Immunoinformatique

Marie-Paule Lefranc

Laboratoire d'ImmunoGénétique Moléculaire
Université Montpellier, UPR CNRS 1142, IGH
Institut Universitaire de France

Master recherche 2, TC3, Universités Montpellier I et II, 5 octobre 2006

When and why was IMGT created?

IMGT birth and objectives

IMGT was created in June 1989, following the Human Genome Meeting HGM, at New Haven, USA.

* ***Thomas Shows and Ken Kidd, meeting organizers***
* ***Lap Chee Tsui, chromosome 7 responsable***
* ***Claude Boucheix (CD nomenclature)***

1. At New haven, entry of the 16 human T cell receptor gamma genes (TRG) in the genome database, Genome Databank (GDB).

2. IMGT was created to answer the need to manage the T cell receptor (TR) and immunoglobulin (IG) genes and to enter them in the generalist databases.

Why was IMGT necessary?

1. Unique structure of the immunoglobulin (IG) and T cell receptor (TR) genes.

Genomic DNA in germline configuration

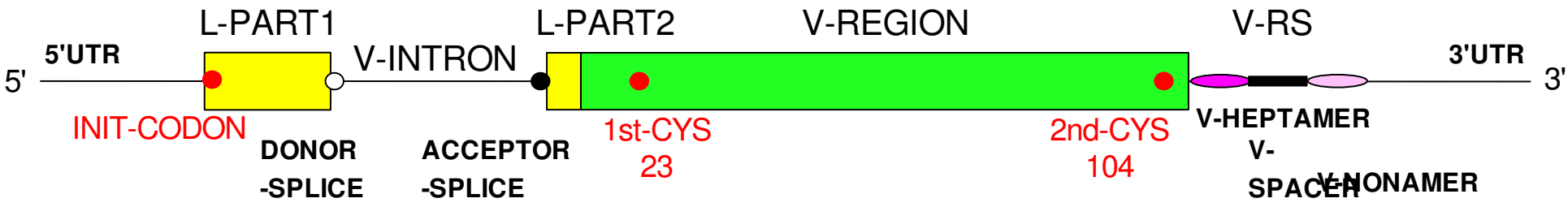
http://imgt.cines.fr

V-GENE

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

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

tgaaa
acccacatcc

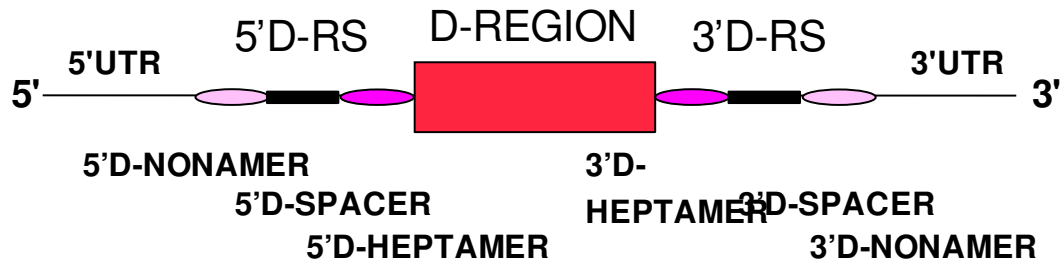


Genomic DNA in germline configuration

Im Muno Gene Tics Information system®
http://imgt.cines.fr

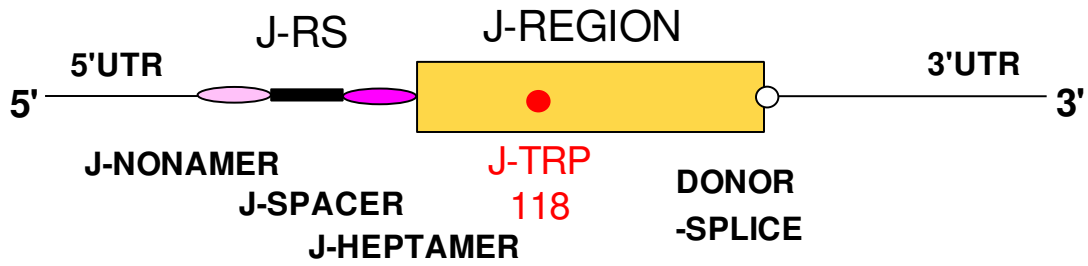
D-GENE

ggtttttggc tgagctgaga accactgtgc taactgggga cacagtcatt
ggcagctcta caaaaacc



J-GENE

ggtttct gtgcccctgg ctcagggctg actcaccgtg gctgaatact
tccagcactg gggccagggc accctggtca ccgtctcctc ag

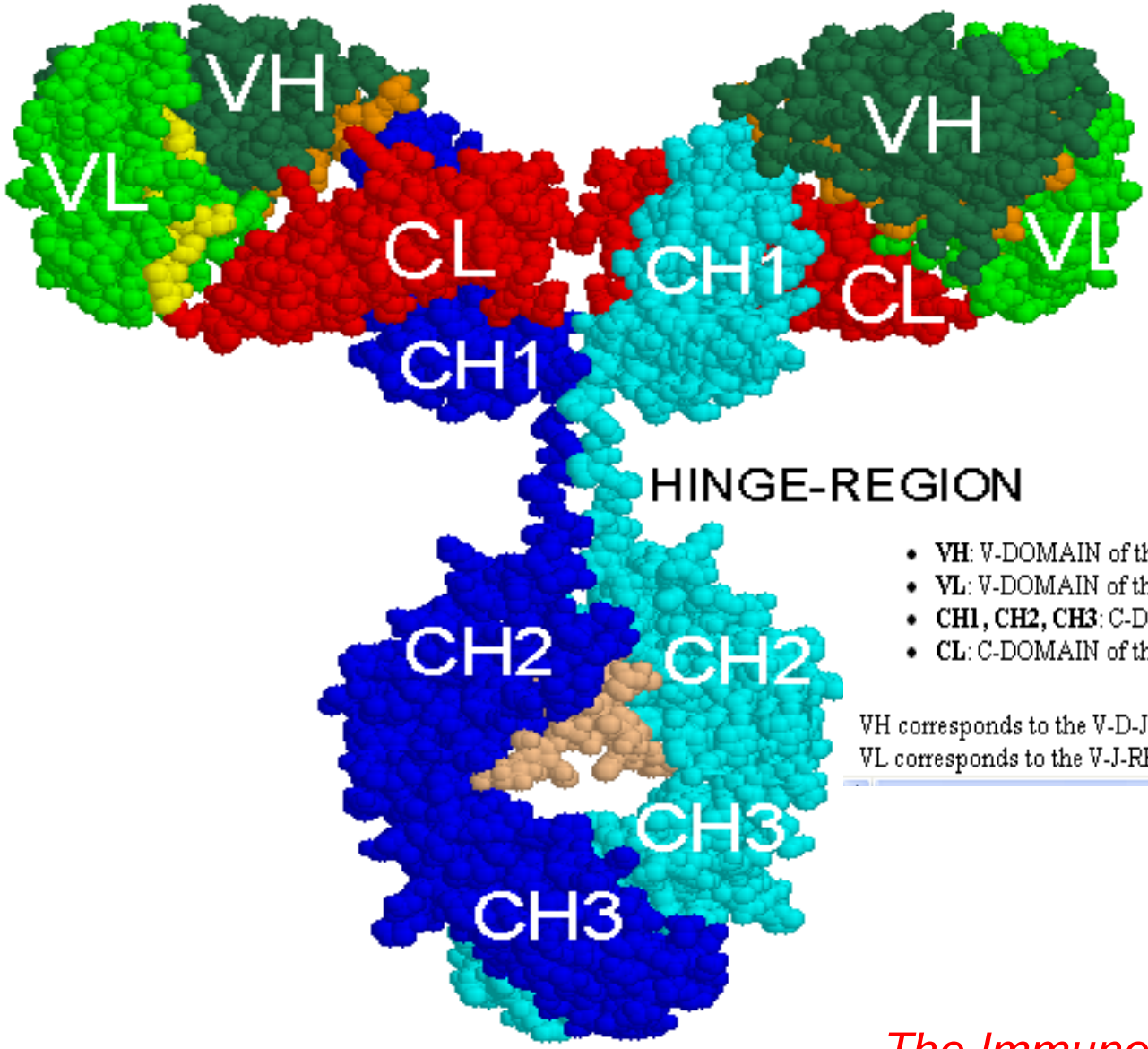


Why was IMGT necessary?

2. Huge diversity of the immunoglobulin and T cell receptor repertoires.

Spacefill 3D representation of an IgG

http://imgt.cines.fr



- **VH**: V-DOMAIN of the immunoglobulin heavy chain
- **VL**: V-DOMAIN of the immunoglobulin light chain
- **CH1, CH2, CH3**: C-DOMAIN of the immunoglobulin heavy chain
- **CL**: C-DOMAIN of the immunoglobulin light chain

VH corresponds to the V-D-J-REGION (in green (V), orange (DJ)) of the heavy chain.
VL corresponds to the V-J-REGION (in green (V) and yellow (J)) of the light chain.

The Immunoglobulin FactsBook, 2001

Immunoglobulin (IG)

T cell receptor (TR)

http://imgt.cines.fr

V-DOMAIN

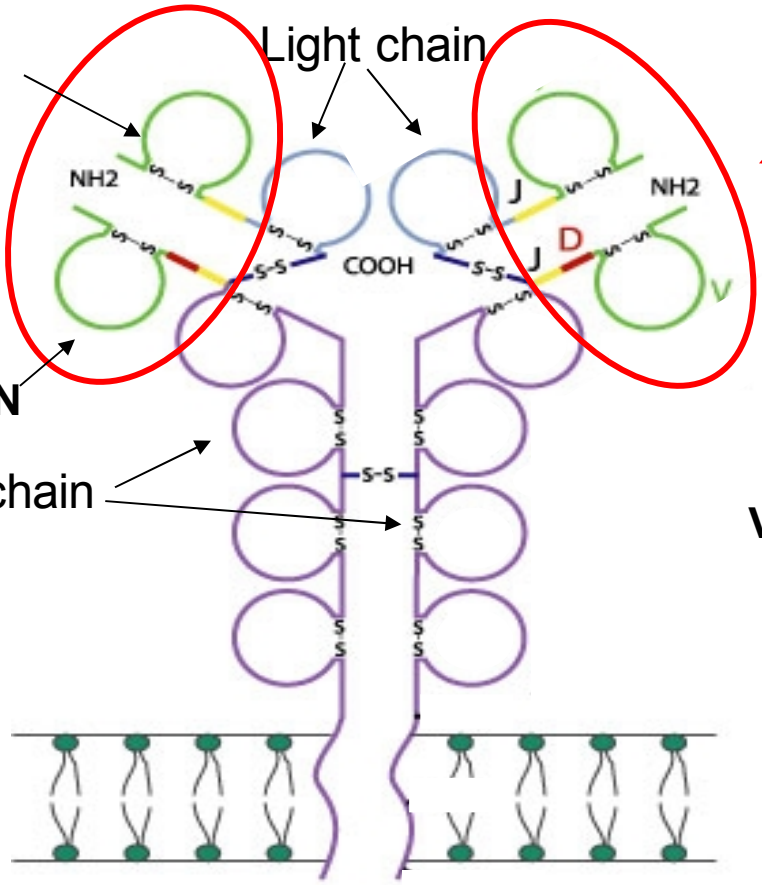
V-J-REGION

V-DOMAIN

V-D-J-REGION

Light chain

Heavy chain



Membrane IgM

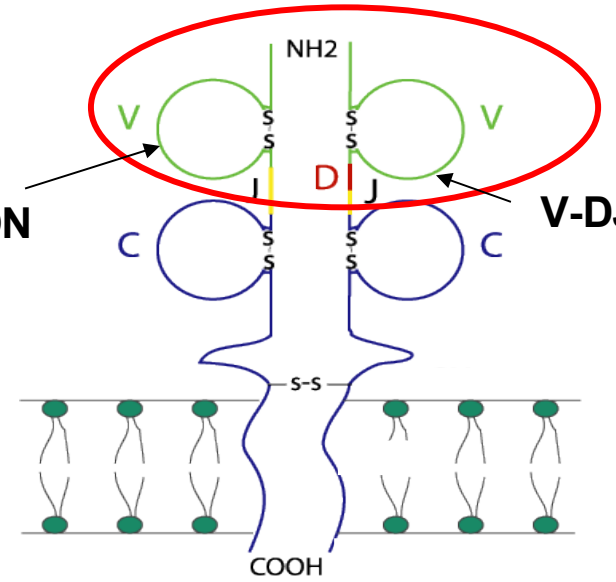
Contribution of the
2 V-DOMAINs
to the antigen binding site

Alpha - Beta

Gamma - Delta

V-J-REGION

V-DJ-REGION

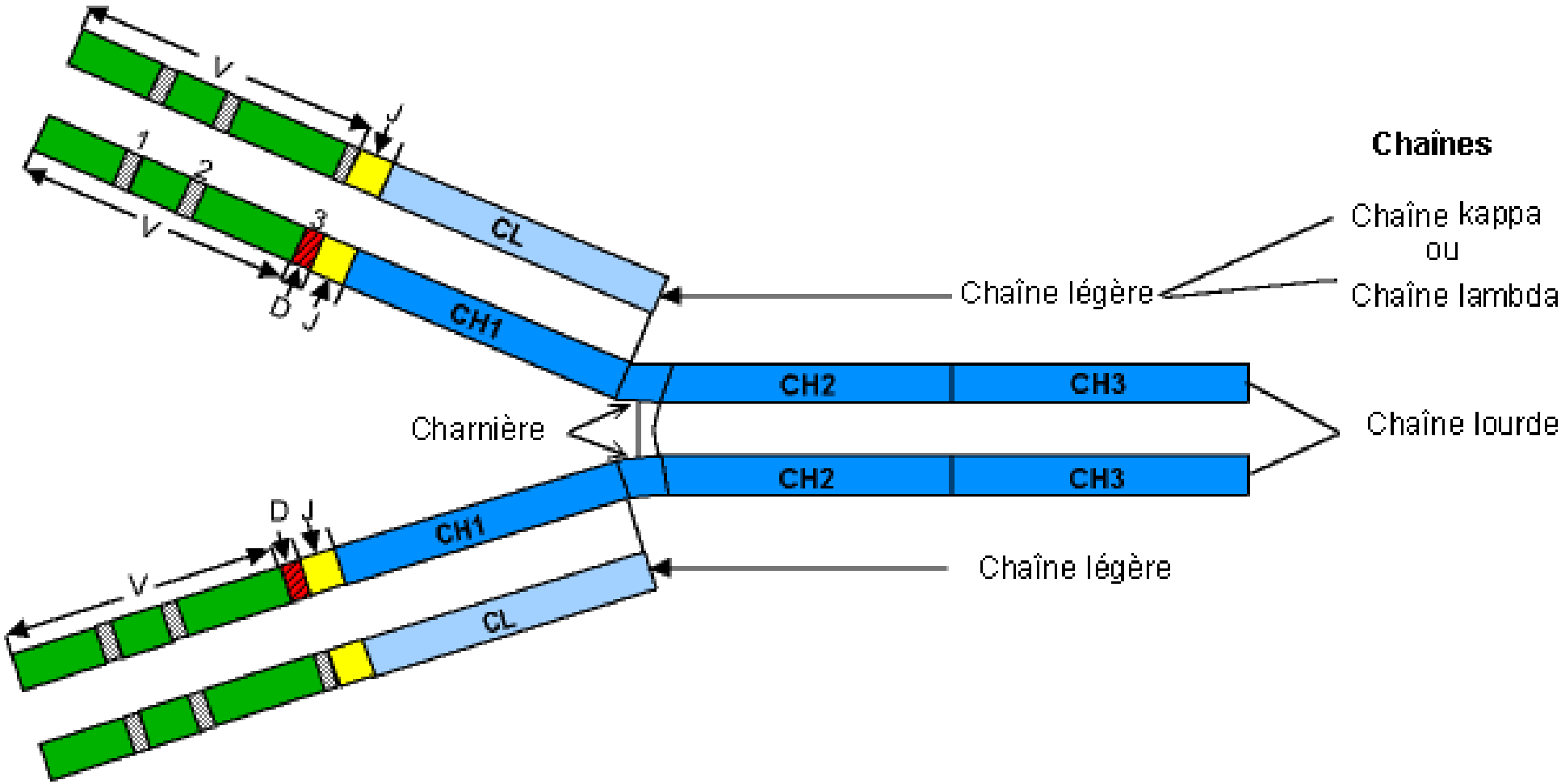


T cell receptor

The Immunoglobulin FactsBook, 2001

Immunoglobulin IgG

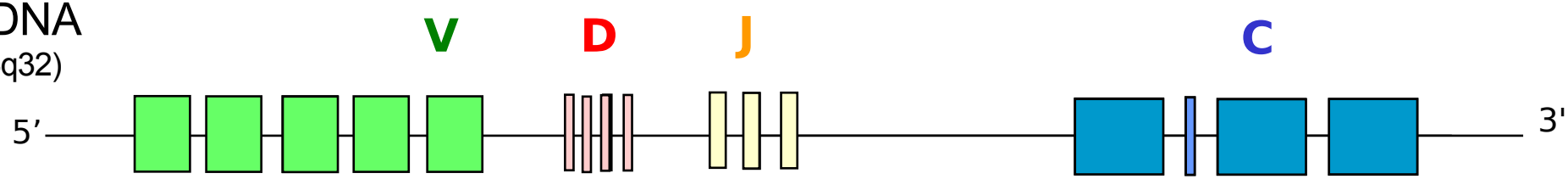
http://imgt.cines.fr



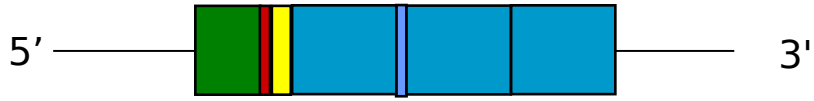
Immunoglobulin (IG) synthesis

http://imgt.cines.fr

genomic DNA
(IGH Locus 14q32)

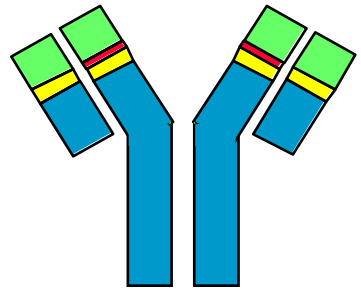


rearranged DNA



mRNA

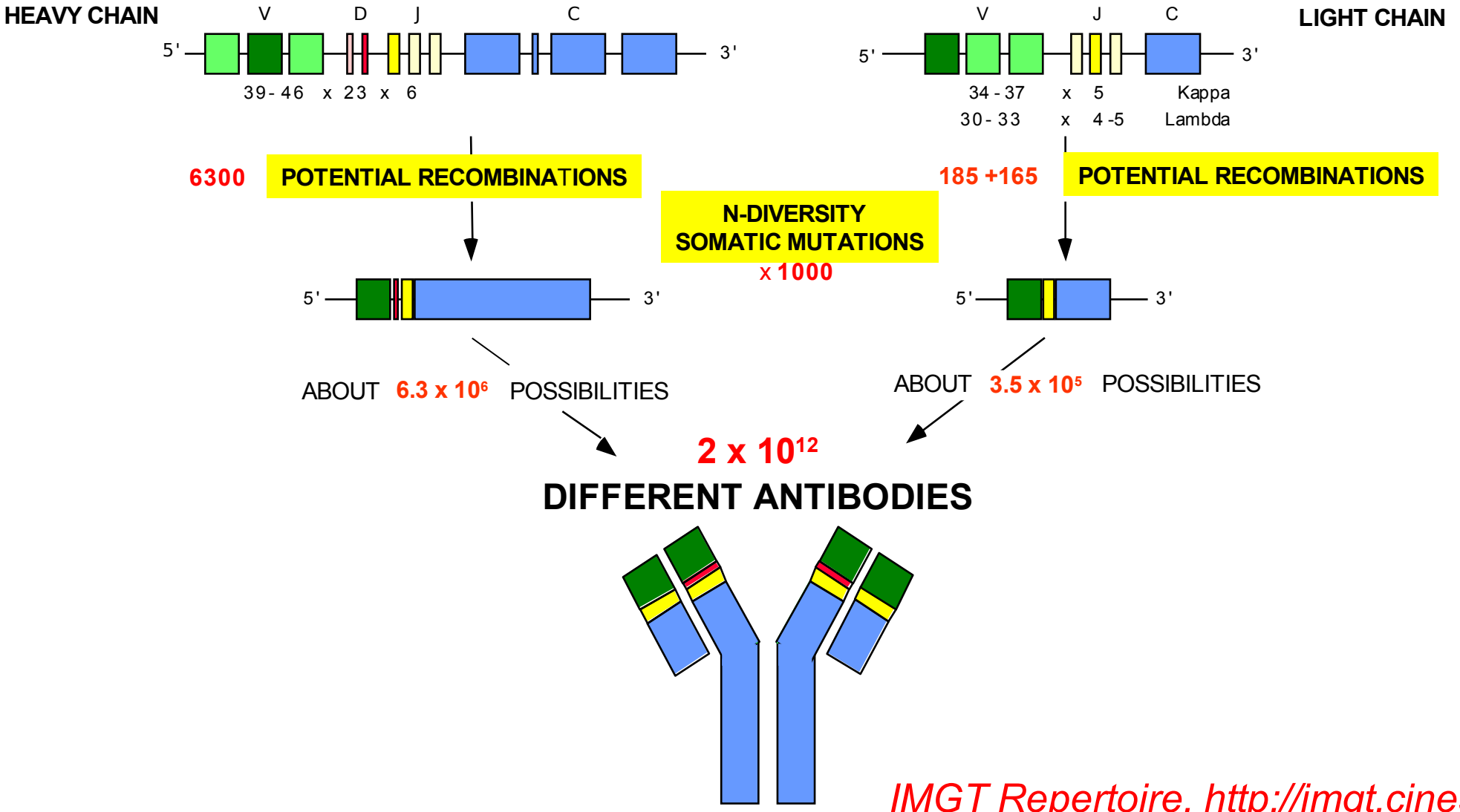
2 x 10^{12} different IG per individual



Immunoglobulin (IG) synthesis

http://imgt.cines.fr

150
FUNCTIONAL IG GENES



What was the first goal?

