	FBbt	fly_anatomy.obo 🐾	2008/06/13
Environment Ontology	environment	ENVO	envo.obo 🐾	2008/08/07
Evidence codes	experiments	ECO	evidence_code.obo	2008/05/29
Fly taxonomy	taxonomy	FBsp	fly_taxonomy.obo 🐾	2007/04/10
Foundational Model of Anatomy (subset)	anatomy	FMA	fma_obo.obo 🐾	
Fungal gross anatomy	anatomy	FAO	fungal_anatomy.obo	2005/08/10
Human developmental anatomy, abstract version	anatomy	EHDAA	human-dev-anat-abstract.obo	2007/06/17
Human developmental anatomy, timed				

Quick Links

- Mappings between ontologies
- Download alternate formats
- About the OBO Foundry
- 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)

Done

The Open Biomedical Ontologies - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.obofoundry.org/

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

Mass spectrometry	experiments	MS	psi-ms.obo	
Medaka fish anatomy and development	anatomy	MFO	medaka_ontology.obo	2005/11/08
Molecular function	biological function	GO	gene_ontology_edit.obo	2008/09/27
Mosquito gross anatomy	anatomy	TGMA	mosquito_anatomy.obo	2008/08/20
Mosquito insecticide resistance	environment	MIRO	mosquito_insecticide_resistance.obo	2008/09/27
Mouse adult gross anatomy	anatomy	MA	adult_mouse_anatomy.obo	2008/09/26
Mouse gross anatomy and development	anatomy	EMAP	EMAP.obo	2007/06/17
Mouse pathology	health	MPATH	mouse_pathology.obo	2005/08/10
NMR-instrument specific component of metabolomics investigations	experiments	NMR	NMR.owl	
OBO relationship types	all	OBO_REL	ro.obo	2008/05/06
Ontology for biomedical investigations	experiments	OBI	obi.owl	
Pathogen transmission	health	TRANS	transmission_process.obo	2008/02/06
Phenotypic quality	phenotype	PATO	quality.obo	2008/07/30
Plant growth and developmental stage	anatomy	PO	po_temporal.obo	2008/06/11
Plant structure	anatomy	PO	po_anatomy.obo	2008/09/12
Protein modification	proteins	MOD	PSI-MOD.obo	
protein ontology	proteins	PRO	pro.obo	2008/03/31
Protein-protein interaction	experiments	MI	psi-mi.obo	
Sequence types and features	biological sequence	SO	so.obo	2008/09/19
Spatial Ontology	anatomy	BSPO	spatial.obo	2008/03/18
Spider Ontology	anatomy	SPD	spider_comparative_biology.obo	2008/09/24
Subcellular anatomy ontology	anatomy	SAO	SAO.owl	
Suggested Ontology for Pharmacogenomics	health	SOPHARM	sopharm	
Systems Biology	biochemistry	SBO	SBO_OBO.obo	
Teleost anatomy and development	anatomy	TAO	teleost_anatomy.obo	2008/09/24
Teleost taxonomy	taxonomy	TTO	teleost_taxonomy.obo	2008/09/22
Tick gross anatomy	anatomy	TADS	tick_anatomy.obo	2007/12/13
Units of measurement	phenotype	UO	unit.obo	2007/12/20
Xenopus anatomy and development	anatomy	XAO	xenopus_anatomy.obo	2008/09/06
Yeast phenotypes	phenotype	YPO	yeast_phenotype.obo	2008/09/10

Find: sequence Next Previous Highlight all Match case

Done

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).

the Gene Ontology - Mozilla Firefox

File Edit View History Bookmarks Tools Help

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

the Gene Ontology

Search gene or protein name go!

- Open menus
- Home
- FAQ
- **Downloads**
- **Tools**
- **Documentation**
- **About GO**
- **Projects**
- Contact GO
- Site Map

Gene Ontology Home

The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism. *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.

GO website

- The latest news and views in the GO newsletter
- GO downloads, including ontology files, annotations and the GO database
- Tools for using GO, including OBO-Edit downloads, AmiGO, and the GO Online SQL Environment.
- Request new terms or ontology changes or get help with new term submission
- Documentation on all aspects of the GO project and the GO FAQ
- Projects within the GO consortium, including Reference Genomes and immune system annotation
- Gene Ontology mailing lists and contact details

Done

AmiGO: Term Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

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

Filter search results

Ontology

All
biological process
cellular component
molecular function

Set filters

Remove all filters

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

Select all | Clear all | Perform an action with the selected terms... | Go!

rel ↓	Accession , Term		Ontology
☐	GO:0007067 : **mitosis** [hide def] Progression through mitosis, the division of the eukaryotic cell nucleus to produce two daughter nuclei that, usually, contain the identical chromosome complement to their mother.	687 gene products view in tree	biological process
☐	GO:0051337 : a**mitosis** [show def]	0 gene products view in tree	biological process
☐	GO:0007113 : endomitotic cell cycle [show def] Query matches synonym "endo**mitosis**" [exact synonym]	5 gene products view in tree	biological process
☐	GO:0010458 : exit from **mitosis** [show def]	87 gene products view in tree	biological process
☐	GO:0007088 : regulation of **mitosis** [show def]	215 gene products view in tree	biological process

Done

AmiGO: Tree Browser - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://amigo.geneontology.org/cgi-bin/amigo/browse.cgi?action=set-tree&term=GO

go ontology

Filter tree view

Filter by ontology

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

Filter Gene Product Counts

Data source: All, CGD, dictyBase, EcoCyc

Species: All, Anaplasma phagocy..., Arabidopsis thaliana, Bacillus anthraci...

View Options

Tree view: Full, Compact

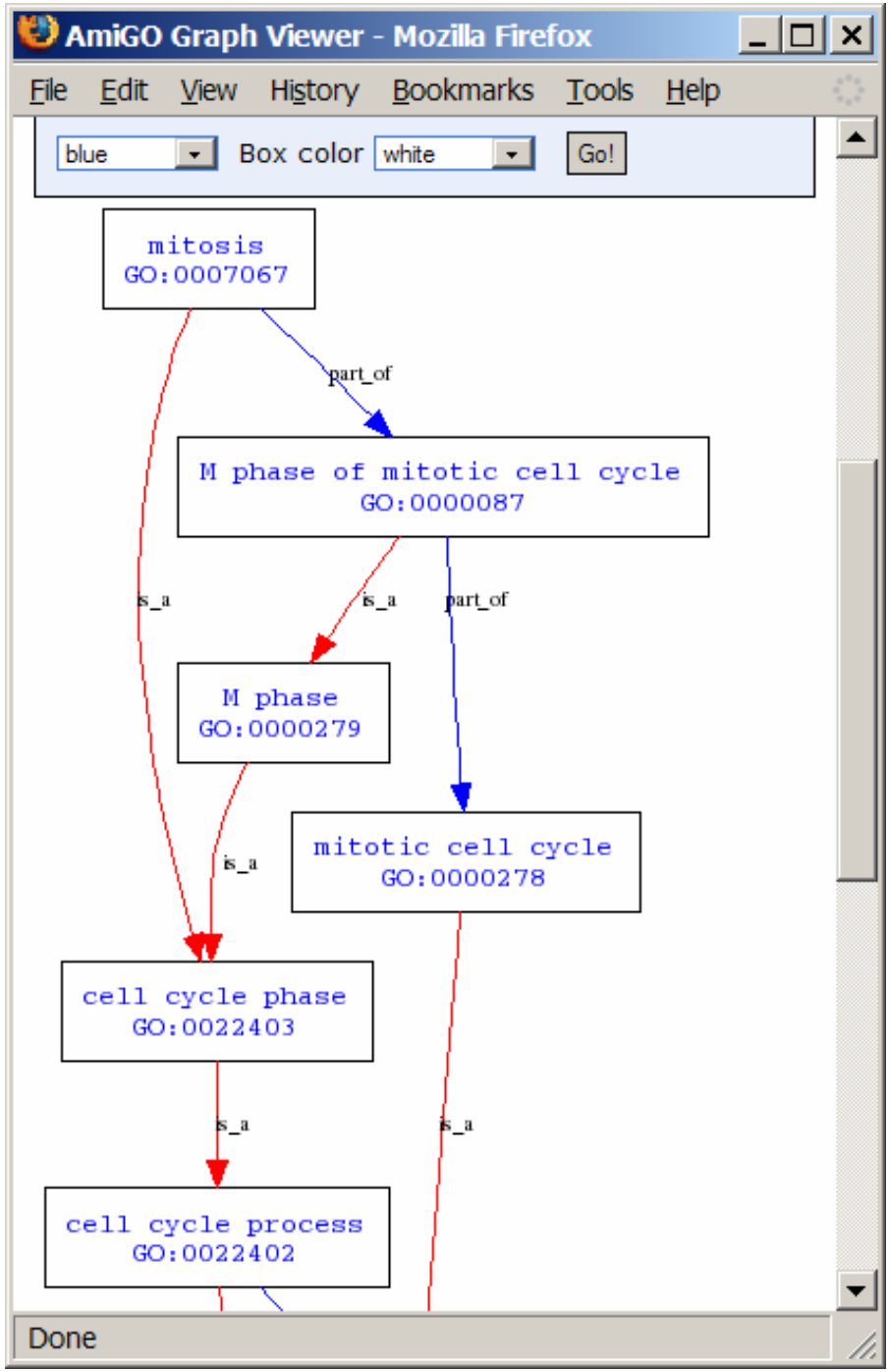
Set filters

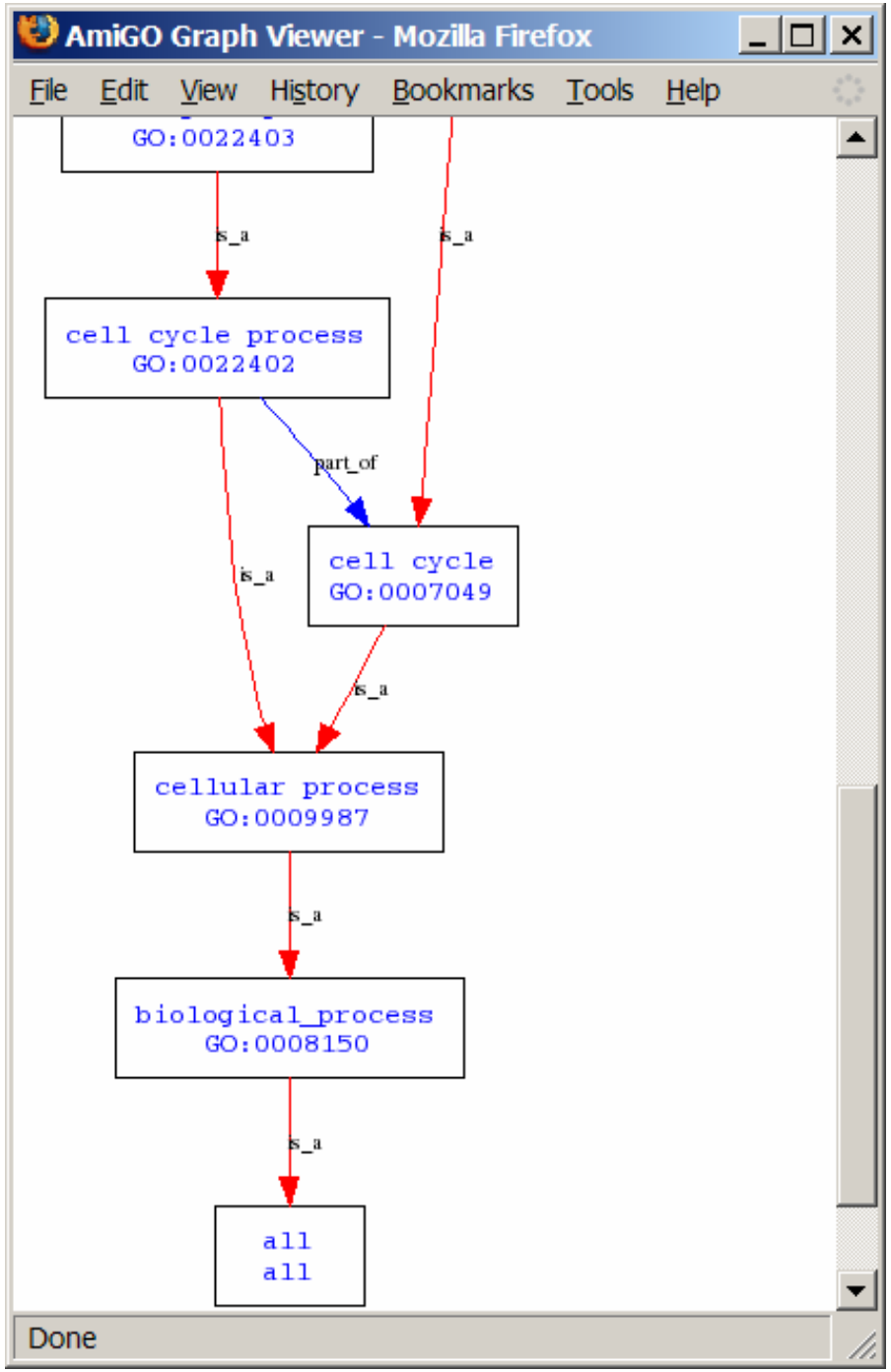
Remove all filters

- all : all [246223 gene products]
  - GO:0008150 : biological_process [164796 gene products]
    - GO:0009987 : cellular process [79232 gene products]
      - GO:0007049 : cell cycle [3620 gene products]
        - GO:0022402 : cell cycle process [2721 gene products]
          - GO:0022403 : cell cycle phase [2229 gene products]
            - GO:0000279 : M phase [1825 gene products]
              - GO:0000087 : M phase of mitotic cell cycle [705 gene products]
              - GO:0007067 : mitosis [687 gene products]
            - GO:0007067 : mitosis [687 gene products]
        - GO:0000278 : mitotic cell cycle [1865 gene products]
          - GO:0000087 : M phase of mitotic cell cycle [705 gene products]
            - GO:0007067 : mitosis [687 gene products]
      - GO:0022402 : cell cycle process [2721 gene products]
        - GO:0022403 : cell cycle phase [2229 gene products]
          - GO:0000279 : M phase [1825 gene products]
            - GO:0000087 : M phase of mitotic cell cycle [705 gene products]
              - GO:0007067 : mitosis [687 gene products]
        - GO:0007067 : mitosis [687 gene products]

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

Done


AmiGO Graph Viewer - Mozilla Firefox
File Edit View History Bookmarks Tools Help
blue
Box color
white
Go!
mitosis
GO:0007067
part_of
M phase of mitotic cell cycle
GO:0000087
is_a
is_a
part_of
M phase
GO:0000279
is_a
mitotic cell cycle
GO:0000278
cell cycle phase
GO:0022403
is_a
is_a
cell cycle process
GO:0022402
Done


AmiGO Graph Viewer - Mozilla Firefox
File Edit View History Bookmarks Tools Help
GO:0022403
is_a
is_a
cell cycle process
GO:0022402
part_of
is_a
cell cycle
GO:0007049
is_a
cellular process
GO:0009987
is_a
biological_process
GO:0008150
is_a
all
all
Done

the Gene Ontology - Mozilla Firefox

File Edit View History Bookmarks Tools Help

the Gene Ontology

Search

gene or protein name go!

Open menus
Home
FAQ
Downloads
Tools
Documentation
About GO
Projects
Contact GO
Site Map

Gene Ontology Home

The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism. *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.

GO website

- The latest news and views in the GO newsletter
- GO downloads, including ontology files, annotations and the GO database
- Tools for using GO, including OBO-Edit downloads, AmiGO, and the GO Online SQL Environment.

× Find: COMPON Next Previous Highlight all Match case

Done

AmiGO: Term Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

59 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.

