

IMGT databases and tools for the analysis of the B cell repertoire in normal and pathological situations

Marie-Paule Lefranc

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

Colloque d'Oléron, Lymphocytes B normaux et tumoraux,
May 12-15, 2005, Ile d'Oléron, France



The international ImMunoGeneTics information system®
Coordinator: M.-P. Lefranc, Montpellier, France <http://imgt.cines.fr>



When and why was IMGT created?

IMGT birth and objectives

1. 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*

** Paul Boucheix (CD nomenclature), Roland Berger (translocations).*

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.

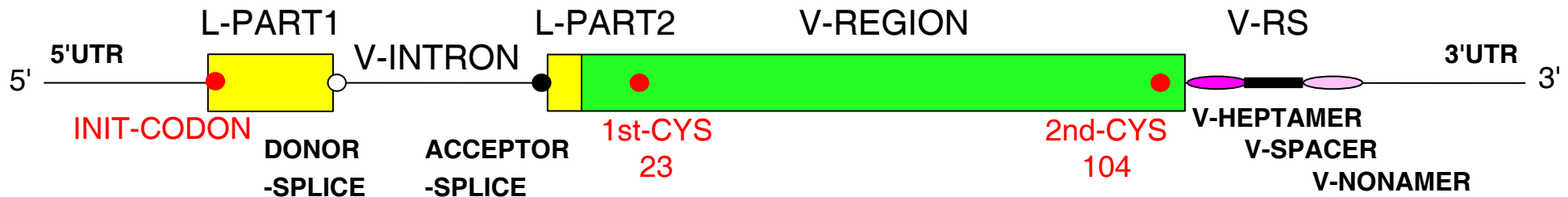
Genomic DNA in germline configuration

V-GENE

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

```

tgagagctcc gttcctcacc atggactgga cctggaggat cctcttcttg gtggcagcag      60
ccacaggaa gaggctcctt agtcccagtg atgagaaaga gattgagtcc agtccagggg      120
gatctcatcc acttctgtgt tctctccaca ca ggagcccact cccaggtgca gctgggtgcag      180
tctggggctg aggtgaagaa gcctggggcc tcagtgaagg tctcctgcaa ggcttctgga      240
tacaccttca ccggctacta tatgcactgg gtgcgacagg cccttggaaca agggcttgag      300
tggatgggat ggatcaacc taacagtggg ggcacaaact atgcacagaa gtttcagggc      360
agggtcacca tgaccagggg cacgtccatc agcacagcct acatggagct gagcaggctg      420
agatctgacg acacggccgt gtattactgt gcgagagaca ca cagtgtgaaa acccacatcc      480
tgagggtgtc agaaacccaa gggaggaggc ag
    
```



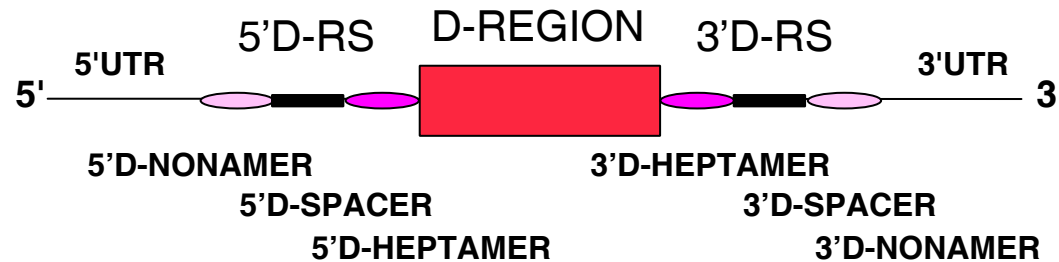
Genomic DNA in germline configuration

D-GENE

>J00256|IGHD7-27*01|*Homo sapiens* D-GENE

```
ccagccgcag ggtttttggc tgagctgaga ac cactgtgc taactgggga cacagtgatt
ggcagctcta caaaaaccat gctccccgg g
```

60



J-GENE

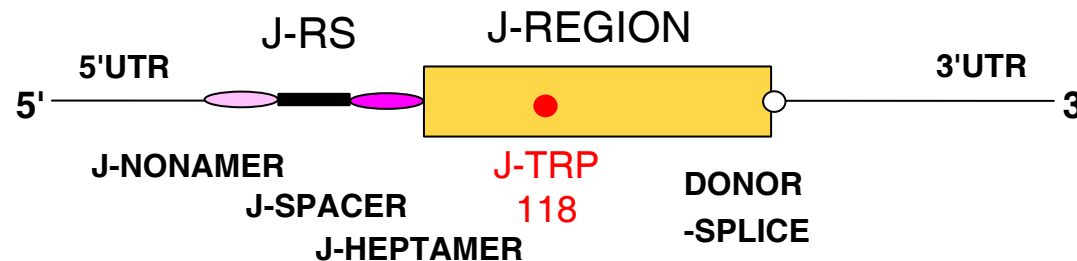
>J00256|IGHJ1*01|*Homo sapiens* J-GENE

```
accccgggct gtgggtttct gtgccctgg ctcagggtg actcaccgtg gctgaatact
tccagcactg gggccagggc accctgggtca ccgtctcctc aggtgagttct gctgtactgg
ggatagcggg gagccatgtg tactgggcca agcaagggtc ttggcttcag
```

60

120

170



Why was IMGT necessary?

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

What was the first goal?

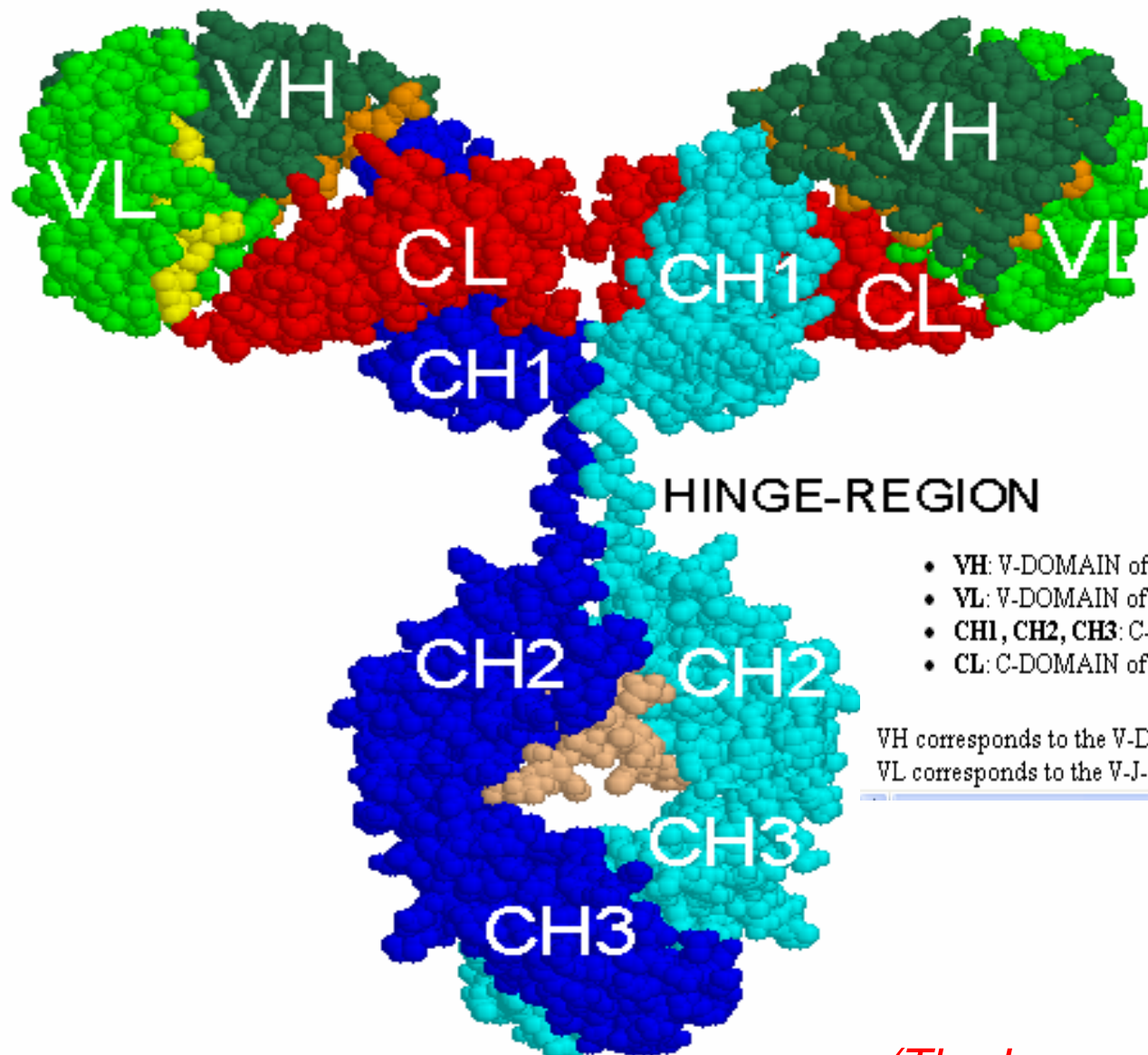
CONTENT

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

Spacefill 3D representation of an IgG



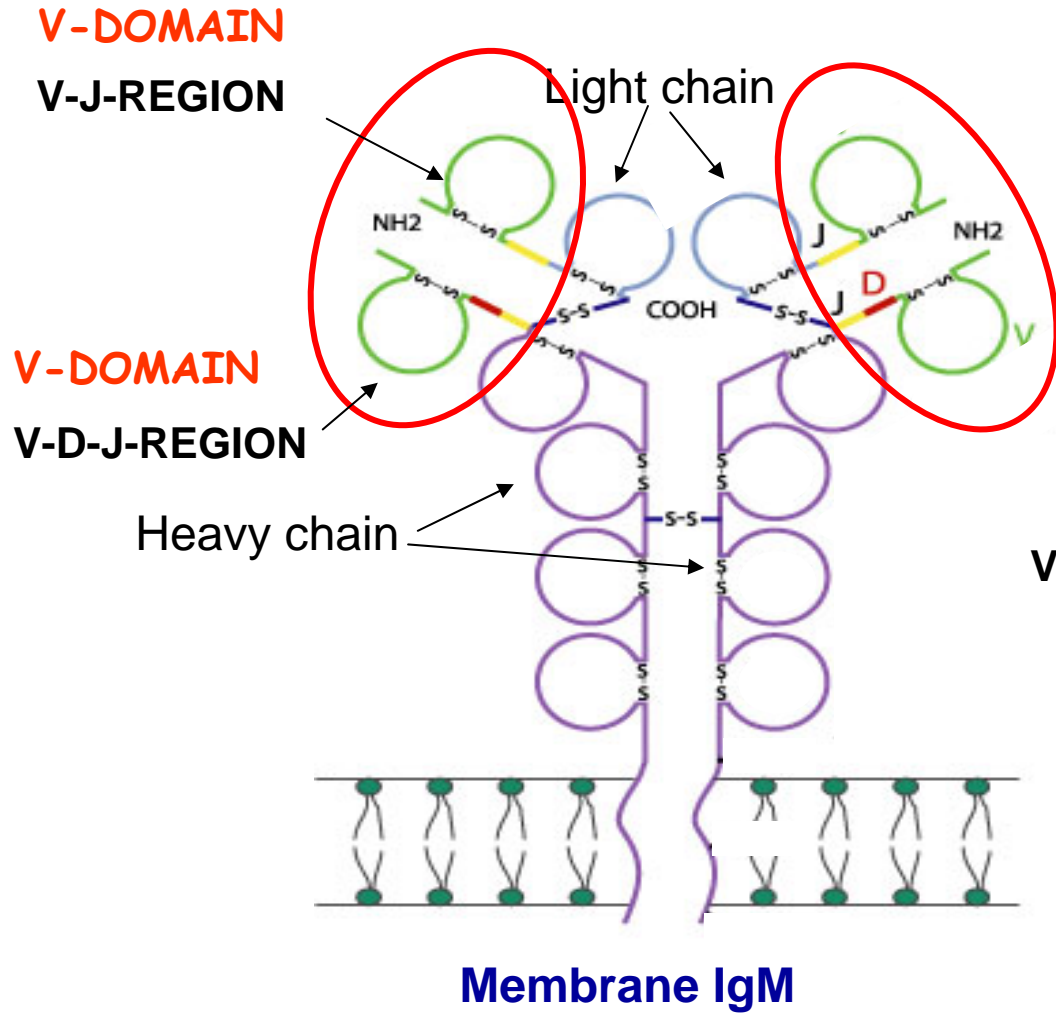
- VH: V-DOMAIN of the immunoglobulin heavy chain
- VL: V-DOMAIN of the immunoglobulin light chain
- CH1, CH2, CH3: C-DOMAIN of the immunoglobulin heavy chain
- CL: C-DOMAIN of the immunoglobulin light chain

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)