CONTENT

1. Immunoglobulins AND T cell receptors
2. Human AND other vertebrate species

DATA

1. Genes AND alleles
2. Sequences (germline AND rearranged)
3. Polymorphisms: genomic AND somatic hypermutations
4. 3D structures

IMGT® Web resources

http://imgt.cines.fr

- Chromosomal localizations
- Locus representations
- Genes tables
- Potential repertoire
- Alignments of alleles
- Protein displays
- IMGT Colliers de Perles…

IMGT Index

IMGT Bloc-notes

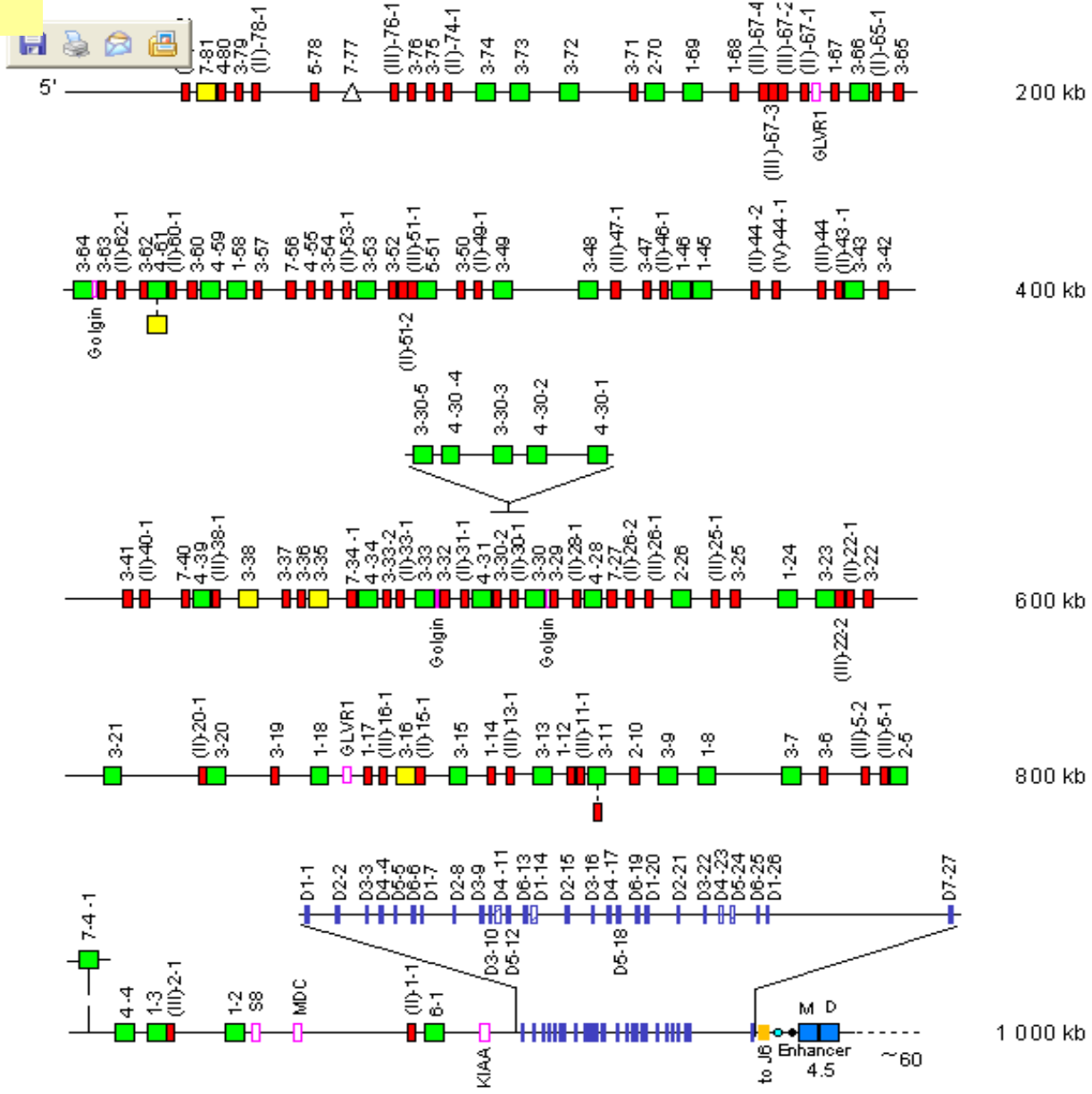
- Interesting links, PubMed,
- Meeting announcements
- IMGT Immunoinformatics page,…

IMGT Education

- Tutorials, IMGT Lexique, …

Human IGH locus

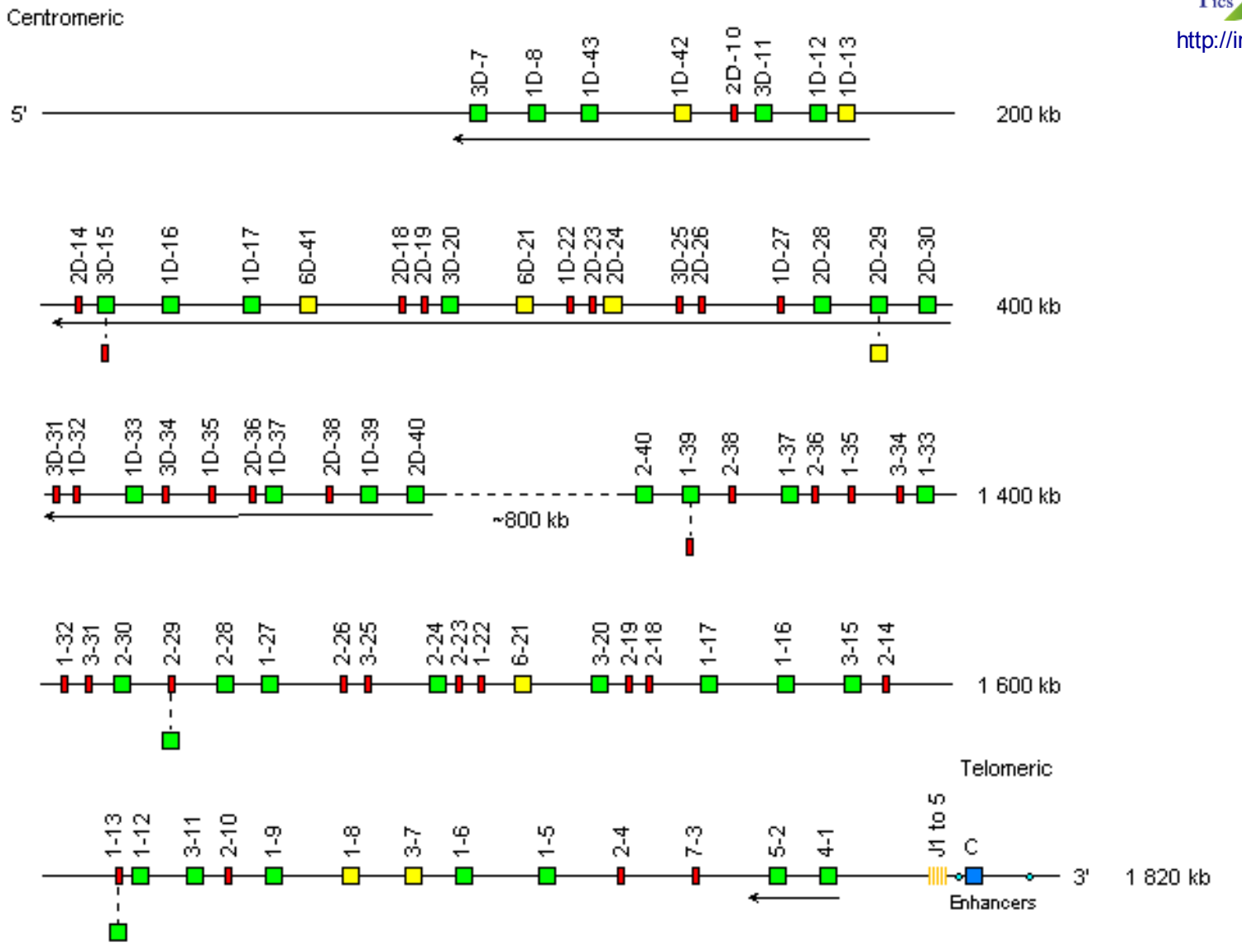
Chromosome 14q32.33



http://imgt.cines.fr

Human IGK locus

Chromosome 2p11.2

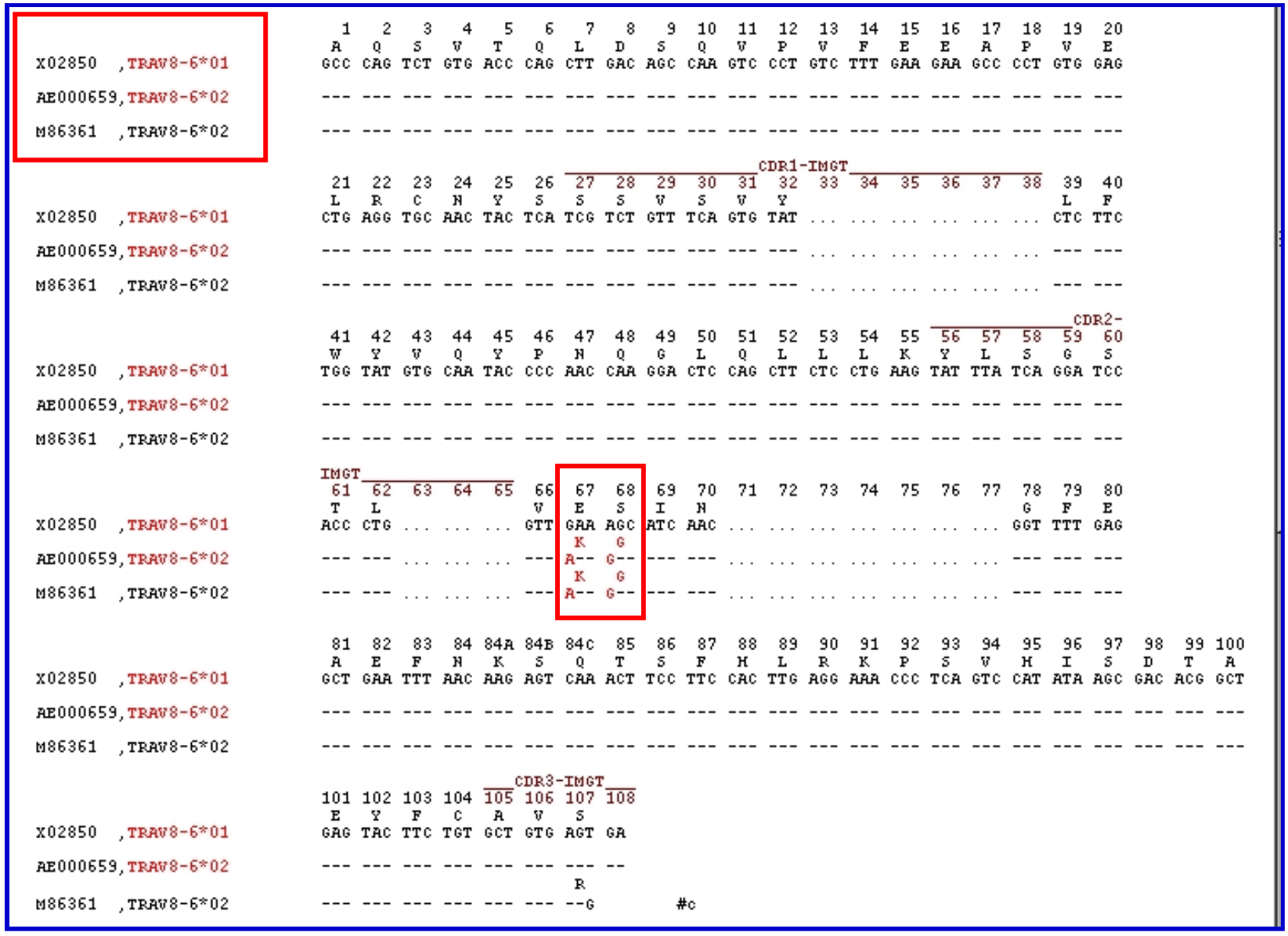


Human IGHV genes on chromosome 14 at 14q32.33

http://imgt.cines.fr

Subgroup	Functional	ORF	Pseudogene	Total
IGHV1	9	-	5	14
IGHV2	3	-	1	4
IGHV3	18-20**(+1)*	3	25(+1)*	47-49**
IGHV4	6-9**(+1)*	(+1)*	2	9-12**
IGHV5	1	-	1	2
IGHV6	1	-	-	1
IGHV7	0-1**	1	4	5-6**
IGHV(II)	-	-	22	22
IGHV(III)	-	-	18	18
IGHV(IV)	-	-	1	1
Total	38-44(+2)*	4(+1)*	79(+2)*	123-129**

IMGT Alignment of alleles

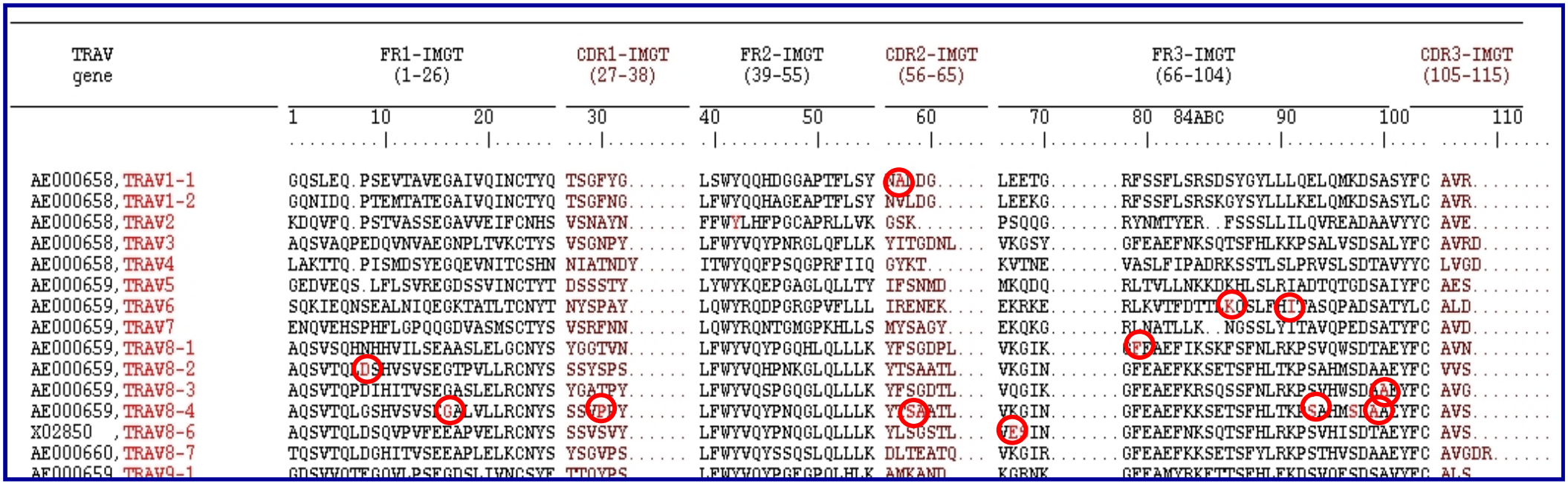


```
                                1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
                                A   Q   S   V   T   Q   L   D   S   Q   V   P   V   F   E   E   A   P   V   E
X02850   ,TRAV8-6*01          GCC CAG TCT GTG ACC CAG CTT GAC AGC CAA GTC CCT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

http://imgt.cines.fr

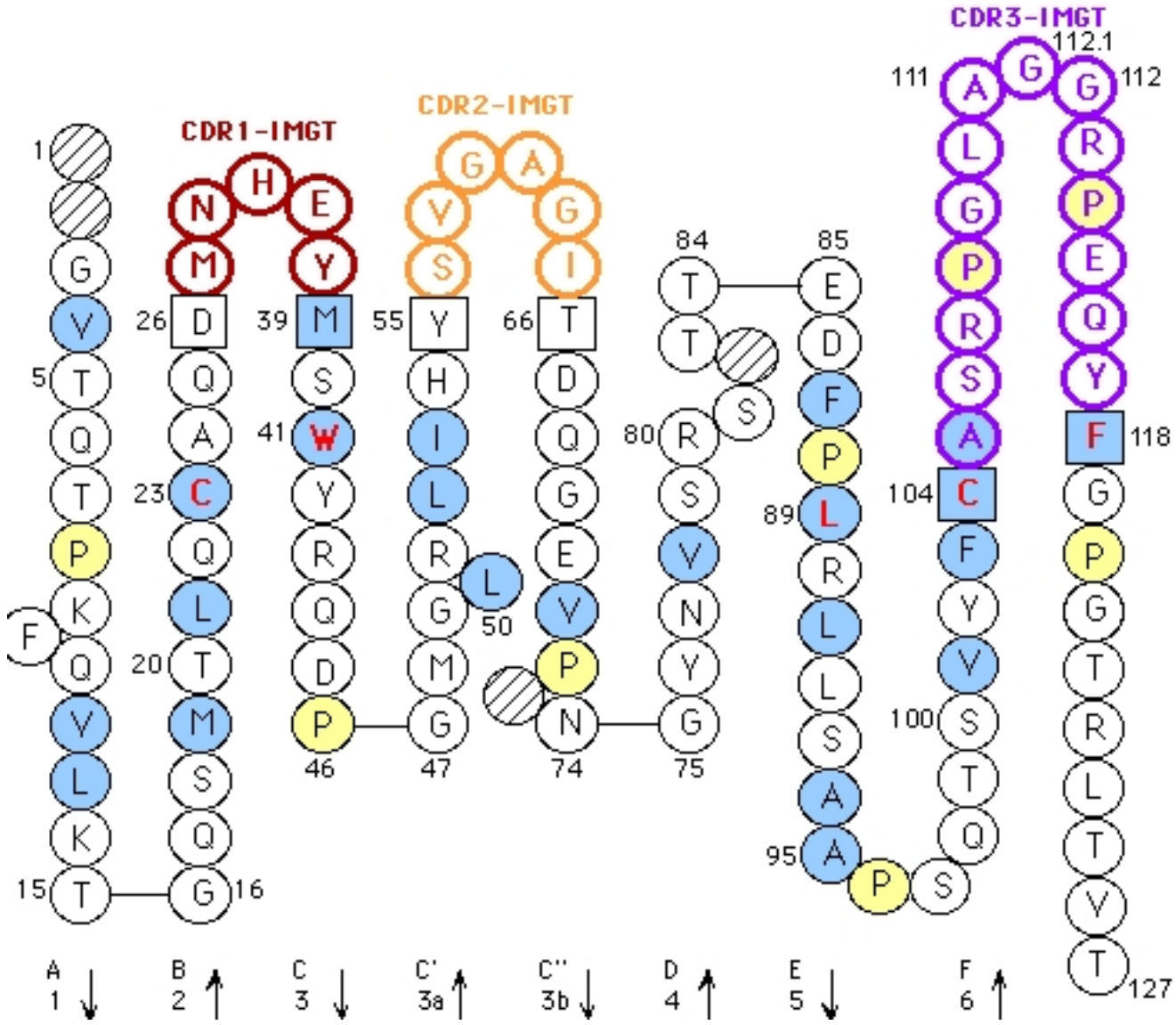
IMGT Protein display

http://imgt.cines.fr

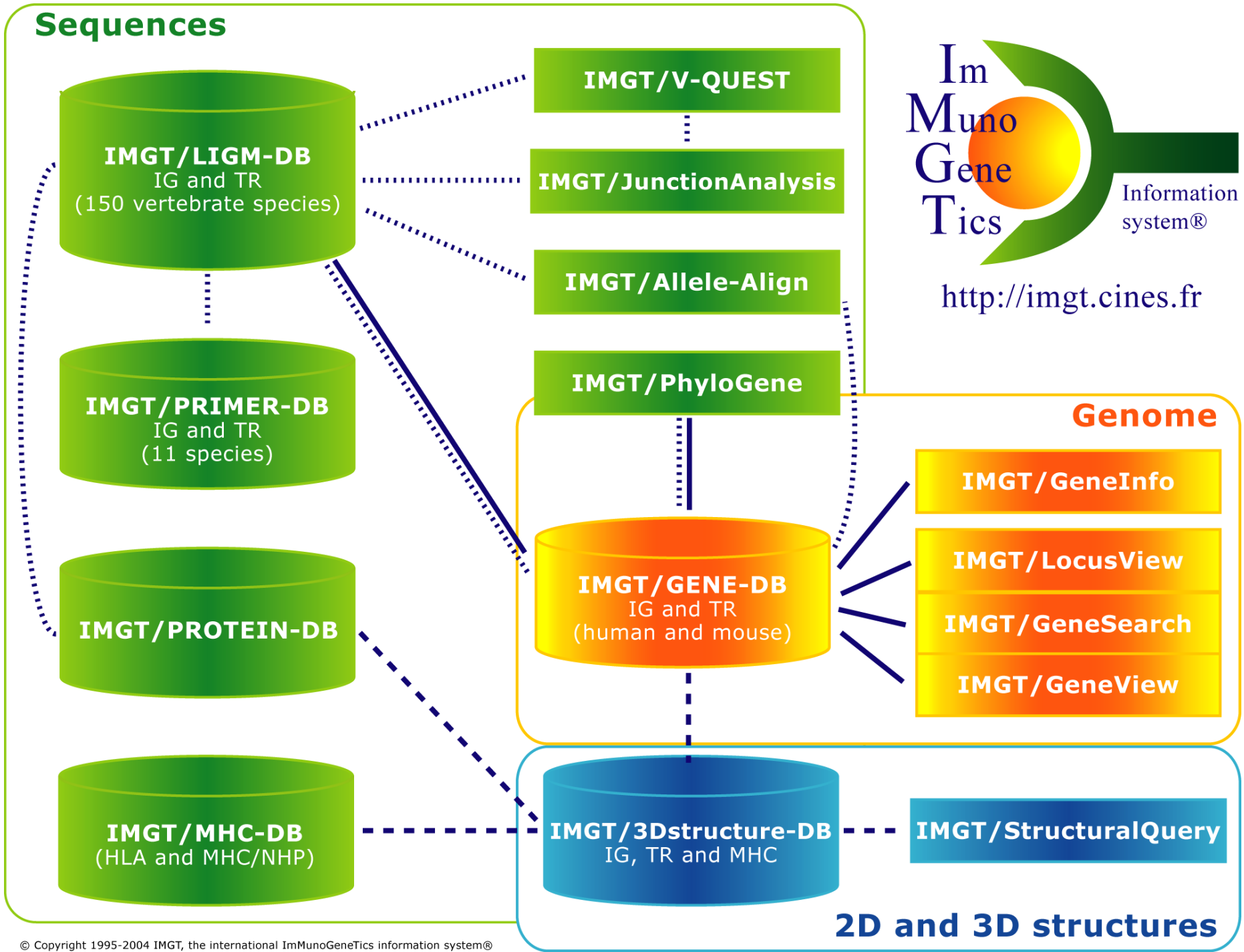


IMGT Collier de Perles

http://imgt.cines.fr



IMGT databases and tools



© Copyright 1995-2004 IMGT, the international ImMunoGeneTics information system®

IMGT Marie-Paule page: IMGT repertoire - Netscape

Fichier Edition Affichier Aller Communicator Aide

Locus representation: Human IGL

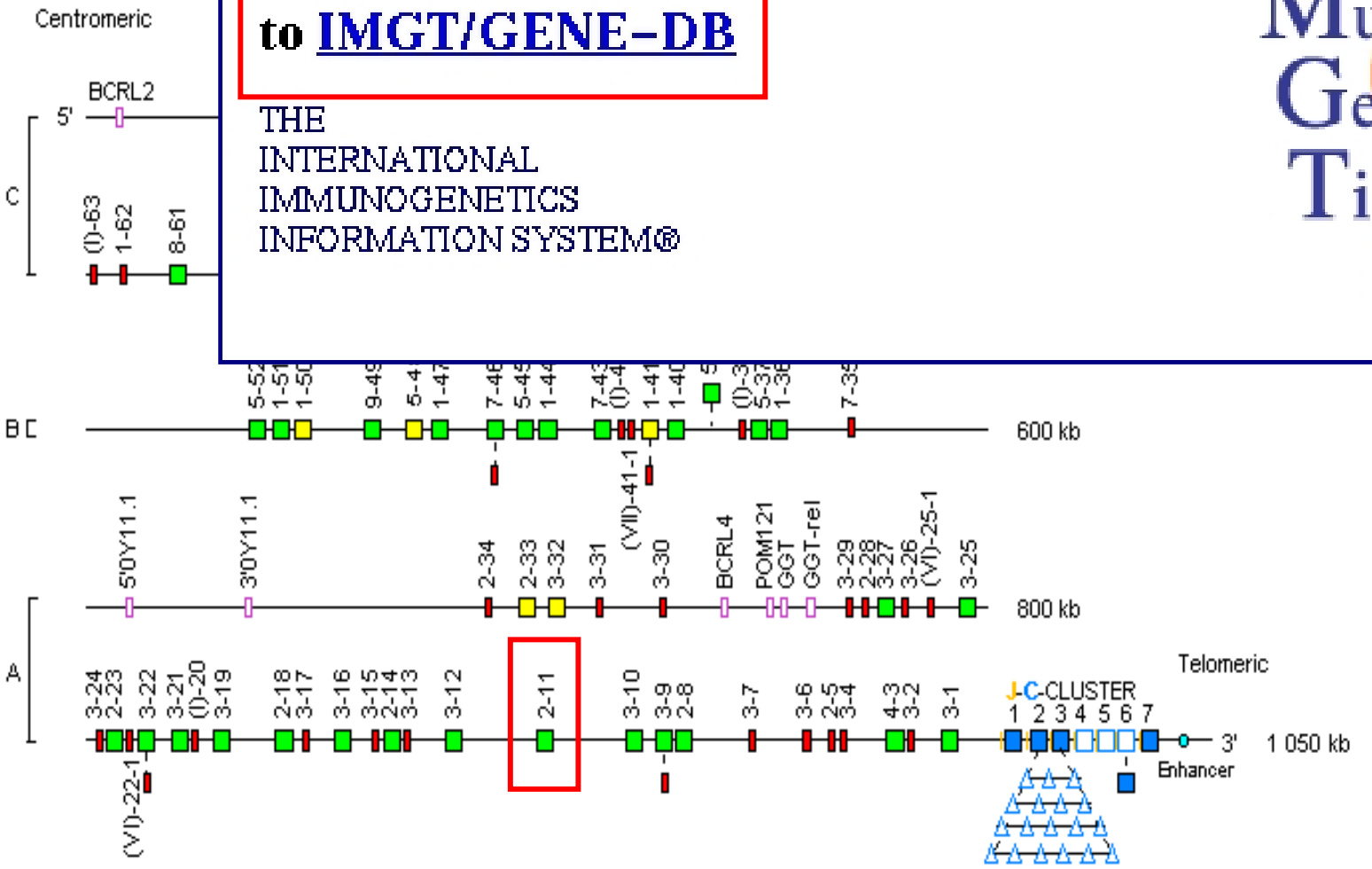
Human IGL 22q11.2

WELCOME !
to IMGT/GENE-DB

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

Im Muno Gene Tics Information system®

http://imgt.cines.fr



Giudicelli V. et al. Nucleic Acids Res. 33, D256-D261

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
FT
FT
FT
FT
FT
FT
FT   V-REGION
FT
FT
FT
FT
FT
FT   FR1-IMGT
FT
FT
FT   1st-CYS
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
```