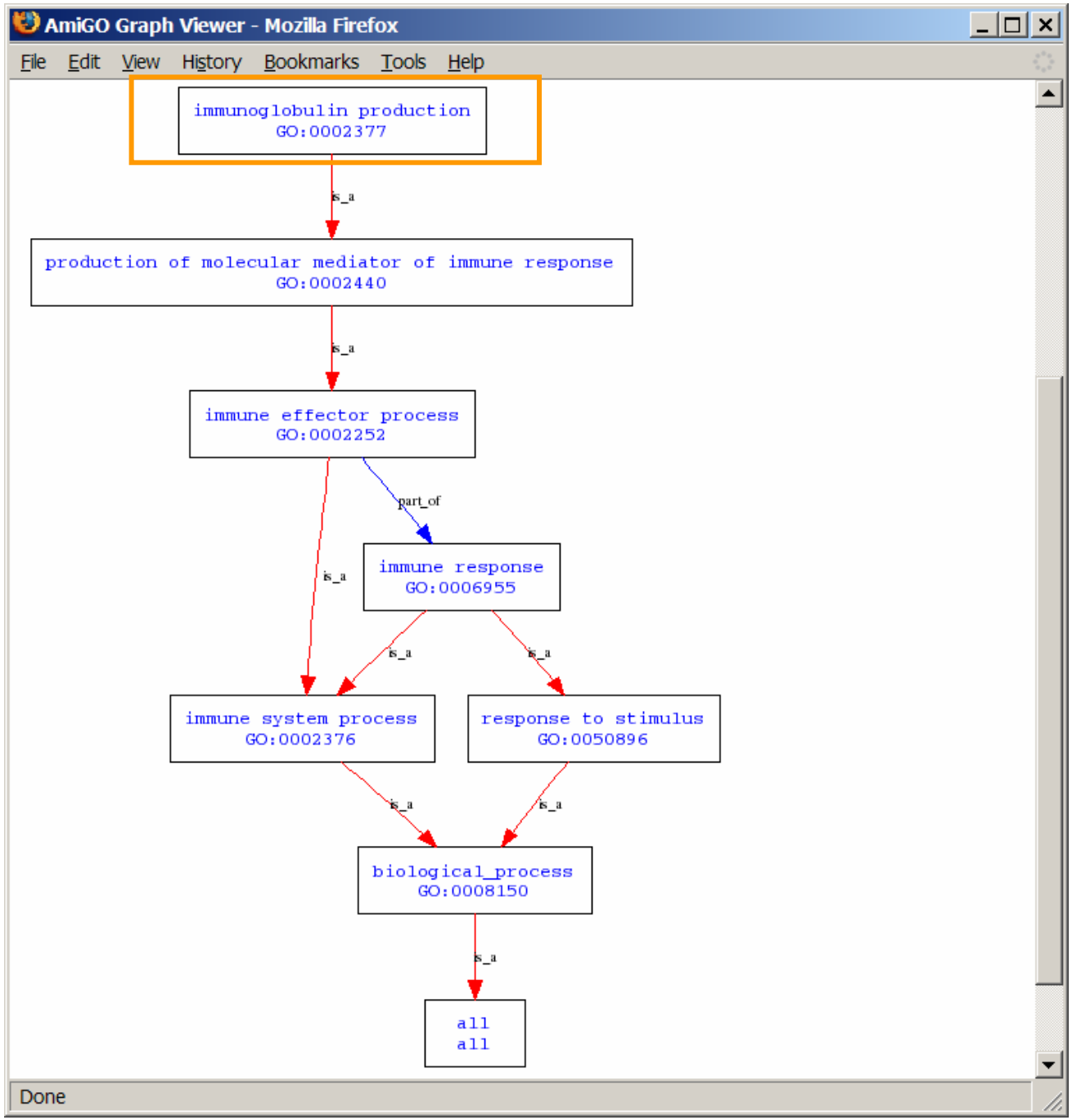
1 2 View all results

Select all | Clear all | Perform an action with the selected terms... | Go!

rel ↓	Accession , Term		Ontology
☐	GO:0019814 : **immunoglobulin** complex [show def]	18 gene products view in tree	cellular component
☐	GO:0019865 : **immunoglobulin** binding [show def]	41 gene products view in tree	molecular function
☐	GO:0048305 : **immunoglobulin** secretion [show def]	21 gene products view in tree	biological process
☐	GO:0002377 : **immunoglobulin** production [show def]	115 gene products view in tree	biological process
☐	GO:0033152 : **immunoglobulin** V(D)J recombination [show def] Query matches synonyms "**immunoglobulin** V(D)J joining" [exact synonym], and 4 more	8 gene products view in tree	biological process
☐	GO:0034987 : **immunoglobulin** receptor binding [show def]	0 gene products view in tree	molecular function

× Find: COMPON Next Previous Highlight all Match case

Done


AmiGO Graph Viewer - Mozilla Firefox
File Edit View History Bookmarks Tools Help
immunoglobulin production
GO:0002377
is_a
production of molecular mediator of immune response
GO:0002440
is_a
immune effector process
GO:0002252
part_of
is_a
immune response
GO:0006955
is_a
is_a
immune system process
GO:0002376
response to stimulus
GO:0050896
is_a
is_a
biological_process
GO:0008150
is_a
all
all
Done

AmiGO: immunoglobulin production Details - Mozilla Firefox

File Edit View History Bookmarks Tools Help

immunoglobulin production

Term information ⬇ Term lineage ⬇ External references ⬇ 115 gene product associations ➡

Term Information

Accession GO:0002377

Ontology **biological process**

Synonyms **exact**: antibody production

Definition The appearance of immunoglobulin due to biosynthesis or secretion following a cellular stimulus, resulting in an increase in its intracellular or extracellular levels. [source: GOC:add, ISBN:0781735149]

Comment None

Subset None

Back to top

Term Lineage

Switch to viewing term parents, siblings and children

Filter tree view

Filter Gene Product Counts

Data source: All, CGD, dictyBase, EcoCyc

Species: All, Anaplasma phagocy..., Arabidopsis thaliana, Bacillus anthraci...

View Options

Tree view: Full / Compact

Set filters

Remove all filters

- all : all [246223 gene products]
  - GO:0008150 : biological_process [164796 gene products]
    - GO:0002376 : immune system process [2714 gene products]
      - GO:0002252 : immune effector process [588 gene products]
        - GO:0002440 : production of molecular mediator of immune response [181 gene products]
          - **GO:0002377 : immunoglobulin production [115 gene products]**
      - GO:0006955 : immune response [1682 gene products]
        - GO:0002252 : immune effector process [588 gene products]

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

Done

AmiGO: immunoglobulin production Details - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Filter tree view

Filter Gene Product Counts

Data source: All, CGD, dictyBase, EcoCyc

Species: All, Anaplasma phagocy..., Arabidopsis thaliana, Bacillus anthraci...

View Options

Tree view: Full / Compact

Set filters

Remove all filters

- GO:0002440 : production of molecular mediator of immune response [181 gene products]
  - GO:0002775 : antimicrobial peptide production [45 gene products]
  - GO:0002367 : cytokine production during immune response [40 gene products]
  - **GO:0002377 : immunoglobulin production [115 gene products]**
    - GO:0002378 : immunoglobulin biosynthetic process [4 gene products]
    - **GO:0002381 : immunoglobulin production during immune response [78 gene products]**
    - GO:0048305 : immunoglobulin secretion [21 gene products]
    - GO:0002638 : negative regulation of immunoglobulin production [0 gene products]
    - GO:0002639 : positive regulation of immunoglobulin production [6 gene products]
    - GO:0002637 : regulation of immunoglobulin production [8 gene products]
    - GO:0016445 : somatic diversification of immunoglobulins [84 gene products]
      - GO:0002564 : alternate splicing of immunoglobulin genes [0 gene products]
      - GO:0002206 : gene conversion of immunoglobulin genes [0 gene products]
      - GO:0002570 : somatic diversification of immunoglobulin genes by N region addition [0 gene products]
      - GO:0002208 : somatic diversification of immunoglobulins during immune response [69 gene products]
      - GO:0016446 : somatic hypermutation of immunoglobulin genes [15 gene products]
        - GO:0002205 : somatic hypermutation of immunoglobulin genes during immune response [0 gene products]
      - GO:0016447 : somatic recombination of immunoglobulin gene segments [79 gene products]
        - GO:0002452 : B cell receptor editing [0 gene products]
        - GO:0033152 : immunoglobulin V(D)J recombination [8 gene products]
        - GO:0002570 : somatic diversification of immunoglobulin genes by N region addition [0 gene products]
        - GO:0002204 : somatic recombination of immunoglobulin genes during immune response [69 gene products]
          - GO:0045190 : isotype switching [69 gene products]
  - GO:0002701 : negative regulation of production of molecular mediator of immune response [11 gene products]
  - GO:0002702 : positive regulation of production of molecular mediator of immune response [20 gene products]
  - GO:0002700 : regulation of production of molecular mediator of immune response [32 gene products]

Actions...
Last action: Closed
GO:0002381
Graphical View
Reset tree
View in tree browser

Done

the Gene Ontology - Mozilla Firefox

File Edit View History Bookmarks Tools Help

the Gene Ontology

Search

gene or protein name go!

- Open menus
- Home
- FAQ
- **Downloads**
- **Tools**
- **Documentation**
- **About GO**
- **Projects**
- Contact GO
- Site Map

Gene Ontology Home

The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism. *Read more about the Gene Ontology...*

Search the Gene Ontology Database

Search for genes, proteins or GO terms using AmiGO:

RAG GO!

gene or protein name GO term or ID

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

GO website

- The latest news and views in the GO newsletter
- GO downloads, including ontology files, annotations and the GO database
- Tools for using GO, including OBO-Edit downloads, AmiGO, and the GO Online SQL Environment.
- Request new terms or ontology changes or get help with new term submission
- Documentation on all aspects of the GO project and the GO FAQ
- Projects within the GO consortium, including Reference Genomes and immune system annotation
- Gene Ontology mailing lists and contact details

Done

AmiGO: Gene Product Search Results - Mozilla Firefox

File Edit View History Bookmarks Tools Help

	branched-chain amino acid transport protein BrAG		aeruginosa PAO1
☐	CRAG Calmodulin-binding protein related to a Rab3 GDP/GTP exchange protein	2 associations BLAST	**gene** from *Drosophila melanogaster*
☐	frag fragmented mitochondria	1 association	**gene** from *Drosophila melanogaster*
☐	Rag1 recombination activating gene 1	5 associations BLAST	**gene** from *Mus musculus*
☐	Rag1 recombination activating gene 1	8 associations BLAST	**gene** from *Rattus norvegicus*
☐	RAG1 Seed allergenic protein RAG1 precursor	3 associations	**protein** from *Oryza sativa*
☐	RAG1 V(D)J recombination-activating protein 1	3 associations BLAST	**protein** from *Homo sapiens*
☐	Rag2 recombination activating gene 2	9 associations BLAST	**gene** from *Mus musculus*
☐	Rag2 recombination activating gene 2	11 associations BLAST	**gene** from *Rattus norvegicus*
☐	RAG2 Seed allergenic protein RAG2 precursor	3 associations	**protein** from *Oryza sativa*
☐	RAG2 V(D)J recombination-activating protein 2	4 associations BLAST	**protein** from *Homo sapiens*
☐	ragA	4 associations	**gene** from *Dictyostelium*

Done

AmiGO: RAG2 Associations - Mozilla Firefox

File Edit View History Bookmarks Tools Help

gene association format RDF-XML

Filter associations displayed

Filter Associations

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

Evidence Code: All, IC, IDA, EXP

Set filters

Remove all filters

Select all | Clear all | Perform an action with the selected terms... | Go!

Accession, Term		Ontology	Qualifier	Evidence	Reference	Assigned by
GO:0016445 : somatic diversification of immunoglobulins	84 gene products view in tree	biological process		NAS	UniProtKB:P55895	UniProtKB
GO:0005634 : nucleus	16803 gene products view in tree	cellular component		NAS	UniProtKB:P55895	UniProtKB
GO:0003677 : DNA binding	9895 gene products view in tree	molecular function		NAS	UniProtKB:P55895	UniProtKB
GO:0004519 : endonuclease activity	655 gene products view in tree	molecular function		NAS	UniProtKB:P55895	UniProtKB

Select all | Clear all | Perform an action with the selected terms... | Go!

Back to top

Done

Dans le domaine de l'immunogénétique

Problèmes et limites de GO:

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

- Quels sont les motifs constitutifs importants?

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

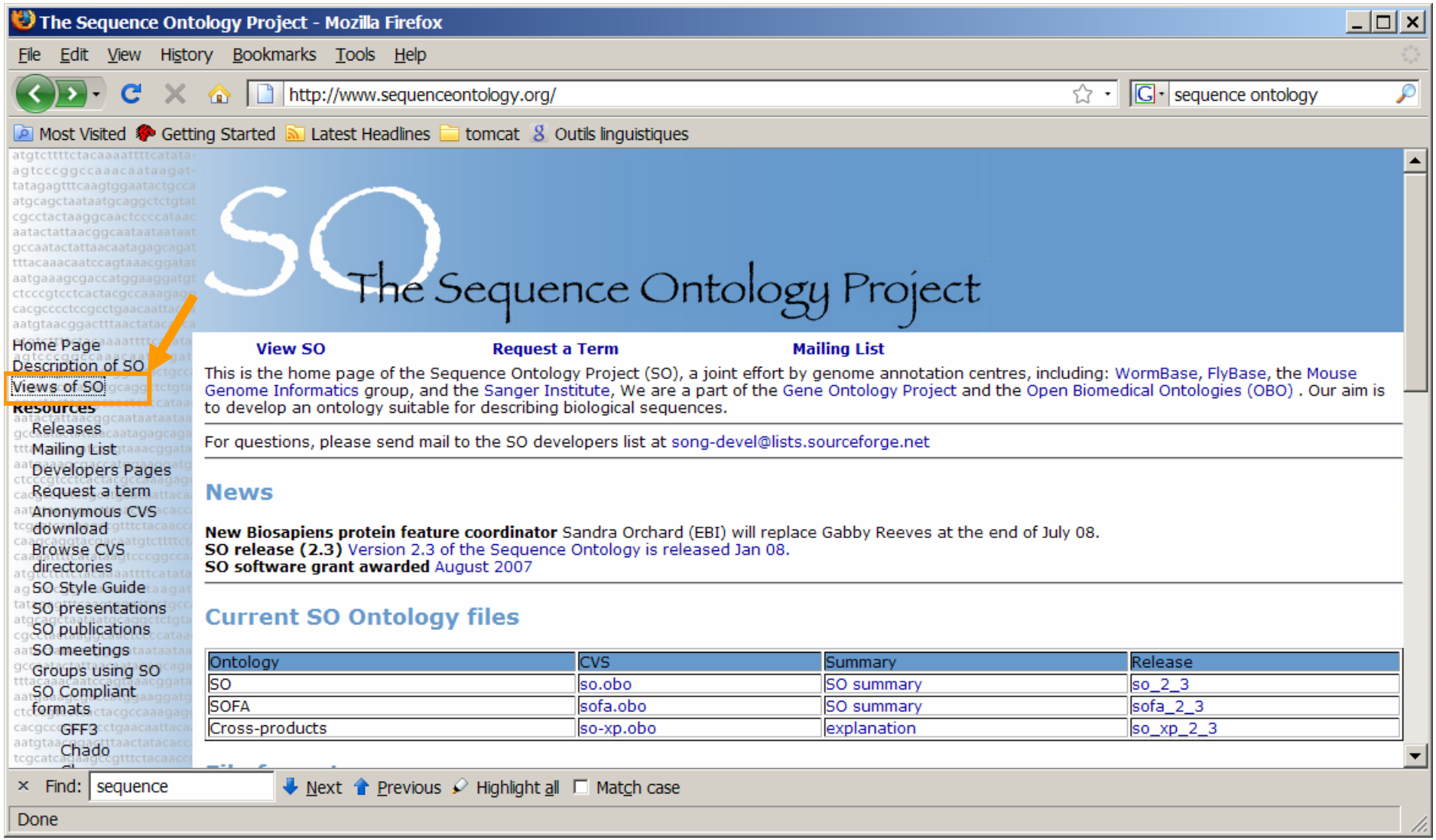
Sequence Ontology

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

- proposer une représentation structurée des annotations

- vocabulaire pour la description des mutations

Sequence Ontology:
une ontologie pour décrire les séquences biologiques



The Sequence Ontology Project - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.sequenceontology.org/view_SO.shtml

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

SO The Sequence Ontology Project

View SO Request a Term Mailing List

Home Page
Description of SO
Views of SO
Resources
Releases
Mailing List
Developers Pages
Request a term
Anonymous CVS download
Browse CVS directories
SO Style Guide
SO presentations
SO publications
SO meetings
Groups using SO
SO Compliant formats
GFF3
Chado

Viewing the Sequence Ontology

miSO

Browse the terms in the Sequence Ontology.