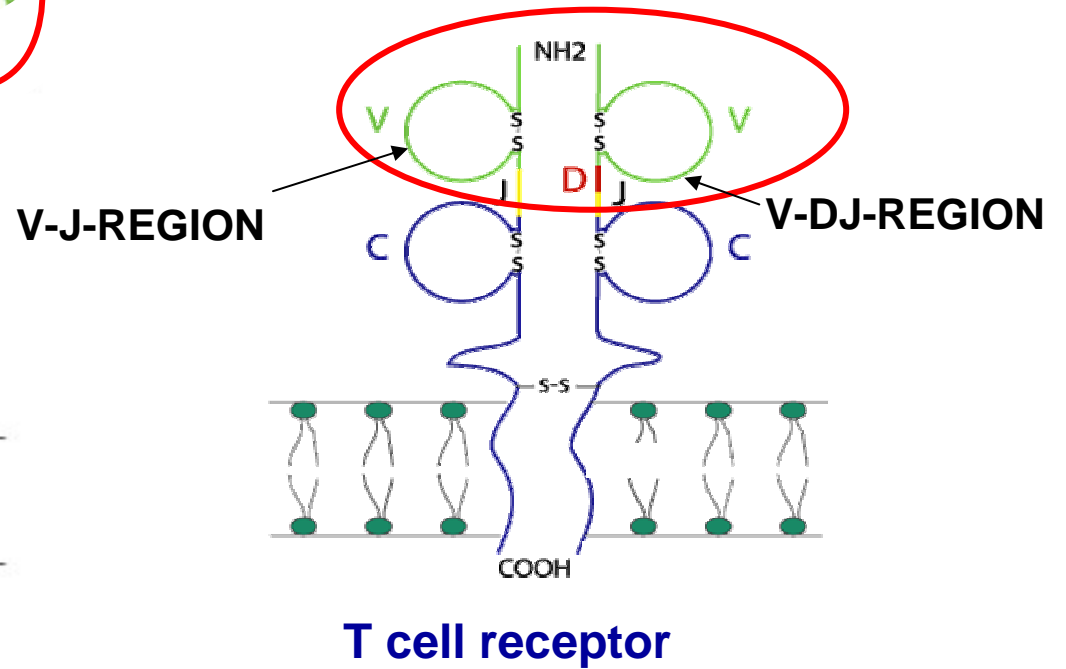
Immunoglobulin (IG)

T cell receptor (TR)

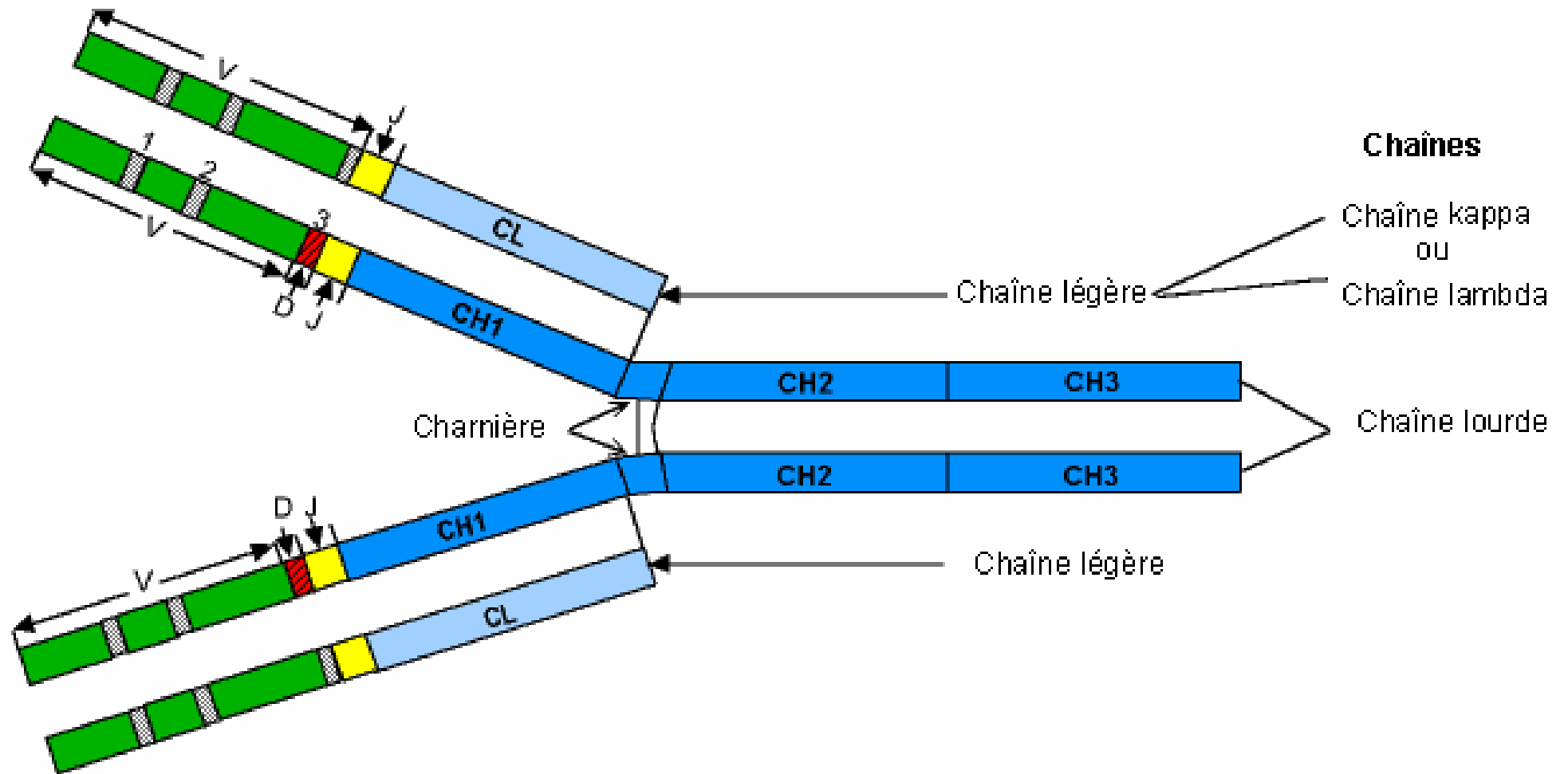


Contribution of the
2 V-DOMAINS
to the antigen binding site

Alpha - Beta
Gamma - Delta

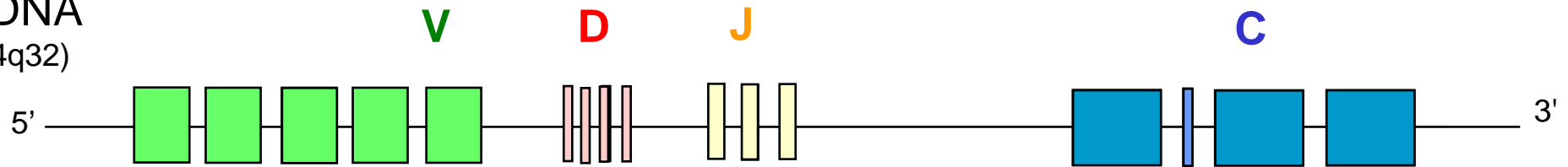


Immunoglobulin IgG



Immunoglobulin (IG) synthesis

genomic DNA
(IGH Locus 14q32)



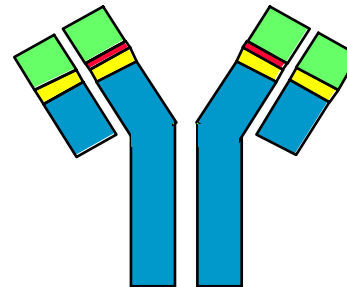
rearranged
DNA



mRNA



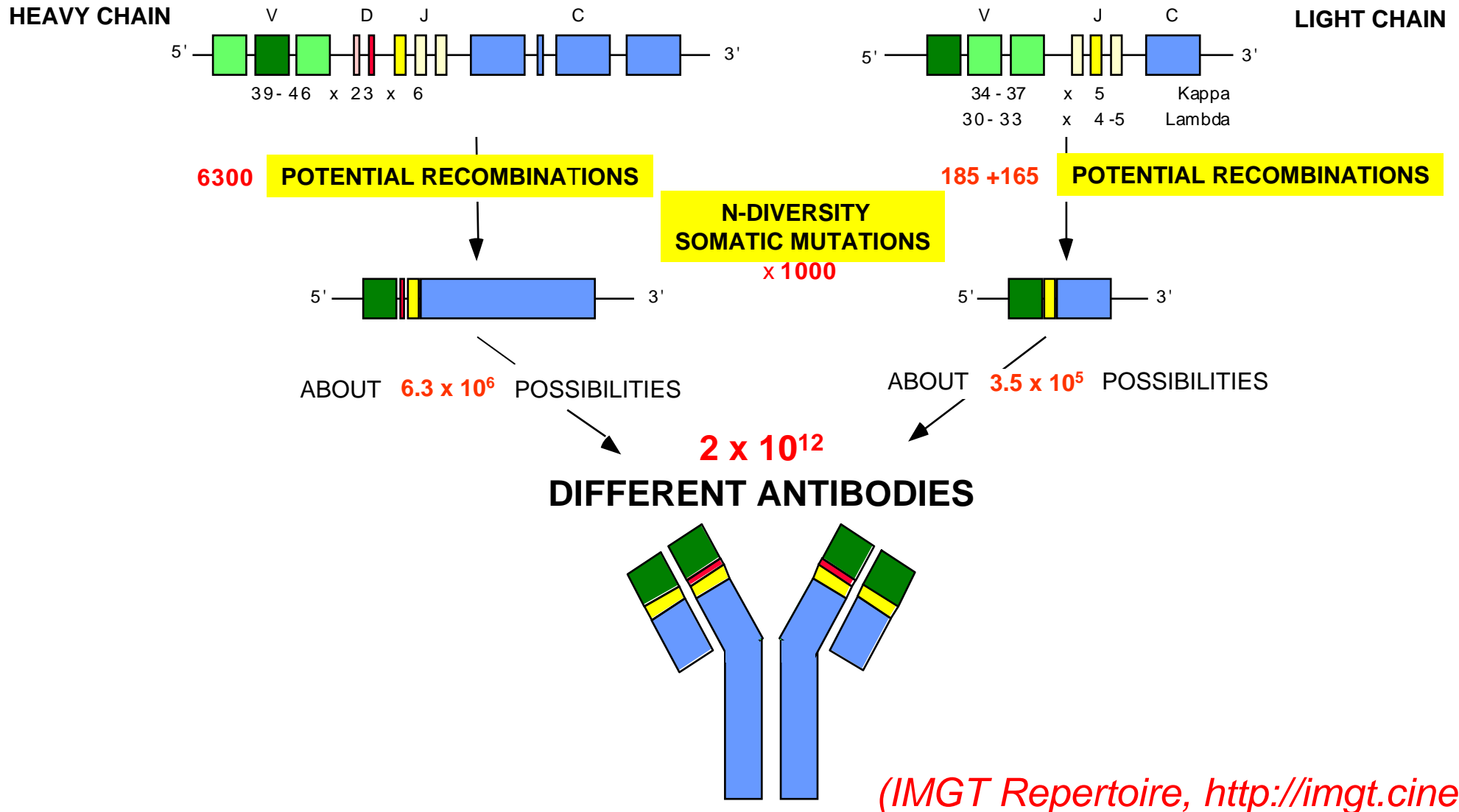
2×10^{12} different IG
per individual



(IMGT Repertoire, <http://imgt.cines.fr>)

Immunoglobulin (IG) synthesis

150 FUNCTIONAL IG GENES



IMGT® Web resources

IMGT Repertoire **8,000 pages HTML**

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 IGHV genes on chromosome 14 at 14q32.33

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

(IMGT Repertoire, <http://imgt.cines.fr>)

IMGT Alignment of alleles



<http://imgt.cines.fr>

X02850 , TRAV8-6*01	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			
	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																						
X02850 , TRAV8-6*01	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			
	CTG	AGG	TGC	ARC	TAC	TCA	TCG	TCT	GTT	TCA	GTG	TAT	CTC	TTC			
AE000659, TRAV8-6*02	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---			
M86361 , TRAV8-6*02	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---			
	CDR2-																						
X02850 , TRAV8-6*01	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			
	TGG	TAT	GTG	CAA	TAC	CCC	ARC	CAA	GGA	CTC	CAG	CTT	CTC	CTG	ARG	TAT	TTA	TCA	GGA	TCC			
AE000659, TRAV8-6*02	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---			
M86361 , TRAV8-6*02	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---			
	IMGT																						
X02850 , TRAV8-6*01	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				
	ACC	CTG	GTT	GAA	AGC	ATC	AAC	GGT	TTT	GAG				
AE000659, TRAV8-6*02	---	---	---	---	---	---	A	G			---	---	---	---	---	---	---	---	---	---			
M86361 , TRAV8-6*02	---	---	---	---	---	---	A	G			---	---	---	---	---	---	---	---	---	---			
	IMGT																						
X02850 , TRAV8-6*01	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
	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																						
X02850 , TRAV8-6*01	101	102	103	104	105	106	107	108															
	E	Y	F	C	A	V	S																
	GAG	TAC	TTC	TGT	GCT	GTG	AGT	GA															
AE000659, TRAV8-6*02	---	---	---	---	---	---	---	---															
M86361 , TRAV8-6*02	---	---	---	---	---	---	---	---															
							R																
							---	G															
									#c														

(IMGT Repertoire, <http://imgt.cines.fr>)