Help | IMGT Home page | IMGT Marie-Paule page | New search

Created by Marie-Paule Lefranc (CNRS, Montpellier II University, France)
http://imgt.cines.fr

IMGT/LIGM-DB ON LINE, HERE YOU ARE !

Five types of search are available : select one by clicking on the button

Catalogue — accession number, mnemonic, definition, creation date, length, annotation level

Denys Chaume

Document : chargé

cDNA (in databases: mRNA!)

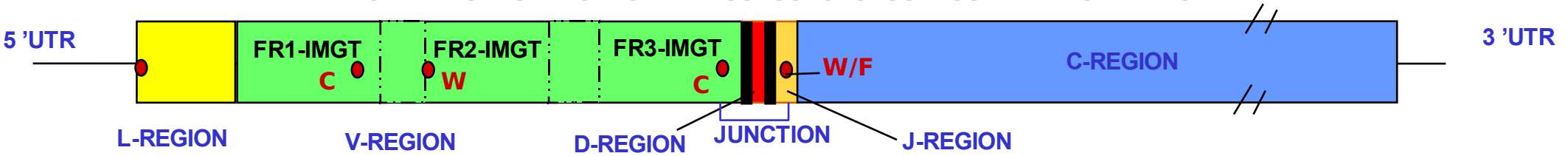
Im Muno Gene Tics Information system®
http://imgt.cines.fr

```
……..gagga ttcaccatgg aactggggct ccgctgggtt ttccttgttg ctttttaga 120
aggtgtccag tgtgaggtgc aactggtgga gtctggggga ggcctggtca agccgggggg 180
gtccctgaga ctctcctgtg cagcctctgg attaagcttc agtacctatg ccatgaactg 240
ggtccgccag gctccaggga aggggctgga atgggtctca agtattagta gtagaagtga 300
ttacatatac tatagagact cagtgaaggg ccgattcacc atctccagag acaacgccaa 360
gaattcactg tatctgcaaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgcgagagat tcttgtaatg gtgctatatg ttatggtttc agtccctggg gccagggaac 480
cctggtcacc gtctcctcag catccccgac cagccccaag gtcttcccgc tgagcctctg 540
cagcacccag ccagatggga acgtggtcat cgcctgcctg gtccagggct tcttccccca 600
ggagccactc agtgtgacct ggagcgaaag cggacagggc gtgaccgcca gaaacttccc 660
acccagccag gatgcctccg gggacctgta caccacgagc agccagctga ccctgccggc 720
cacacagtgc ctagccggca agtccgtgac atgccacgtg aagcactaca cgaatcccag 780
ccaggatgtg actgtgccct gcccagttcc ctcaactcca cctaccccat ctccctcaac 840
tccacctacc ccatctccct catgctgcca cccccgactg tcactgcacc gaccggccct 900
cgaggacctg ctcttaggtt cagaagcgaa cctcacgtgc acactgaccg gcctgagaga 960
tgcctcgggt gtcaccttca cctggacgcc ctcaagtggg aagagcgctg ttcaaggacc 1020
acctgagcgt gacctctgtg gctgctacag cgtgtccagt gtcctgccgg gctgtgccga 1080
gccatggaac catgggaaga ccttcacttg cactgctgcc taccccgagt ccaagacccc 1140
gctaaccgcc accctctcaa aatccggaaa cacattccgg cccgaggtcc acctgctgcc 1200
gccgccgtcg gaggagctgg ccctgaacga gctggtgacg ctgacgtgcc tggcacgtgg 1260
cttcagcccc aaggacgtgc tggttcgctg gctgcagggg tcacaggagc tgccccgcga 1320
gaagtacctg acttgggcat cccggcagga gcccagccag ggcaccacca ccttcgctgt 1380
gaccagcata ctgcgcgtgg cagccgagga ctggaagaag ggggacacct tctcctgcat 1440
ggtgggccac gaggccctgc cgctggcctt cacacagaag accatcgacc gcttggcggg 1500
taaacccacc catgtcaatg tgtctgttgt catggcggag gtggacggca cctgctactga 1560 …..
```

cDNA

Im Muno Gene Tics Information system®
http://imgt.cines.fr

```
.....gagga ttcaccatgg aactggggct ccgctgggtt ttccttgttg ctttttttaga 120
aggtgtccag tgtgaggtgc aactggtgga gtctgggggga ggcctggtca agccgggggg 180
gtccctgaga ctctcctgtg cagcctctgg attaagcttc agtacctatg ccatgaactg 240
ggtccgccag gctccaggga aggggctgga atgggtctca agtattagta gtagaagtga 300
ttacatatac tatagagact cagtgaaggg ccgattcacc atctccagag acaacgccaa 360
gaattcactg tatctgcaaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgcgagaga  t tcttgtaatg gtgctatatg ttatggtttc agt ccctggg gccagggaac 480
cctggtcacc gtctcctcag catccccgac cagccccaag gtcttcccgc tgagcctctg 540
cagcacccag ccagatggga acgtggtcat cgcctgcctg gtccagggct tcttccccca 600
ggagccactc agtgtgacct ggagcgaaag cggacagggc gtgaccgcca gaaacttccc 660
acccagccag gatgcctccg gggacctgta caccacgagc agccagctga ccctgccggc 720
cacacagtgc ctagccggca agtccgtgac atgccacgtg aagcactaca cgaatcccag 780
ccaggatgtg actgtgccct gcccagttcc ctcaactcca cctaccccat ctccctcaac 840
tccacctacc ccatctccct catgctgcca cccccgactg tcactgcacc gaccggccct 900
cgaggacctg ctcttaggtt cagaagcgaa cctcacgtgc acactgaccg gcctgagaga 960
tgcctcgggt gtcaccttca cctggacgcc ctcaagtggg aagagcgctg ttcaaggacc 1020
acctgagcgt gacctctgtg gctgctacag cgtgtccagt gtcctgccgg gctgtgccga 1080
gccatggaac catgggaaga ccttcacttg cactgctgcc taccccgagt ccaagacccc 1140
gctaaccgcc accctctcaa aatccggaaa cacattccgg cccgaggtcc acctgctgcc 1200
gccgccgtcg gaggagctgg ccctgaacga gctggtgacg ctgacgtgcc tggcacgtgg 1260
cttcagcccc aaggacgtgc tggttcgctg gctgcagggg tcacaggagc tgccccgcga 1320
gaagtacctg acttgggcat cccggcagga gcccagccag ggcaccacca ccttcgctgt 1380
gaccagcata ctgcgcgtgg cagccgagga ctggaagaag ggggacacct tctcctgcat 1440
ggtgggccac gaggccctgc cgctggcctt cacacagaag accatcgacc gcttggcggg 1500
taaacccacc catgtcaatg tgtctgttgt catggcggag gtggacggca cctgctactga 1560 ....
```



The IMGT-ONTOLOGY concepts

http://imgt.cines.fr

IDENTIFICATION

IMGT/LIGM-DB Consultation module v3 - Netscape

Fichier Edition Afficher Aller Communicator Aide

```
ID   HSCOMFRAA 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.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;
OC   Catarrhini; Hominidae; 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   EMBL; L39956.
XX
FH   Key                 Location/Qualifiers
FH
FT   L-V-D-J-C-SEQUENCE  <1..375>
FT                       /partial
```

Document : chargé

The IMGT-ONTOLOGY concepts

http://imgt.cines.fr

DESCRIPTION

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..375>
FT                        /partial
FT                        /db_xref="taxon:9606"
FT                        /cell_type="B-cell hybridoma 2F7"
FT                        /IMGT_note="automatically annotated with IMGT tools"
FT                        /organism="Homo sapiens"
FT   V-D-J-REGION         1..375
FT                        /translation="QVHLVESGGAVFHPGRSLRLSRAASGFTFSSYGMHWVRQAP
FT                        AKGLEWVAVIWYDGSNKYYADSVKGRFTISRDNSKNTLYL
FT                        AKHVTIAAAGRRGAGMDVWGQGTTVTVSS"
FT   V-REGION             1..296
FT                        /allele="IGHV3-33*01, putative"
FT                        /gene="IGHV3-33"
FT                        /CDR_length="[8,8,18]"
FT                        /putative_limit="3' side"
FT                        /translation="QVHLVESGGAVFHPGRSLRLSRA
FT                        AKGLEWVAVIWYDGSNKYYADSVKGRFTISRDNSKNT
FT                        AK"
FT   FR1-IMGT             1..75
FT                        /AA_IMGT="1 to 26, AA 10 is missing"
FT                        /translation="QVHLVESGGAVFHPGRSLRLSRAAS"
FT   CDR1-IMGT            76..99
FT                        /AA_IMGT="27 to 34"
FT                        /translation="GFTFSSYG"
FT   FR2-IMGT             100..150
FT                        /AA_IMGT="39 to 55"
FT                        /translation="MHWVRQAPAKGLEWVAV"
FT   CONSERVED-TRP        106..108
FT   CDR2-IMGT            151..174
FT                        /AA_IMGT="56 to 63"
FT                        /translation="IWYDGSNK"
FT   FR3-IMGT             175..288
FT                        /AA_IMGT="66 to 104, AA 73 is missing"
FT                        /translation="YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYC"
```

Done

OBTENTION

CLASSIFICATION

NUMEROTATION

IMGT/V-QUEST

IMGT/V-QUEST

http://imgt.cines.fr

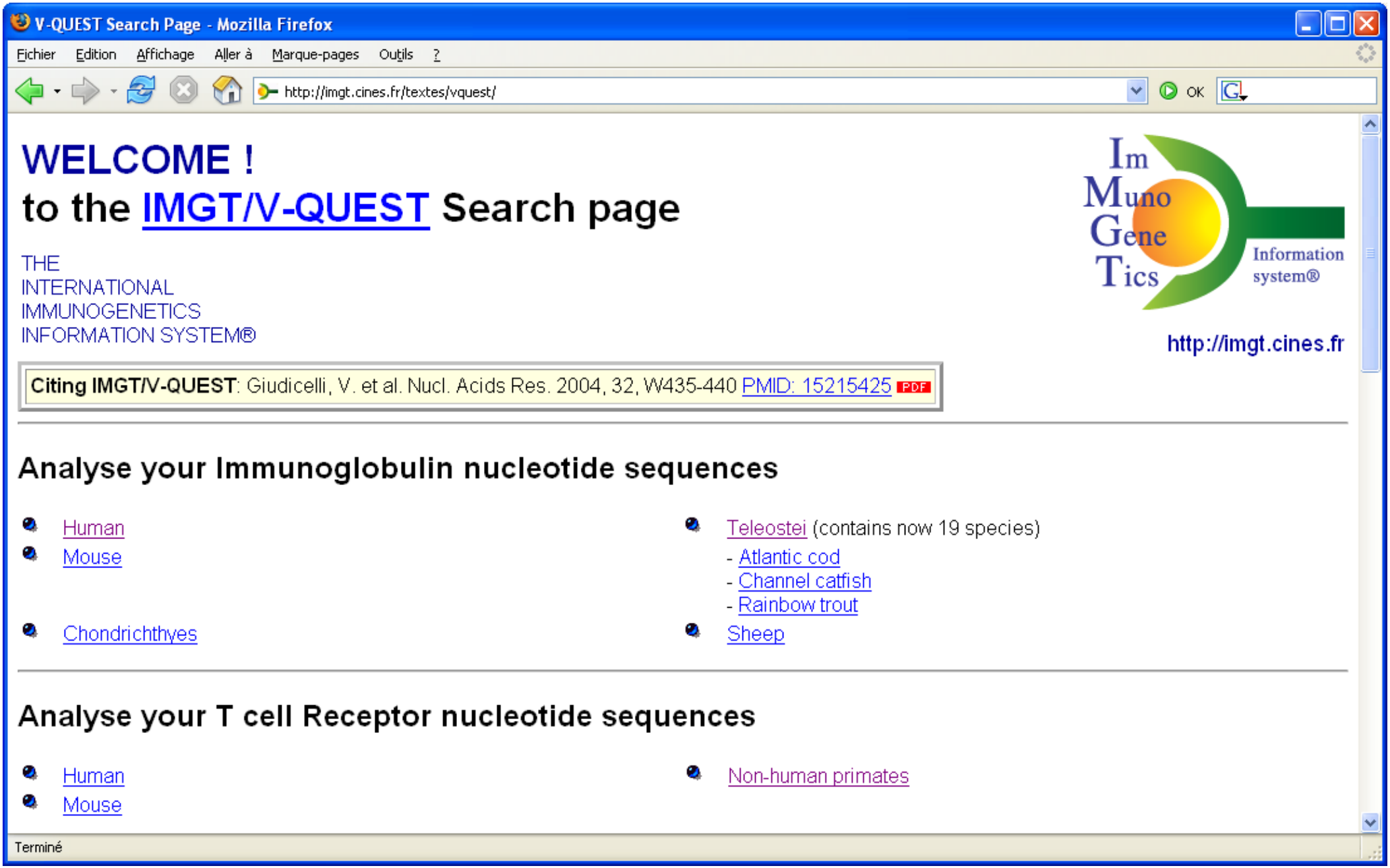
Alignment for V-GENE

```
AF402940                    score GTGCAGCTGCTCGAGCAGTCTGGGGCT___GAGGTGAGCAAGCCTGGGGCCTCAGTAAAGGTTTCCTGCA
X62109   IGHV1-3*01          1146 CA.GTC.A...T.T............___.......AG.................G.............
X62107   IGHV1-3*02          1110 CA.GTT.A...G.T............___.......AG.................G.............
M99637   IGHV1-8*01           957 CA.GT..A...G.T............___.......AG.................G.....C.......
L06612   IGHV1-46*03          948 CA.GT..A...G.T............___.......AG.................G.............
X92343   IGHV1-46*01          948 CA.GT..A...G.T............___.......AG.................G.............
```

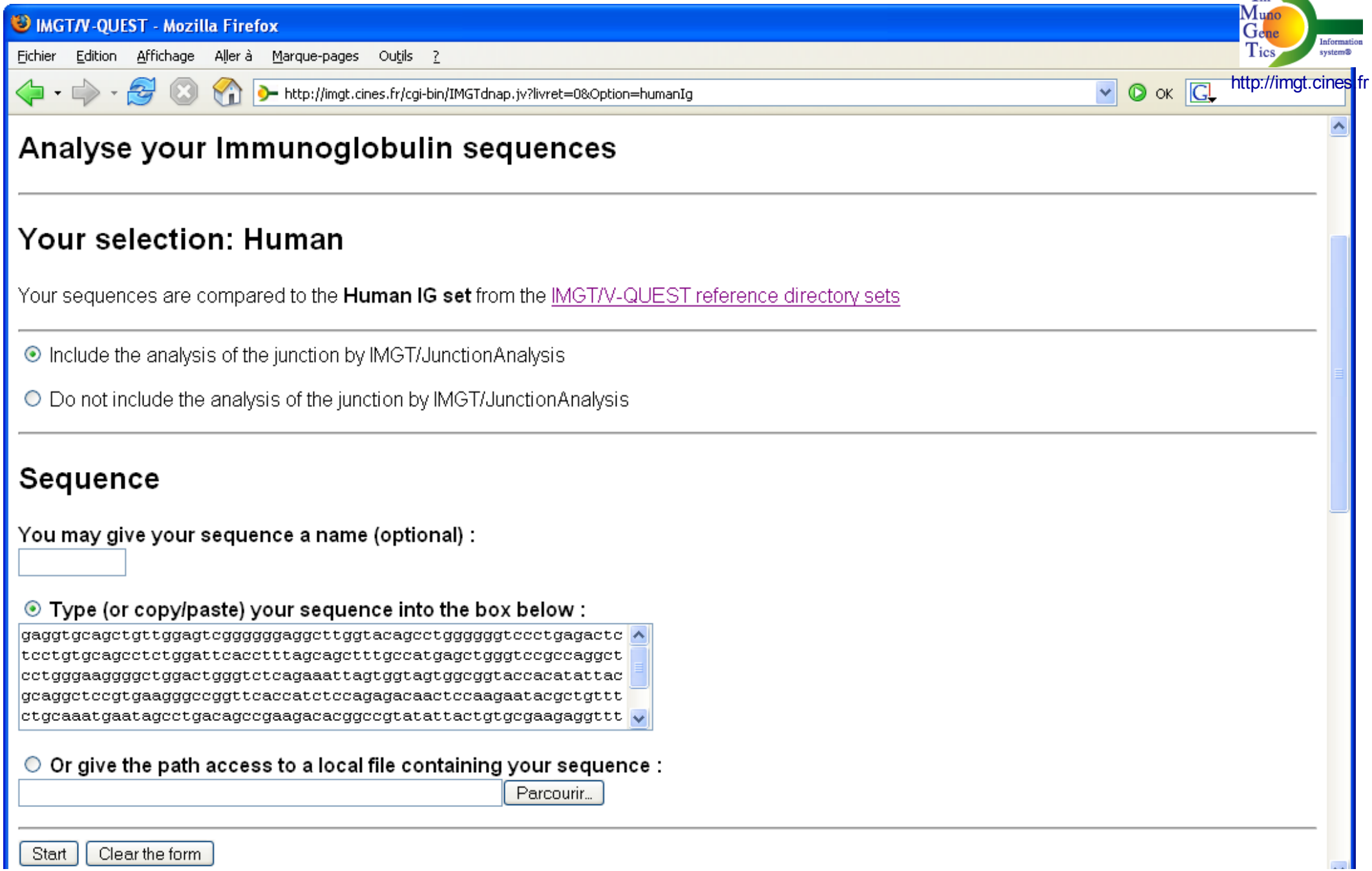
Alignment for J-GENE