OBO-Edit

Download OBO-Edit to view and edit OBO ontologies.

SVG Graphs

An SVG graph of SO provided by Mary Dolan, University of Maine. The SVG images can be panned and zoomed - see support documentation

SO Term Finder

A term finder for SO at the Broad Institute.

Find: sequence Next Previous Highlight all Match case

Done

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

miSO sequence ontology term browser

| SO home | Releases | CVS | Mailing List | Trackers |

Browse SO

miSO provides a graphical means to browse the terms in the Sequence Ontology. The view is organised via a parent/child approach. The terms are all linked to the SO term tables web page which displays more details such as the relationship types, the definition, the synonyms and the cross references. We have provided the option to search either the latest upload of SO SO from the CVS repository (the cutting edge version), or the latest release of SO (the more stable version).

- The latest CVS update of SO
  - sequence_collection
  - sequence_variant
  - sequence_replication_mode
  - chromosome_variation
  - sequence_feature
  - sequence_attribute
  - sequence_operation
  - mutation
  - relationships
- The most recent SO release
  - Sequence_Ontology
    - chromosome_variation
    - genome
    - mutation
    - operation
    - replication_mode
    - sequence_attribute
    - sequence_feature
    - sequence_variant
  - relationships

created from: ontology/archives/SO_current.obo
format-version: 1.2
date: 30:01:2008 17:03
saved-by: kareneilbeck
default-namespace: sequence
remark: autogenerated-by\: DAG-Edit version 1.417\nsaved-by\: eilbeck\ndate\: Tue May 11 15\:18\:44 PDT 2004\nversion\: $Revision\: 1.45 $

File created on Mon Sep 8 12:12:39 2008

Done

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

miSO sequence ontology term browser

| SO home | Releases | CVS | Mailing List | Trackers |

- sequence_alteration
- sequence_variant
- relationships
  - position_of
  - sequence_of
  - derives_from
  - associated_with
  - member_of
  - variant_of
  - has_quality
  - genome_of
  - adjacent_to
  - part_of
  - similar_to
  - has_part
  - has_origin

part_of

part_of	part_of
Definition: undefined	
Aspect: root	
DBxref:	
Synonyms:	
Parent relationships:	**Child relationships:**

File created on Mon Sep 8 12:12:42 2008

Done

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

miSO

sequence ontology term browser

| SO home | Releases | CVS | Mailing List | Trackers |

- region
  - CpG_island
  - QTL
  - assembly
  - assembly_component
  - binding_site
  - cap
  - chromosome
  - chromosome_part
  - consensus_region
  - conserved_region
  - deletion
  - engineered_region
  - flanking_region
  - gene
  - gene_component_region
  - **gene_group**
    - gene
      - allele
      - engineered_gene
      - epigenetically_modified_gene
      - foreign_gene
      - fusion_gene
      - gene_cassette
      - gene_component_region

gene_group

gene_group	**SO:0005855**
Definition: A collection of related genes.	
Aspect: sequence_feature	
DBxref:SO:ma	
Synonyms:	
Parent relationships: gene_group **is_a** region	**Child relationships:** gene_subarray **is_a** gene_group operon **is_a** gene_group gene_cassette_array **is_a** gene_group gene_group_regulatory_region **member_of** gene_group regulon **is_a** gene_group gene_array **is_a** gene_group gene **member_of** gene_group

File created on Mon Sep 8 12:12:52 2008

Done

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

miSO sequence ontology term browser

| SO home | Releases | CVS | Mailing List | Trackers |

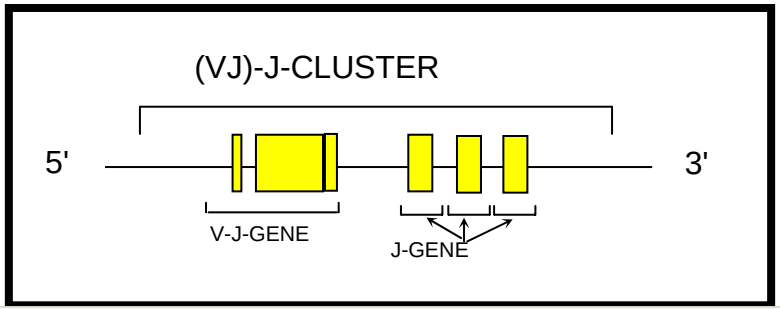
- sequence_feature
  - junction
  - region
    - CpG_island
    - QTL
    - assembly
    - assembly_component
    - binding_site
    - cap
    - chromosome
    - chromosome_part
    - consensus_region
    - conserved_region
    - deletion
    - engineered_region
    - flanking_region
    - gene
    - gene_component_region
    - gene_group
      - gene
        - allele
        - engineered_gene
        - epigenetically_modified_gene
        - foreign_gene
        - fusion_gene
        - gene_cassette
        - gene_component_region
        - gene_member_region

allele

allele	**SO:0001023**
Definition: An allele is one of a set of coexisting sequence variants of a gene.	
Aspect: sequence_feature	
DBxref:SO:immuno_workshop	
Synonyms:	
Parent relationships: allele **is_a** sequence_variant allele **variant_of** gene	**Child relationships:**

File created on Mon Sep 8 12:13:11 2008

Done



Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

miSO sequence ontology term browser

| SO home | Releases | CVS | Mailing List | Trackers |

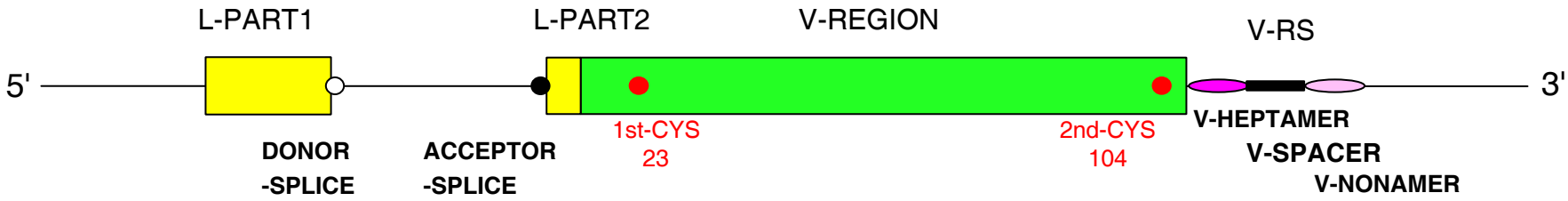
- V_J_C_cluster
- V_J_cluster
- V_cluster
- vertebrate_immunoglobulin_T_cell_receptor_rearra
  - DJ_C_cluster
  - DJ_J_C_cluster
  - DJ_J_cluster
  - D_DJ_C_cluster
  - D_DJ_J_C_cluster
  - D_DJ_J_cluster
  - D_DJ_cluster
  - VDJ_C_cluster
  - VDJ_J_C_cluster
  - VDJ_J_cluster
  - VJ_C_cluster
  - VJ_J_C_cluster
  - **VJ_J_cluster**
  - V_DJ_C_cluster
  - V_DJ_J_C_cluster
  - V_DJ_J_cluster
  - V_DJ_cluster

VJ_J_cluster

VJ_J_cluster	**SO:0000491**
Definition: Genomic DNA of immunoglobulin/T-cell receptor gene in rearranged configuration including at least one VJ-gene and one J-gene.	
Aspect: sequence_feature	
DBxref:URL:http://imgt.cines.fr/ligmdb/LIGMlect?query=7	**IMGT**
Synonyms: [(VJ)-J-CLUSTER]	
Parent relationships: VJ_J_cluster **is_a** vertebrate_immunoglobulin_T_cell_receptor_rearranged_gene_cluster	**Child relationships:** VJ_gene **part_of** VJ_J_cluster; J_gene **part_of** VJ_J_cluster

Done

V-GENE



Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited | Getting Started | Latest Headlines | tomcat | Outils linguistiques

miSO sequence ontology term browser

| SO home | Releases | CVS | Mailing List | Trackers |

- macronucleus_destined_segment
- specific_recombination_site
  - recombination_feature_of_rearranged_gene
    - vertebrate_immune_system_gene_recombinatior
      - recombinationally_rearranged_vertebrate_immu
      - vertebrate_immune_system_gene_recombina
        - D_gene_recombination_feature
        - J_gene_recombination_feature
        - **V_gene_recombination_feature**
        - heptamer_of_recombination_feature_of_ve
        - nonamer_of_recombination_feature_of_ve
      - vertebrate_immune_system_gene_recombina
      - vertebrate_immunoglobulin_T_cell_receptor_
      - vertebrate_immunoglobulin_T_cell_receptor_
      - vertebrate_immunoglobulin_T_cell_receptor_
      - vertebrate_immunoglobulin_T_cell_receptor_
  - site_specific_recombination_target_region
- remark
- repeat_component
- repeat_region
- etron

V_gene_recombination_feature

V_gene_recombination_feature	**SO:0000538**

Definition: Recombination signal including V-heptamer, V-spacer and V-nonamer in 3' of V-region of a V-gene or V-sequence of an immunoglobulin/T-cell receptor gene.

Aspect: sequence_feature

DBxref:URL:http://imgt.cines.fr/ligmdb/LIGMlect?query=7 **IMGT**

Synonyms: [V-RS]

Parent relationships:
V_gene_recombination_feature **is_a** vertebrate_immune_system_gene_recombination_signal_feature

Child relationships:
V_spacer **part_of** V_gene_recombination_f
V_nonamer **part_of** V_gene_recombination_f
V_heptamer **part_of** V_gene_recombination_f

Done

Dans le domaine de l'immunogénétique

Limites de SO:

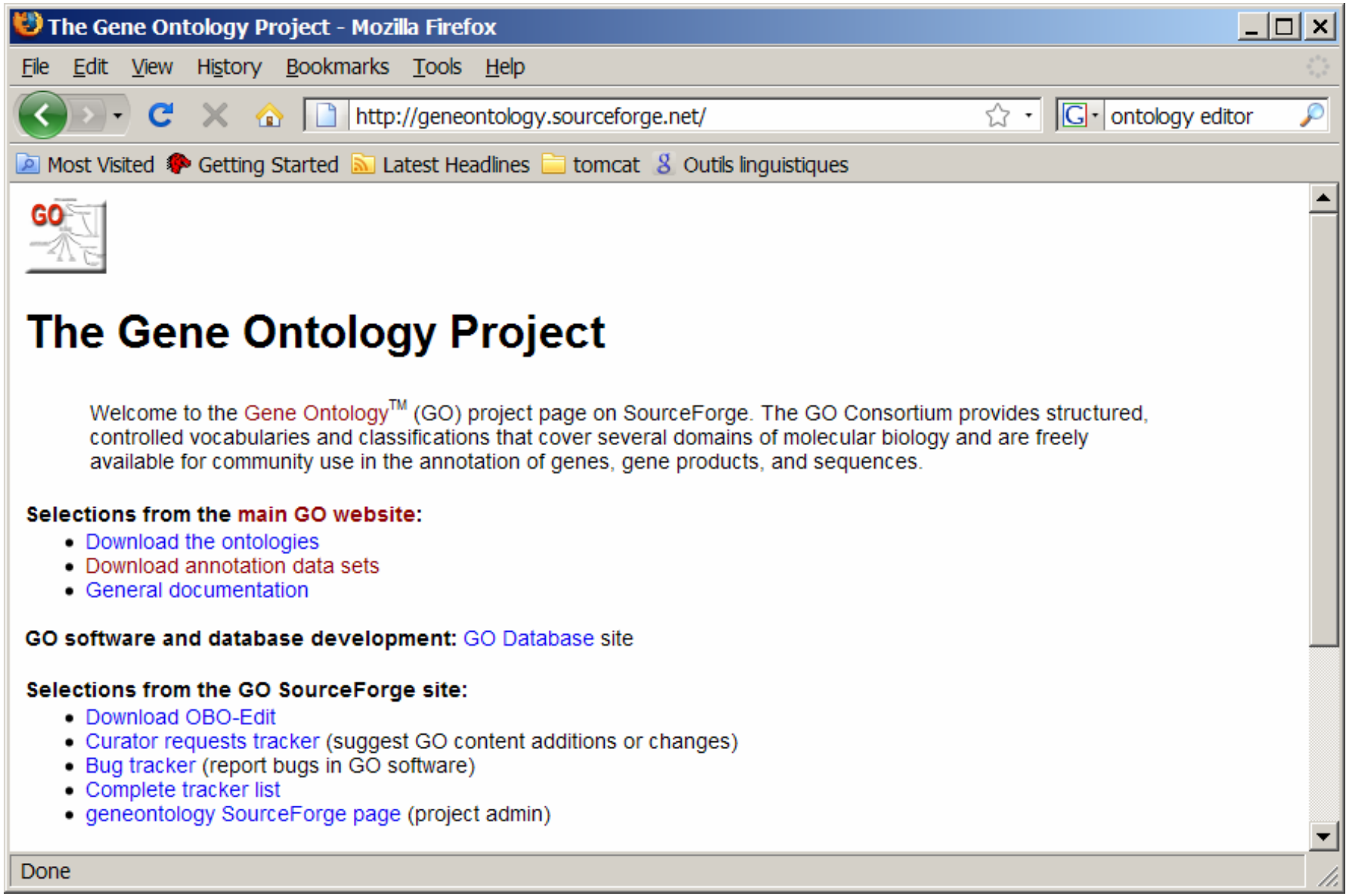
- Il est nécessaire de prendre en compte d'autres informations:
  type de gènes, type de chaînes, …

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

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

=> nécessité d'une ontologie spécifique

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



Exemple d'éditeur d'ontologie: Protégé

The Protégé Ontology Editor and Knowledge Acquisition System - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://protege.stanford.edu/ protege editor

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

protégé

Search:

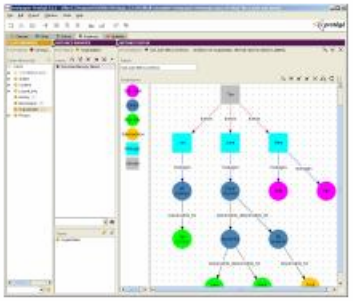
HOME | OVERVIEW | DOCUMENTATION | DOWNLOADS | SUPPORT | COMMUNITY | WIKI | ABOUT US

welcome to protégé

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)



go to protégé-frames

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.

news

22nd Aug. 2008
Protégé 4.0 goes beta!
(more details)
(download)

community	
Registered Users	104,744
protege-users list members	17,246
protege-discussion list members	3,572

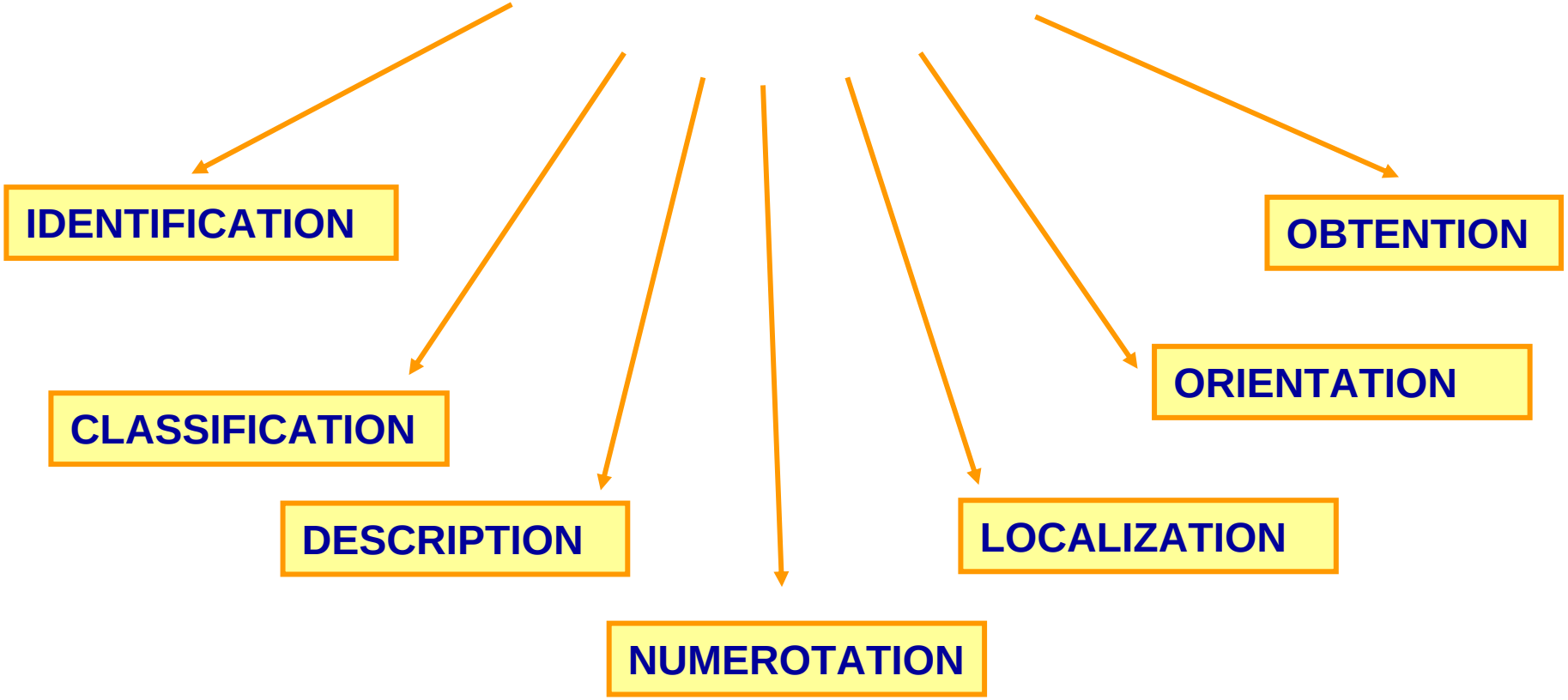
Done

IMGT-ONTOLOGY

http://imgt.cines.fr

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

IMGT-ONTOLOGY repose sur 7 axiomes:



IMGT-ONTOLOGY

http://imgt.cines.fr

IMGT Scientific chart :

IMGT-ONTOLOGY en *langage naturel* pour les biologistes

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

En cours : édition avec Protégé

pour faciliter la représentation et le partage

(agents humains et logiciels)

"IDENTIFICATION"

http://imgt.cines.fr

Species

human
mouse..