IMGT Protein display

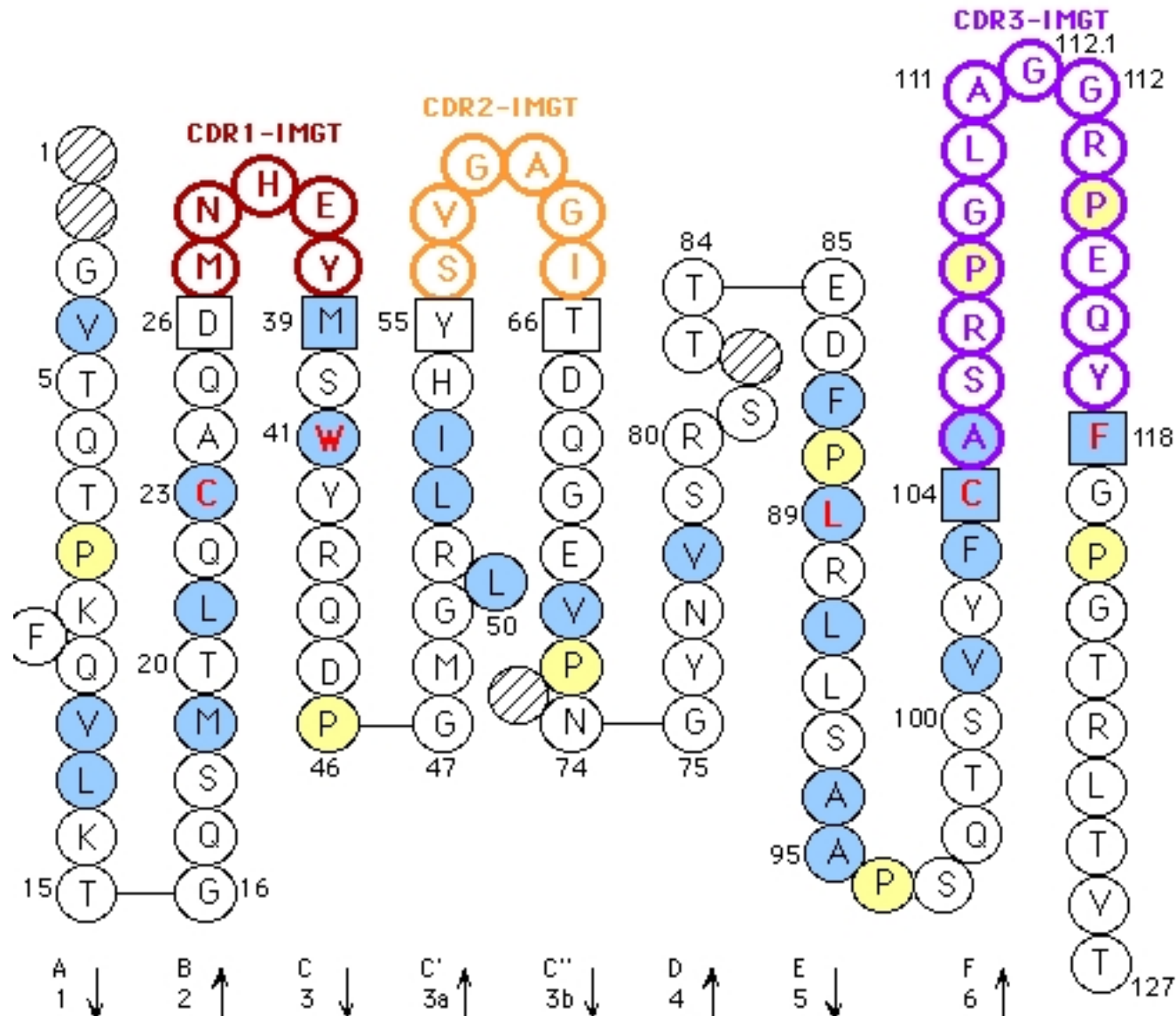


<http://imgt.cines.fr>

TRAV gene	FR1-IMGT (1-26)			CDR1-IMGT (27-38)			FR2-IMGT (39-55)			CDR2-IMGT (56-65)			FR3-IMGT (66-104)			CDR3-IMGT (105-115)		
	1	10	20	30	40	50	60	70	80	84ABC	90	100	110					
AE000658, TRAV1-1	GQSLEQ	PSEVTAVEGAI	VQINCTYQ	TSGFYG	LSWYQQHDGG	APTFLSY	NADG	LEETG	RFSSFLSR	SDSYGYLL	QELQMKDS	SASYFC	AVR					
AE000658, TRAV1-2	GQNIDQ	PTEMTATEGAI	VQINCTYQ	TSGFNG	LFWYQQHAGE	APTFLSY	NVLDG	LEEKG	RFSSFLSR	SKGYSYLL	LKELQMKD	SASYLC	AVR					
AE000658, TRAV2	KDQVFQ	PSTVASSEGAV	VEIFCNHS	VSNAYN	FFWYLHFPG	CAPRLLVK	GSK	PSQQG	RYNMTYER	FSSLLILQ	VREADAAV	YVC	AVE					
AE000658, TRAV3	AQSVAQPEDQ	VNVAEGNPL	TVKCTYS	VSGNPY	LFWYVQYP	NRGLQFLK	YITGDNL	VKGSY	GFEAEFNK	SQTSFHLK	KPSALV	SDSALYFC	AVRD					
AE000658, TRAV4	LAKTTQ	PISMSYEGQ	EVNITCSHN	NIATNDY	ITWYQQFP	SQGRFIIQ	GYKT	KVTNE	VASLFI	PADRKS	SSTLSL	PRVSL	SDTAVYYC	LVGD				
AE000659, TRAV5	GEDVEQS	LFLSVREGD	SSVINCTYT	DSSSTY	LYWYKQEP	CAGLQLLTY	IFSNMD	MKQDQ	RLTVLLN	KKDKHL	SLRIADT	QTGDSAI	YFC	AES				
AE000659, TRAV6	SQKIEQNSE	ALNIQEK	TATLTCNYT	NYSPAY	LQWYRQDP	GRGPVFLLL	IRENEK	EKRKE	RLKVT	FDTT	KLSL	IL	ASQPAD	SATYLC	ALD			
AE000659, TRAV7	ENQVEHSP	HFLGPQQG	DVASMSCTYS	VSRFNN	LQWYRQNT	GMGPKLLS	MYSAGY	EKQKG	RLNATL	LK	NGSSLY	ITAVQ	PEDSATYFC	AVD				
AE000659, TRAV8-1	AQSVSQHN	HVILSEA	ASLELGCNYS	YGGTVN	LFWYVQYP	QGHLQLLLK	YFSGDPL	VKGIK	GFEAEF	IKSKFS	FNLRK	PSVQ	SDTAEYFC	AVN				
AE000659, TRAV8-2	AQSVTQ	DSHVS	VSECTPVLLRCNYS	SSYSPS	LFWYVQH	PNKGLQLLLK	Y TSAATL	VKGIN	GFEAEF	KKSET	SFHLTK	PSAH	MSDAAEYFC	VVS				
AE000659, TRAV8-3	AQSVTQ	PDIHIT	VSEGASLELRCNYS	YGATPY	LFWYVQ	SPQGLQLLLK	YFSGDTL	VQGIK	GFEAEF	KRSQ	SSFNLR	KPSV	HWSDAEYFC	AVG				
AE000659, TRAV8-4	AQSVTQ	LDSQV	VFEEAPVELRCNYS	SSVFPY	LFWYVQ	YPNQGLQLLLK	Y TSAATL	VKGIN	GFEAEF	KKSET	SFHLTK	PSAH	MSDAAEYFC	AVS				
X02850, TRAV8-6	AQSVTQ	LDSQV	VFEEAPVELRCNYS	SSVSVY	LFWYVQ	YPNQGLQLLLK	YLSGSTL	VEIN	GFEAEF	NK	SQTSFHL	RKPSV	HISD	TAEYFC	AVS			
AE000660, TRAV8-7	TQSVTQ	LDSGHIT	VSEEAPLELRCNYS	YSGVPS	LFWYVQ	YSSQSLQLLLK	DLTEATQ	VKGIR	GFEAEF	KKSET	SFYL	RKPS	THVSD	AAEYFC	AVGDR			
AE000659, TRAV9-1	QDSWVQTE	GGWLP	SEGD	SLIVNCSYF	TTQVPS	LFWYVQ	YVPGEG	QDLHLK	AMKAMD	KGRNK	GFEAMY	RKETT	SEHLEK	DSVQ	SDSAVYFC	ALS		

(IMGT Repertoire, <http://imgt.cines.fr>)

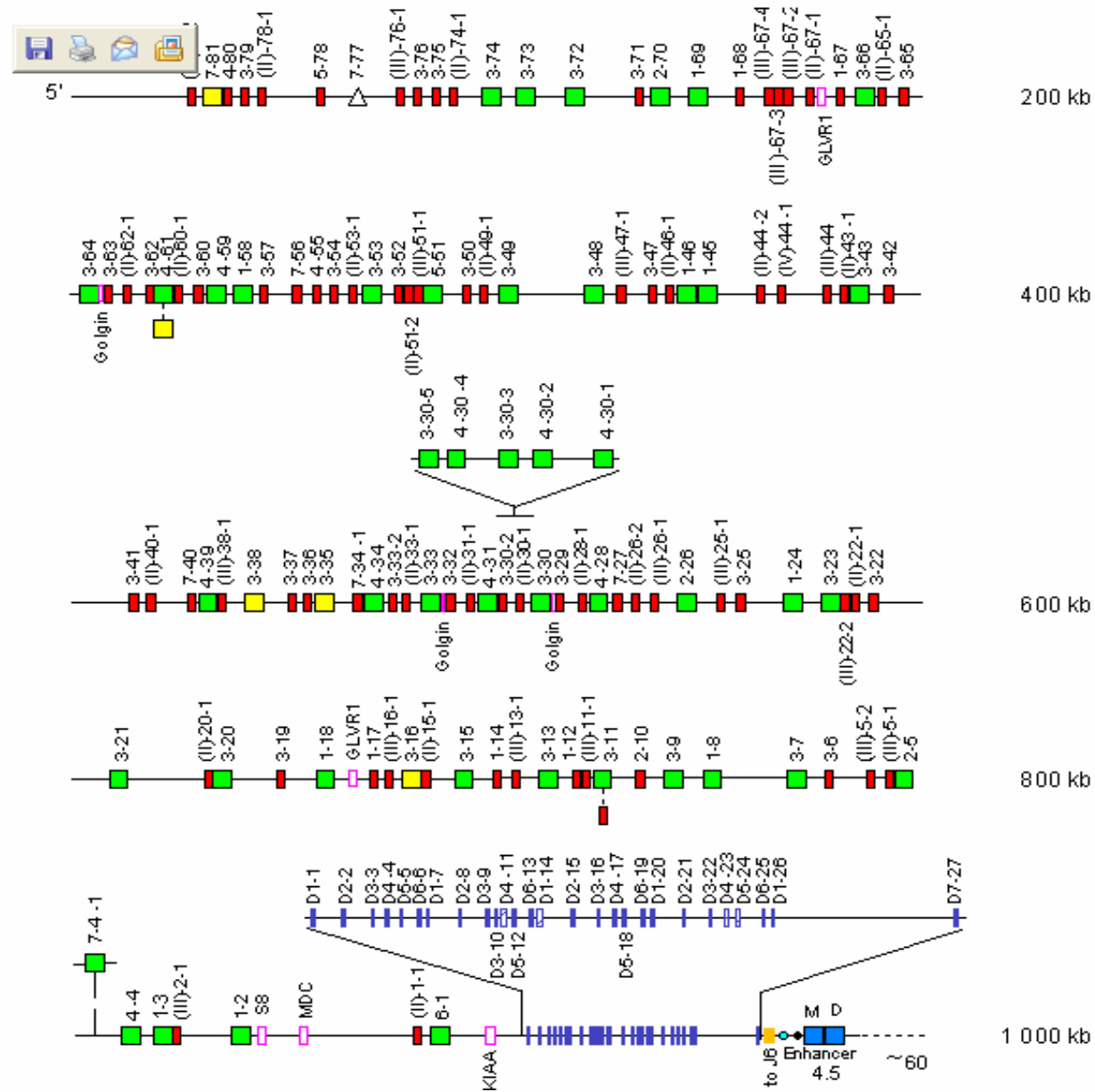
IMGT Collier de Perles



(IMGT Repertoire, <http://imgt.cines.fr>)