```
AF402948                    score CTTCACGGGGCGGGACGCTTTGGACGTCTGGGGCCAAGGGACCACGGTCACCGTCTCCTCA
J00256   IGHJ3*01             181 _________________.T.....T..T.................A.T...........T...G
X86355   IGHJ6*02             179 T.A.TACTACTACT...G.A.....................................
X86355   IGHJ3*02             172 _________________.T.....T..TA................A.T...........T...G
```

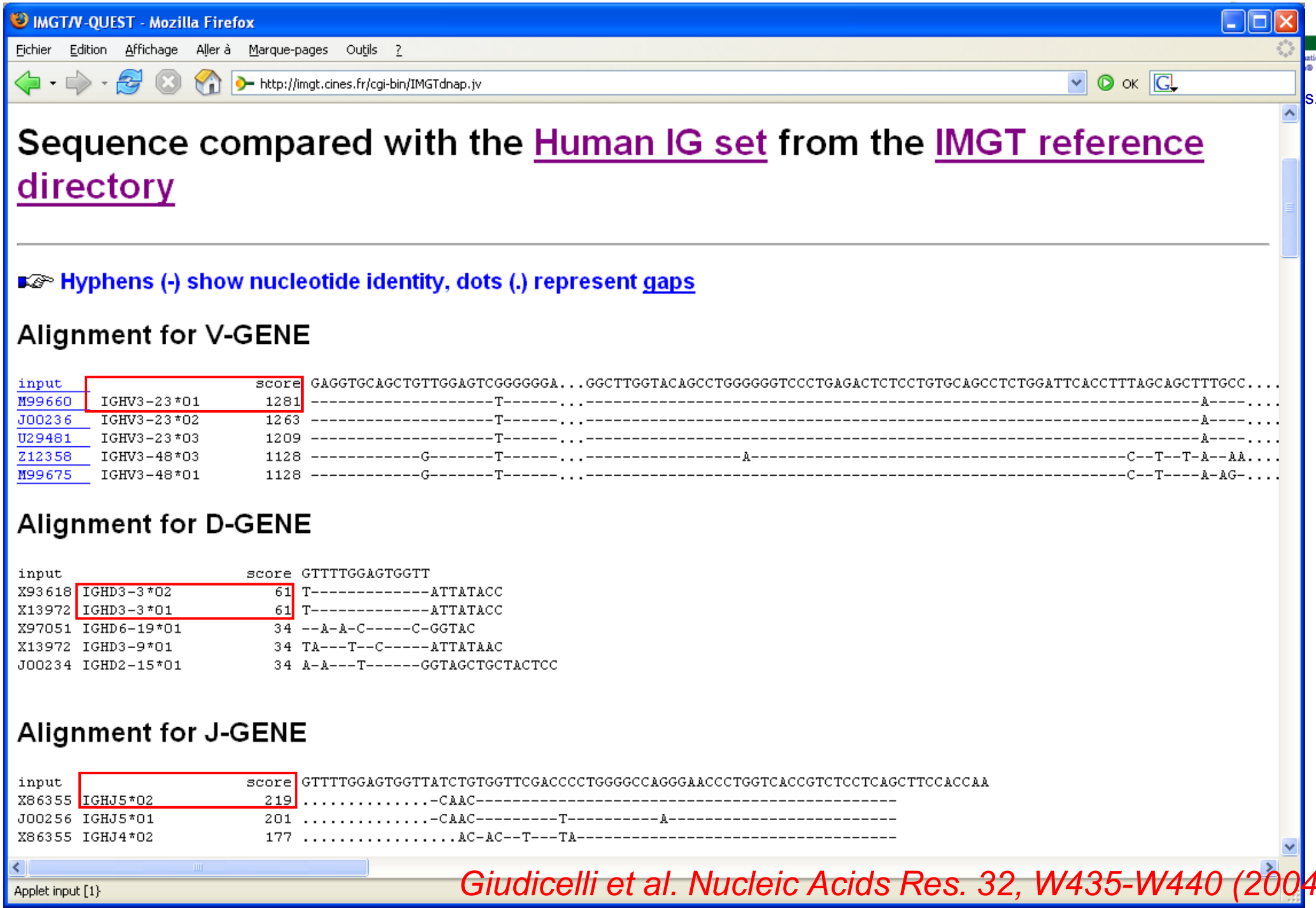
Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



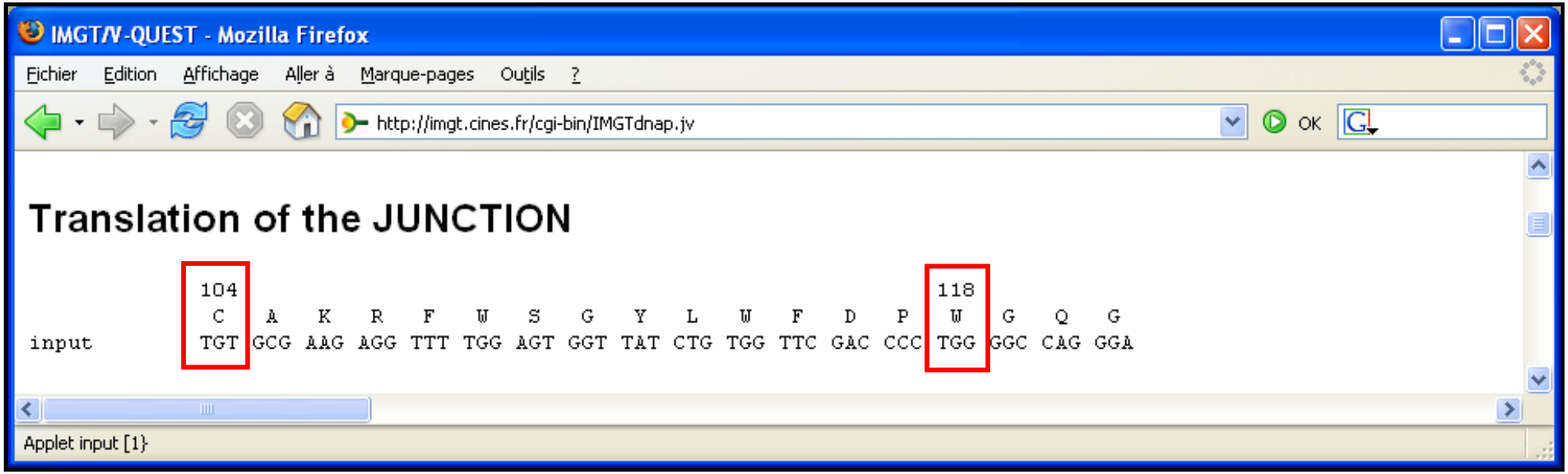
Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



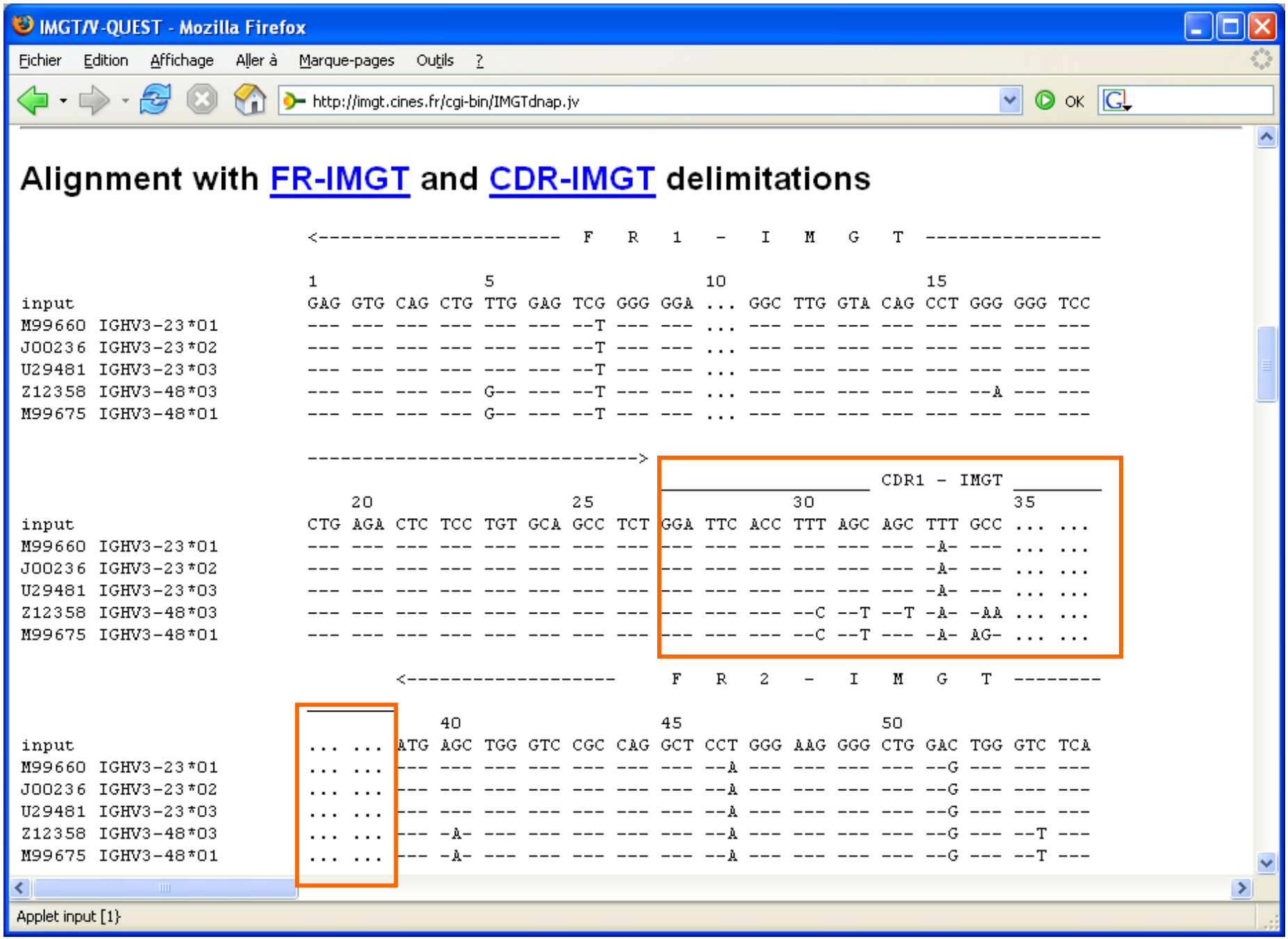
Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)

IMGT/V-QUEST

http://imgt.cines.fr

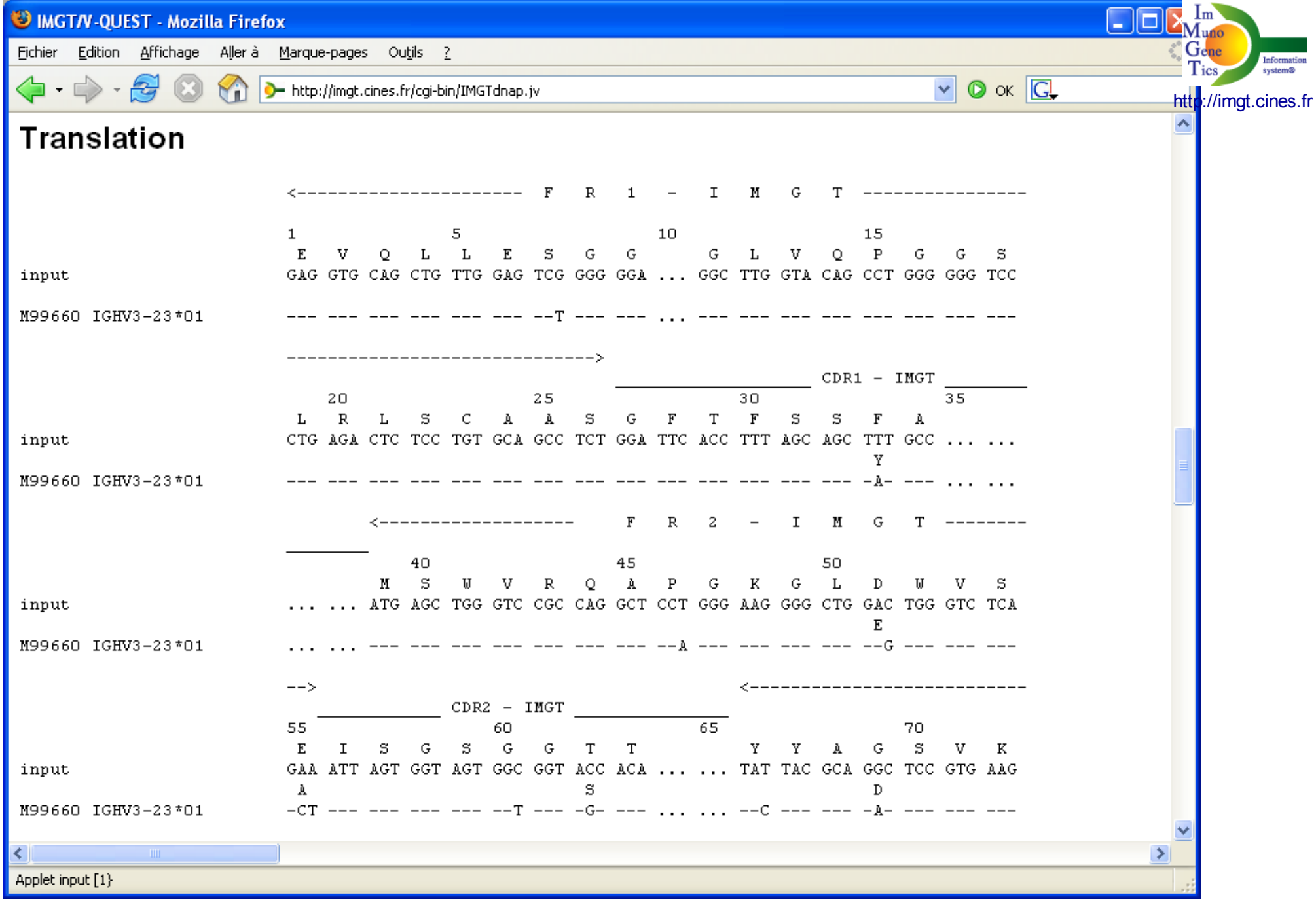


Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



http://imgt.cines.fr

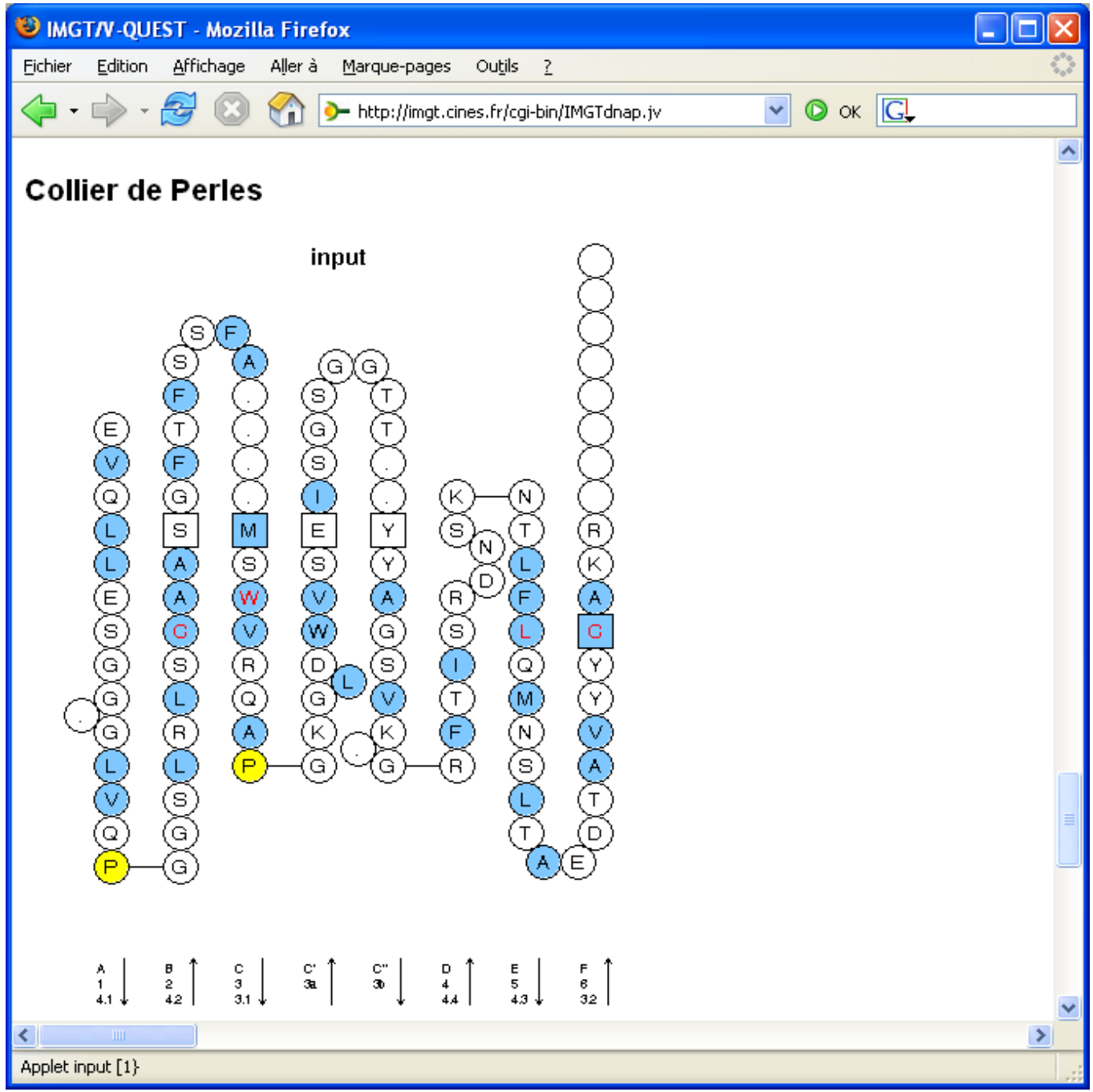
Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



Im Muno Gene Tics Information system®

http://imgt.cines.fr

Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



Im Muno Gene Tics Information system®
http://imgt.cines.fr

Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)

IMGT/V-QUEST

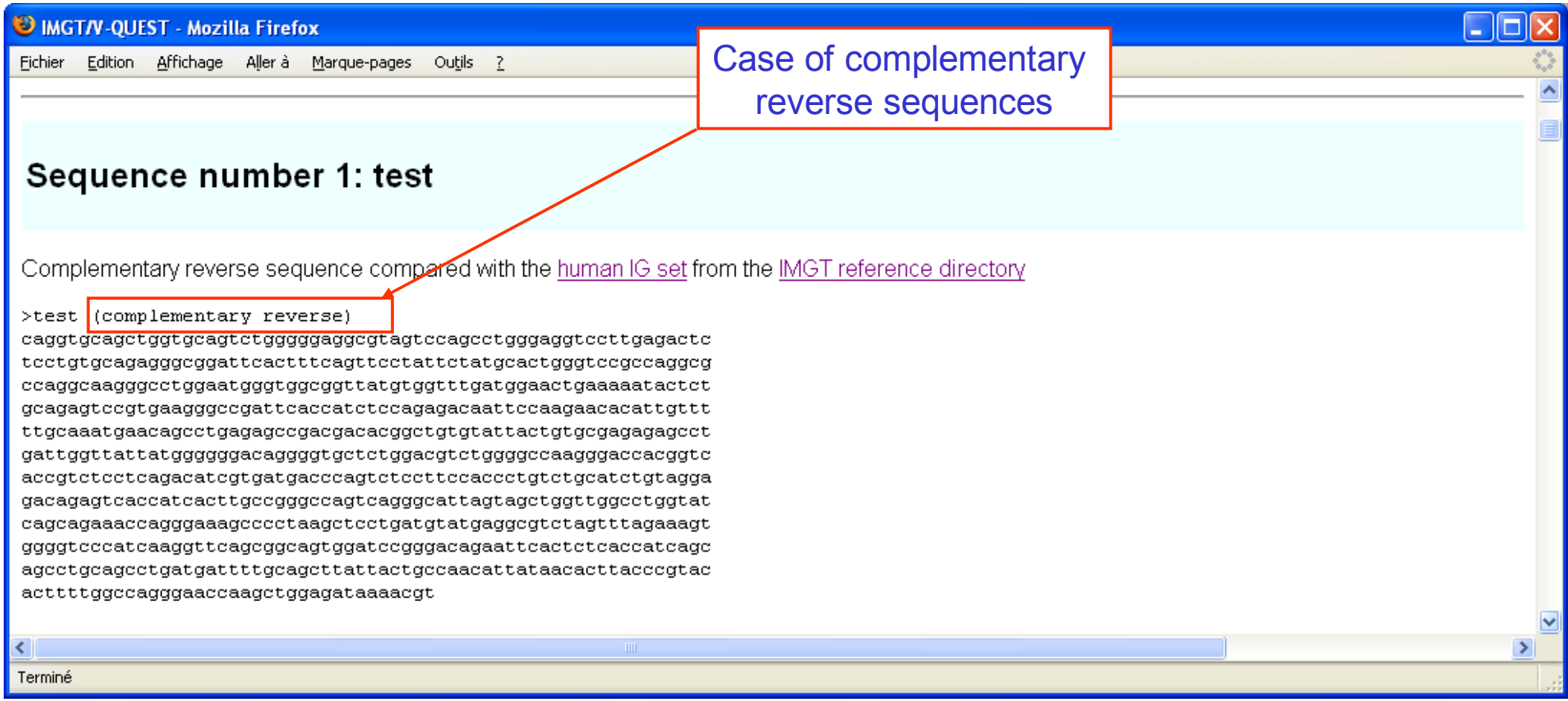
Some improvements for the next release

Analysis of several sequences (in FASTA format) in a single run

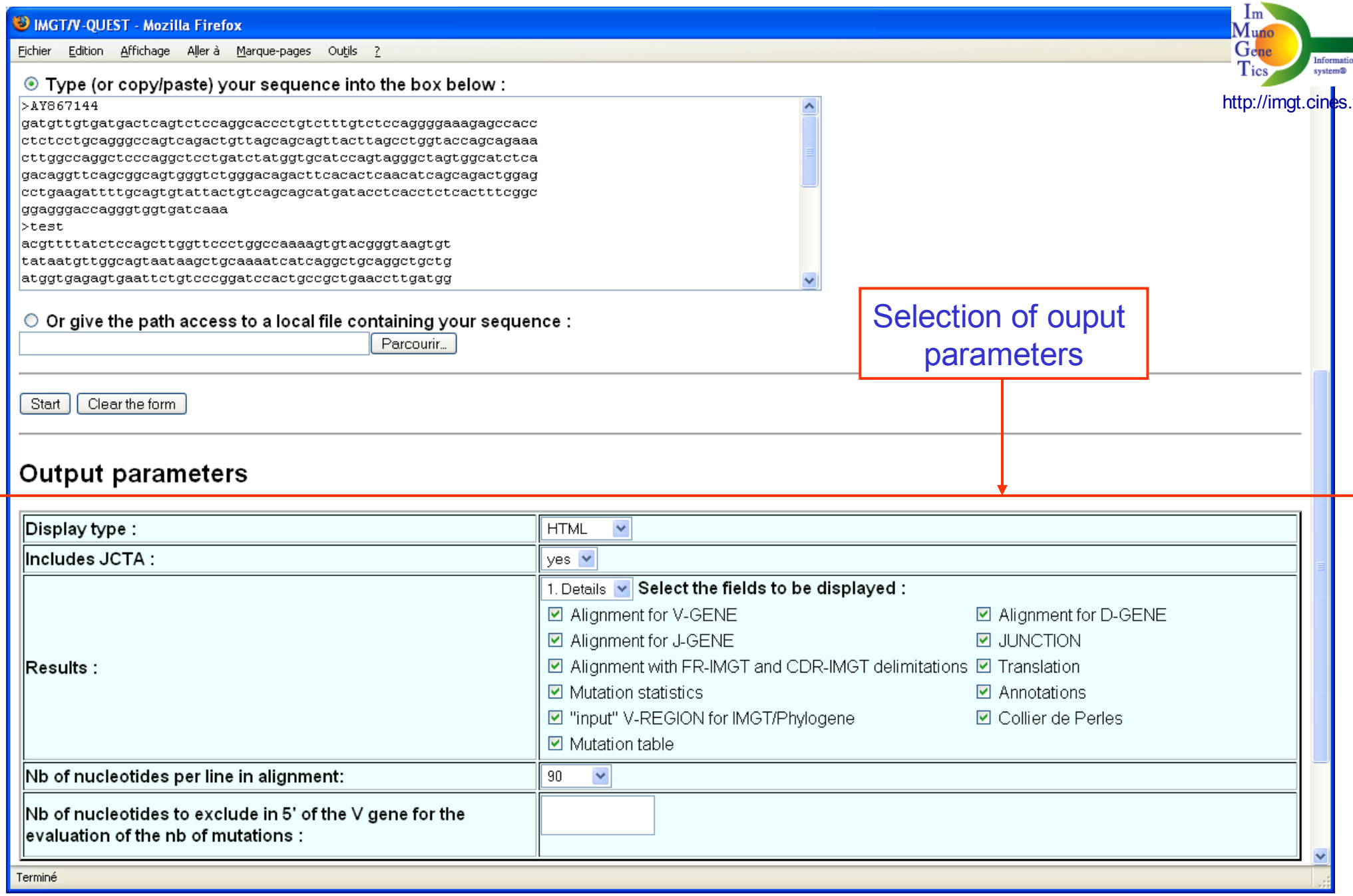
Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)

IMGT/V-QUEST

http://imgt.cines.fr



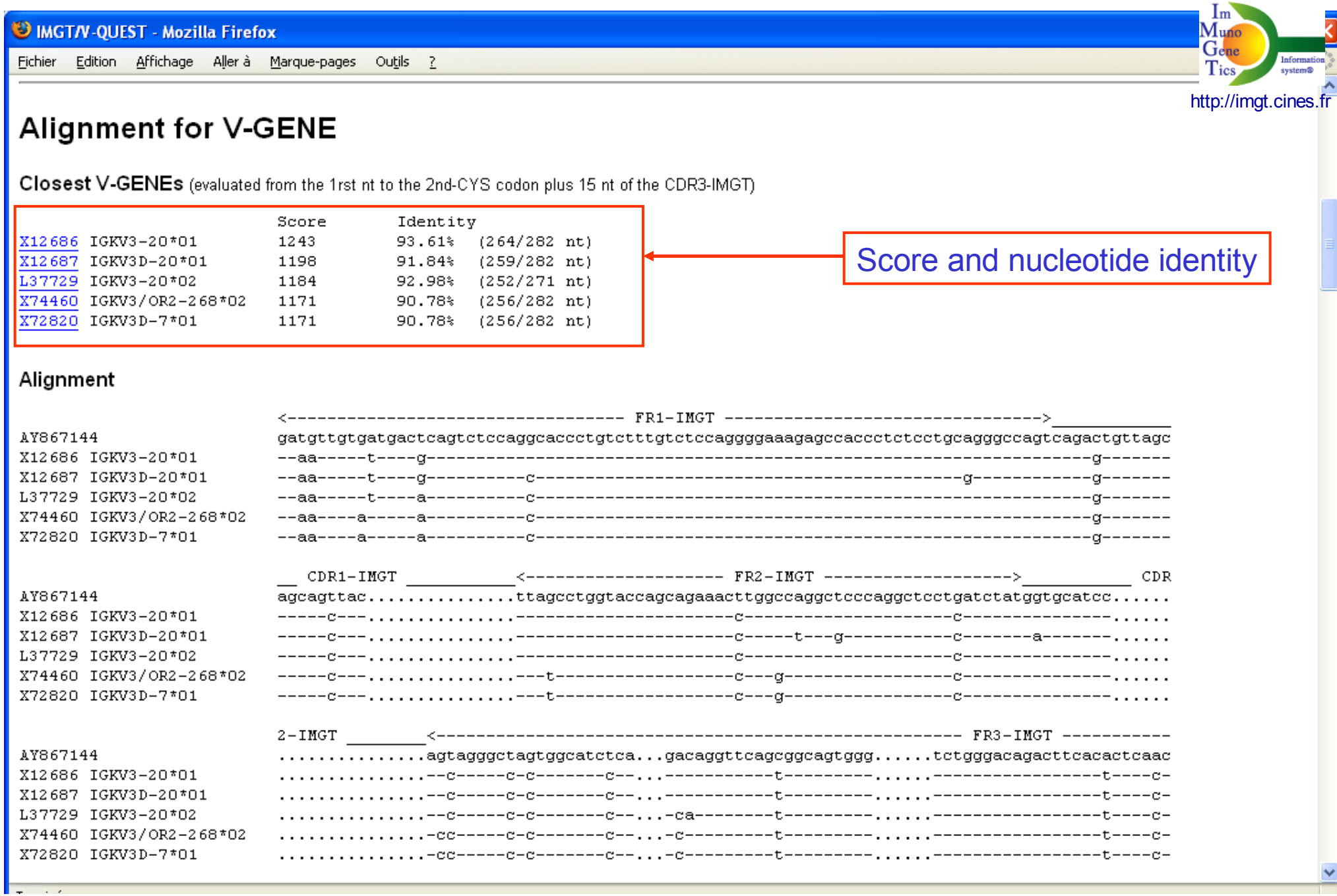
Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



Im Muno Gene Tics Information system®

http://imgt.cines.fr

Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)



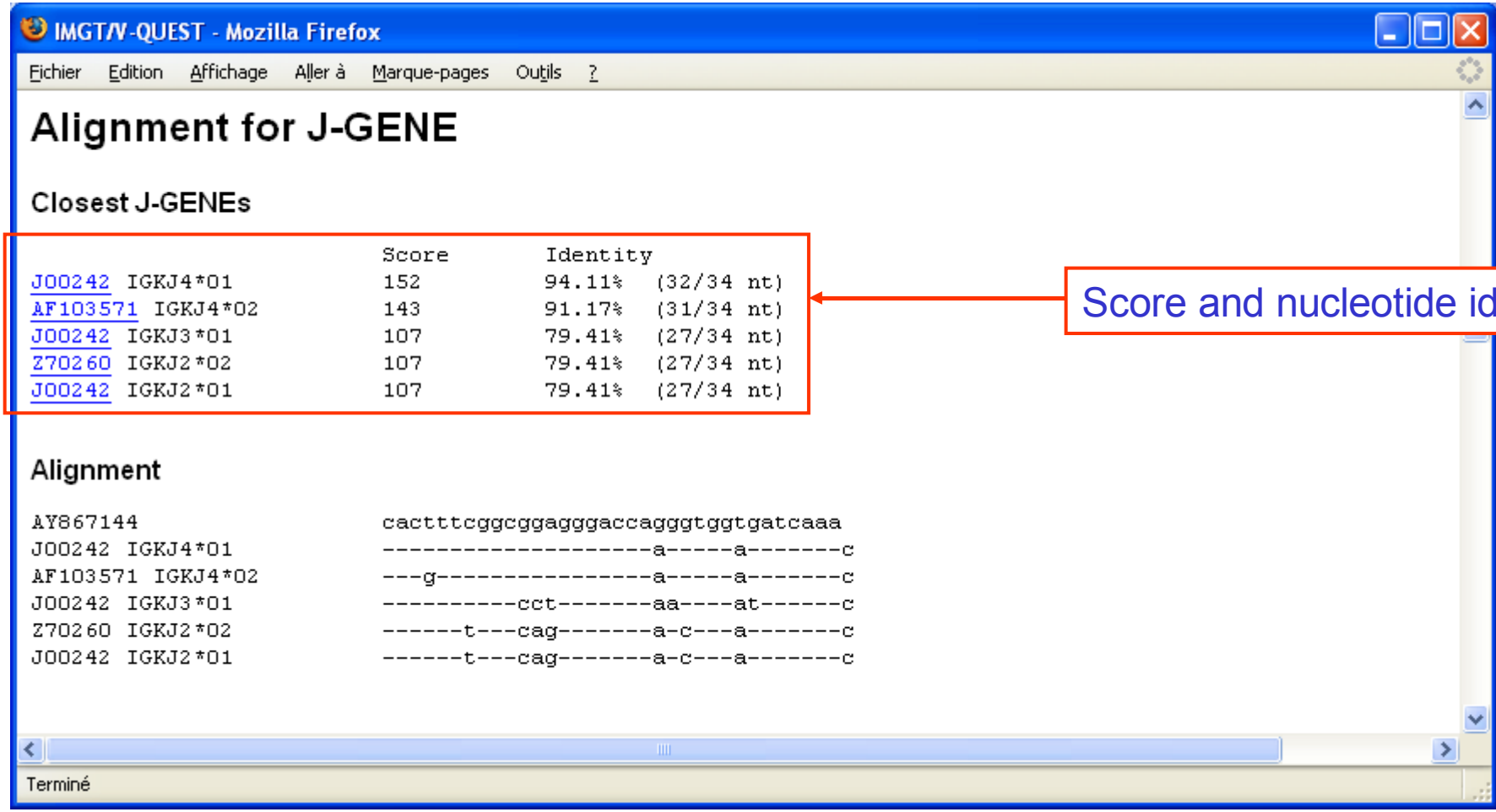
		Score	Identity	
X12686	IGKV3-20*01	1243	93.61%	(264/282 nt)
X12687	IGKV3D-20*01	1198	91.84%	(259/282 nt)
L37729	IGKV3-20*02	1184	92.98%	(252/271 nt)
X74460	IGKV3/OR2-268*02	1171	90.78%	(256/282 nt)
X72820	IGKV3D-7*01	1171	90.78%	(256/282 nt)