MoleculeType

genomic DNA
cDNA
protein..

Gene Type
variable
diversity
joining
constant
conventional

Configuration Type
germline
eearranged
undefined

Chain type

Ig-Heavy
Ig-Light-Lambda
TcR-Alpha
TcR-Beta...

Structure type

regular
translocated ...

Functionality

functional
ORF
pseudogene
productive

Specificity

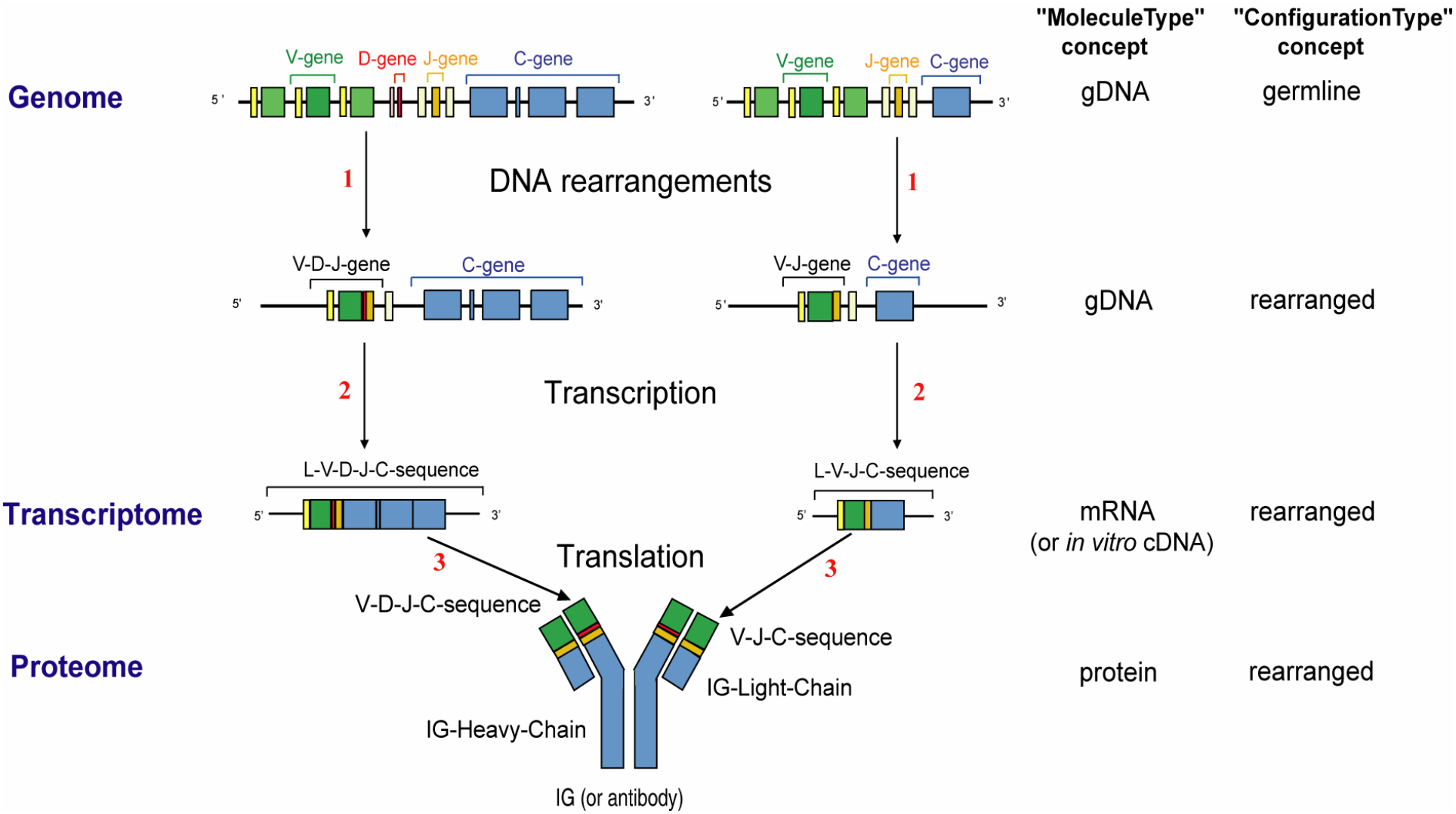
Anti-DNA
Anti-HIV ...

Molecular ReceptorType
IgA
TcR gamma-delta

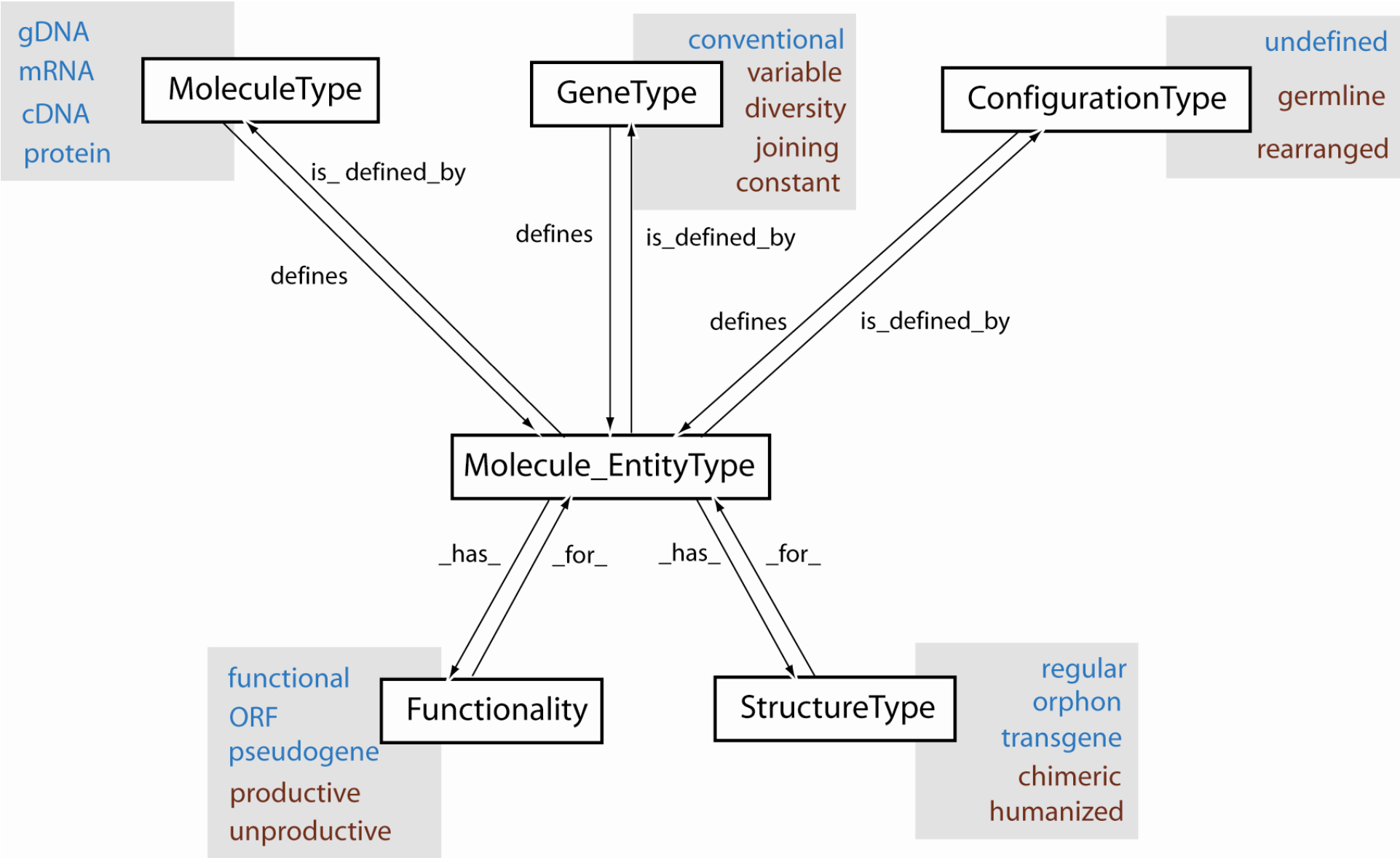
+ Molecule_EntityType

Synthèse des immunoglobulines

http://imgt.cines.fr



http://imgt.cines.fr



IMGT/LIGM-DB - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

IMGT FLAT-FILE

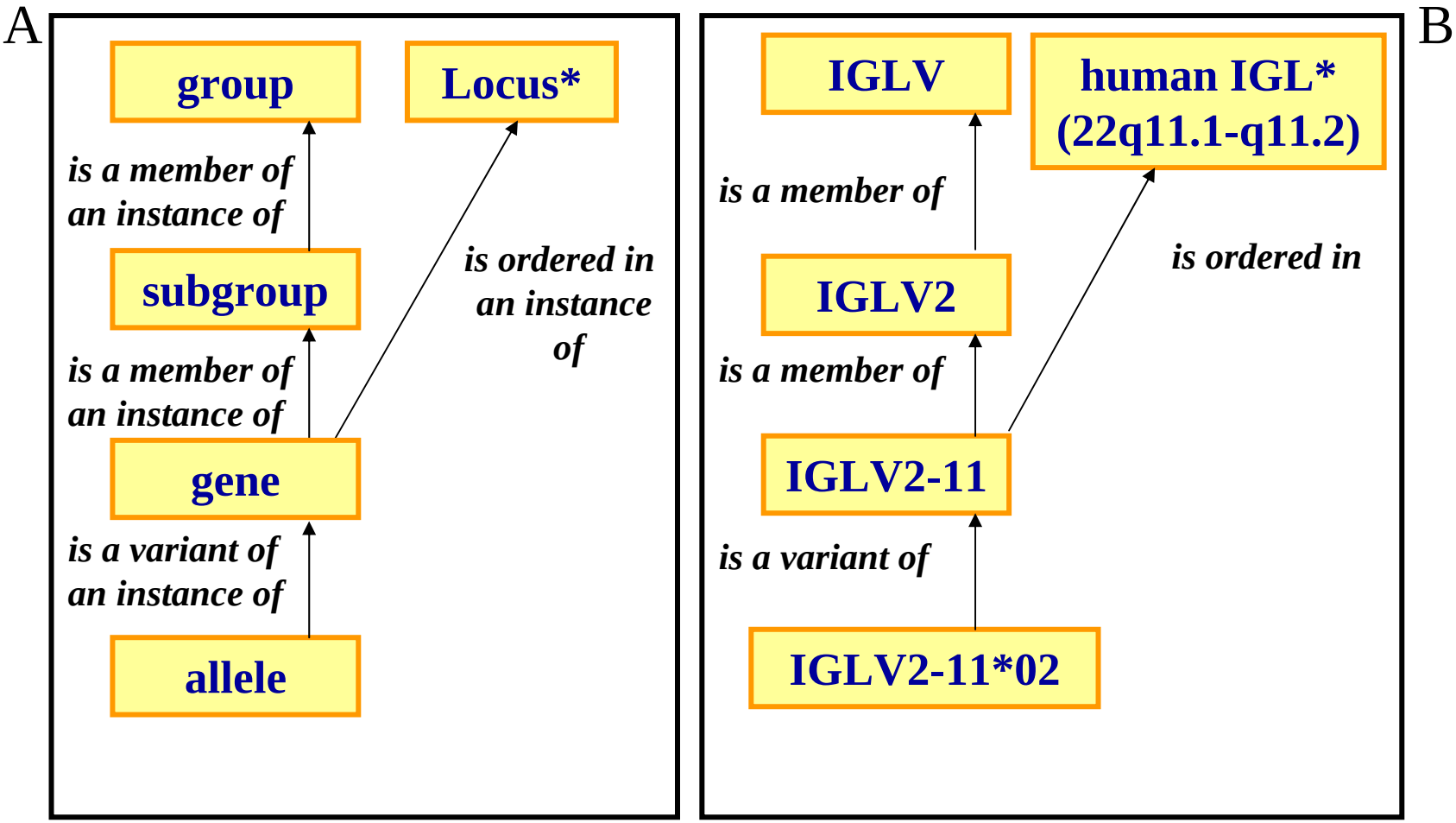
```
ID   L39956 IMGT/LIGM annotation : automatic; RNA; HUM; 375 BP.
XX
AC   L39956;
XX
DT   28-SEP-2001 (Rel. 200139-5, arrived in LIGM-DB )
DT   13-NOV-2001 (Rel. 200146-2, Last updated, Version 2)
XX
DE   Homo sapiens monospecific anti-ssDNA antibody heavy chain variable region
DE   mRNA, complementarity determining regions  1-3 and framework regions 1-4.
DE   ;
DE   RNA; rearranged configuration; Ig-Heavy; regular; functionality
DE   productive; group IGHV; subgroup HV3; specificity anti-DNA single-stranded
DE   (ss) [human].
XX
KW   antigen receptor; immunoglobulin superfamily; IG; IG-Heavy; variable;
KW   diversity; joining; hybridoma; immunoglobulin; cDNA; undefined;
KW   rearranged; L-V-D-J-C-sequence; partner.
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   1-375
RX   MEDLINE; 96195158.
RA   Mitamura K., Suenaga R., Wilson K.B., Abdou N.I.;
RT   "V gene sequences of human anti-ssDNA antibodies secreted by lupus-derived
RT   CD5-negative B cell hybridomas";
RL   Clin. Immunol. Immunopathol. 78(2):152-160(1996).
XX
DR   IMGT/LIGM:partner; L39957;
DR   EMBL; L39956.
```