Human IGH locus

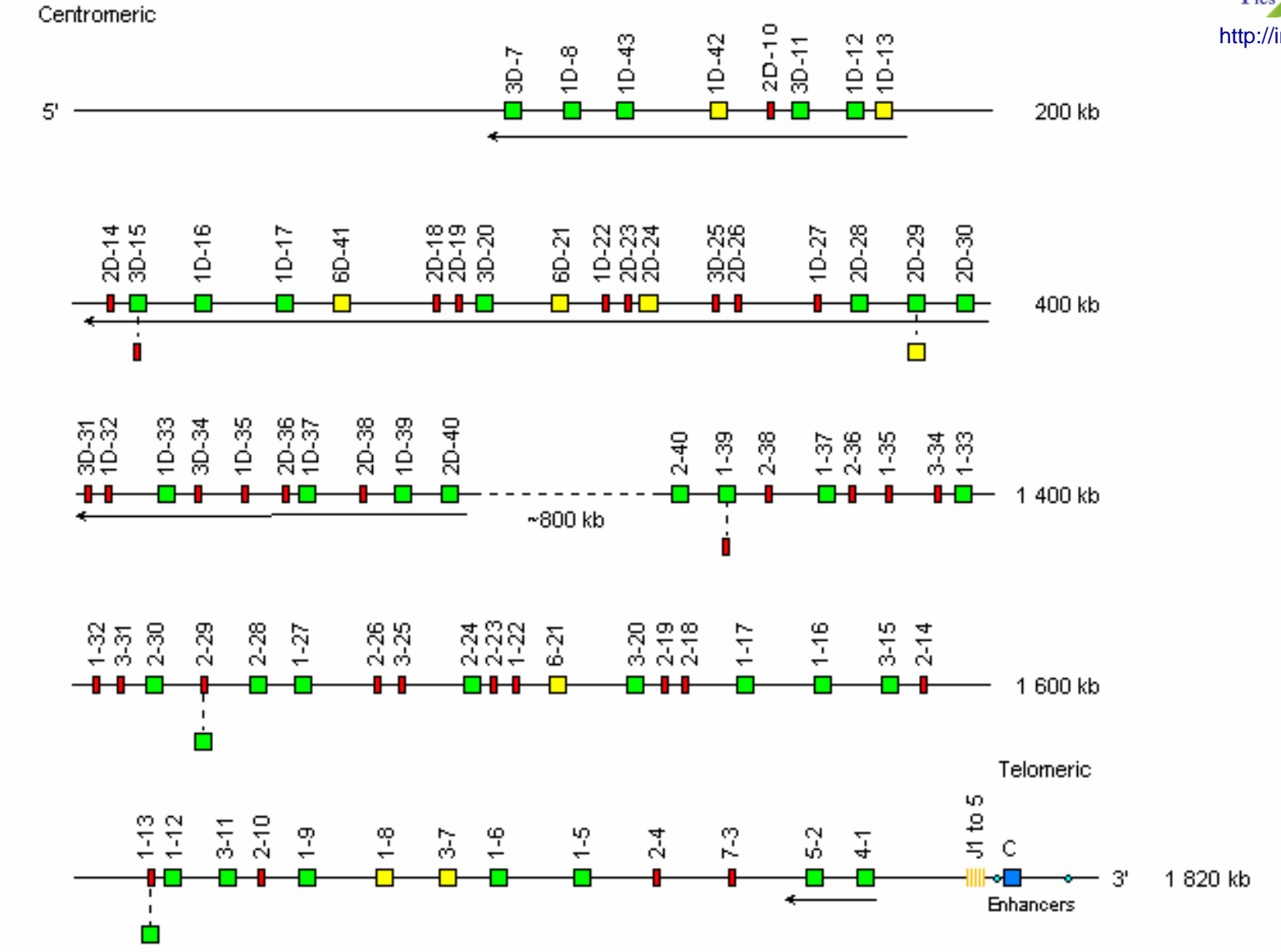
Chromosome
14q32.33



(IMGT Repertoire, <http://imgt.cines.fr>)

Human IGK locus

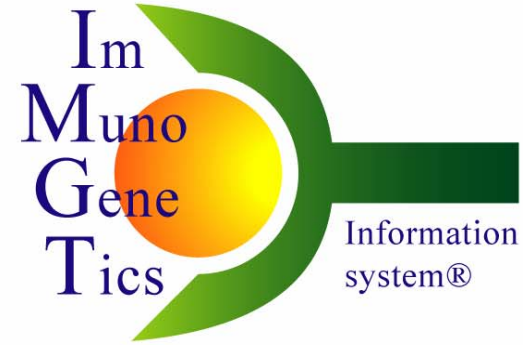
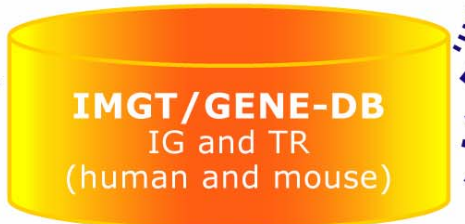
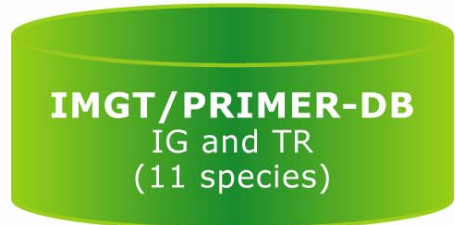
Chromosome 2p11.2



(IMGT Repertoire, <http://imgt.cines.fr>)

IMGT databases and tools

Sequences



<http://imgt.cines.fr>

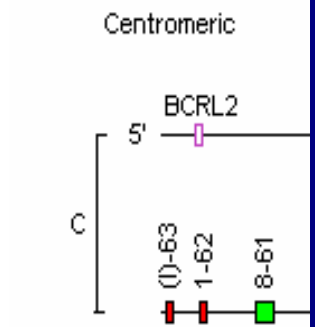
Genome



2D and 3D structures

Locus representation: Human IGL

Human IGL 22q11.2

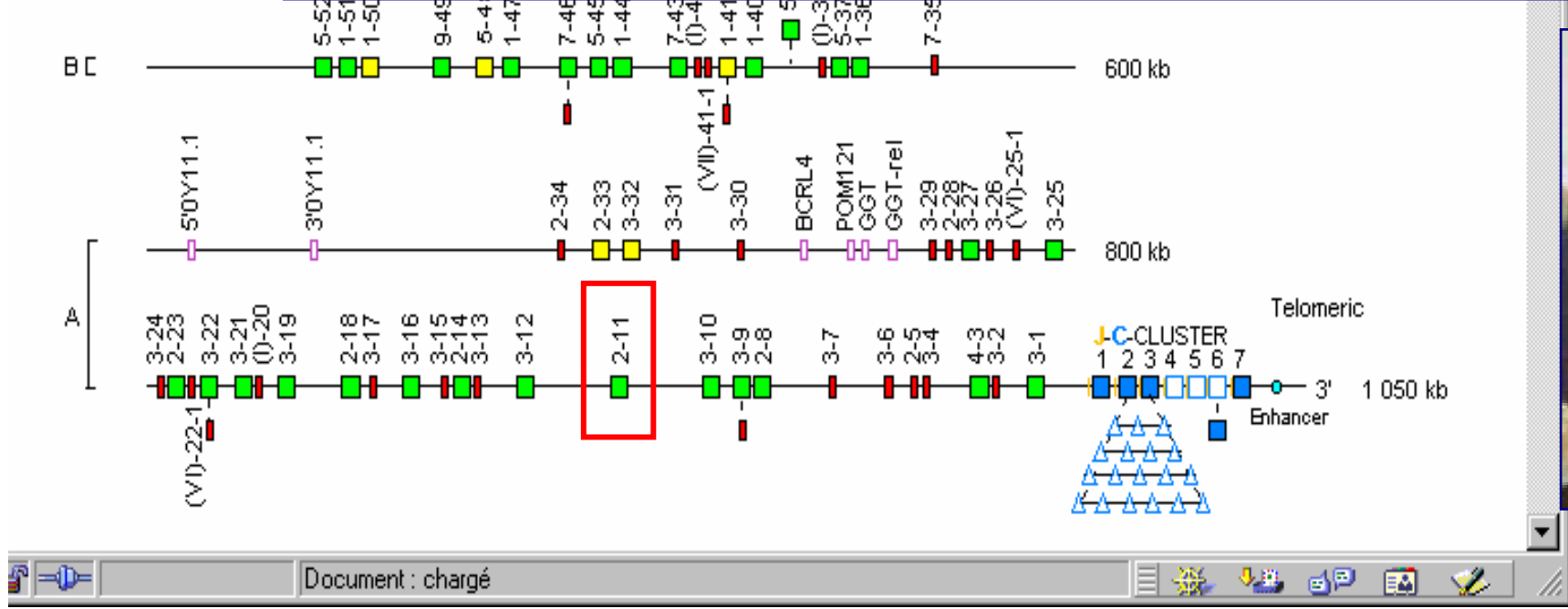


WELCOME !
to IMGT/GENE-DB

THE
 INTERNATIONAL
 IMMUNOGENETICS
 INFORMATION SYSTEM®

Information system®

<http://imgt.cines.fr>



[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

[accession number, mnemonic, definition, creation date, length, annotation level](#)

```

FT  V-GENE
FT
FT
FT
FT
FT
FT
FT
FT
FT
FT
FT  V-REGION
FT
FT
FT
FT
FT
FT
FT  FR1-IMGT
FT
FT
FT
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="VSWYQQHPGKAPKLMIIY"
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
    tctctactca gaaccacacg tcatattgat gattataact atctctctca ctaccaacac      120
  
```



cDNA (in databases: mRNA!)

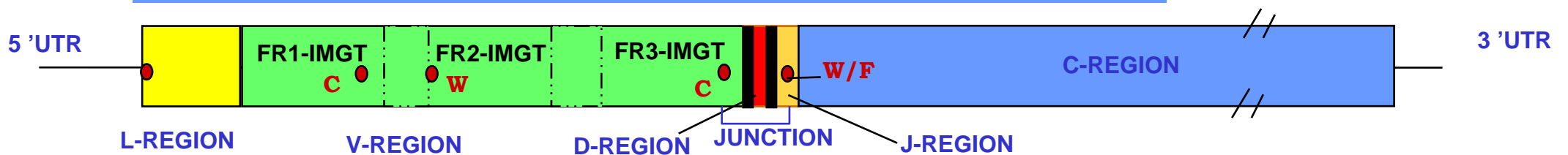
```
..... .gagga ttcacccatgg aactggggct ccgctggggt ttccttggtg cttttttaga 120
aggtgtccag tgtgaggtgc aactggtgga gtctggggga ggcctgggtca agccgggggg 180
gtccctgaga ctctcctgtg cagcctctgg attaagcttc agtacctatg ccatgaactg 240
ggtcgccag gctccagga aggggctgga atgggtctca agtattagta gtagaagtga 300
ttacatatac tatagagact cagtgaaggg ccgattcacc atctccagag acaacgcca 360
gaattcactg tatctgcaaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgcgagagat tcttgtaatg gtgctatatg ttatggtttc agtccctggg gccaggggaa 480
cctggtcacc gtctcctcag catccccgac cagccccaag gtcttcccgc tgagcctctg 540
cagcaccag ccagatggga acgtgggtcat cgctgcctg gtccagggct tcttccccca 600
ggagccactc agtgtgacct ggagcgaaag cggacagggc gtgaccgcca gaaacttccc 660
accagccag gatgcctccg gggacctgta caccacgagc agccagctga ccctgccggc 720
cacacagtgc ctagccggca agtccgtgac atgccacgtg aagcactaca cgaatcccag 780
ccaggatgtg actgtgccct gccagttcc ctcaactcca cctaccccat ctccctcaac 840
tccacctacc ccatctccct catgctgcca cccccgactg tcaactgcacc gaccggccct 900
cgaggacctg ctcttaggtt cagaagcgaa cctcacgtgc aactgaccg gcctgagaga 960
tgcctcgggt gtcaccttca cctggacgcc ctcaagtggg aagagcgctg ttcaaggacc 1020
acctgagcgt gacctctgtg gctgctacag cgtgtccagt gtccctgccg gctgtgccga 1080
gccatggaac catgggaaga cttcacttg cactgctgcc taccgaggt 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 gccagccag ggcaccacca ccttcgctgt 1380
gaccagcata ctgcgctgg cagccgagga ctggaagaag ggggacacct tctcctgcat 1440
ggtgggccac gaggccctgc cgctggcctt cacacagaag accatcgacc gcttggcggg 1500
taaaccacc catgtcaatg tgtctgttgt catggcggag gtggacggca cctgctactga 1560 ....
```

cDNA

```

..... .gagga ttcaccatgg aactggggct ccgctgggtt ttccttggtg cttttttaga 120
aggtgtccag tgtgaggtgc aactgggtgga gtctggggga ggctgggtca agccgggggg 180
gtccctgaga ctctctgtg cagcctctgg attaagcttc agtacctatg cdatgaadtg 240
gggccgag gtcaccagga aggggctgga atgggtctca agtattagta gtagaagtga 300
ttacatatac tatagagact cagtgaaggg ccgattcacc atctccagag acaacgccaa 360
gaattcactg tatctgcaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgcgagagat tctttgtaatg gtgctatatg ttatggtttc agtccctggg gccagggaac 480
cctggtcacc gtctctcag catccccgac cagccccaaag gtcttcccgc tgagcctctg 540
cagcaccag ccagatggga acgtgggtcat cgctgacctg gtccagggtc tcttcccca 600
ggagccactc agtgtgacct ggagcgaaag cggacagggc gtgaccgcca gaaacttccc 660
accagccag gatgcctccg gggacctgta caccacgagc agccagctga ccctgccggc 720
cacacagtgc ctagccggca agtccgtgac atgccacgtg aagcactaca cgaatcccag 780
ccaggatgtg actgtgccct gccagttcc ctcaactcca cctaccccat ctccctcaac 840
tccacctacc ccatctccct catgctgcc accccgactg tcaactgacc gaccggcct 900
cgaggacctg ctcttaggtt cagaagcgaa cctcacgtgc aactgaccg gcctgagaga 960
tgctcgggt gtcacctca cctggacgcc ctcaagtggg aagagcgctg tcaaggacc 1020
acctgagcgt gacctctgtg gctgctacag cgtgtccagt gtctgcccgg gctgtgccga 1080
gcatggaac catgggaaga ccttcaactg cactgctgcc taccgaggt ccaagacccc 1140
gctaaccgcc accctctcaa aatccgaaa cacattccgg cccgaggtcc acctgctgcc 1200
gccgccgtcg gaggagctgg ccctgaacga gctggtgacg ctgacgtgcc tggcacgtgg 1260
cttcagccc aaggacgtgc tggttcgtg gctgcagggg tcacaggagc tgccccgca 1320
gaagtacctg acttgggcat cccggcagga gccagccag ggcaccacca ccttcgctgt 1380
gaccagcata ctgctcgtgg cagccgagga ctggaagaag ggggacacct tctcctgcat 1440
gggtggccac gaggccctgc cgctggcctt cacacagaag accatcgacc gcttggcggg 1500
taaaccacc catgtcaatg tgtctgttgt catggcggag gtggacggca cctgctactga 1560 ....