```
                          <---------------------------------- FR1-IMGT ---------------------------------->___________
AY867144                  gatgttgtgatgactcagtctccaggcaccctgtctttgtctccaggggaaagagccaccctctcctgcagggccagtcagactgttagc
X12686 IGKV3-20*01        --aa-----t----g-------------------------------------------------------------------------g-------
X12687 IGKV3D-20*01       --aa-----t----g----------c-----------------------------------------------g------------g-------
L37729 IGKV3-20*02        --aa-----t----a----------c-------------------------------------------------------------g-------
X74460 IGKV3/OR2-268*02   --aa----a-----a----------c-------------------------------------------------------------g-------
X72820 IGKV3D-7*01        --aa----a-----a----------c-------------------------------------------------------------g-------

                          __ CDR1-IMGT ___________<-------------------- FR2-IMGT ------------------->___________ CDR
AY867144                  agcagttac..............ttagcctggtaccagcagaaacttggccaggctcccaggctcctgatctatggtgcatcc......
X12686 IGKV3-20*01        -----c---..............----------------------c---------------------c---------------......
X12687 IGKV3D-20*01       -----c---..............----------------------c-----t---g----------c-------a-------......
L37729 IGKV3-20*02        -----c---..............----------------------c---------------------c---------------......
X74460 IGKV3/OR2-268*02   -----c---................---t-----------------c---g-----------------c---------------......
X72820 IGKV3D-7*01        -----c---................---t-----------------c---g-----------------c---------------......