Done One active download (1 minute remaining)

http://imgt.cines.fr

"CLASSIFICATION"

http://imgt.cines.fr



* Axiom localization

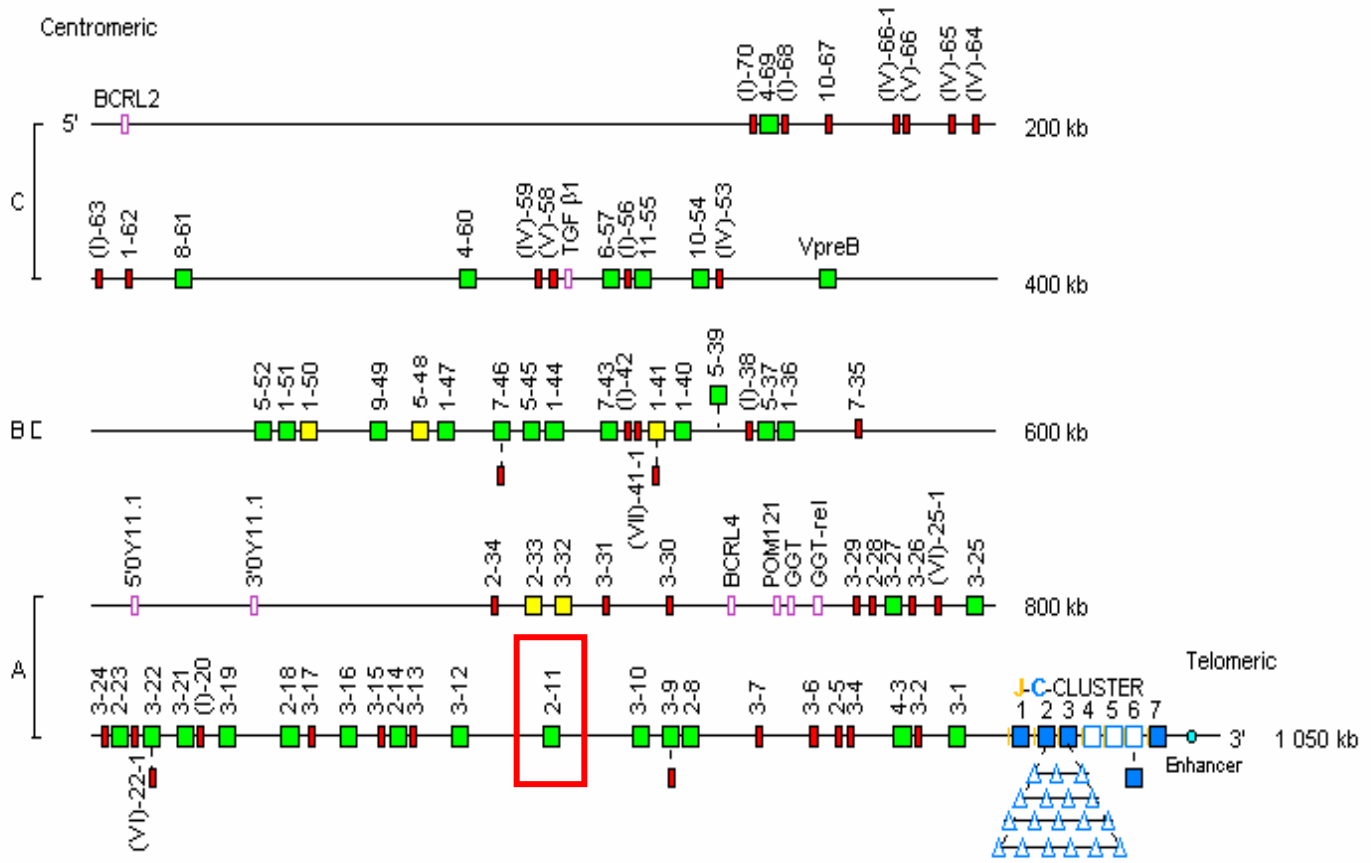
http://imgt.cines.fr

IMGT Marie-Paule page: IMGT repertoire - Netscape

Fichier Edition Afficher Aller Communicator Aide

Locus representation: Human IGL

Human IGL 22q11.2



Document : chargé

Entrez Gene: IGLV2-11 immunoglobulin lambda variable 2-11 [Homo sapiens] - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Most Visited Getting Started Latest Headlines tomcat Outils linguistiques

NCBI Entrez Gene

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search Gene for IGLV2-11 Go Clear Save Search

Limits Preview/Index History Clipboard Details

Display Full Report Show 20 Sort by Relevance Send to

All: 1 Current Only: 1 Genes Genomes: 1 SNP GeneView: 0

1: IGLV2-11 immunoglobulin lambda variable 2-11 [*Homo sapiens*]

GeneID: 28816 updated 28-Aug-2008

Summary

Official Symbol	IGLV2-11 provided by HGNC
Official Full Name	immunoglobulin lambda variable 2-11 provided by HGNC
Primary source	HGNC:5887
See related	IMGT/GENE-DB:IGLV2-11
Gene type	other
RefSeq status	REVIEWED
Organism	*Homo sapiens*
Lineage	*Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo*
Also known as	V1-3; IGLV211

Entrez Gene Home

Table Of Contents

- Summary
- Genomic regions, transcripts...
- Genomic context
- Bibliography
- General gene information
- Reference Sequences
- Related Sequences
- Additional Links

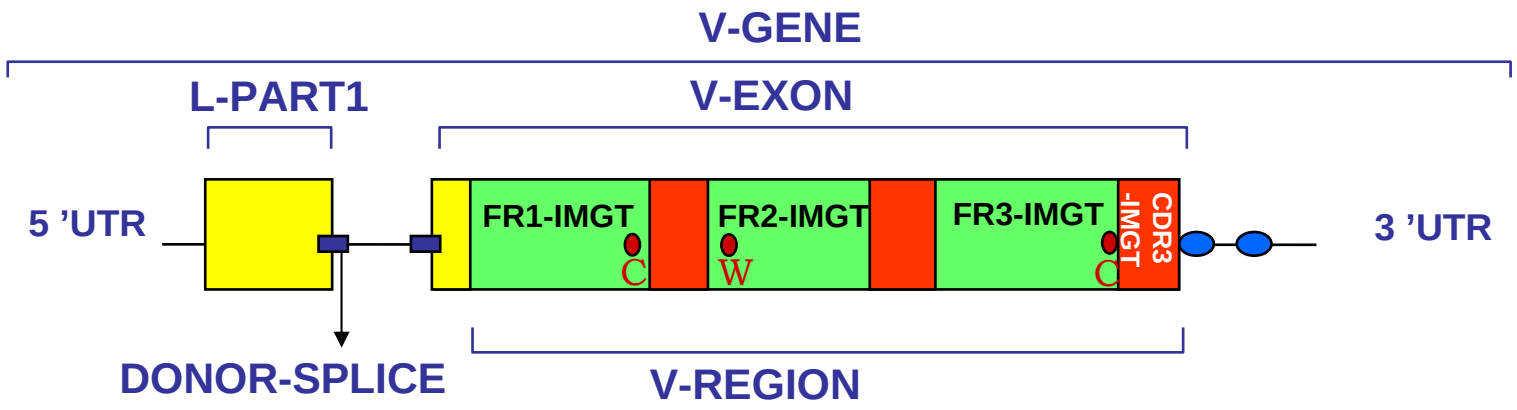
Links Explain

- Conserved Domains
- Genome
- GEO Profiles
- Map Viewer
- Nucleotide
- Protein
- PubMed
- SNP: Genotype

Done

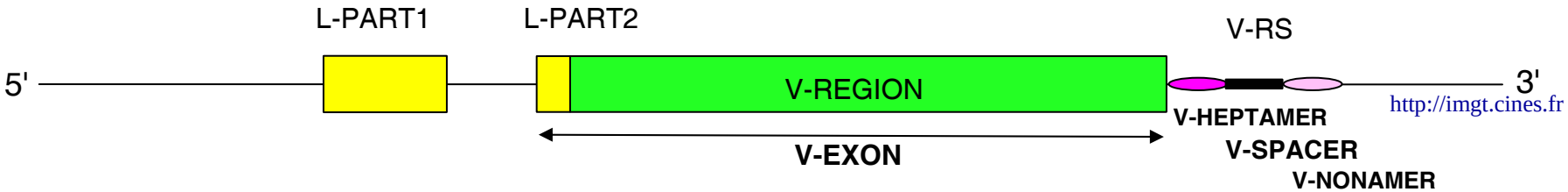
"DESCRIPTION"

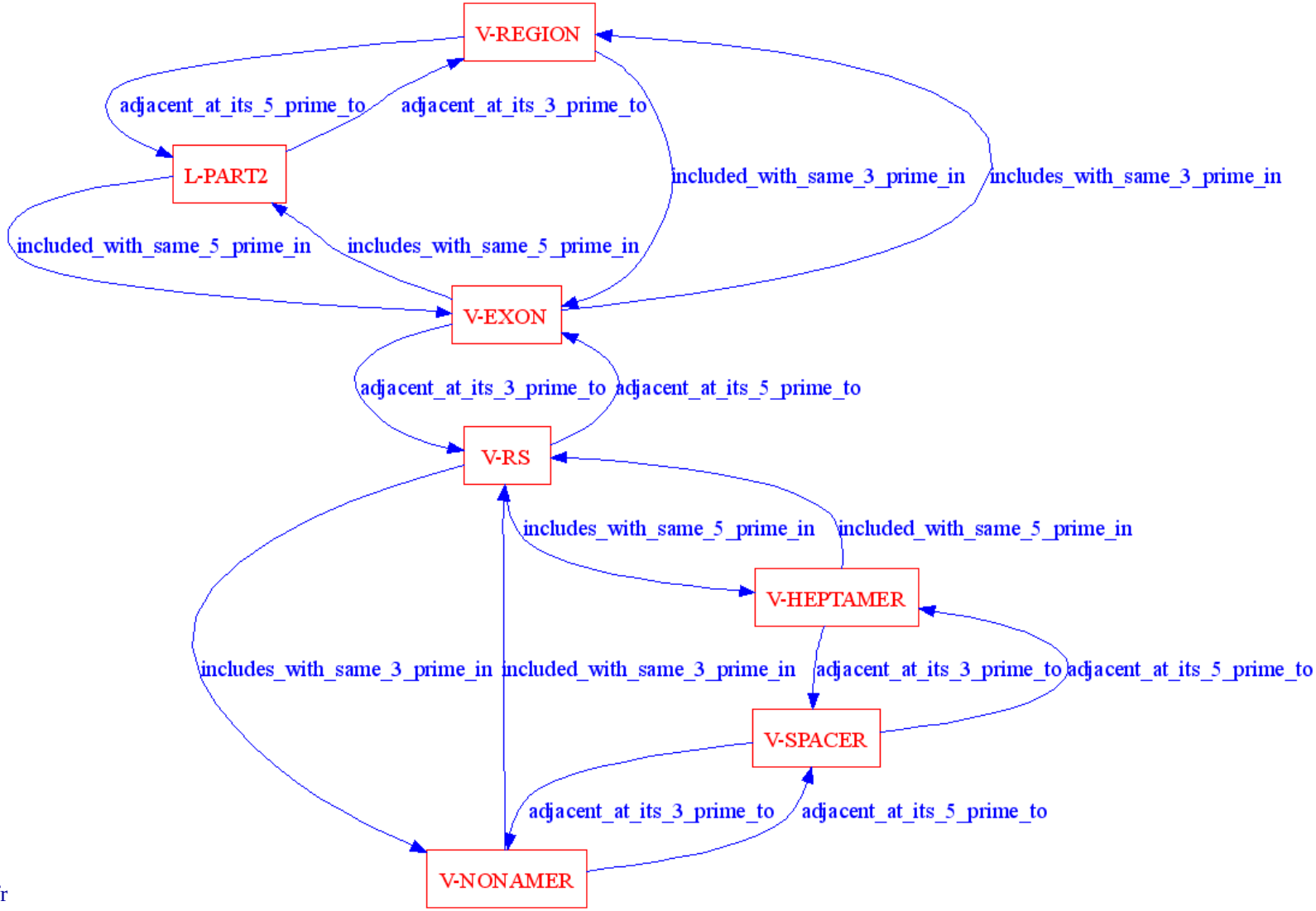
http://imgt.cines.fr



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

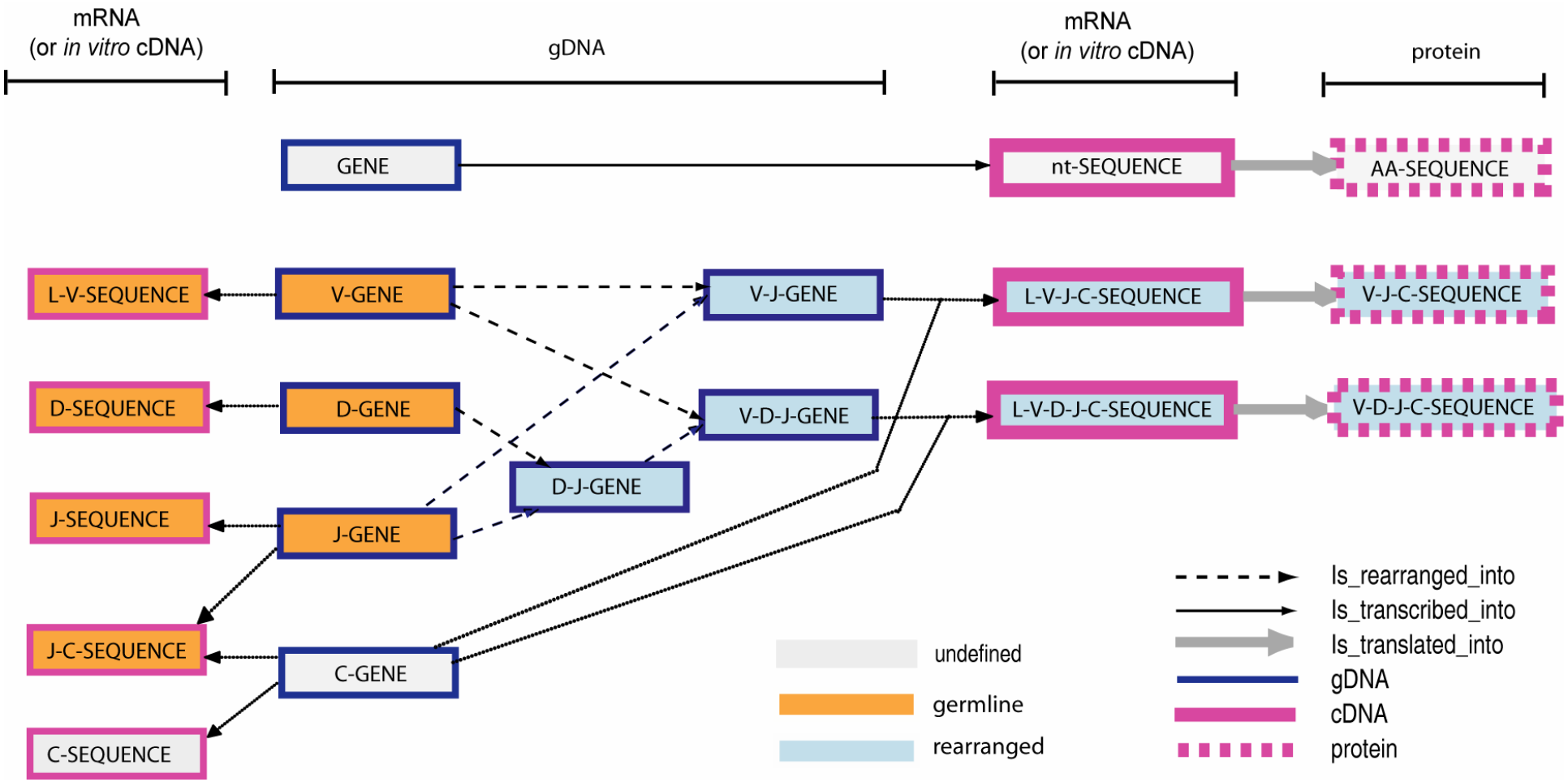




http://imgt.cines.fr

Molecule_EntityPrototype

http://imgt.cines.fr



IMGT/LIGM-DB Consultation module v3 - Netscape

Fichier Edition Afficher Aller Communicator Aide

Signets Adresse : http://ligm.igh.cnrs.fr:8104/cgi-bin/IMGTlect.jv Infos connexes

```
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 tgatattgat gattataact atgtctcctg gtaccaacag       120
```