```



The IMGT-ONTOLOGY concepts

I
D
E
N
T
I
F
I
C
A
T
I
O
N

```
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
FT

Document : chargé
```

The IMGT-ONTOLOGY concepts

**D
E
S
C
R
I
P
T
I
O
N**

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

OBTENTION

CLASSIFICATION

NUMEROTATION

IMGT/V-QUEST

IMGT/V-QUEST

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

AF402940		score	CTTCACGGGGCGGGACGCTTTGGACGTCTGGGGCCAAGGGACCACGGTCACCGTCTCCTCA
J00256	IGHJ3*01	181T.....T..T.....A.T.....T...G
X86355	IGHJ6*02	179	T.A.TACTACTACT...G.A.....
X86355	IGHJ3*02	172T.....T..TA.....A.T.....T...G



WELCOME ! to the [IMGT/V-QUEST](http://imgt.cines.fr) Search page

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>

Citing [IMGT/V-QUEST](#): Giudicelli, V. et al. Nucl. Acids Res. 2004, 32, W435-440 [PMID: 15215425](#) [PDF](#)

Analyse your Immunoglobulin nucleotide sequences

- [Human](#)
- [Mouse](#)
- [Chondrichthyes](#)
- [Teleostei](#) (contains now 19 species)
 - [Atlantic cod](#)
 - [Channel catfish](#)
 - [Rainbow trout](#)
- [Sheep](#)

Analyse your T cell Receptor nucleotide sequences

- [Human](#)
- [Mouse](#)
- [Non-human primates](#)

Analyse your Immunoglobulin sequences

Your selection: Human

Your sequences are compared to the **Human IG set** from the [IMGTV-QUEST reference directory sets](#)

- Include the analysis of the junction by IMGT/JunctionAnalysis
- Do not include the analysis of the junction by IMGT/JunctionAnalysis

Sequence

You may give your sequence a name (optional) :

- Type (or copy/paste) your sequence into the box below :

```
gagggtgcagctggtggagtcggggggaggcttgggtacagcctgggggggtccctgagactc
tccgtgtgcagcctctggattcacctttagcagctttgccatgagctgggtccgccaggct
cctgggaaggggctggactgggtctcagaaatagtggtagtggcggtaccacatatc
gcaggctccgtgaagggccgggtcaccatctccagagacaactccaagaatacgtgttt
ctgcaaatgaatagcctgacagccgaagacacggccgtatattactgtgcaagaggttt
```

- Or give the path access to a local file containing your sequence :

Sequence compared with the Human IG set from the IMGT reference directory

Hyphens (-) show nucleotide identity, dots (.) represent gaps

Alignment for V-GENE

input		score	GAGGTGCAGCTGTTGGAGTCGGGGGA...GGCTTGGTACAGCCTGGGGGGTCCCTGAGACTCTCTGTGCAGCCTCTGGATTACCTTTAGCAGCTTTGCC...
M99660	IGHV3-23*01	1281	-----T-----A-----
J00236	IGHV3-23*02	1263	-----T-----A-----
U29481	IGHV3-23*03	1209	-----T-----A-----
Z12358	IGHV3-48*03	1128	-----G-----T-----A-----C--T--T-A--AA...
M99675	IGHV3-48*01	1128	-----G-----T-----C--T--A-AG--...

Alignment for D-GENE

input		score	GTTTTGGAGTGGTT
X93618	IGHD3-3*02	61	T-----ATTATACC
X13972	IGHD3-3*01	61	T-----ATTATACC
X97051	IGHD6-19*01	34	--A-A-C-----C-GGTAC
X13972	IGHD3-9*01	34	TA---T--C-----ATTATAAC
J00234	IGHD2-15*01	34	A-A---T-----GGTAGCTGCTACTCC

Alignment for J-GENE

input		score	GTTTTGGAGTGGTTATCTGTGGTTCGACCCCTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCAGCTTCCACCAA
X86355	IGHJ5*02	219-CAAC-----
J00256	IGHJ5*01	201-CAAC-----T-----A-----
X86355	IGHJ4*02	177AC-AC--T--TA-----

IMGT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

← → ↻ × 🏠 <http://imgt.cines.fr/cgi-bin/IMGTdnajv> OK

Translation of the JUNCTION

	104																			118	
	C	A	K	R	F	W	S	G	Y	L	W	F	D	P	W	G	Q	G			
input	TGT	GCG	AAG	AGG	TTT	TGG	AGT	GGT	TAT	CTG	TGG	TTC	GAC	CCC	TGG	GGC	CAG	GGA			

Applet input [1}

IMGT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

<http://imgt.cines.fr/cgi-bin/IMGTdnap.jv>

Alignment with FR-IMGT and CDR-IMGT delimitations

```

<----- F R 1 - I M G T -----
input      1           5           10           15
M99660 IGHV3-23*01 GAG GTG CAG CTG TTG GAG TCG GGG GGA ... GGC TTG GTA CAG CCT GGG GGG TCC
J00236 IGHV3-23*02 --- --- --- --- --- --- --T --- --- --- --- --- --- --- --- ---
U29481 IGHV3-23*03 --- --- --- --- --- --- --T --- --- --- --- --- --- --- --- ---
Z12358 IGHV3-48*03 --- --- --- G--- --- --T --- --- --- --- --- --- --A --- ---
M99675 IGHV3-48*01 --- --- --- G--- --- --T --- --- --- --- --- --- --- --- ---

----->
input      20           25           30           35
M99660 IGHV3-23*01 CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTT AGC AGC TTT GCC ... ..
J00236 IGHV3-23*02 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
U29481 IGHV3-23*03 --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Z12358 IGHV3-48*03 --- --- --- --- --- --- --- --- --- --- --C --T --T -A -AA ... ..
M99675 IGHV3-48*01 --- --- --- --- --- --- --- --- --- --- --C --T --- -A -AG- ... ..

----- F R 2 - I M G T -----
input      40           45           50
M99660 IGHV3-23*01 ... .. ATG AGC TGG GTC CGC CAG GCT CCT GGG AAG GGG CTG GAC TGG GTC TCA
J00236 IGHV3-23*02 ... .. --- --- --- --- --- --- --- --A --- --- --- --G --- ---
U29481 IGHV3-23*03 ... .. --- --- --- --- --- --- --- --A --- --- --- --G --- ---
Z12358 IGHV3-48*03 ... .. --- --- --- --- --- --- --- --A --- --- --- --G --- --T ---
M99675 IGHV3-48*01 ... .. --- --- --- --- --- --- --- --A --- --- --- --G --- --T ---

```

Applet input [1]

Translation

```

<----- F R 1 - I M G T -----
1           5           10           15
input      E V Q L L E S G G           G L V Q P G G S
           GAG GTG CAG CTG TTG GAG TCG GGG GGA ... GGC TTG GTA CAG CCT GGG GGG TCC
M99660 IGHV3-23*01
           --- --T --- ... ---
    
```

```

----->
                CDR1 - IMGT
           20           25           30           35
input      L R L S C A A S G F T F S S F A
           CTG AGA CTC TCC TGT GCA GCC TCT GGA TTC ACC TTT AGC AGC TTT GCC ... ..
           Y
M99660 IGHV3-23*01
           --- --A--- ... ..
    
```

```

<----- F R 2 - I M G T -----
           40           45           50
input      M S W V R Q A P G K G L D W V S
           ... .. ATG AGC TGG GTC CGC CAG GCT CCT GGG AAG GGG CTG GAC TGG GTC TCA
           E
M99660 IGHV3-23*01
           ... .. --A--- --G---
    
```

```

-->                CDR2 - IMGT                <-----
           55           60           65           70
input      E I S G S G G T T           Y Y A G S V K
           GAA ATT AGT GGT AGT GGC GGT ACC ACA ... .. TAT TAC GCA GGC TCC GTG AAG
           A           S           D
M99660 IGHV3-23*01
           -CT --- --T --- -G--- ... .. --C --- --A---
    
```

IMGT/QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

http://imgt.cines.fr/cgi-bin/IMGTdnap.jv

Collier de Perles

input

Diagram illustrating a protein structure (Collier de Perles) with amino acid residues represented by beads. The residues are arranged in a complex, interconnected network. The residues are labeled with their single-letter codes: P, Q, V, L, G, S, E, L, L, S, A, S, M, W, V, R, Q, A, P, S, F, T, G, A, S, A, C, S, L, R, L, S, G, S, F, S, A, S, G, G, I, S, S, V, W, D, G, K, G, G, T, T, Y, A, G, S, V, K, G, R, K, N, D, S, T, L, F, Q, M, N, S, L, T, A, E, R, K, A, C, Y, V, A, T, D.

Legend for the diagram:

A	B	C	C'	C''	D	E	F
1	2	3	3a	3b	4	5	6
↓	↑	↓	↑	↓	↑	↓	↑
4.1	4.2	3.1			4.4	4.3	3.2

Applet input [1}

IMGT/V-QUEST

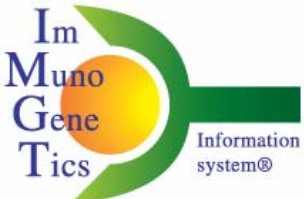
Some improvements for the next release

IMGTV-QUEST - Mozilla Firefox

Echier Edition Affichage Aller à Marque-pages Outils ?

WELCOME ! to IMGTV-QUEST

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



(none)

Citing IMGTV-QUEST: Giudicelli, V. et al. Nucl. Acids Res. 2004, 32, W435-440 [PMID: 15215425](#) PDF

Analyse your Immunoglobulin sequences

Your selection: Human

Your sequences are compared to the **Human IG set** from the [IMGTV-QUEST reference directory sets](#)

Nucleotide sequences

Enter your sequence(s) in FASTA format :

Type (or copy/paste) your sequence into the box below :

```
>AY867144
gatggttgatgactcagtcctccaggcaccctgtctttgtctccaggggaaagagccacc
ctctcctgcagggccagtcagactgttagcagcagttacttagcctggtaccagcagaaa
cttggccaggctcccaggctcctgatctatggtgcatccagtagggctagtggcatctca
gacaggttcagcggcagtggtctgggacagacttcacactcaacatcagcagactggag
cctgaagattttgcagtgat tactgtcagcagcatgat accctaccctctcactttcggc
ggagggaccagggtggtgatcaa
>test
acgttttatctccagcttggtccctggccaaaagtgtacgggtaagtgt
tataatggtggcagtaataagctgcaaaatcatcaggctgcaggctgctg
atggtgagagtgaattctgtccccggatccactgccgctgaaccttgatgg
```

Or give the path access to a local file containing your sequence :

Terminé

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

Case of complementary reverse sequences

Sequence number 1: test

Complementary reverse sequence compared with the [human IG set](#) from the [IMG T reference directory](#)

```
>test (complementary reverse)
caggtgcagctggtgcagctctgggggaggcgtagtccagcctgggaggtccttgagactc
tctgtgcagaggggggattcactttcagttcctattctatgcactgggtccgccaggcg
ccaggcaagggcctggaatgggtggcggttatgtggtttgatggaactgaaaaatactct
gcagagtccgtgaagggccgattcaccatctccagagacaattccaagaacacattgttt
ttgcaaatgaacagcctgagagccgacgacacggctgtgtattactgtgcgagagagcct
gattggttattatggggggacaggggtgctctggacgtctggggccaagggaccacggtc
accgtctcctcagacatcgtgatgaccagtcctcctccaccctgtctgcacatctgtagga
gacagagtaccatcacttgccgggcccagtcagggcattagtagctgggtggcctggtat
cagcagaaaccagggaaagcccctaagctcctgatgtatgaggegtctagtttagaaagt
ggggtcccatcaaggttcagcggcagtggtccgggacagaattcactctcaccatcagc
agcctgcagcctgatgatatttgagcttattactgccaacattataaacacttacccgtac
acttttgccaggggaaccaagctggagataaaaact
```

Terminé

Type (or copy/paste) your sequence into the box below :

```
>AY867144
gatgttgtgatgactcagtcctccagggaccctgtctttgtctccaggggaaagagccacc
cttcctgcagggccagtcagactgttagcagcagttacttagcctggtaccagcagaaa
cttggccaggtcccaggtcctgatctatggtgcatccagtagggttagtggtcctca
gacaggttcagcggcagtggtctgggacagacttcacactcaacatcagcagactggag
cctgaagattttgcagtgatattactgtcagcagcatgatacctcacctctcacttcggc
ggagggaccaggtggtgatcaaa
>test
acgttttatctccagcttggttcctggccaaaagtgtacgggtaagtgt
tataatgttggcagtaataagctgcaaaatcatcaggtgcaggtgctg
atggtgagagtgaattctgtcccggatccactgcccgtgaaccttgatgg
```

Or give the path access to a local file containing your sequence :