                          2-IMGT ________<------------------------------------------------------ FR3-IMGT -----------
AY867144                  ..............agtagggctagtggcatctca...gacaggttcagcggcagtggg......tctgggacagacttcacactcaac
X12686 IGKV3-20*01        ..............--c-----c-c-------c--...----------t---------.....----------------t----c-
X12687 IGKV3D-20*01       ..............--c-----c-c-------c--...----------t---------.....----------------t----c-
L37729 IGKV3-20*02        ..............--c-----c-c-------c--...-ca--------t---------.....----------------t----c-
X74460 IGKV3/OR2-268*02   ..............-cc-----c-c-------c--...-c---------t---------.....----------------t----c-
X72820 IGKV3D-7*01        ..............-cc-----c-c-------c--...-c---------t---------.....----------------t----c-
```

Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)

IMGT/V-QUEST

http://imgt.cines.fr



Score and nucleotide identity

Giudicelli et al. Nucleic Acids Res. 32, W435-W440 (2004)

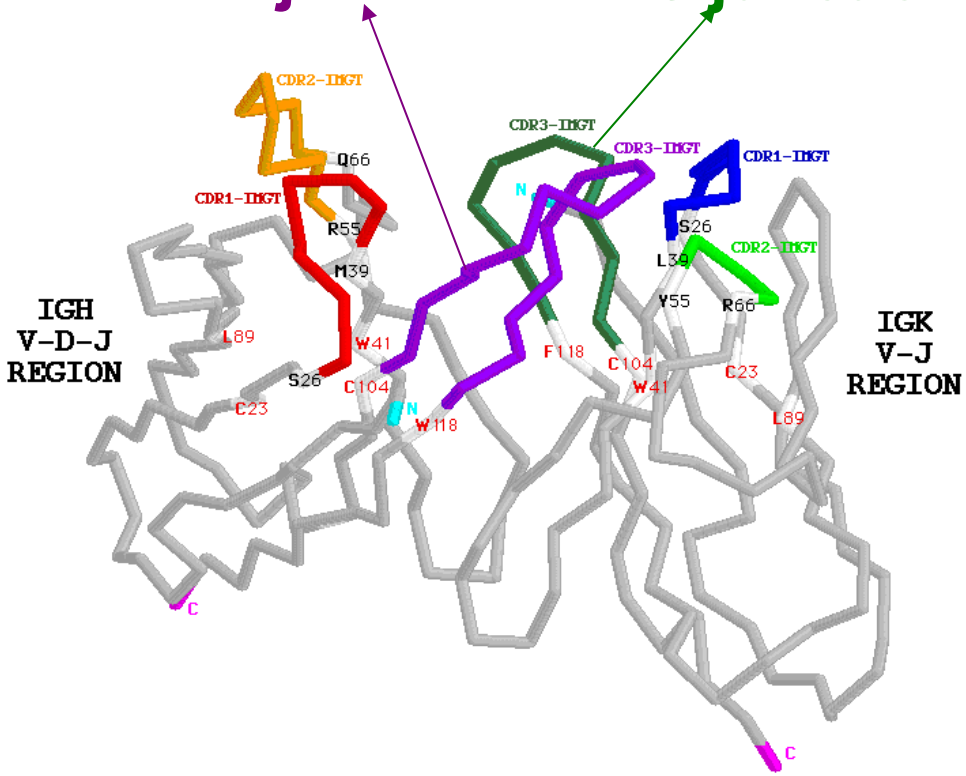
IMGT/JunctionAnalysis

Junctions of the V-DOMAINs

http://imgt.cines.fr

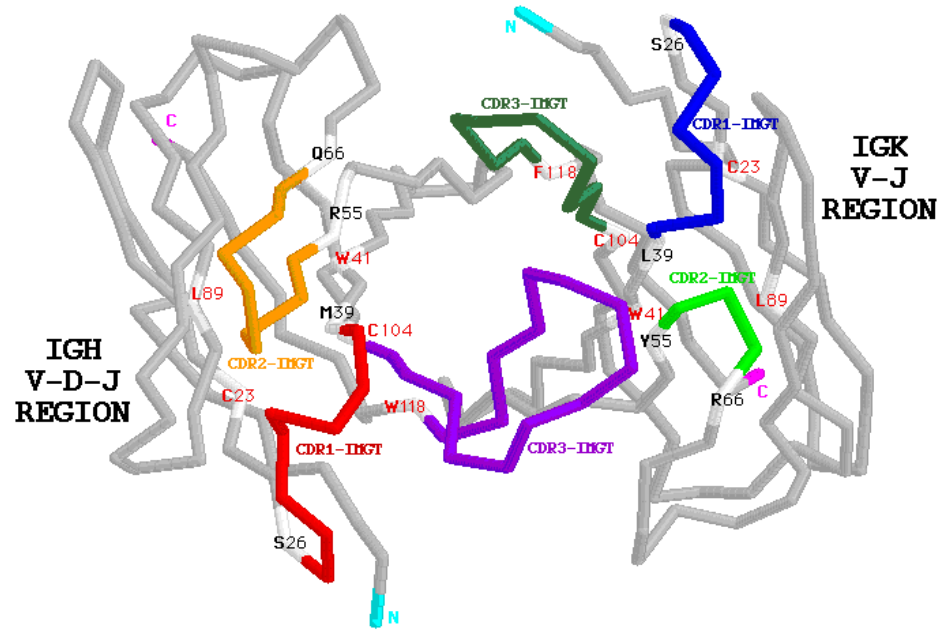
VH V-KAPPA

V-D-J junction V-J junction



Side view

VH V-KAPPA



View from above

Mouse(Mus musculus) E5.2Fv

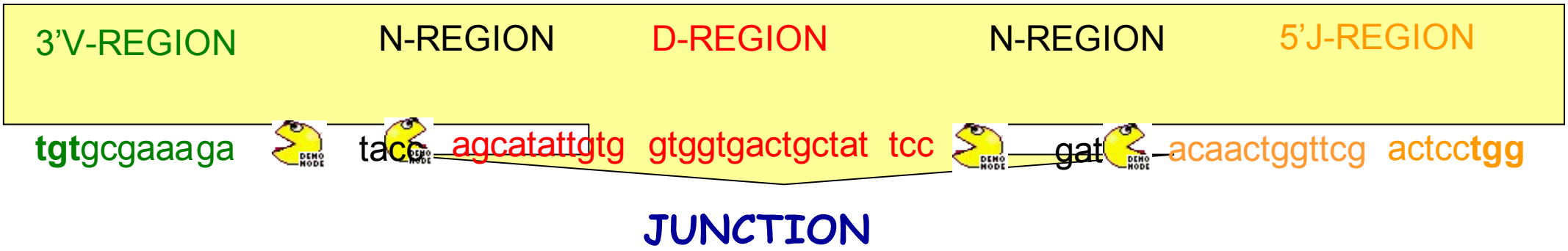
CDR3-IMGT= Complementarity determining region (105-117)

V-J junction (104-118)

V-D-J junction (104-118)

Generation of the JUNCTION diversity

http://imgt.cines.fr



	C	A	P	Y	R	G	D	T	Y	D	Y	S
V	tgt	gcg	cca	tac	cgg	ggt	gac	act	tat	gat	tac	tcc tgg

THANK YOU
for using IMGT/JunctionAnalysis

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

IMGT Information system®

http://imgt.cines.fr

Analysis of the JUNCTIONs

	Input	V name	V-REGION	N1	D-REGION	N2
#1	M62724	IGHV7-4-1*02	tgtgcgagaga	aga	.tagcaatggctacaa....	aata
#2	Z47269	IGHV1-69*06	tgtgcgagag.	ggggggctaagg	....tcgaatttttggagtggtt.....	tcatgggt

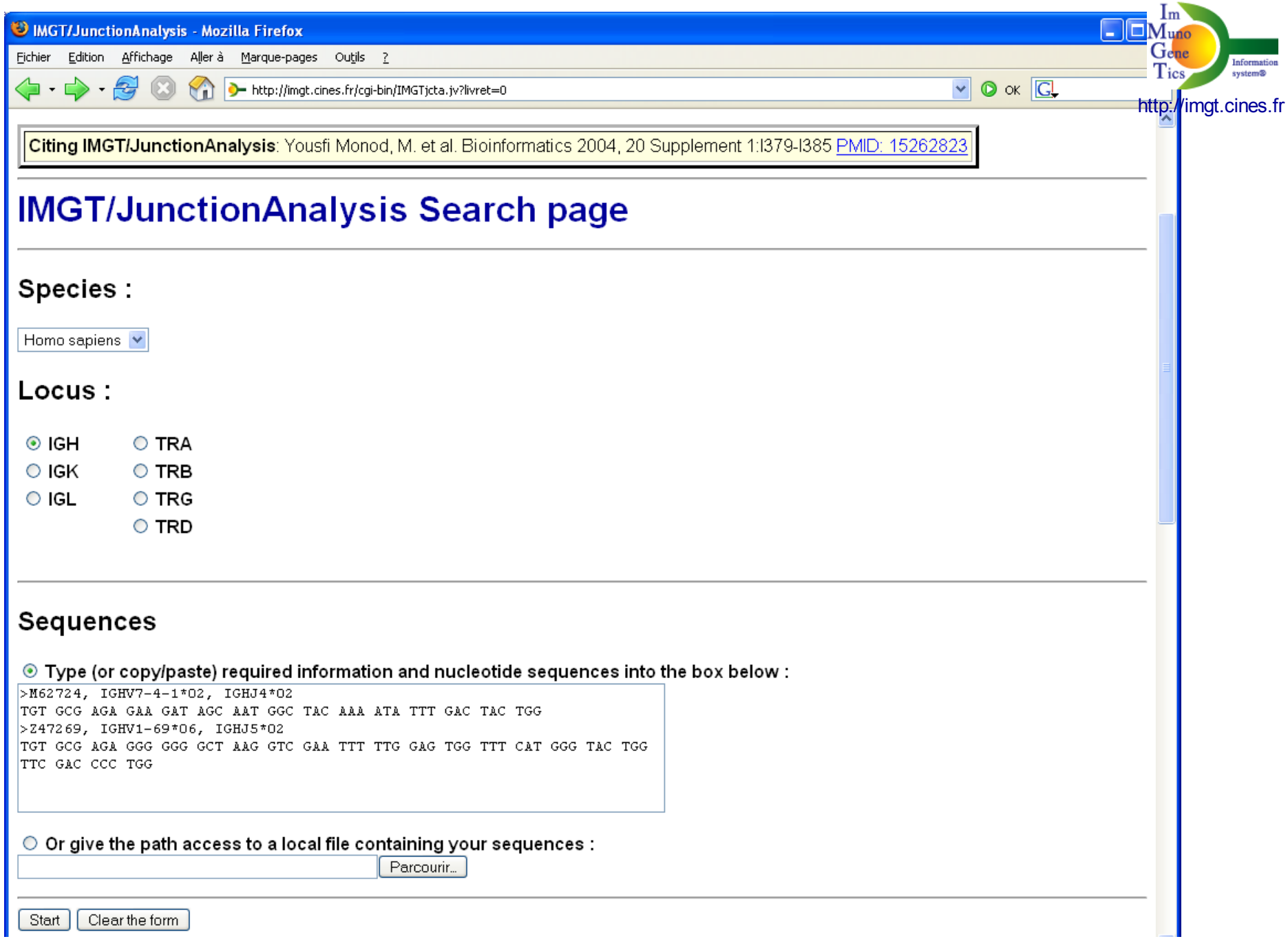
	Input	J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
#1	M62724	.....tttgactactgg	IGHJ4*02	IGHD5-24*01	0	2	0	1/7
#2	Z47269	...actggttcgacccctgg	IGHJ5*02	IGHD3-3*02	0	2	0	13/20

Translation of the JUNCTIONs

		104	105	106	107	108	109	110	111	111.1	111.2	111.3	112.4	112.3	112.2	112.1	112	113	114	115	116	117	118	frame	CDR3-IMGT length
			C	A	R	E	D	S	N	G									Y	K	I	F	D	Y	W		
#1	M62724		tgt	gcg	aga	gaa	gat	agc	aat	ggc									tac	aaa	ata	ttt	gac	tac	tgg	+	13
			C	A	R	G	G	A	K	V	F	F	L	F	W	F	H	G	Y	W	F	D	P	W		
#2	Z47269		tgt	gcg	aga	ggg	ggg	gct	aag	gtc	gaa	ttt	ttg	gag	tgg	ttt	cat	ggg	tac	tgg	ttc	gac	ccc	tgg	+	20

-> IMGT/JunctionAnalysis Search page
-> IMGT/JunctionAnalysis Documentation

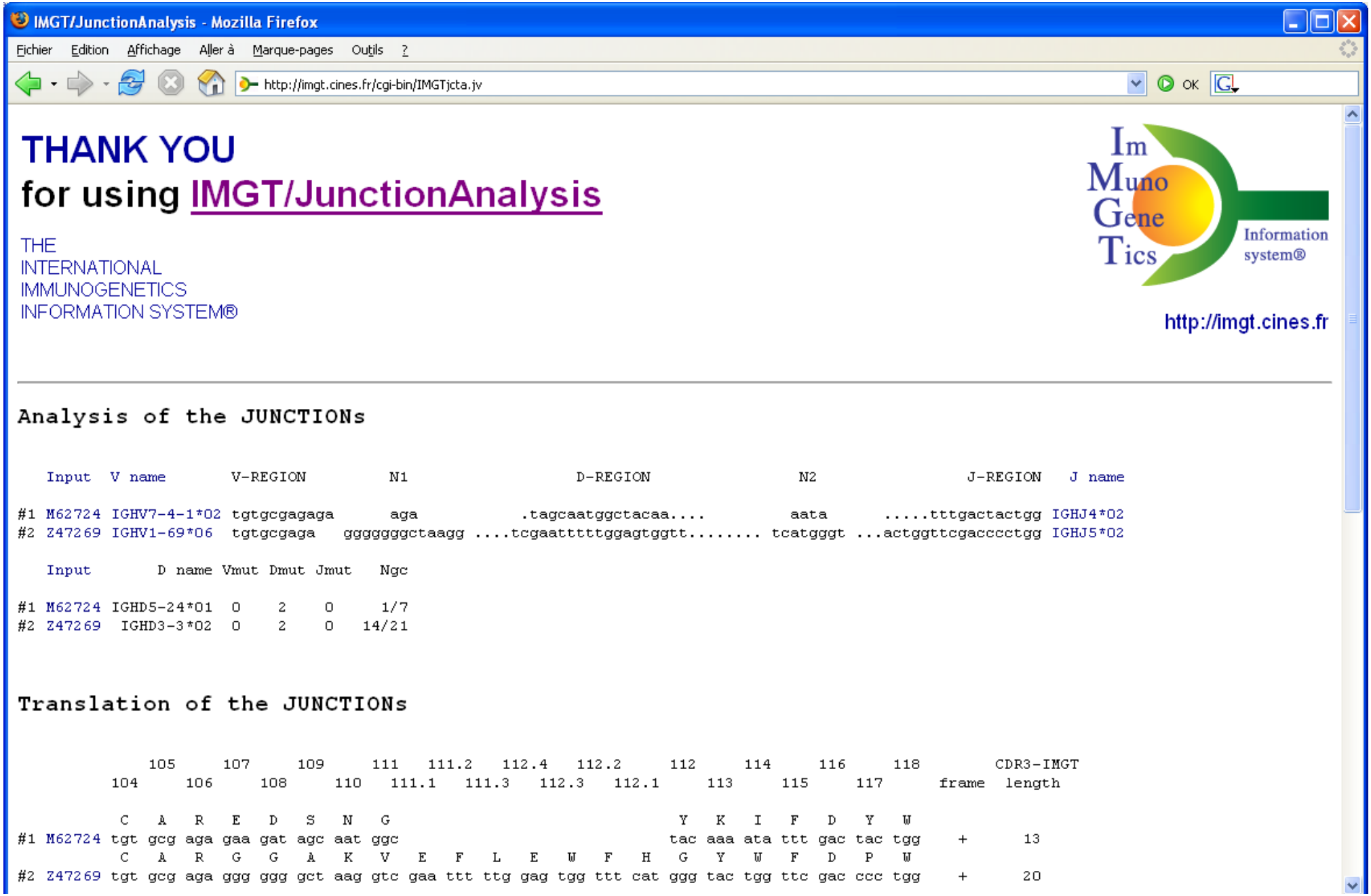
Yousfi Monod et al. Bioinformatics, 20, I379-I385 (2004)



Im Muno Gene Tics Information system®

http://imgt.cines.fr

Yousfi Monod et al. Bioinformatics, 20, I379-I385 (2004)



Yousfi Monod et al. Bioinformatics, 20, I379-I385 (2004)

Im Muno Gene Tics Information system®

http://imgt.cines.fr

IMGT/JunctionAnalysis Results

Locus IGH

Species Homo sapiens

IMGTrepertoire link Locus representation

Maximum number of mutations :
V-REGION : 2; D-REGION : 4; J-REGION : 2
Deletion criterium : Using patterns
Best D gene choice for a same score : Less mutations

Description of the JUNCTIONs

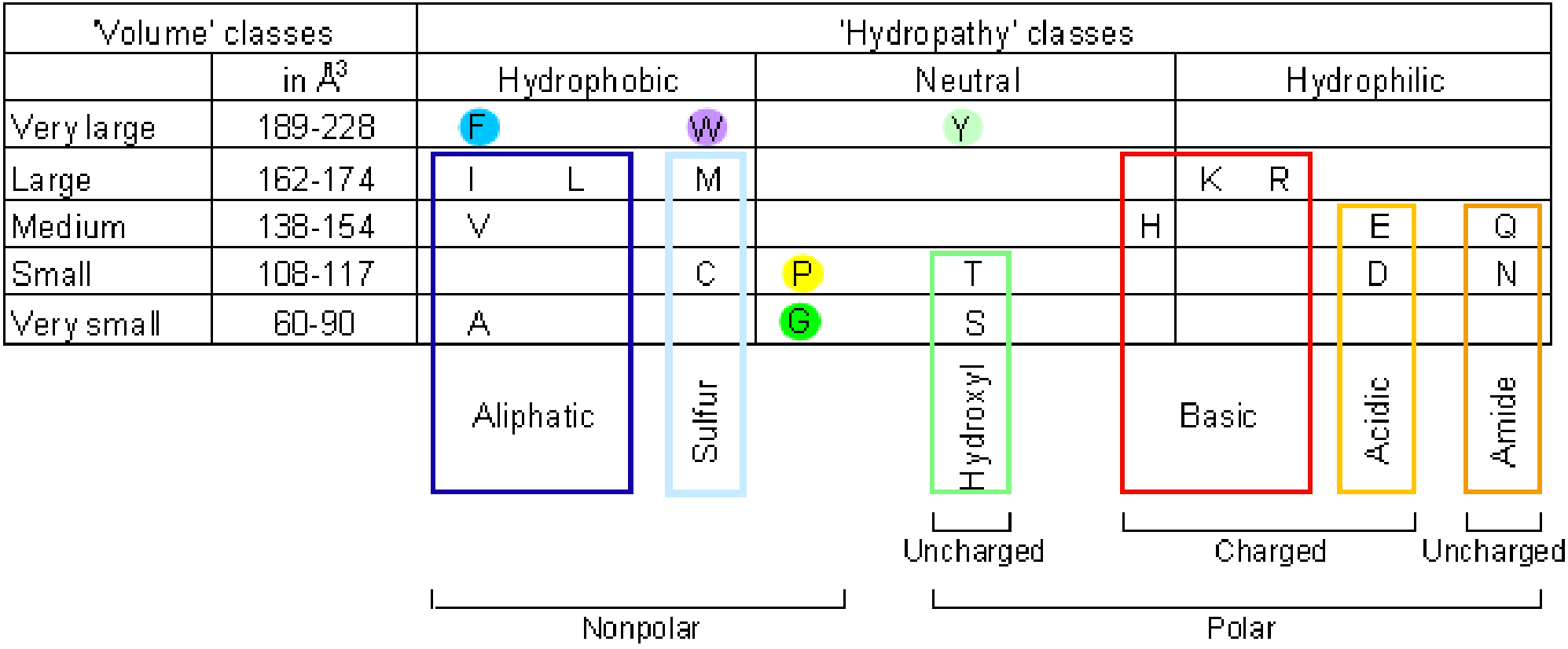
Click on mutated (underscored) nucleotid to see the original one: c

	Input	V name	V-REGION	P	N1	D-REGION	N2	P	J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
#1	Z70256	IGHV2-26*01	tgtgtacg.....		tgttgt	.....gcagcgcctggtac	ccaaatatc		...actttgaccactgg	IGHJ4*02	IGHD6-13*01	1	2	1	5/15
#2	Z70257	IGHV3-7*02	tgtgcgag.		ggatgg	.....cagctcttatgcc	cgccc		ctactggtacttcgatctctgg	IGHJ2*01	IGHD2-2*01	0	2	0	9/11
#3	Z70606	IGHV4-31*03	tgtgcgagag.		c	.gactacg.....	cact		..atgcttttgatgtctgg	IGHJ3*01	IGHD4-17*01	0	0	0	3/5
#4	Z70608	IGHV4-39*05	tgtgc.		cagagta	.....acgatttttggagtggttatt.....	ccccgggggga		..atgcttttgatatctgg	IGHJ3*02	IGHD3-3*01	0	0	0	12/17
#5	Z70610	IGHV4-34*09	tgtgcgagag.		tcgggag	......cgatttttggagtggttatt.....	cccga	ca	tgatgcttttgatatctgg	IGHJ3*02	IGHD3-3*01	0	0	0	9/12
#6	Z70611	IGHV4-59*01	tgtgcgaga..		ca	.....tggtaactataa.	tgccggcgttg		...actggttcgacccctgg	IGHJ5*02	IGHD3-9*01	0	2	0	9/13
#7	Z70613	IGHV4-59*01	tgtgcgagag.			.....cagcagctggtac	ctccct		....ctttgactactgg	IGHJ4*02	IGHD6-13*01	0	0	0	4/6
#8	Z70614	IGHV4-59*01	tgtgcgaga..		cactataa	.....ttcggggacttat.....	cccctc		.....gactactgg	IGHJ4*02	IGHD3-16*01	0	2	0	7/14
#9	Z70615	IGHV4-59*01	tgtgcgagag.		ggctg	gtaaagaggg.....	tttcggaa		.tactggtacttcgatctctgg	IGHJ2*01	IGHD5-24*01	0	2	0	7/13
#10	Z70616	IGHV4-34*01	tgtgcgagag.		cgg	.....gtttggg.....	ttccc		...actggttcgacccctgg	IGHJ5*02	IGHD3-16*01	0	0	0	6/8
#11	Z70620	IGHV4-30-4*01	tgtgcgagaga		cc	.....ggggcgggatggtt.....	cgg		.gatgcttttgatatctgg	IGHJ3*02	IGHD3-16*01	1	4	0	5/5
#12	Z70621	IGHV4-39*01	tgtgcgagaca		ccacgatttatgg	.....ttcggggagtt.....	tgaccccc		.....ttgactactgg	IGHJ4*02	IGHD3-16*01	0	1	0	12/21
#13	Z70622	IGHV4-39*06	tgtgcgagaga	t	tgccccgctcctgccaaaat	gtattactatggttcgggga.....	tatgtacg		.....tttgactactgg	IGHJ4*03	IGHD3-10*01	0	0	0	15/28

Yousfi Monod et al. Bioinformatics, 20, I379-I385 (2004)

The eleven IMGT amino acid classes according to the physico-chemical properties

Im Muno Gene Tics Information system®
http://imgt.cines.fr



Pommié et al. J. Mol Recognit. ***17****, 17-32, 2004*

IMGT/JunctionAnalysis - Mozilla Firefox
Fichier Edition Affichage Aller à Marque-pages Outils ?

JUNCTION alignments

http://imgt.cines.fr

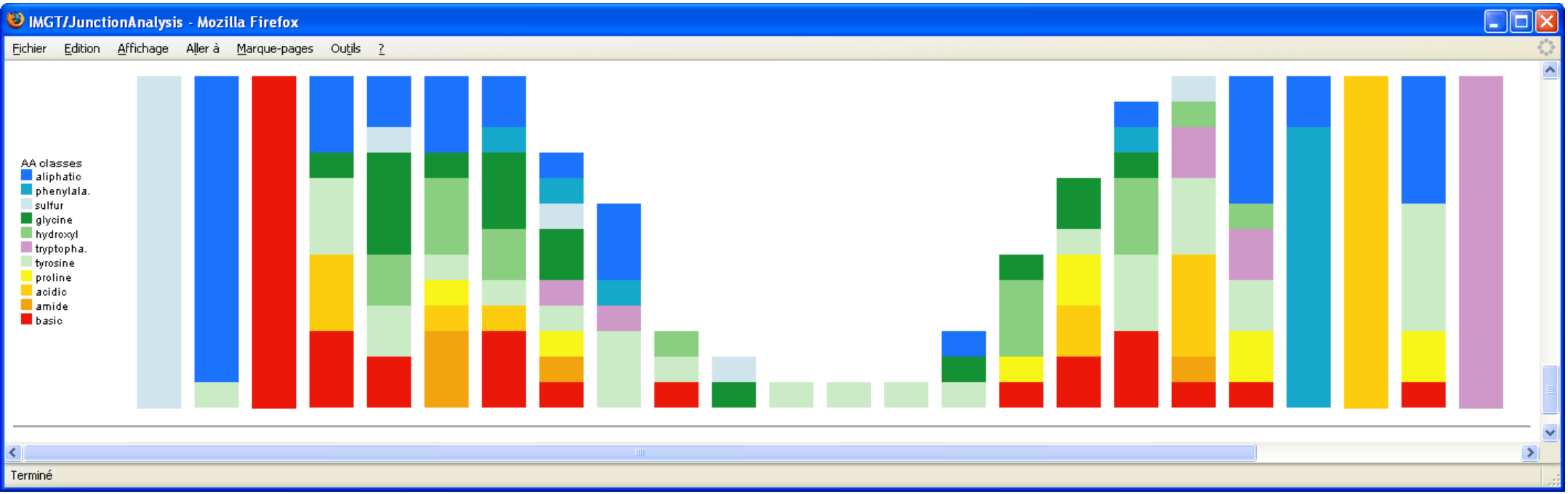
Click on mutated (underscored) amino acid to see the original one:

		104	105	106	107	108	109	110	111	111.1	111.2	111.3	111.4	112.5	112.4	112.3	112.2	112.1	112	113	114	115	116	117	118	Frame	CDR3-IMGT length	Molecular mass
		C	V	R	V	V	Q	R	L	V								P	K	Y	H	F	D	H	W			
#1	Z70256	tgt	gta	cgt	gtt	gtg	cag	cgc	ctg	gta	...	...	...	...	...	...	...	ccc	aaa	tat	cac	ttt	gac	cac	tgg	+	15	2,438.76