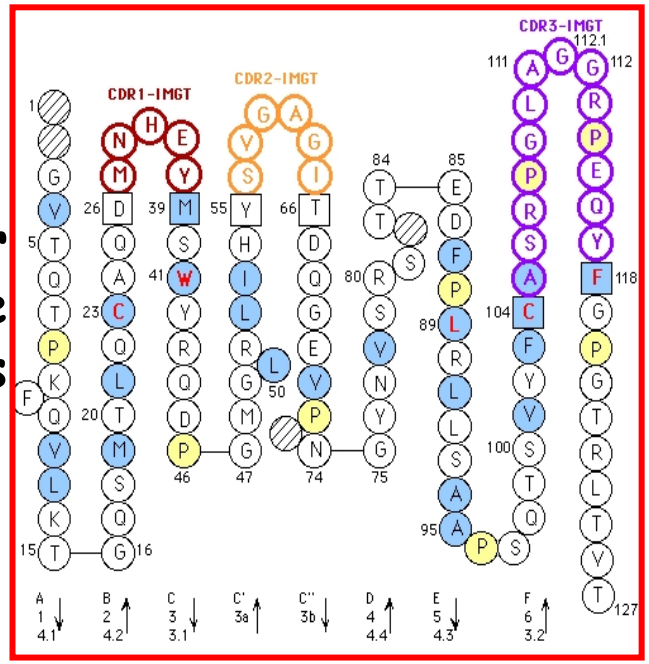
Document : chargé

http://imgt.cines.fr

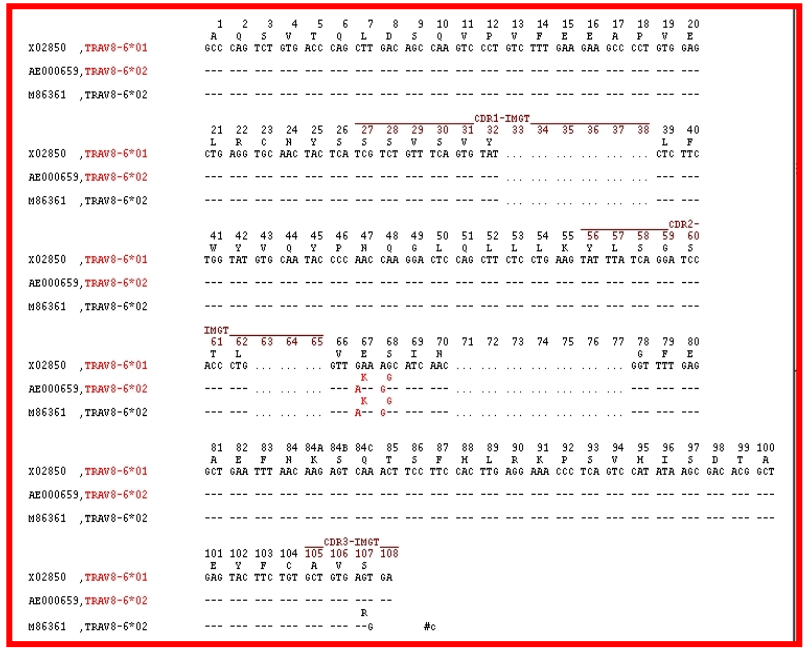
"NUMEROTATION"

http://imgt.cines.fr

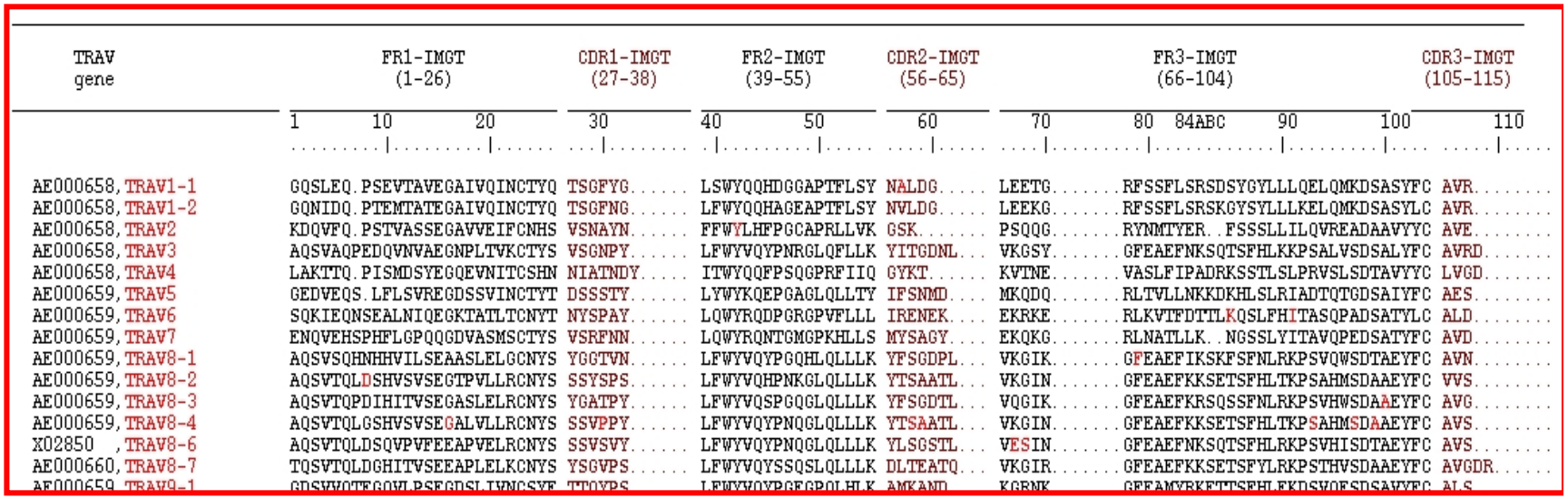
Collier De Perles



Alignment of alleles



Protein Display



"OBTENTION"

http://imgt.cines.fr

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

origine

methodologie

The "OBTENTION" concepts specify the origin and methodology

"ORIENTATION"

http://imgt.cines.fr

IMGT Repertoire (RPI) - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

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

Rechercher : cluster Occurrence suivante Occurrence précédente Surligner tout Respecter la casse

Annotation and IMGT-ONTOLOGY concepts

http://imgt.cines.fr

IDENTIFICATION

IMGT/LIGM-DB Consultation module v3 - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