Selection of ouput parameters

Output parameters

Display type :	HTML												
Includes JCTA :	yes												
Results :	<p>1. Details Select the fields to be displayed :</p> <table border="0"> <tr> <td><input checked="" type="checkbox"/> Alignment for V-GENE</td> <td><input checked="" type="checkbox"/> Alignment for D-GENE</td> </tr> <tr> <td><input checked="" type="checkbox"/> Alignment for J-GENE</td> <td><input checked="" type="checkbox"/> JUNCTION</td> </tr> <tr> <td><input checked="" type="checkbox"/> Alignment with FR-IMGT and CDR-IMGT delimitations</td> <td><input checked="" type="checkbox"/> Translation</td> </tr> <tr> <td><input checked="" type="checkbox"/> Mutation statistics</td> <td><input checked="" type="checkbox"/> Annotations</td> </tr> <tr> <td><input checked="" type="checkbox"/> "input" V-REGION for IMGT/Phylogene</td> <td><input checked="" type="checkbox"/> Collier de Perles</td> </tr> <tr> <td><input checked="" type="checkbox"/> Mutation table</td> <td></td> </tr> </table>	<input checked="" type="checkbox"/> Alignment for V-GENE	<input checked="" type="checkbox"/> Alignment for D-GENE	<input checked="" type="checkbox"/> Alignment for J-GENE	<input checked="" type="checkbox"/> JUNCTION	<input checked="" type="checkbox"/> Alignment with FR-IMGT and CDR-IMGT delimitations	<input checked="" type="checkbox"/> Translation	<input checked="" type="checkbox"/> Mutation statistics	<input checked="" type="checkbox"/> Annotations	<input checked="" type="checkbox"/> "input" V-REGION for IMGT/Phylogene	<input checked="" type="checkbox"/> Collier de Perles	<input checked="" type="checkbox"/> Mutation table	
<input checked="" type="checkbox"/> Alignment for V-GENE	<input checked="" type="checkbox"/> Alignment for D-GENE												
<input checked="" type="checkbox"/> Alignment for J-GENE	<input checked="" type="checkbox"/> JUNCTION												
<input checked="" type="checkbox"/> Alignment with FR-IMGT and CDR-IMGT delimitations	<input checked="" type="checkbox"/> Translation												
<input checked="" type="checkbox"/> Mutation statistics	<input checked="" type="checkbox"/> Annotations												
<input checked="" type="checkbox"/> "input" V-REGION for IMGT/Phylogene	<input checked="" type="checkbox"/> Collier de Perles												
<input checked="" type="checkbox"/> Mutation table													
Nb of nucleotides per line in alignment:	90												
Nb of nucleotides to exclude in 5' of the V gene for the evaluation of the nb of mutations :	<input type="text"/>												

IMGT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

Alignment for J-GENE

Closest J-GENES

	Score	Identity
J00242 IGKJ4*01	152	94.11% (32/34 nt)
AF103571 IGKJ4*02	143	91.17% (31/34 nt)
J00242 IGKJ3*01	107	79.41% (27/34 nt)
Z70260 IGKJ2*02	107	79.41% (27/34 nt)
J00242 IGKJ2*01	107	79.41% (27/34 nt)

Score and nucleotide identity

Alignment

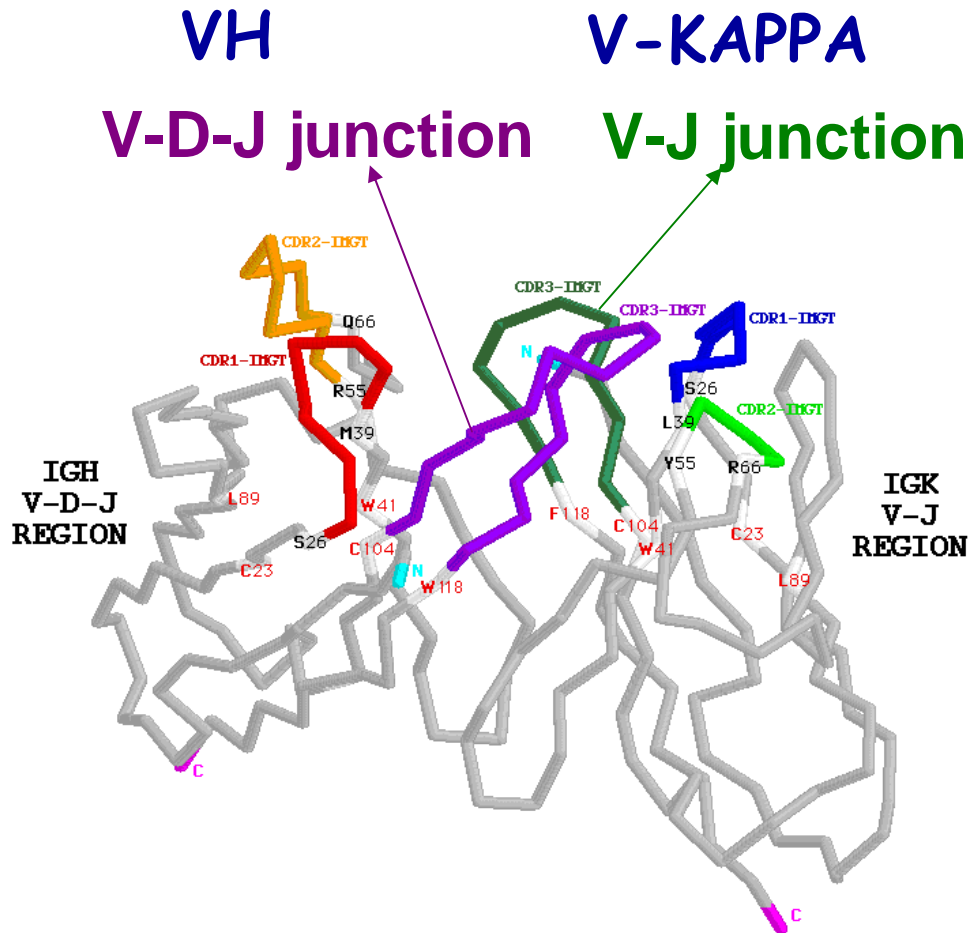
```

AY867144          cactttcggcggaggaccagggtggtgatcaaa
J00242 IGKJ4*01  -----a-----a-----c
AF103571 IGKJ4*02 ---g-----a-----a-----c
J00242 IGKJ3*01  -----cct-----aa----at-----c
Z70260 IGKJ2*02  -----t--cag-----a-c--a-----c
J00242 IGKJ2*01  -----t--cag-----a-c--a-----c
  
```

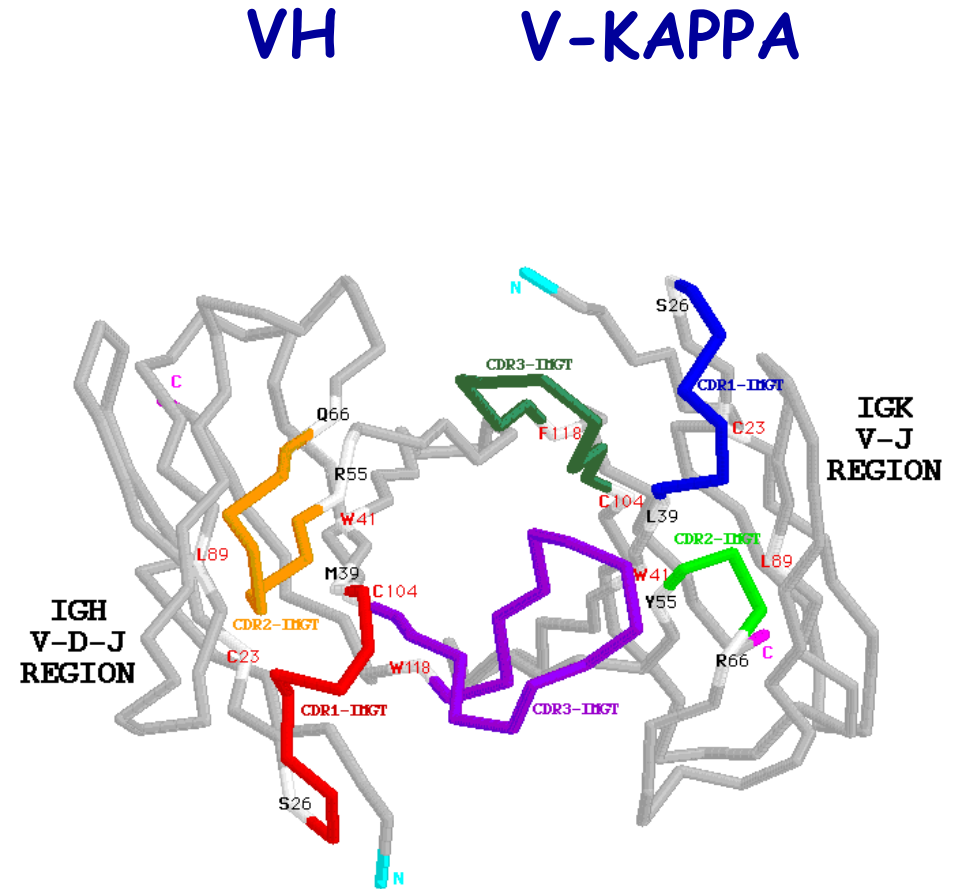
Terminé

IMGT/JunctionAnalysis

V-DOMAINS: VH and V-KAPPA



Side view of the V-DOMAINS



View from above the CDRs

Mouse (*Mus musculus*) E5.2Fv

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

V-J junction (104-118)

V-D-J junction (104-118)

Immunoglobulin V-D-J generation of sequence diversity



JUNCTION

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

THANK YOU for using [IMGT/JunctionAnalysis](#)

THE
INTERNATIONAL
IMMUNOGENETICS
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	tgttccgagaga	aga	.tagcaatggctacaa....	aata
#2 Z47269	IGHV1-69*06	tgtgcgagag.	gggggggctaaggtogaatttttggagtggtt.....	tcatgggt

Input	J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
#1 M62724tttaactactag	IGHJ4*02	IGHD5-24*01	0	2	0	1/7
#2 Z47269	...actggttcgaccctgg	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	112.2	112.1	112	113	114	115	116	117	118	frame	CDR3-IMGT length	
#1 M62724	C	A	R	E	D	S	N	G							Y	K	I	F	D	Y	W	+	13	
#2 Z47269	C	A	R	G	G	A	K	V	E	F	L	E	W	F	H	G	V	W	F	D	P	W	+	20

-> [IMGT/JunctionAnalysis Search page](#)
-> [IMGT/JunctionAnalysis Documentation](#)

IMGT/JunctionAnalysis - Mozilla Firefox

Eichier Edition Affichage Aller à Marque-pages Outils ?

http://imgt.cines.fr/cgi-bin/IMGTjcta.jv?livret=0

<http://imgt.cines.fr>

Citing IMGT/JunctionAnalysis: Yousfi Monod, M. et al. Bioinformatics 2004, 20 Supplement 1:1379-1385 [PMID: 15262823](https://pubmed.ncbi.nlm.nih.gov/15262823/)

IMGT/JunctionAnalysis Search page

Species :

Homo sapiens

Locus :

IGH TRA
 IGK TRB
 IGL TRG
 TRD

Sequences

Type (or copy/paste) required information and nucleotide sequences into the box below :

```
>M62724, IGHV7-4-1*02, IGHJ4*02
TGT GCG AGA GAA GAT AGC AAT GGC TAC AAA ATA TTT GAC TAC TGG
>Z47269, IGHV1-69*06, IGHJ5*02
TGT GCG AGA GGG GGG GCT AAG GTC GAA TTT TTG GAG TGG TTT CAT GGG TAC TGG
TTC GAC CCC TGG
```

Or give the path access to a local file containing your sequences :

Parcourir...

Start Clear the form

Terminé

THANK YOU for using IMGT/JunctionAnalysis

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>

Analysis of the JUNCTIONS

Input	V name	V-REGION	N1	D-REGION	N2	J-REGION	J name
#1	M62724	IGHV7-4-1*02	tgtgcgagaga	aga	.tagcaatggctacaa....	aatatttgactactgg IGHJ4*02
#2	Z47269	IGHV1-69*06	tgtgcgaga	gggggggctaaggtcgaatttttggagtggtt.....	tcatgggt	...actggttcgacccctgg IGHJ5*02

Input	D name	Vmut	Dmut	Jmut	Ngc	
#1	M62724	IGHD5-24*01	0	2	0	1/7
#2	Z47269	IGHD3-3*02	0	2	0	14/21

Translation of the JUNCTIONS

	105	107	109	111	111.2	112.4	112.2	112	114	116	118	CDR3-IMGT													
	104	106	108	110	111.1	111.3	112.3	112.1	113	115	117	frame 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	E	F	L	E	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

Terminé

IMGT/JunctionAnalysis Results

Locus IGH
Species Homo sapiens
IMGT repertoire 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:

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		tgtgtgtgcagcgcctgttac	ccaaatatc		...acttgaccactgg	IGHJ4*02	IGHD6-13*01	1	2	1	5/15
#2	Z70257	IGHV3-7*02		ggatggcagctcttatgcc	cgccc		ctactggtacttcgatctctgg	IGHJ2*01	IGHD2-2*01	0	2	0	9/11
#3	Z70606	IGHV4-31*03		c	.gactacg.....	cact		..atgcttttgatgtctgg	IGHJ3*01	IGHD4-17*01	0	0	0	3/5
#4	Z70608	IGHV4-39*05		cagagtaacgatttttgagtggttatt....	ccccggggga		..atgcttttgatctctgg	IGHJ3*02	IGHD3-3*01	0	0	0	12/17
#5	Z70610	IGHV4-34*09		tcgggagcgatttttgagtggttatt....	cccga	ca	tgatgctttgatctctgg	IGHJ3*02	IGHD3-3*01	0	0	0	9/12
#6	Z70611	IGHV4-59*01		catggttaactataa.	tgccggcggtg		...actggttcgaccctgg	IGHJ5*02	IGHD3-9*01	0	2	0	9/13
#7	Z70613	IGHV4-59*01		cagcagctgttac	ctccct		...cttgactactgg	IGHJ4*02	IGHD6-13*01	0	0	0	4/6
#8	Z70614	IGHV4-59*01		cactataattcggggacttat.....	cccctc	gactactgg	IGHJ4*02	IGHD3-16*01	0	2	0	7/14
#9	Z70615	IGHV4-59*01		ggctg	gtaaagaggg.....	tttcggaa		.tactggtacttcgatctctgg	IGHJ2*01	IGHD5-24*01	0	2	0	7/13
#10	Z70616	IGHV4-34*01		cgggtttggg.....	ttccc		...actggttcgaccctgg	IGHJ5*02	IGHD3-16*01	0	0	0	6/8
#11	Z70620	IGHV4-30-4*01		ccggggcgggatggtt....	cgg		.gatgctttgatctctgg	IGHJ3*02	IGHD3-16*01	1	4	0	5/5
#12	Z70621	IGHV4-39*01		ccacgattatggttcggggagtt.....	tgaccccc	ttgactactgg	IGHJ4*02	IGHD3-16*01	0	1	0	12/21
#13	Z70622	IGHV4-39*06	t	tgccccctcctgccccaaat	gtattactatggttcgggga.....	tatgtacg	ttgactactgg	IGHJ4*03	IGHD3-10*01	0	0	0	15/28