		C	A	R	D	G	S	S	Y	A							R	P	Y	W	Y	F	D	L	W			
#2	Z70257	tgt	gcg	agg	gat	ggc	agc	tct	tat	gcc	...	...	...	...	...	...	cgc	ccc	tac	tgg	tac	ttc	gat	ctc	tgg	+	16	2,256.49
		C	A	R	A	T	T	H												Y	A	F	D	V	W			
#3	Z70606	tgt	gcg	aga	gcg	act	acg	cac	...	...	...	...	...	...	...	...	...	...	...	tat	gct	ttt	gat	gtc	tgg	+	11	1,604.77
		C	A	R	V	T	I	F	G	V	V					I	P	R	G	N	A	F	D	I	W			
#4	Z70608	tgt	gcc	aga	gta	acg	att	ttt	gga	gtg	gtt	...	...	...	...	att	ccc	cgg	ggg	aat	gct	ttt	gat	atc	tgg	+	18	2,426.78
		C	A	R	V	G	S	D	F	W	S	G				Y	S	R	H	D	A	F	D	I	W			
#5	Z70610	tgt	gcg	aga	gtc	ggg	agc	gat	ttt	tgg	agt	ggt	...	...	...	tat	tcc	cga	cat	gat	gct	ttt	gat	atc	tgg	+	19	2,539.73
		C	A	R	H	G	N	Y	N	A								G	V	D	W	F	D	P	W			
#6	Z70611	tgt	gcg	aga	cat	ggt	aac	tat	aat	gcc	...	...	...	...	...	...	...	ggc	gtt	gac	tgg	ttc	gac	ccc	tgg	+	15	2,072.21
		C	A	R	A	A	A	G											T	S	L	F	D	Y	W			
#7	Z70613	tgt	gcg	aga	gca	gca	gct	ggt	...	...	...	...	...	...	...	...	...	...	acc	tcc	ctc	ttt	gac	tac	tgg	+	12	1,531.71
		C	A	R	H	Y	N	S	G										T	Y	P	L	D	Y	W			
#8	Z70614	tgt	gcg	aga	cac	tat	aat	tcg	ggg	...	...	...	...	...	...	...	...	...	act	tat	ccc	ctc	gac	tac	tgg	+	13	1,846.01
		C	A	R	G	L	V	K	R	V							S	E	Y	W	Y	F	D	L	W			
#9	Z70615	tgt	gcg	aga	ggg	ctg	gta	aag	agg	gtt	...	...	...	...	...	...	tcg	gaa	tac	tgg	tac	ttc	gat	ctc	tgg	+	16	2,419.75
		C	A	R	A	G	L	G											S	H	W	F	D	P	W			
#10	Z70616	tgt	gcg	aga	gcg	ggt	ttg	ggt	...	...	...	...	...	...	...	...	...	...	tcc	cac	tgg	ttc	gac	ccc	tgg	+	12	1,602.8
		C	A	R	D	R	G	G	M									V	R	D	A	F	D	I	W			
#11	Z70620	tgt	gcg	aga	gac	cgg	ggc	ggg	atg	...	...	...	...	...	...	...	...	gtt	cgg	gat	gct	ttt	gat	atc	tgg	+	14	1,932.17
		C	A	R	H	H	D	L	W	F							G	E	F	D	P	L	D	Y	W			
#12	Z70621	tgt	gcg	aga	cac	cac	gat	tta	tgg	ttc	...	...	...	...	...	...	ggg	gag	ttt	gac	ccc	ctt	gac	tac	tgg	+	16	2,307.53
		C	A	R	D	C	P	A	P	A	K	M	Y	Y	Y	G	S	G	I	C	T	F	D	Y	W			

Terminé

IMGT/JunctionAnalysis

http://imgt.cines.fr



IMGT/3Dstructure-DB and IMGT/StructuralQuery

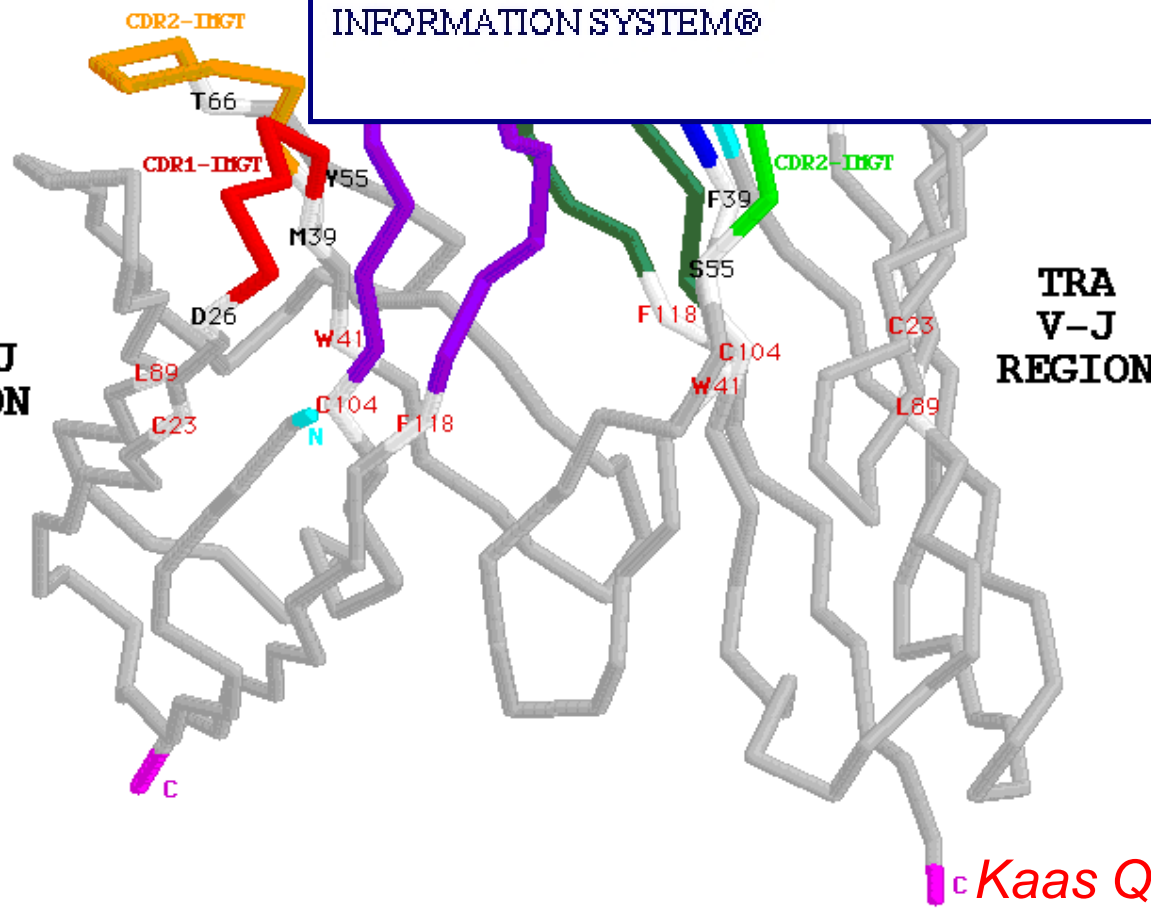
WELCOME !

to **IMGT/3Dstructure-DB**

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

IMGT Information system®

http://imgt.cines.fr



TRB V-D-J REGION

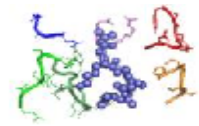
TRA V-J REGION

Kaas Q. et al. NAR 32, D208-D210 (2004)

IMGT/3Dstructure-DB

http://imgt.cines.fr

IMGT/3Dstructure-DB card for : 1u8k



IMGT protein name	IMGT receptor description	Type	Ligand(s)	Species	CC	Chain ID
2F5	FAB-GAMMA-3_KAPPA	IG		*Homo sapiens* (Human)	1	[1u8k_A 1u8k_B]
	Peptide		GP41		1	[1u8k_C]

Experimental technique X-ray diffraction Resolution 2.24 PDB release date 05-OCT-04

Chain ID	1u8k_B
IMGT chain description	VH-CH1
Chain amino acid sequence	[IGHV_1 RITLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAIIYSDDDKRYSPSLNTRLTITKDTSKNQVVLVMTRVSPV][D][IGHJ_1] [IG DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTISSTSTKGPSVFPLAPCSRSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1] GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVNHKPSNTKVDKRVEPKS
V-DOMAIN	IMGT domain description: VH IGHV gene and allele name: IGHV2-5*05 (85.00%), IGHV2-5*06 (85.00%), *Alignment details* IGHJ gene and allele name: IGHJ6*01 (83.33%), IGHJ6*02 (83.33%), *Alignment details* 2D representation: Collier de Perles or Collier de Perles on 2 layers CDR-IMGT lengths: [10.7.24] Sheet composition: [A B D E] [C C" C' F G] [CDR1] [CDR2] RITLKESGP.PLVKPTQTLTLTCSFSGFSLSDFGVG..VGWIRQPPGKALEWLAIIYSDDDK...RYSPSLN.TRLTITKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTI
C-DOMAIN	IMGT domain description: CH1 IGHC gene and allele name: IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%), IGHG3*14 (97.94%), IGHG3*15 (97.94%), IGHG3*16 (97.94%), IGHG3*17 (97.94%), IGHG3*20 (97.94%), *Alignment details* 2D representation: Collier de Perles or Collier de Perles on 2 layers Sheet composition: [A B D E] [C F G] .....STKGPSVFPLAPCSRSTA...GGTAALGCLVKDYFP..EPVTVSWNSGALTS....GVHTFPAVLQSS.....GLYSLSSVVTV PSSSL...GTQTYTCNVNHKP..SNTKVDKRV

Chain ID	1u8k_A
IMGT chain description	L-KAPPA
Chain amino acid sequence	[IGKV_1 ALQLTQSPSSLSASVGDRITITCRASQGVTSALAWYRQKPGSPPQLLIYDASSLESGVPSRFSGSGSGTEFTLTISTLRPEDFATYYCQQ] [IGKJ_1][IGKC_1 LHFYPHTFGGGTRVDVRRTVAAPSVFIFPPSDRQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT] LSKADYEKHKVYECEVTHQGLSSPVTKSFNRGEC
V-DOMAIN	IMGT domain description: V-KAPPA IGKV gene and allele name: IGKV1-13*02 (85.26%) *Alignment details* IGKJ gene and allele name: IGKJ4*01 (63.64%) , IGKJ4*02 (63.64%) *Alignment details* 2D representation: Collier de Perles or Collier de Perles on 2 layers CDR-IMGT lengths: [6.3.9] Sheet composition: [A B D E] [C C" C' F G] [CDR1] [CDR2] ALQLTQSPSSLSASVGDRITITCRASQGVTSA.....LAWYRQKPGSPPQLLIYDAS.....SLESGVP.SRFSGSG..SGTEFTLT [CDR3] ISTLRPEDFATYYCQQLHF....YPHTFGGGTRVDVR
	IMGT domain description: C-KAPPA IGKC gene and allele name: IGKC*01 (98.13%) *Alignment details*

Chain ID	1u8k_C
IMGT chain description	Peptide
Chain amino acid sequence	LELDKWASL

Kaas Q. et al. NAR 32, D208-D210 (2004)

IMGT/3Dstructure-DB

http://imgt.cines.fr

Region identification and delimitation

Chain ID	**1u8k_B**	
IMGT chain description	VH-CH1	
Chain amino acid sequence	[IGHV_1 RITLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAIIYSDDDKRYSPSLNTRLTITKDTSKNQVVLVMTRVSPV][D][IGHJ_1] [IG DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTISSTSTKGPSVFPLAPCS*KSTAGG*TAALGCLVKDYFPEPVTVSWNSGALTS HC_1] GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVNHKPSNTKVDKRVEPKS	
V-DOMAIN	IMGT domain description	VH
	IGHV gene and allele name	IGHV2-5*05 (85.00%), IGHV2-5*06 (85.00%), *Alignment details*
	IGHJ gene and allele name	IGHJ6*01 (83.33%), IGHJ6*02 (83.33%), *Alignment details*
	2D representation	Collier de Perles or Collier de Perles on 2 layers
	CDR-IMGT lengths	[10.7.24]
	Sheet composition	[A B D E] [C C" C' F G]