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

Annotation and IMGT-ONTOLOGY concepts

http://imgt.cines.fr

IMGT/LIGM-DB Consultation module v3 - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

```
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

http://imgt.cines.fr

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://imgt.cines.fr

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://imgt.cines.fr

3 - Automatisation des procédures:

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

Les atouts de IMGT-ONTOLOGY

http://imgt.cines.fr

4 – Intégration des nouvelles connaissances en immunogénétique

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

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

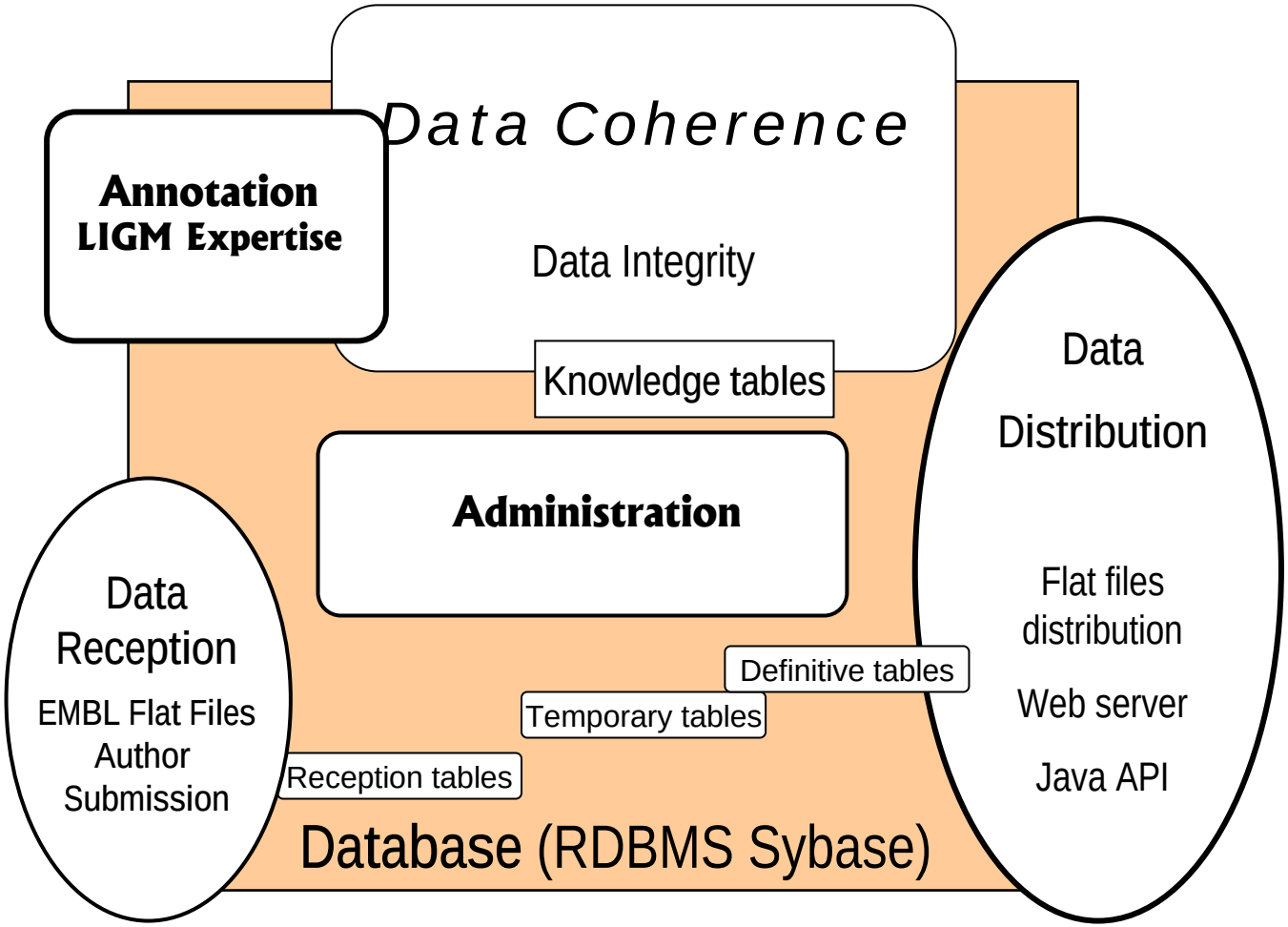
http://imgt.cines.fr

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?

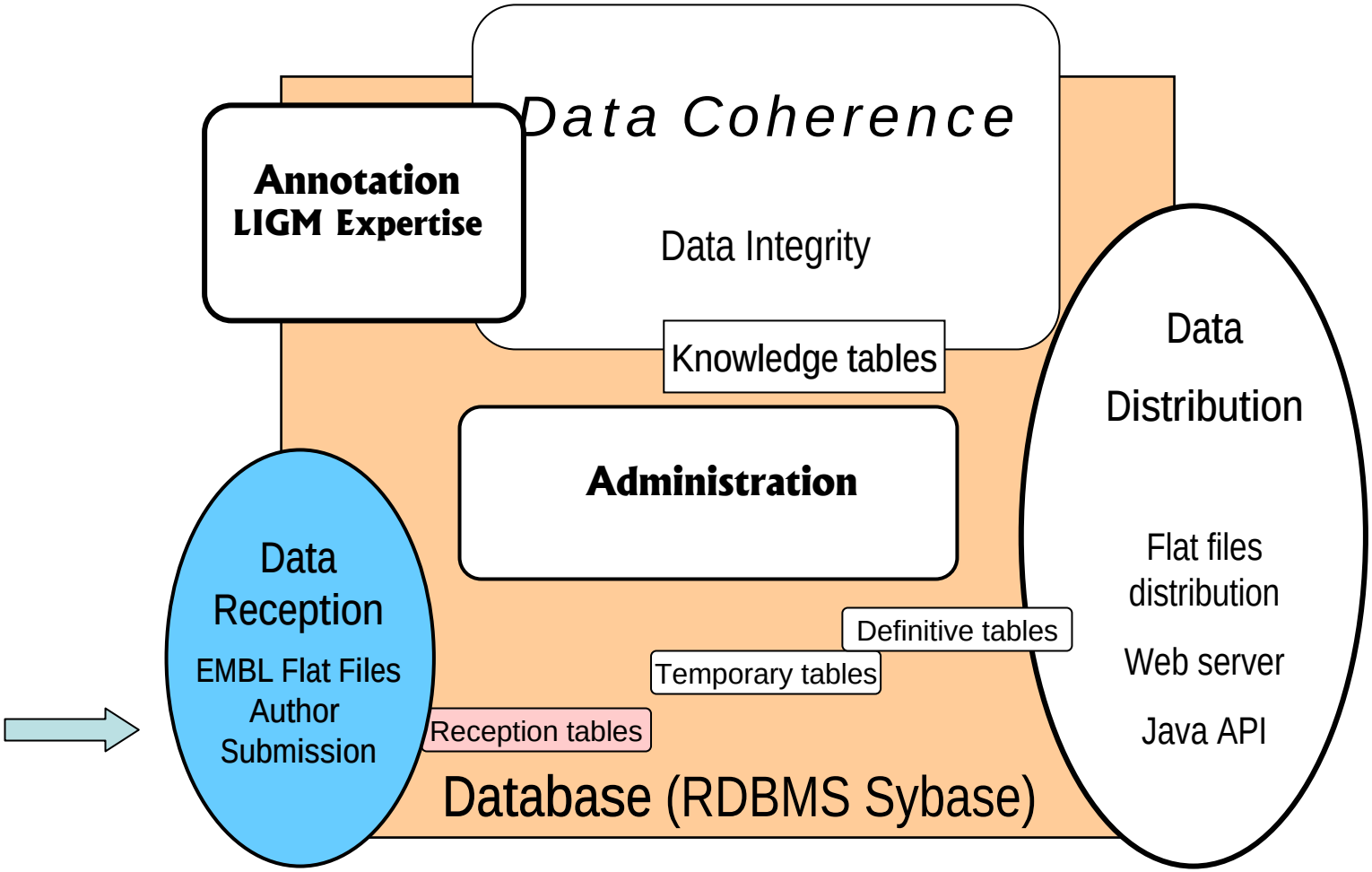
1- Coordination des flux de données

http://imgt.cines.fr



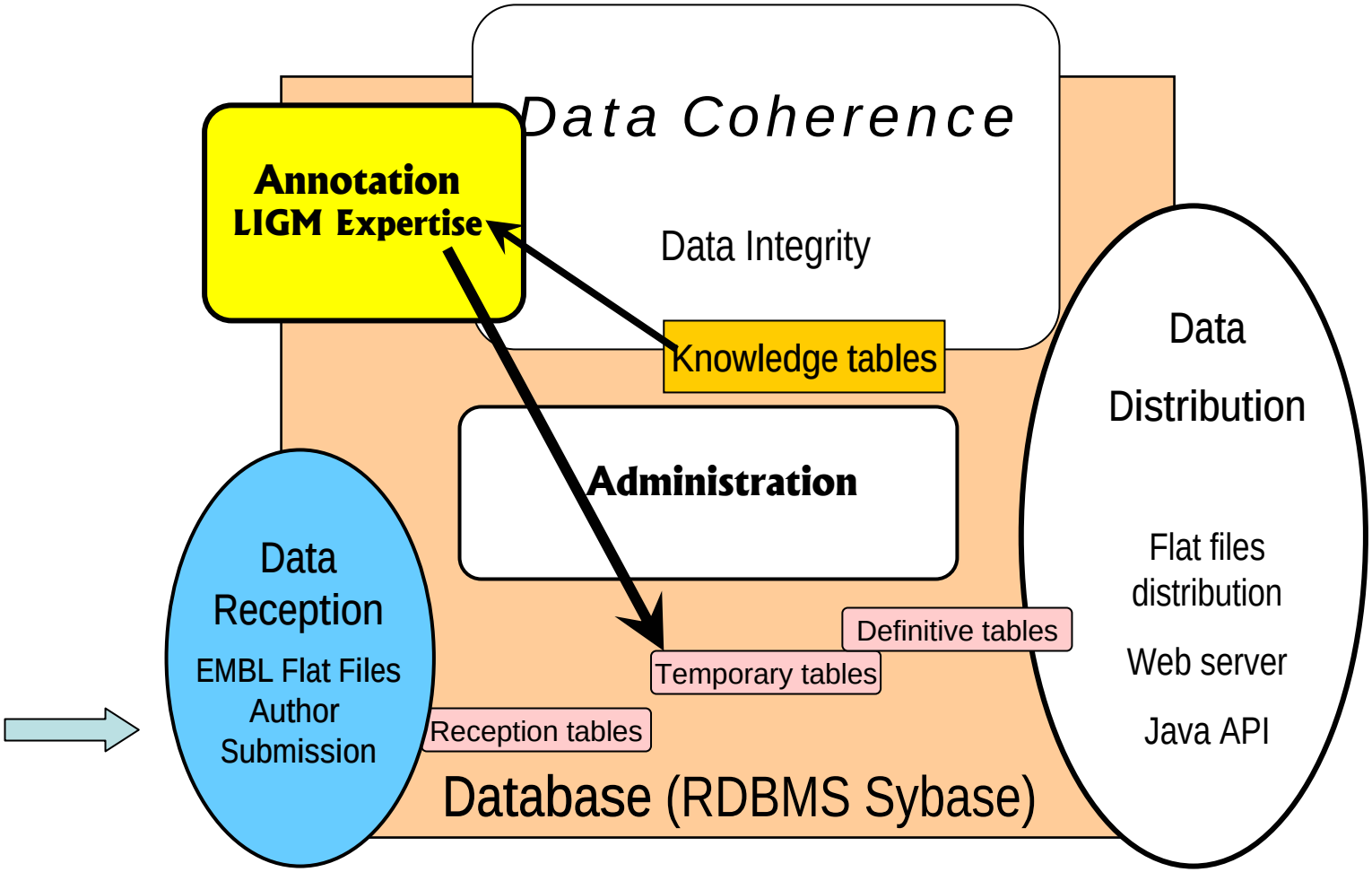
1- Coordination des flux de données

http://imgt.cines.fr



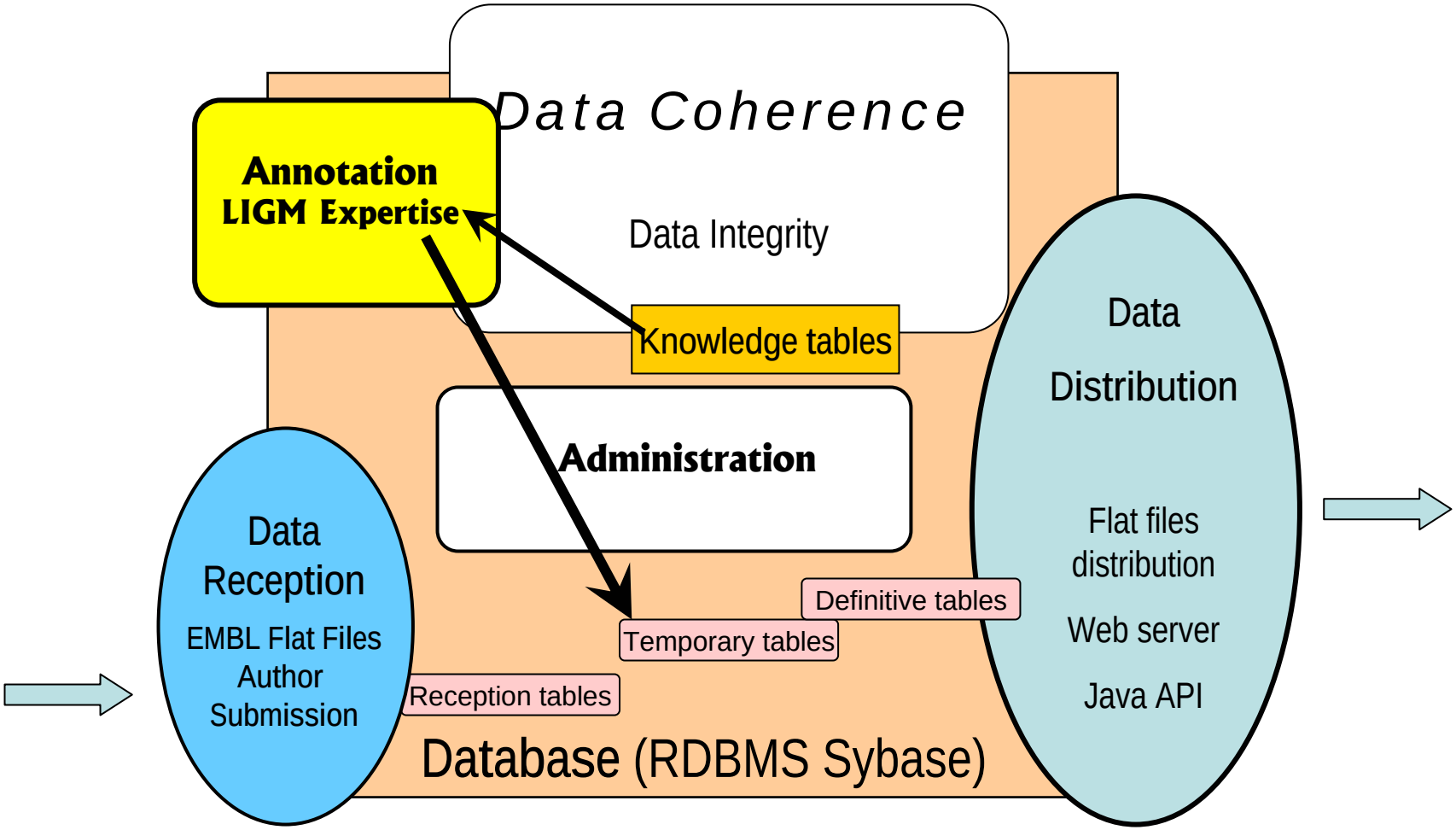
1- Coordination des flux de données

http://imgt.cines.fr



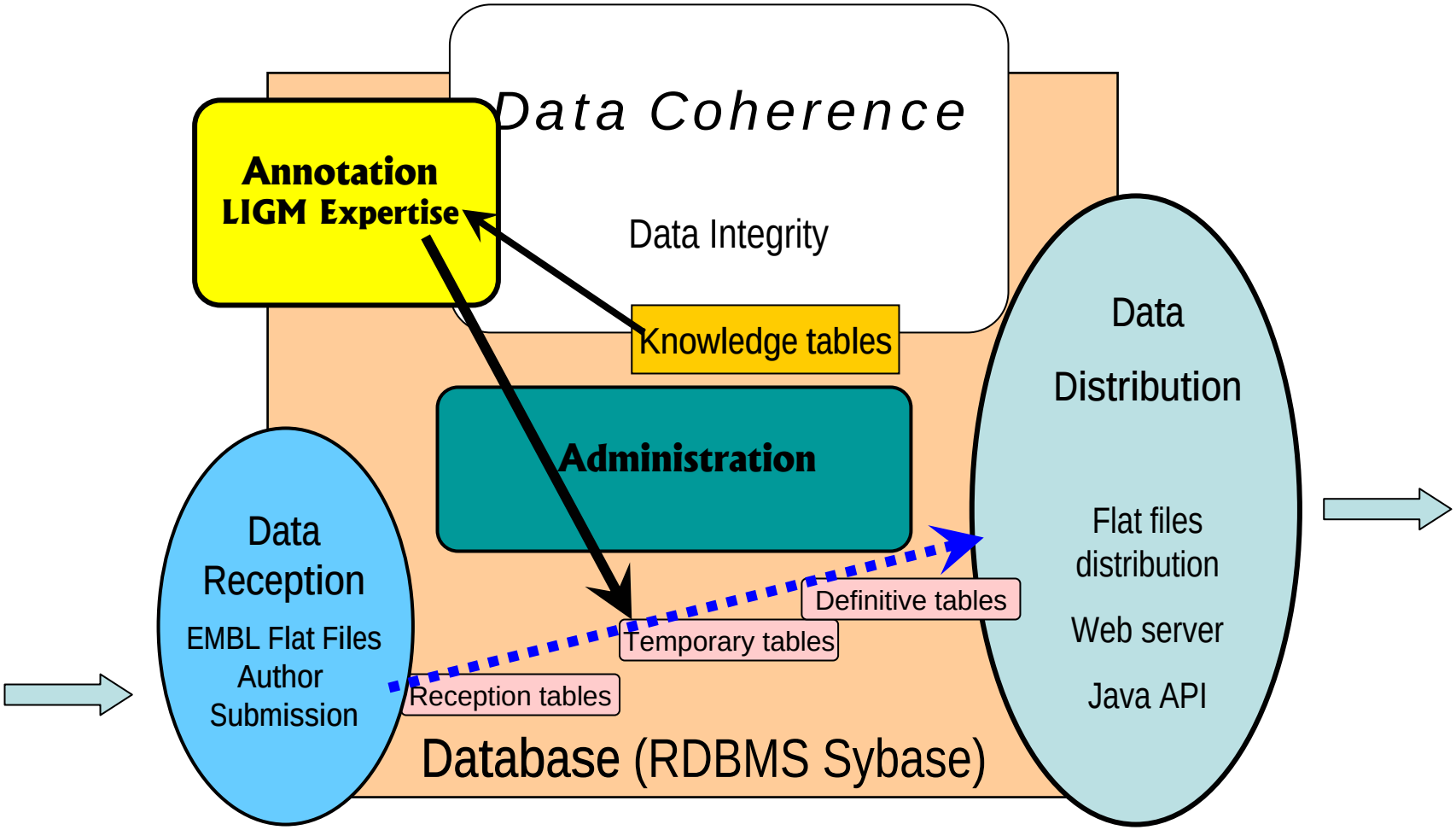
1- Coordination des flux de données

http://imgt.cines.fr



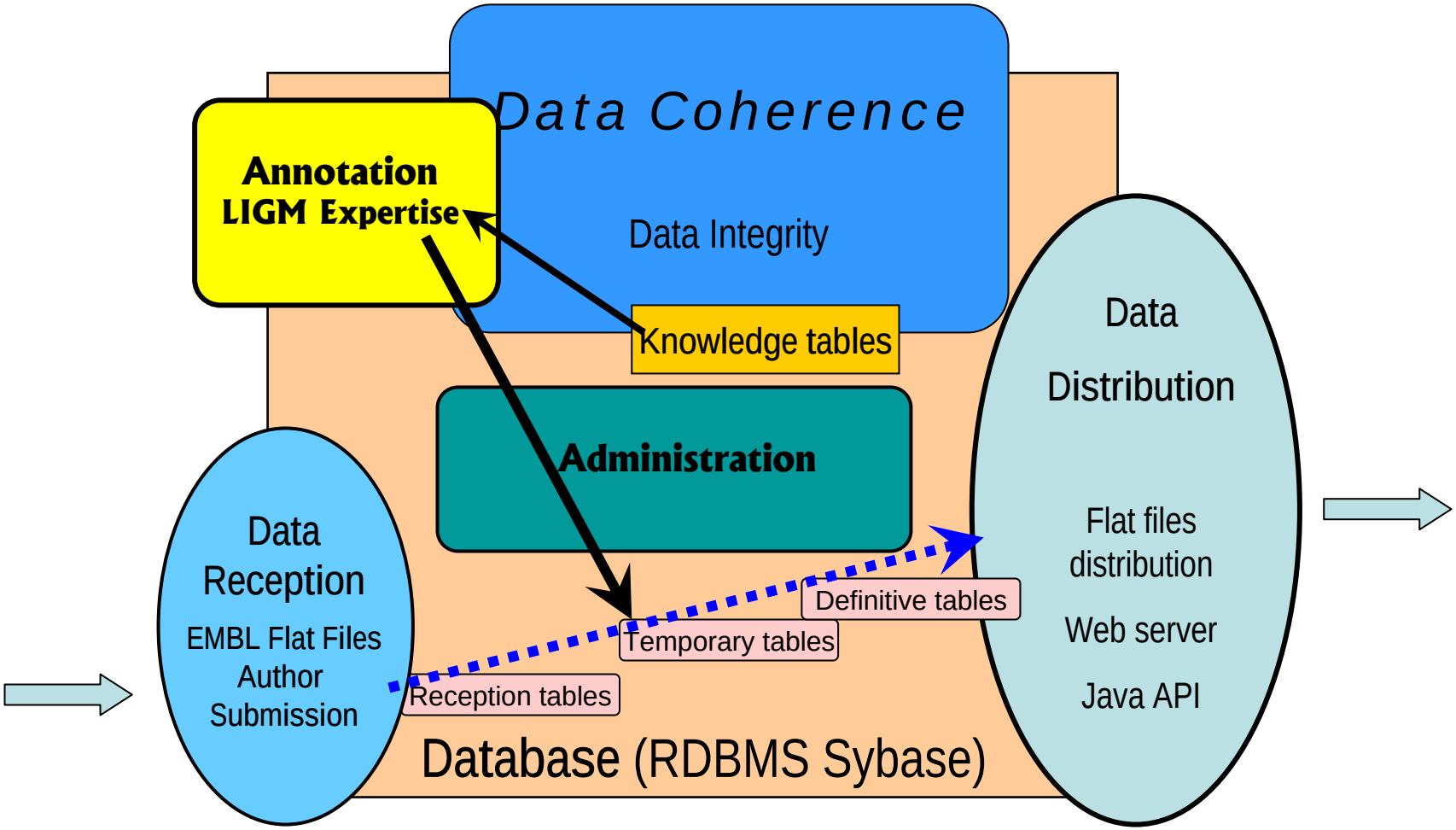