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

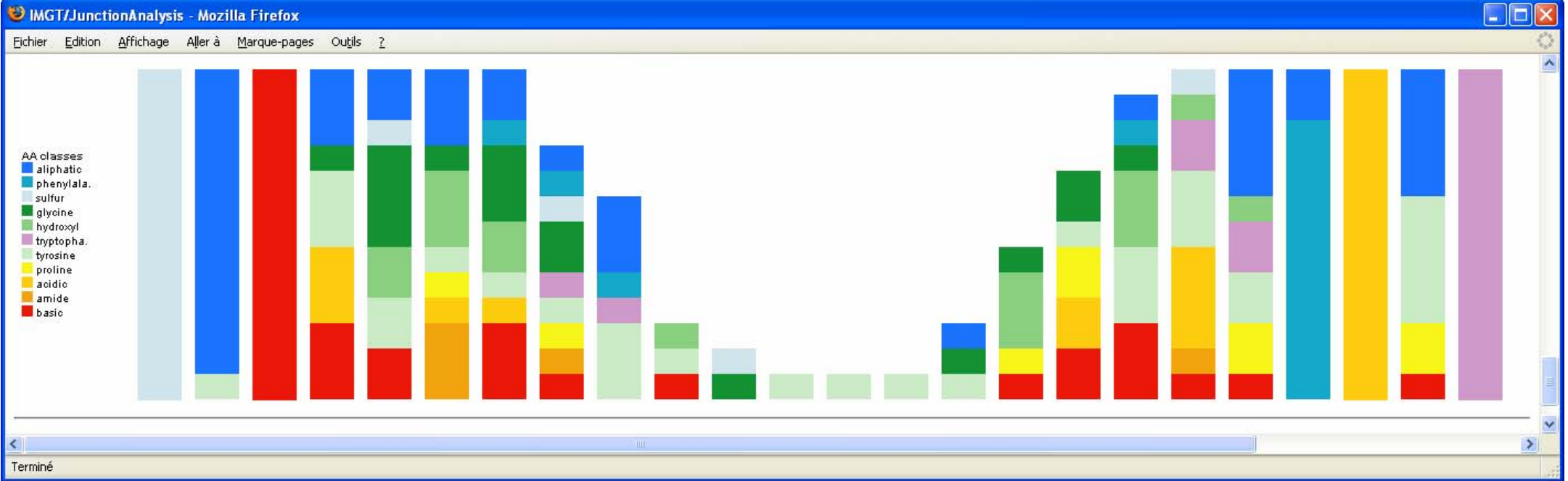
'Volume' classes		'Hydropathy' classes						
	in Å ³	Hydrophobic		Neutral	Hydrophilic			
Very large	189-228	F	W	Y				
Large	162-174	I L	M		H K R			
Medium	138-154	V				E Q		
Small	108-117		C	P	T	D N		
Very small	60-90	A		G	S			
		Aliphatic		Sulfur	Hydroxyl	Basic	Acidic	Amide
		Nonpolar			Uncharged	Charged	Uncharged	
					Polar			

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

JUNCTION alignments

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			
#1 Z70256	C	<u>V</u>	R	V	V	Q	<u>R</u>	L	V									P	K	Y	H	F	D	<u>H</u>	W	+	15	2,438.76		
#2 Z70257	C	A	R	D	G	S	<u>S</u>	Y	A									R	P	Y	W	Y	F	D	L	W	+	16	2,256.49	
#3 Z70606	C	A	R	A	T	T	H												Y	A	F	D	V	W	+	11	1,604.77			
#4 Z70608	C	A	R	V	T	I	F	G	V	V								I	P	R	G	N	A	F	D	I	W	+	18	2,426.78
#5 Z70610	C	A	R	V	G	S	D	F	W	S	G							Y	S	R	H	D	A	F	D	I	W	+	19	2,539.73
#6 Z70611	C	A	R	H	G	<u>N</u>	Y	N	A									G	V	D	W	F	D	P	W	+	15	2,072.21		
#7 Z70613	C	A	R	A	A	A	G												T	S	L	F	D	Y	W	+	12	1,531.71		
#8 Z70614	C	A	R	H	Y	N	<u>S</u>	G											<u>I</u>	Y	P	L	D	Y	W	+	13	1,846.01		
#9 Z70615	C	A	R	G	L	V	<u>K</u>	<u>R</u>	V									S	E	Y	W	Y	F	D	L	W	+	16	2,419.75	
#10 Z70616	C	A	R	A	G	L	G												S	H	W	F	D	P	W	+	12	1,602.8		
#11 Z70620	C	A	R	D	R	G	<u>G</u>	M											<u>V</u>	R	D	A	F	D	I	W	+	14	1,932.17	
#12 Z70621	C	A	R	H	H	D	L	W	F										G	E	F	D	P	L	D	Y	W	+	16	2,307.53
	C	A	R	D	C	P	A	P	A	R	M	Y	Y	Y	G	S	G	I	C	T	F	D	Y	W	+					



IMGT/3Dstructure-DB and IMGT/StructuralQuery

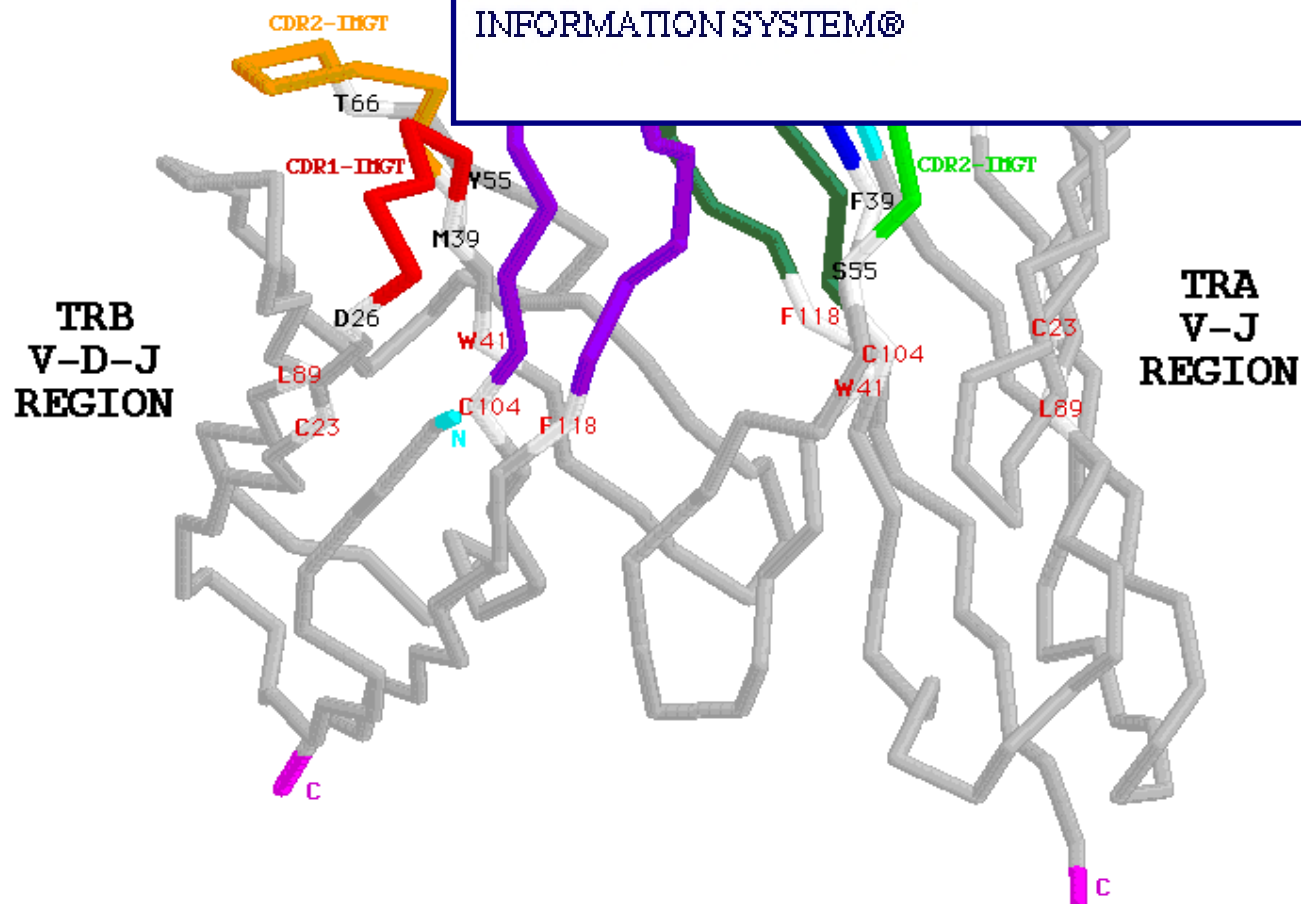
WELCOME !

to **IMGT/3Dstructure-DB**

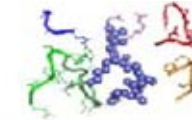
THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<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		<i>Homo sapiens</i> (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	<pre>[IGHV_1 RITLKEGSPPLVKPTQTLTLTCSFSGFSLSDFGVGVGMIRQPPGKALEWLAIIYSDDDKRYSPSINTRLTI TKDTSINQVVLMMTRVSPV][D][IGHJ_1][DTATYFCAHRRGPTTLFGVPIARGPNAMDWVWQGGITVTI SSTSTKGPSVFPPLAPCS RSTAGGTAALGCLWKDYFPEPVTVSWNSGALTS HC_1 GVHTFPAVLQSSGLYSLSSVVTVPSSSLGDTYTNVNHKPSNTRKVDKRVKPS]</pre>	
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
	Sheet composition	[A B D E] [C C' C' F G]
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]
		<pre>.....STKGPSVFPPLAPCS RSTA...GGTAALGCLWKDYFP...EPVTVSWNSGALTS...GVHTFPAVLQSS...GLYSLSSVVTVP PSSSL...GDTYTNVNHK...SNTRKVDKRV</pre>

Chain ID		1u8k_A
IMGT chain description	L-KAPPA	
Chain amino acid sequence	<pre>[IGKV_1 ALQLTQSPSSLSASVGDRI FTI TCRASQGVTSALAWYRQKPGSPQQLLIYDASSLESQVSRFSGSGSGTEFTLTITSLRPEDFATYYQQ][IGKJ_1][LHFYPHTFGGGTRVDVRRVVAAPS VFI FPPSDRQLKSGTASVWQLINNFYPREARVQWQVINALQSGNSQESVTEQDSKDSYSLSTLT] LSKADYEKHKVYEVVTHQGLSSPVTKSFNRGEC]</pre>	
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
	Sheet composition	[A B D E] [C C' C' F G]
C-DOMAIN	IMGT domain description	C-KAPPA
	IGKC gene and allele name	IGKC*01 (98.13%) Alignment details
	2D representation	Collier de Perles or Collier de Perles on 2 layers
	Sheet composition	[A B D E] [C F G]
		<pre>...RIVAAPS VFI FPPSDRQLK...SGTASVWQLINNFYP...REARVQWQVINALQSGNSQESVTEQDSKDS...TYSLSSTLTSL SKADY...EKHKVYEVVTHQGLSSPVTKSFNRGEC</pre>
Chain ID	1u8k_C	
IMGT chain description	Peptide	
Chain amino acid sequence	LELDKWASL	

1u8k_B

Region identification and delimitation

Chain ID	1u8k_B	
IMGT chain description	VH-CH1	
Chain amino acid sequence	<pre>[IGHV_1 RI TLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAI I YSDDDKRYSPSINTRLTI TKDTSKNQVVLVMTRVSPV] [D] [IGHJ_1] [IG DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI SSTSTKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1] GVHTFFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTQNVNHKPSNTKVDKRVPEPKS</pre>	
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]
	<pre> [CDR1] [CDR2] RI TLKESGP . PLVKPTQTLTLTCSFSGFSLSDFGVGV . VGWIRQPPGKALEWLAI I YSDDDK . . RYSPSIN . TRLTI TKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI</pre>	
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]
	<pre>. STKGPSVFPLAPCSKSTA . . GGTAALGCLVKDYFP . . EPVTVSWNSGALTS GVHTFFPAVLQSS GLYSLSSVVTV PSSSL . . . GTQTYTQNVNHK . . . SNTKVDKRV</pre>	