	[CDR1] [CDR2] RITLKESGP.PLVKPTQTLTLTCSFSGFSLSDFGVG..VGWIRQPPGKALEWLAIIYSDDDK...RYSPSLN.TRLTITKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTI	
	IMGT domain description	CH1
	IGHC gene and allele name	IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%),

Kaas Q. et al. NAR 32, D208-D210 (2004)

IMGT/3Dstructure-DB

http://imgt.cines.fr

Chain ID	**1u8k_B**	
IMGT chain description	VH-CH1	
Chain amino acid sequence	[RITLKESGPPLVKPTQTLTLTCS] [D DTATYFCAHRRGPTTLFGVPIAR HC_1] GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVNHKPSNTKVDKRVEPKS	ITRVSPV IG NSGALTS
V-DOMAIN	IMGT domain description	VH
	IGHV gene and allele name	IGHV2-5*05 (85.00%), IGHV2-5*06 (85.00%), *Alignment details*
	IGHJ gene and allele name	IGHJ6*01 (83.33%), IGHJ6*02 (83.33%), *Alignment details*
	2D representation	Collier de Perles or Collier de Perles on 2 layers
	CDR-IMGT lengths	[10.7.24]
	Sheet composition	[A B D E] [C C" C' F G]
	[CDR1] [CDR2] RITLKESGP.PLVKPTQTLTLTCSFSGFSLSDFGVG..VGWIRQPPGKALEWLAIIYSDDDK...RYSPSLN.TRLTITKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTI	
C-DOMAIN	IMGT domain description	CH1
	IGHC gene and allele name	IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%), IGHG3*14 (97.94%), IGHG3*15 (97.94%), IGHG3*16 (97.94%), IGHG3*17 (97.94%), IGHG3*20 (97.94%), *Alignment details*

Identification of the IMGT gene and the IMGT closest allele(s)

Kaas Q. et al. NAR 32, D208-D210 (2004)

IMGT/3Dstructure-DB

http://imgt.cines.fr

Chain ID	**1u8k_B**	
IMGT chain description	VH-CH1	
Chain amino acid sequence	[IGHV_1 RITLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAI PV][D][IGHJ_1] [IG DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTISSTSTKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1] GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVNHKPSNTKVDKRVEPKS	
V-DOMAIN	IMGT domain description	VH
	IGHV gene and allele name	IGHV2-5*05 (85.00%), IGHV2-5*06 (85.00%
	IGHJ gene and allele name	IGHJ6*01 (83.33%), IGHJ6*02 (83.33%), *Alignment details*
	2D representation	Collier de Perles or Collier de Perles on 2 layers
	CDR-IMGT lengths	[10.7.24]
	Sheet composition	[A B D E] [C C" C' F G]
	[CDR1] RITLKESGP.PLVKPTQTLTLTCSFSGFSLSDFGVG..VGWIRQPPGKALEWLAIIYSDDDK...RYSPSLN.TRLTITKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGITVTI	
	IMGT domain description	CH1
	IGHC gene and allele name	IGHG3*01 (97.94%), IGHG3*03 (97.94%), IGHG3*04 (97.94%), IGHG3*05 (97.94%), IGHG3*06 (97.94%), IGHG3*07 (97.94%), IGHG3*08 (97.94%), IGHG3*09 (97.94%), IGHG3*10 (97.94%), IGHG3*11 (97.94%), IGHG3*12 (97.94%), IGHG3*13 (97.94%),

Description of domains

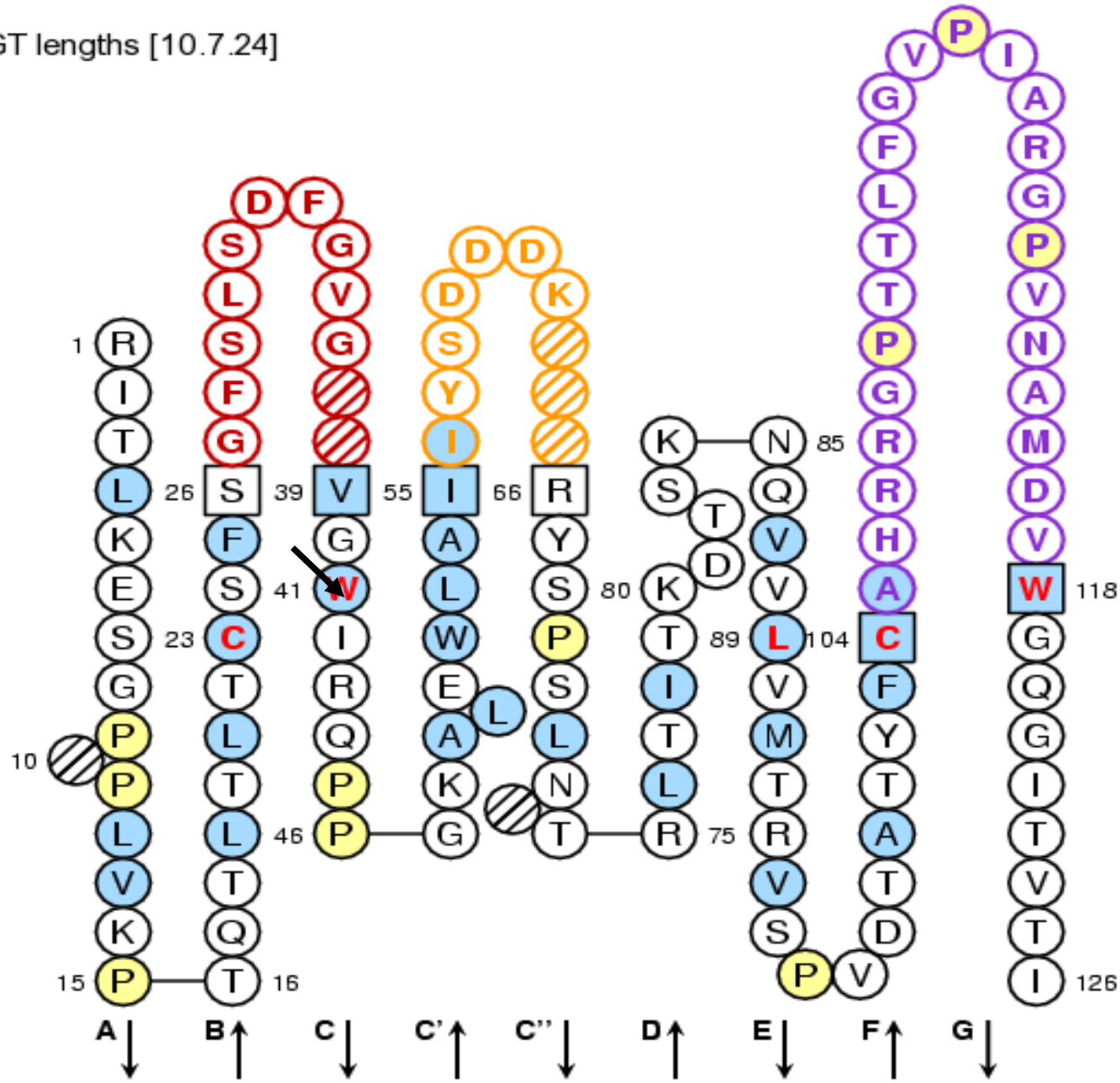
CDR-IMGT lengths

Gaps according to IMGT numbering

Kaas Q. et al. NAR 32, D208-D210 (2004)

IMGT Collier de Perles : *Homo sapiens* (Human) IGHV_1 V-DOMAIN from 2F5 (PDB: 1u8k_B)

CDR-IMGT lengths [10.7.24]



Kaas Q. et al. NAR 32, D208-D210 (2004)

Contact analysis

http://imgt.cines.fr

41V - TRP (W)

chain : 1u8k_B

Secondary structure	Extended conformation
Phi	-122.64
Psi	137.12
ASA	0.0

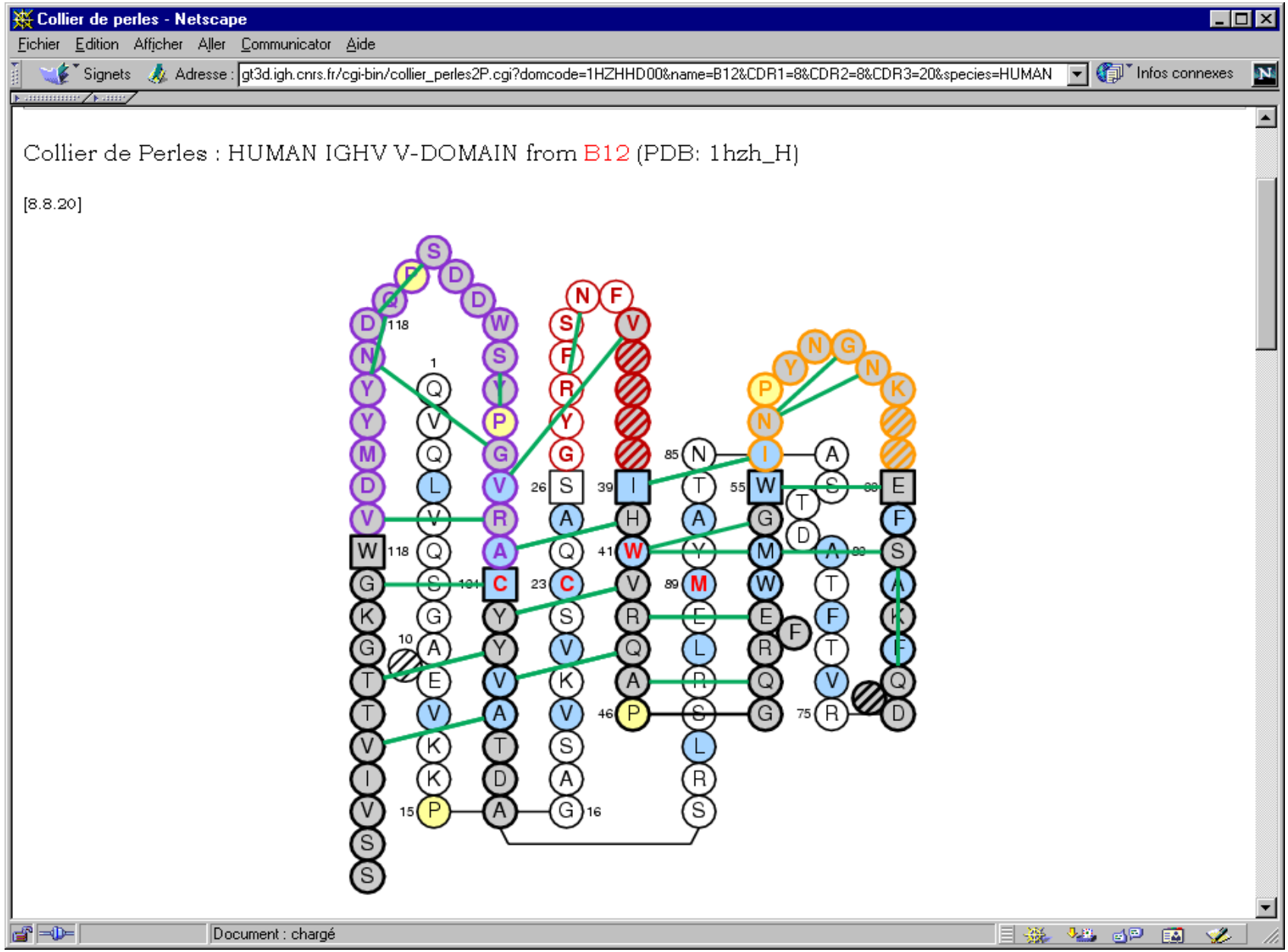
Residue local structure

	IMGT Num	Code 1C	Code 3C	Chain	Domain	Tot	NCov	Pol	HB	NPol	Cov	SS
Info	6V1	E	GLU	1u8k_B	VH	6	6	0	0	6	0	0
Info	21V1	L	LEU	1u8k_B	VH	17	17	0	0	17	0	0
Info	22V1	T	THR	1u8k_B	VH	8	8	0	0	8	0	0
Info	23V1	C	CYS	1u8k_B	VH	10	10	0	0	10	0	0
Info	39V1	V	VAL	1u8k_B	VH	2	2	1	0	1	0	0
Info	43V1	R	ARG	1u8k_B	VH	2	2	1	0	1	0	0
Info	51V1	E	GLU	1u8k_B	VH	2	2	1	0	1	0	0
Info	52V1	W	TRP	1u8k_B	VH	11	11	3	0	8	0	0
Info	53V1	L	LEU	1u8k_B	VH	15	15	1	1	14	0	0
Info	54V1	A	ALA	1u8k_B	VH	18	18	4	2	14	0	0
Info	55V1	I	ILE	1u8k_B	VH	2	2	1	0	1	0	0
Info	78V1	I	ILE	1u8k_B	VH	5	5	0	0	5	0	0
Info	87V1	V	VAL	1u8k_B	VH	11	11	1	0	10	0	0
Info	88V1	V	VAL	1u8k_B	VH	6	6	2	0	4	0	0
Info	89V1	L	LEU	1u8k_B	VH	33	33	1	0	32	0	0
Info	102V1	Y	TYR	1u8k_B	VH	5	5	0	0	5	0	0
Info	103V1	F	PHE	1u8k_B	VH	18	18	2	0	16	0	0
Info	104V1	C	CYS	1u8k_B	VH	26	26	0	0	26	0	0
Info	105V1	A	ALA	1u8k_B	VH	1	1	1	0	0	0	0

Tot	Total number of atomic pair contacts
NCo	Number of non covalent atomic
Pol	Number of polar atomic pair contacts
HB	Number of hydrogen bonds
NPol	Number of non polar atomic pair contacts
Cov	Number of covalent links (other than chain covalent links)
SS	Number of disulfide bridges

Kaas Q. et al. NAR 32, D208-D210 (2004)

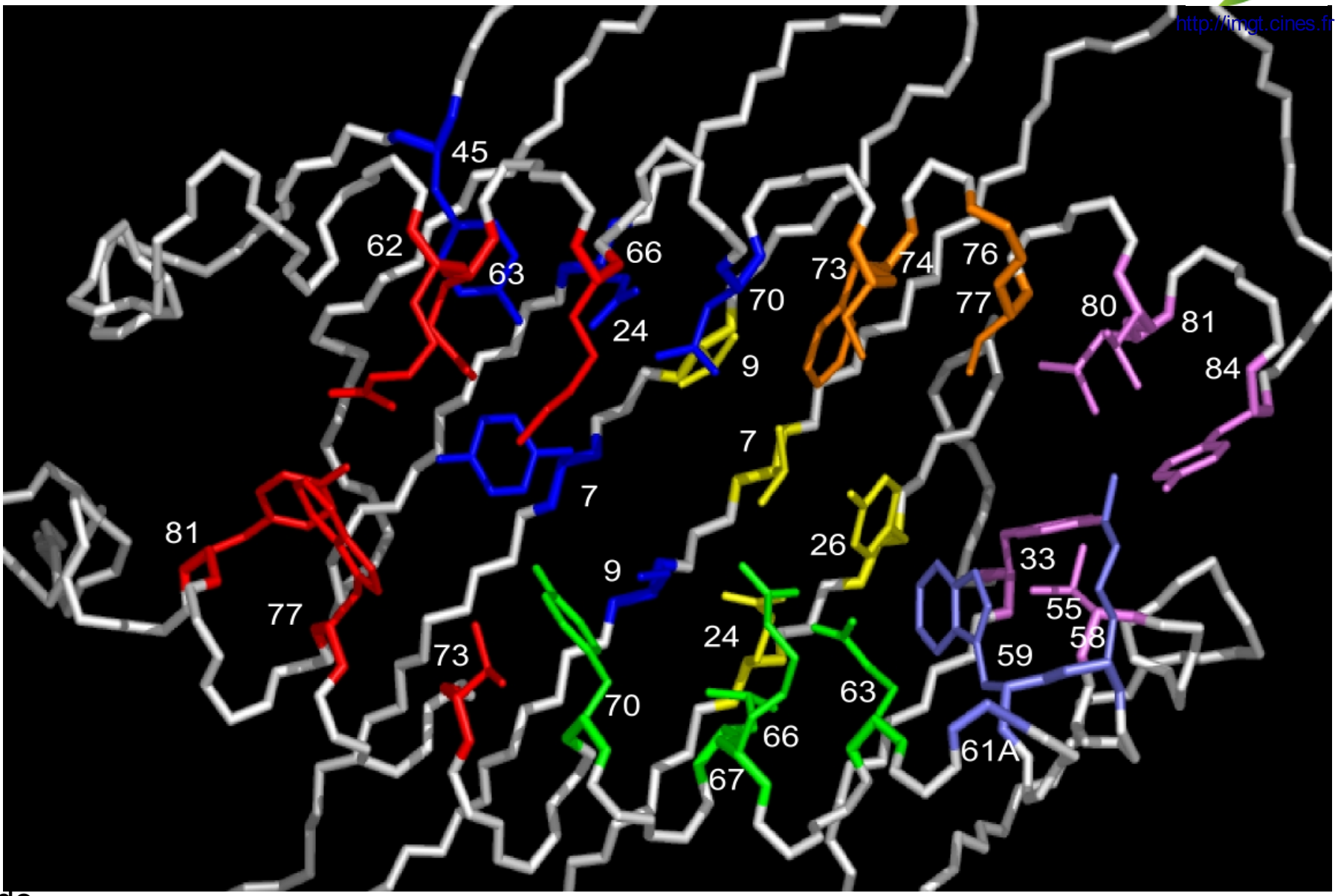
IMGT Collier de Perles



Kaas Q. et al. NAR 32, D208-D210 (2004)

MHC class I IMGT contact sites

http://imgt.cines.fr



C1		1 E
C2		--
C3		2 Q
C4		3 Y
C5		(4 K)
C6		5 F
C7		--
C8		--
C9		6 Y
C10		7 S
C11		8 V

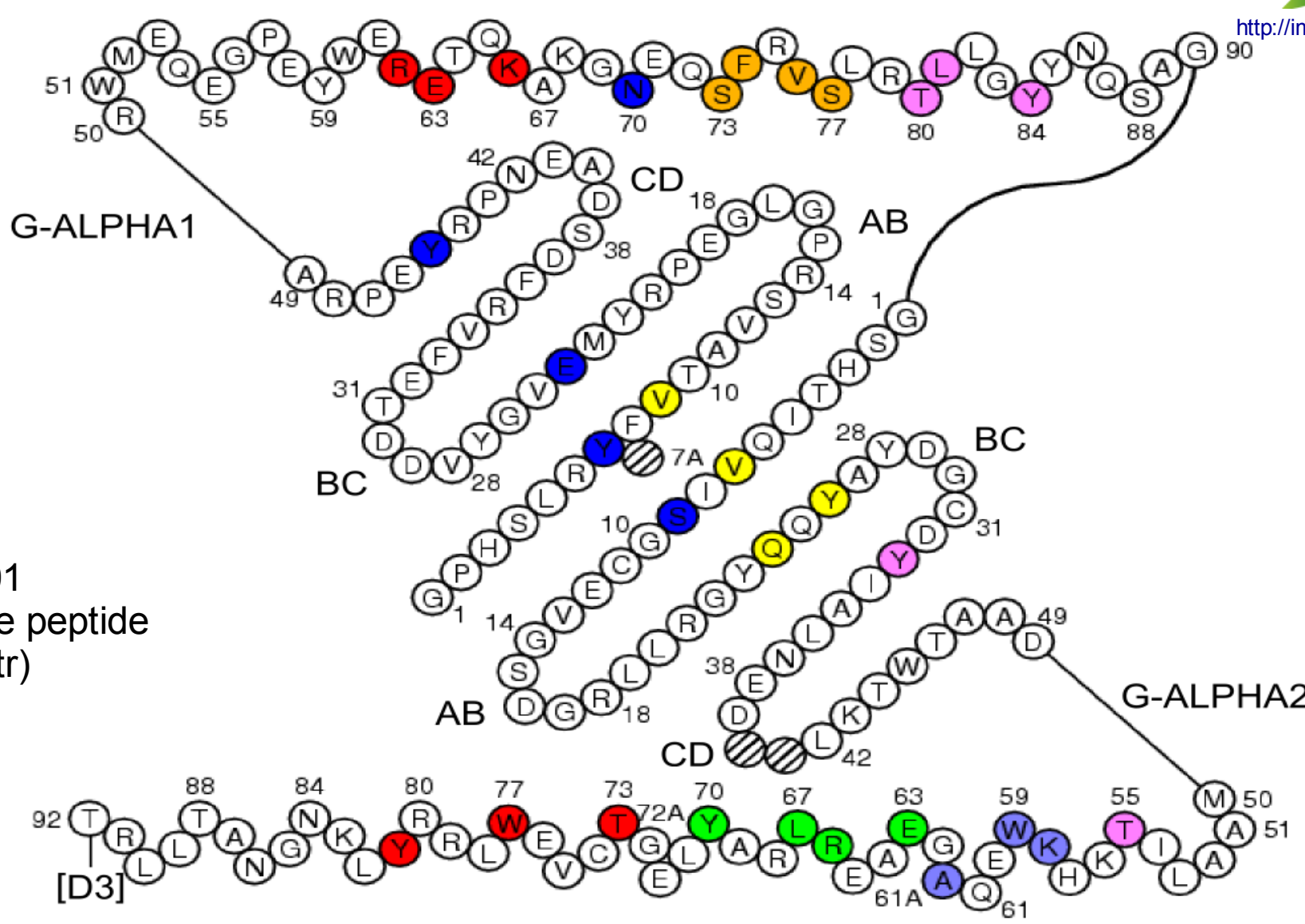
H2-K1*01
(code 1jtr)
8 residue peptide

Kaas and Lefranc In Silico Biology 2005

MHC class I IMGT contact sites

http://imgt.cines.fr

C1	1 E
C2	--
C3	2 Q
C4	3 Y
C5	(4 K)
C6	5 F
C7	--
C8	--
C9	6 Y
C10	7 S
C11	8 V



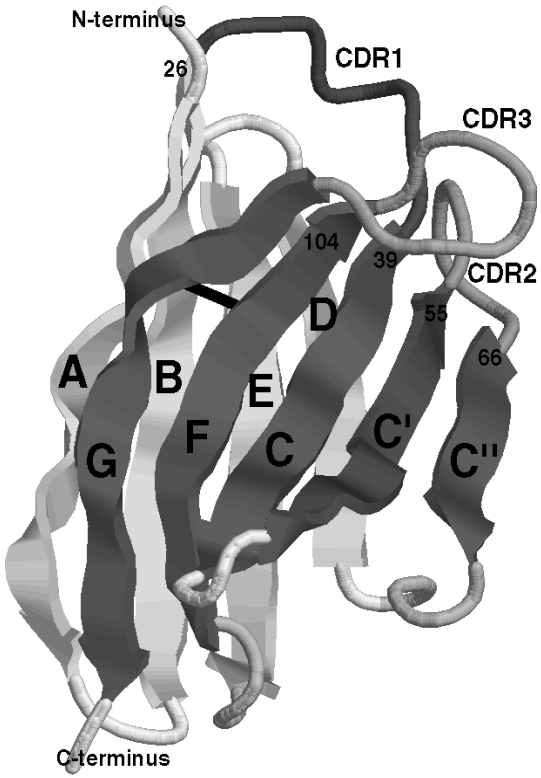
H2-K1*01
8 residue peptide
(code 1jtr)

Lefranc et al. Dev. Comp. Immunol. 29, 917-938 (2005)

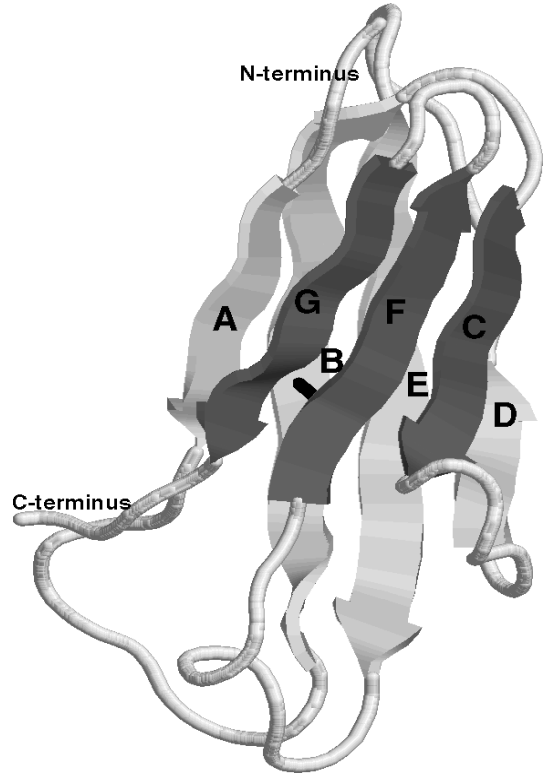
IMGT unique numbering

http://imgt.cines.fr

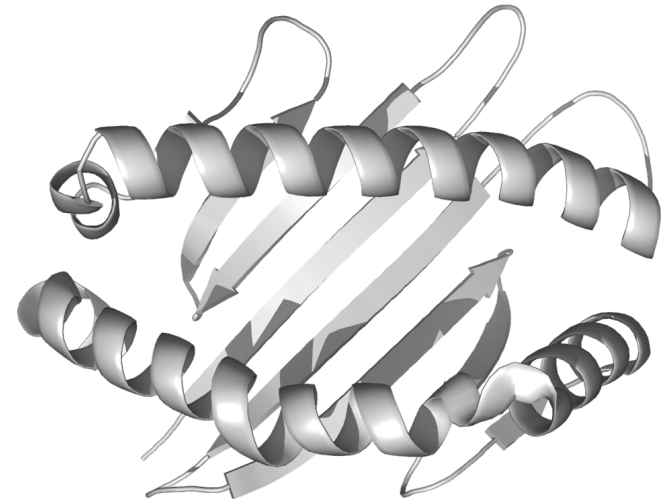
V-DOMAIN (IG,TR)
AND
V-LIKE-DOMAIN
(other than IG,TR)



C-DOMAIN (IG,TR)
AND
C-LIKE-DOMAIN
(other than IG,TR)



G-DOMAIN (MHC)
AND
G-LIKE-DOMAIN
(other than MHC)

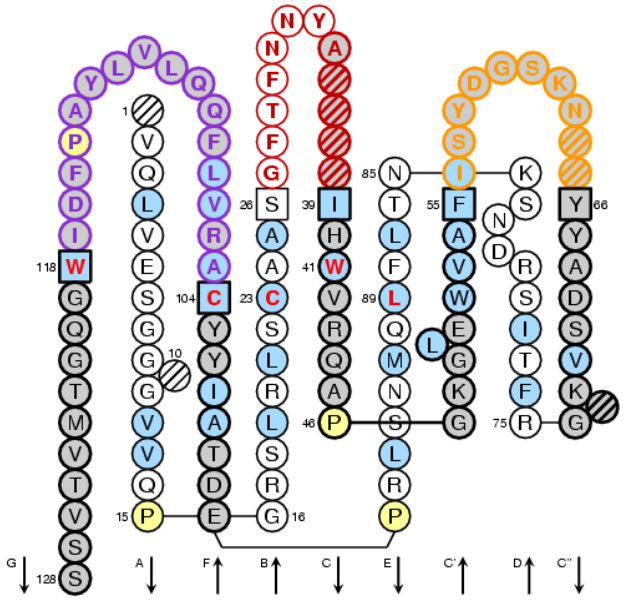


Immunoglobulin superfamily (IgSF)

MHC superfamily (MhcSF)

V-LIKE-DOMAIN

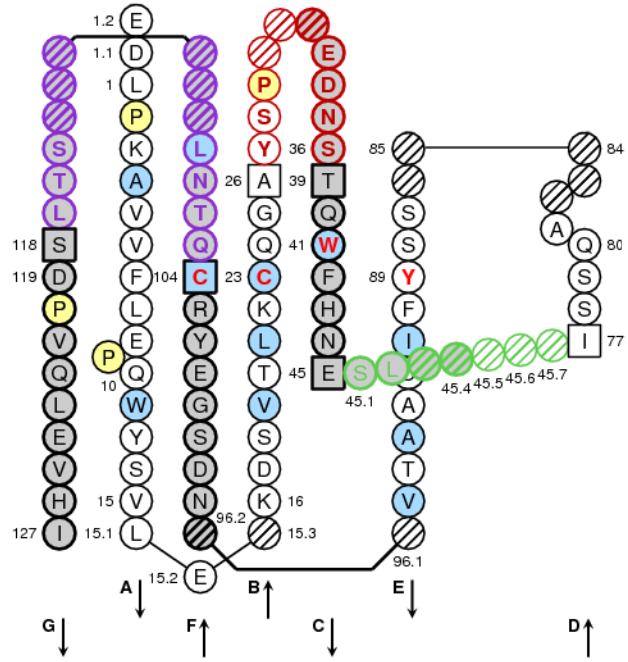
MOG



Duprat, E. et al., Recent Res. Develop. Human Genet., 2, 111-136 (2004)

C-LIKE-DOMAIN

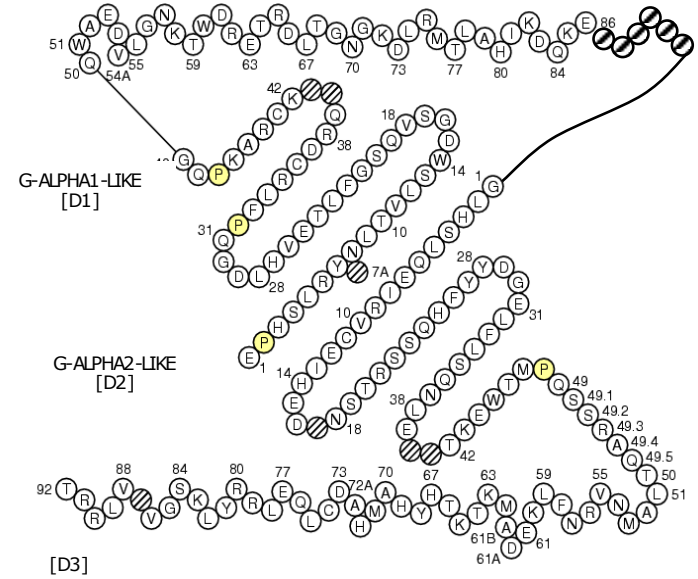
FCGR3B



Bertrand, G. et al., Tissue Antigens, 64, 119-131 (2004)

G-LIKE-DOMAIN

MICA



Frigoul, A. et al., Recent Res. Develop. Human Genet., 3, 95-145 (2005)

Interactions between domains

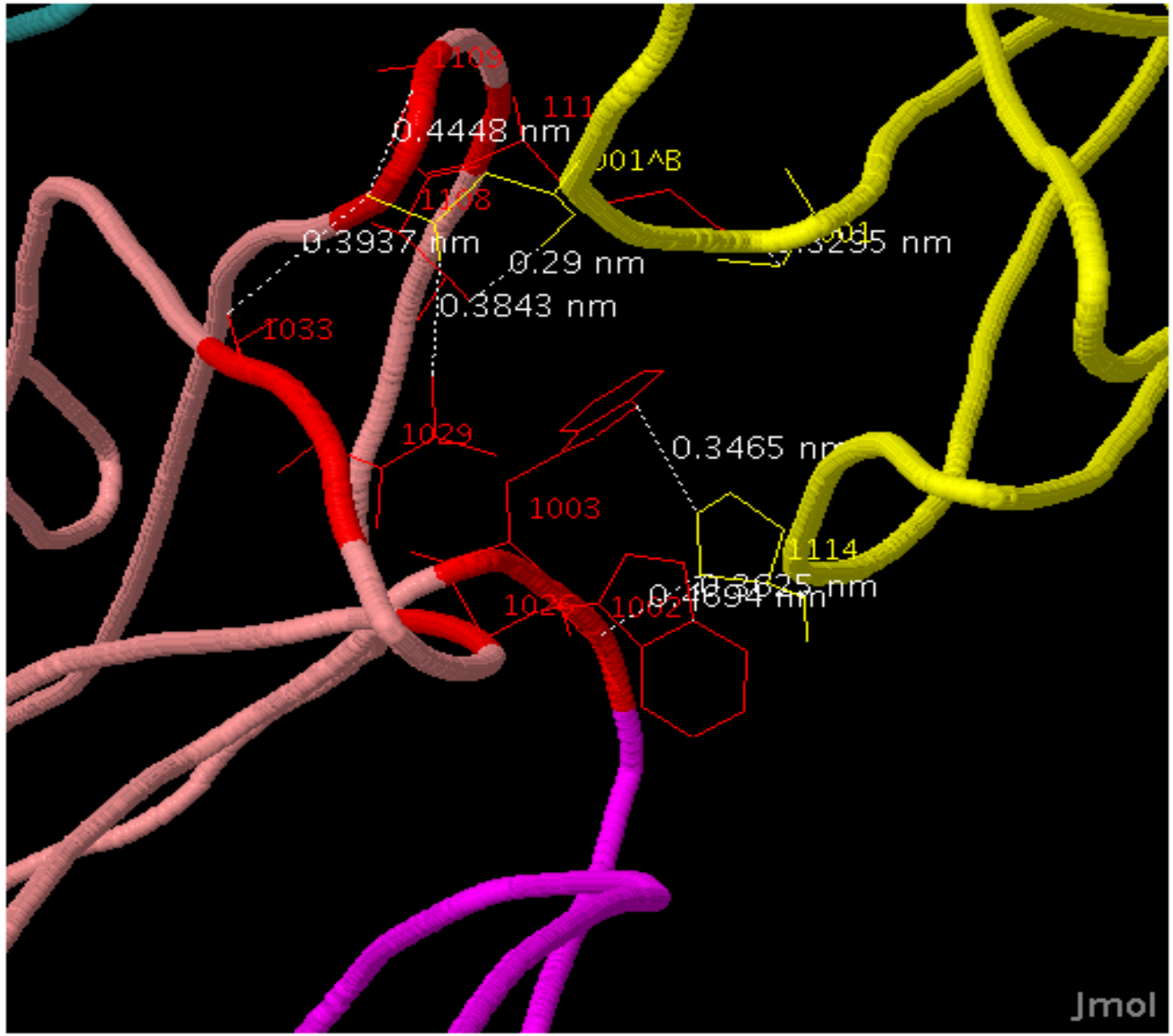
http://imgt.cines.fr

IGHG1
(FC-GAMMA1)

CH2
C-DOMAIN

FCGR3
B

[D2
d-LIKE-
DOMAIN



[D1] C-LIKE-DOMAIN