1- Coordination des flux de données

http://imgt.cines.fr



1- Coordination des flux de données

http://imgt.cines.fr

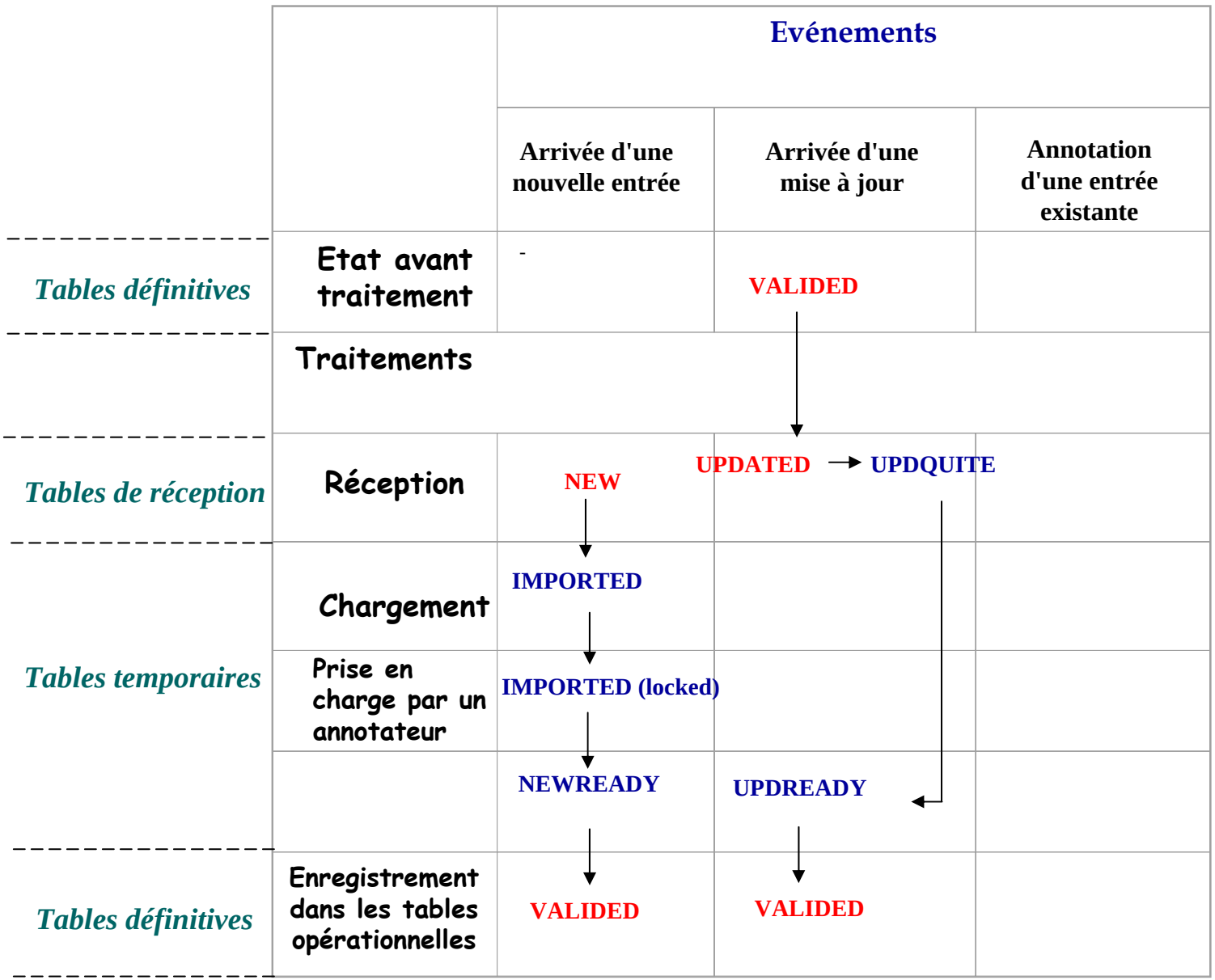


2 - Gestion accès aux données

http://imgt.cines.fr

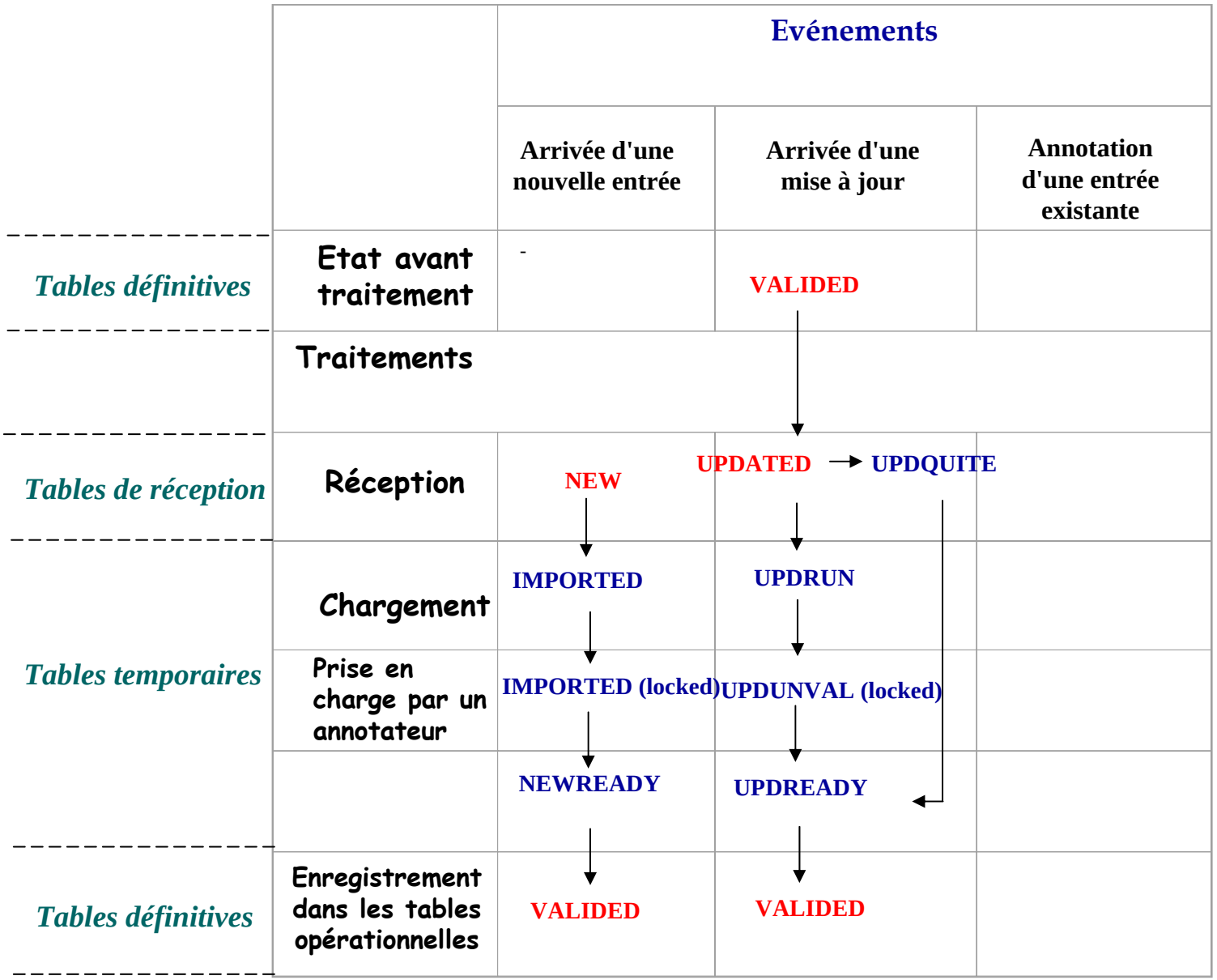
		Evénements		
		Arrivée d'une nouvelle entrée	Arrivée d'une mise à jour	Annotation d'une entrée existante
Tables définitives	Etat avant traitement	-	VALIDED	
	Traitements		↓	
Tables de réception	Réception	NEW	UPDATED	
Tables temporaires	Chargement	IMPORTED		
	Prise en charge par un annotateur	IMPORTED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

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



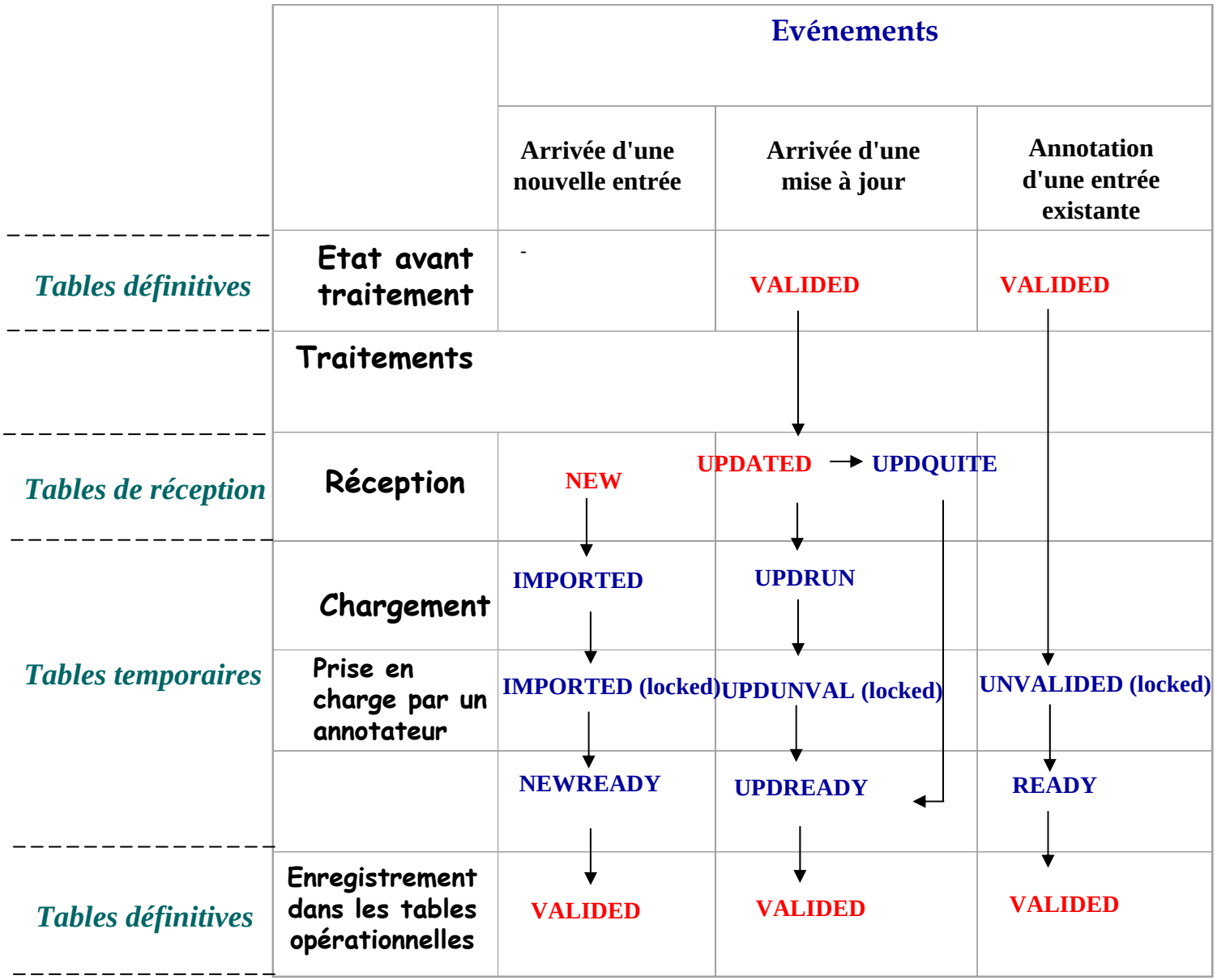
(locked) : indique que l'entrée est prise en charge par un annotateur

http://imgt.cines.fr



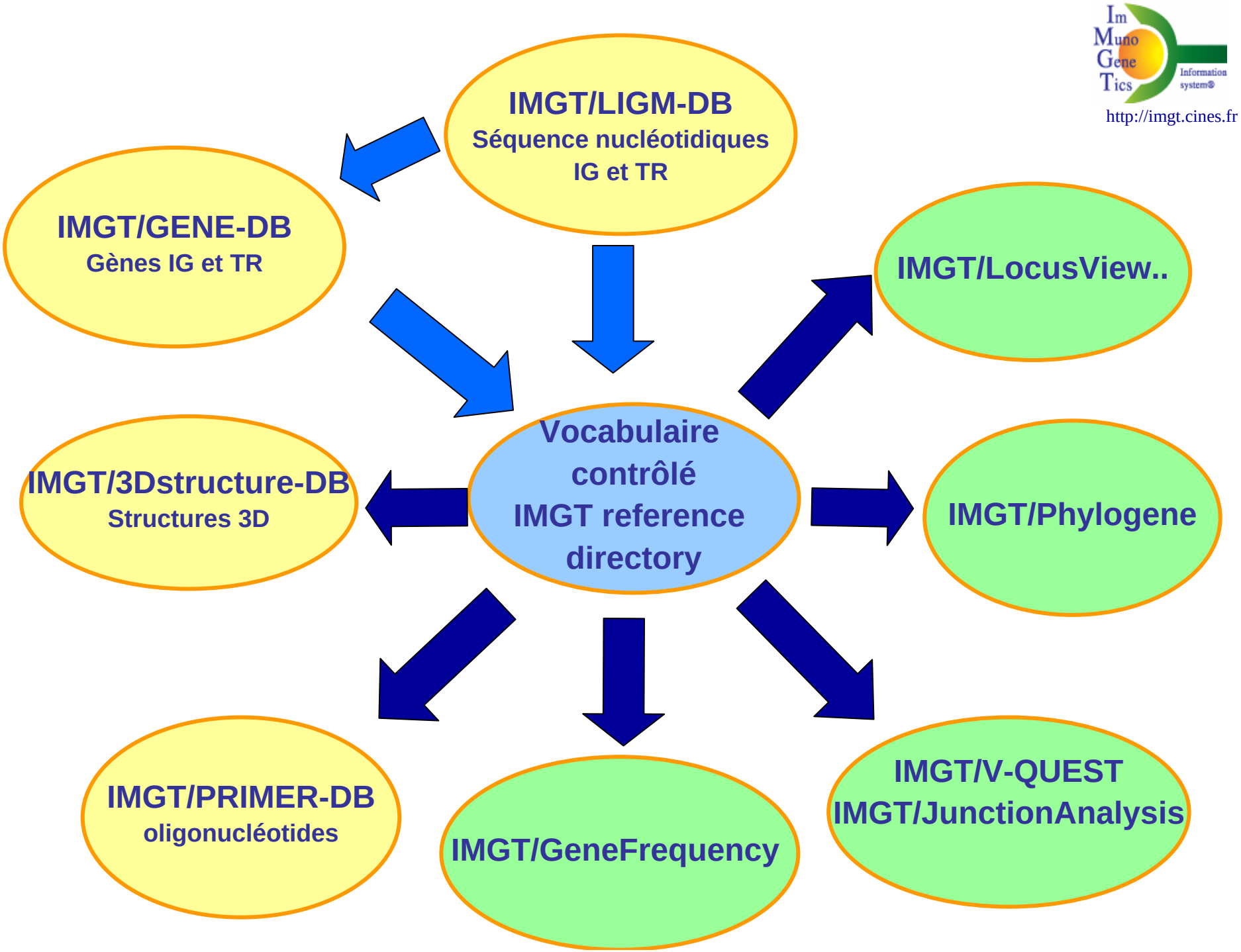
http://imgt.cines.fr

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



http://imgt.cines.fr

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



http://imgt.cines.fr

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, middleware 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 !

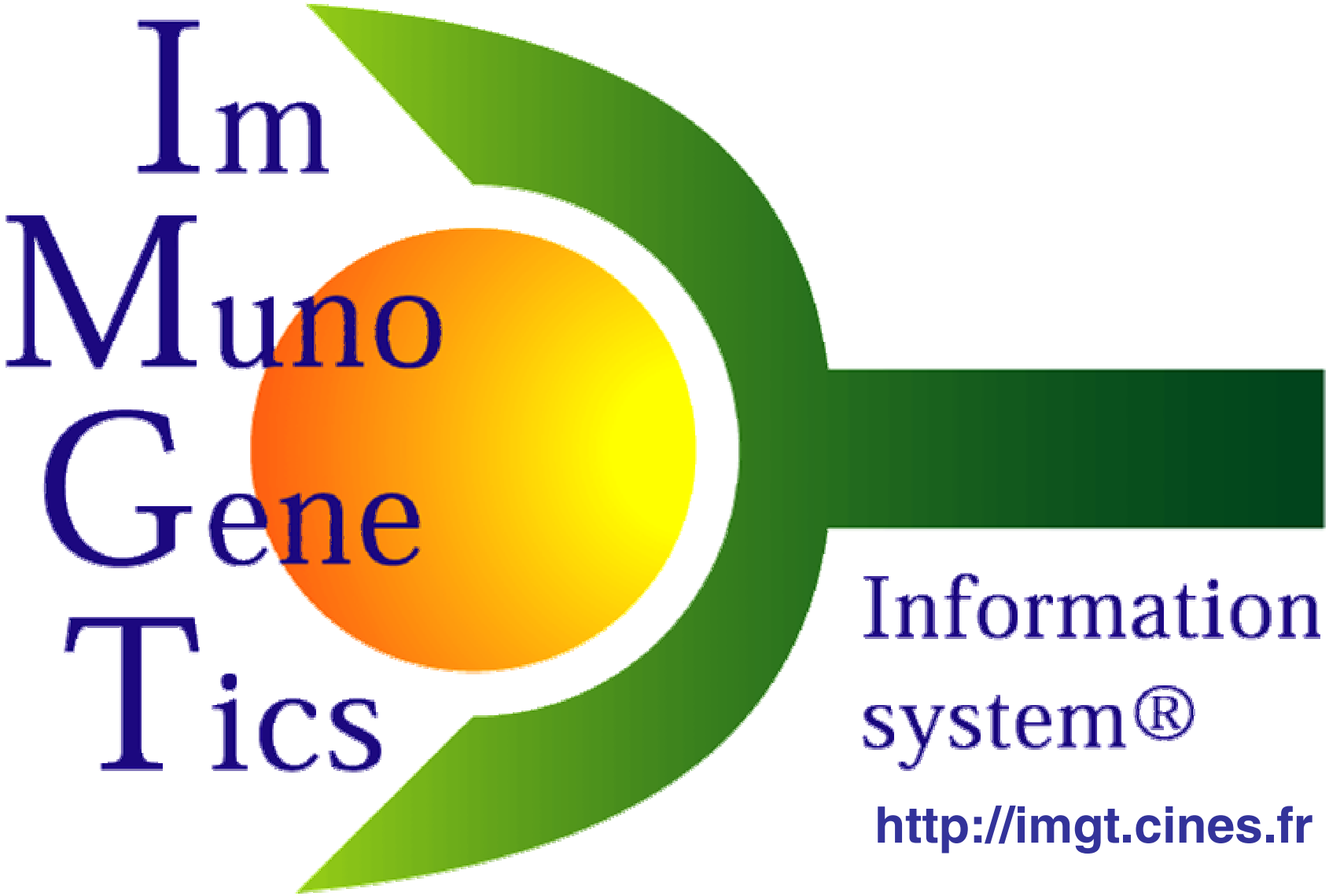
La durée de vie d'un SI dépend de

- 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://imgt.cines.fr