Chain ID	1u8k_B	
IMGT chain description	VH-CH1	
Chain amino acid sequence	<pre> [IGHV_1 RI TLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAI I YSDDDKRYSPSINTRLTI TKDTSKNQVVLVMTRVSPV] [D] [IGHJ_1] [IG DTATYFCAHRRGPTTLFGVPIARGPVMNDVWGQGI TVTI SSTSTKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1 GVHTFPAVLQSSGLYSLSSVTV </pre>	
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]
<pre> [CDR1] [CDR2] RI TLKESGP . PLVKPTQTLTLTCSFSGFSLSDFGVGV . VGWIRQPPGKALEWLAI I YSDDDK . . RYSPSIN . TRLTI TKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVMNDVWGQGI TVTI </pre>		
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]
<pre> STKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS GVHTFPAVLQSS GLYSLSSVTV PSSSL . . . GTQTYTQNVNHKP . . SNTKVDKRV </pre>		

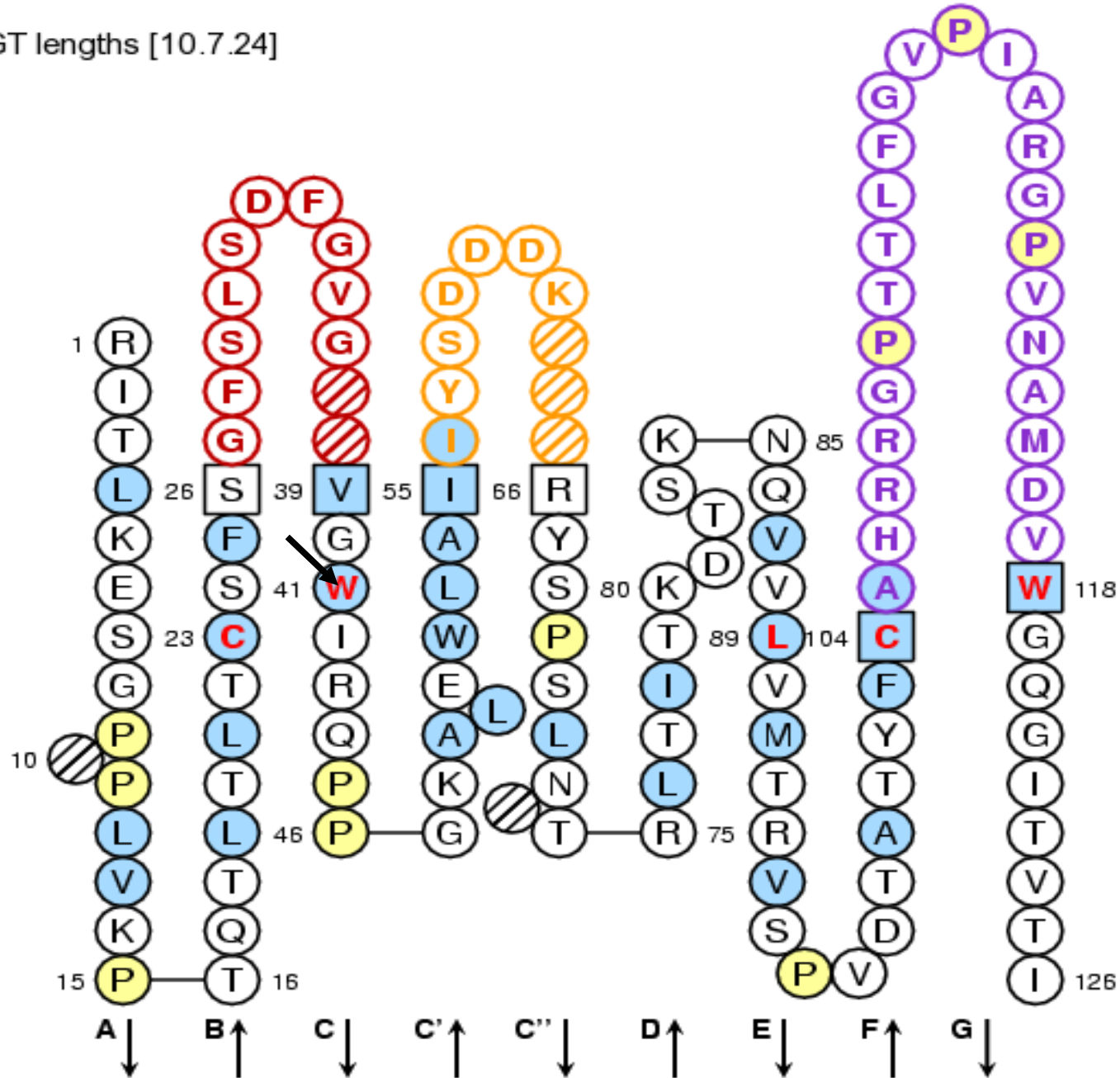
Identification of the IMGT gene and the IMGT closest allele(s)

Chain ID	1u8k_B	
IMGT chain description	VH-CH1	
Chain amino acid sequence	<pre> [IGHV_1 RI TLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAI I YSDDDKRYSPSINTRLTI TKDTSKNQVVLVMTRVSPV] [D] [IGHJ_1] [IG DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI SSTSTKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1] GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTQNVNHKPSNTKVDKRVPEPKS </pre>	

V-DOMAIN	IMGT domain description	VH	Description of domains	
	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]		CDR-IMGT lengths
	Sheet composition	[A B D E] [C C' C' F G]		
		<pre> [CDR1] [CDR2] RI TLKESGP . PLVKPTQTLTLTCSFSGFSLSDFGVGV . . VGWIRQPPGKALEWLAI I YSDDDK . . . RYSPSIN . TRLTI TKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI </pre>	Gaps according to IMGT numbering	

C-DOMAIN	IMGT domain description	CH1	Gaps according to IMGT numbering	
	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]		
		<pre> STKGPSVFPLAPCSKSTA . . . GGTAALGCLVKDYFP . . EPVTVSWNSGALTS GVHTFPAVLQSS GLYSLSSVVTV PSSSL . . . GTQTYTQNVNHK . . . SNTKVDKRV </pre>		

CDR-IMGT lengths [10.7.24]



Contact analysis

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

IMGT Collier de Perles

Collier de perles - Netscape

Fichier Edition Afficher Aller Communicator Aide

Signets Adresse : gt3d.igh.cnrs.fr/cgi-bin/collier_perles2P.cgi?domcode=1HZHHD00&name=B12&CDR1=8&CDR2=8&CDR3=20&species=HUMAN Infos connexes

Collier de Perles : HUMAN IGHV V-DOMAIN from B12 (PDB: 1hzh_H)

[8.8.20]

Document : chargé

IMGT, the international ImMunoGeneTics information system®

Created in 1989 (Université Montpellier II and CNRS)

1) IMGT domains of research:

Immunogenetics -Immunoinformatics

- * *Immunoglobulins (IG), T cell receptors (TR), major histocompatibility complex (MHC) of human and other vertebrates,*
- * *Immunoglobulin superfamily IgSF and MhcSF,*
- * *Related proteins of the immune system (RPI)*

2) IMGT-ONTOLOGY concepts

3) IMGT-Choreography: *Dynamic knowledge management*

Structural domains

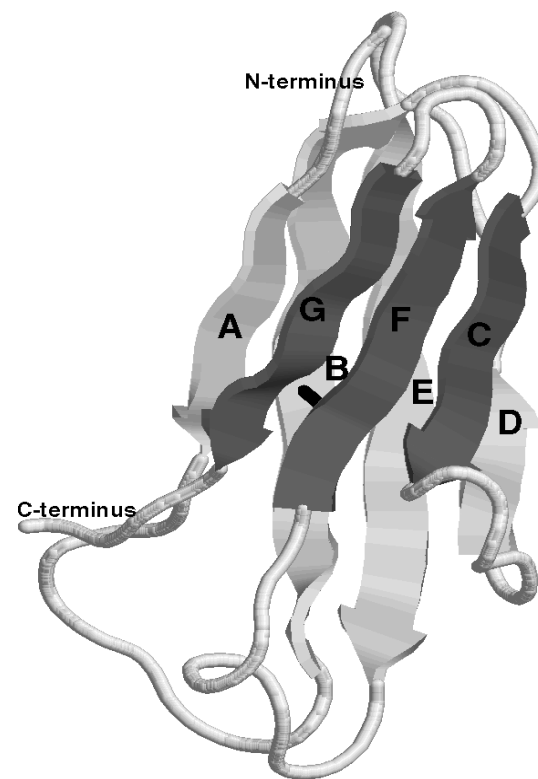
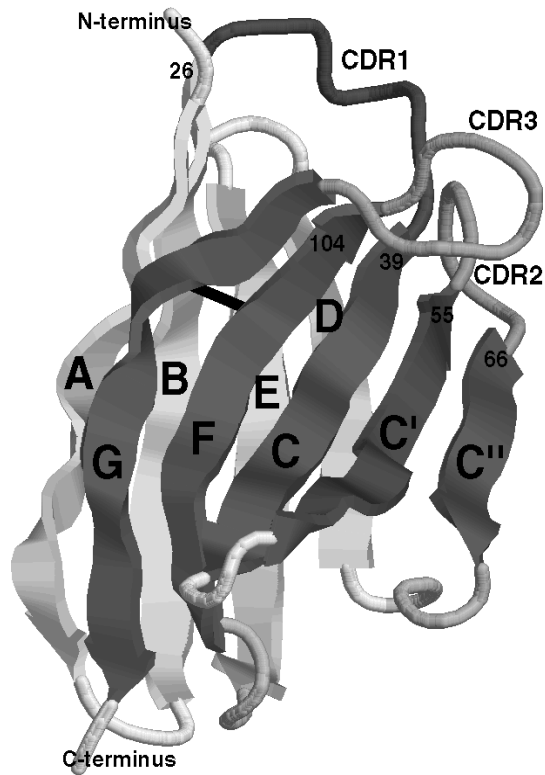
IG and TR

MHC

V-DOMAIN

C-DOMAIN

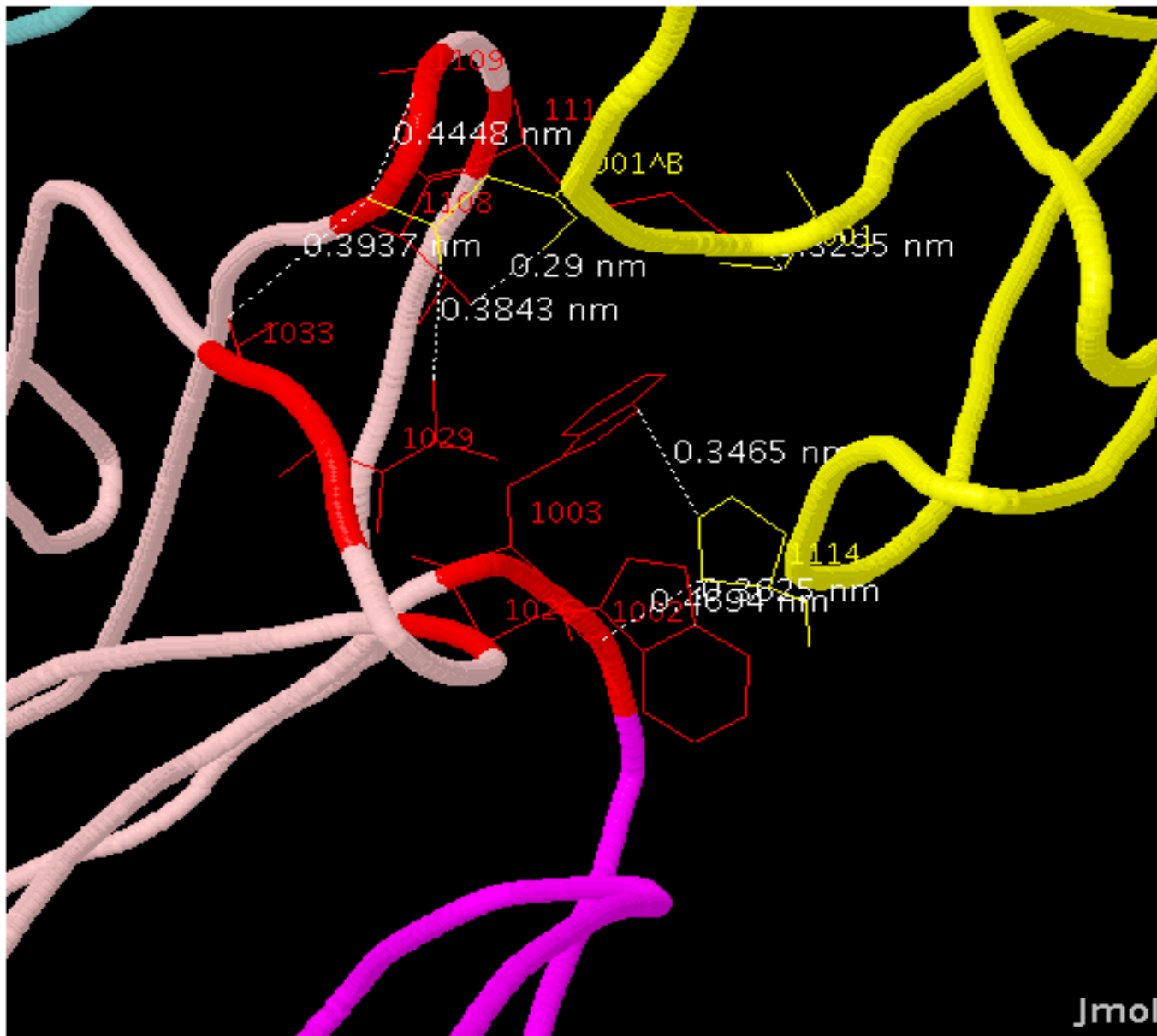
G-DOMAINS



Interactions between domains

FCGR3B

[D2]
C-LIKE-
DOMAIN



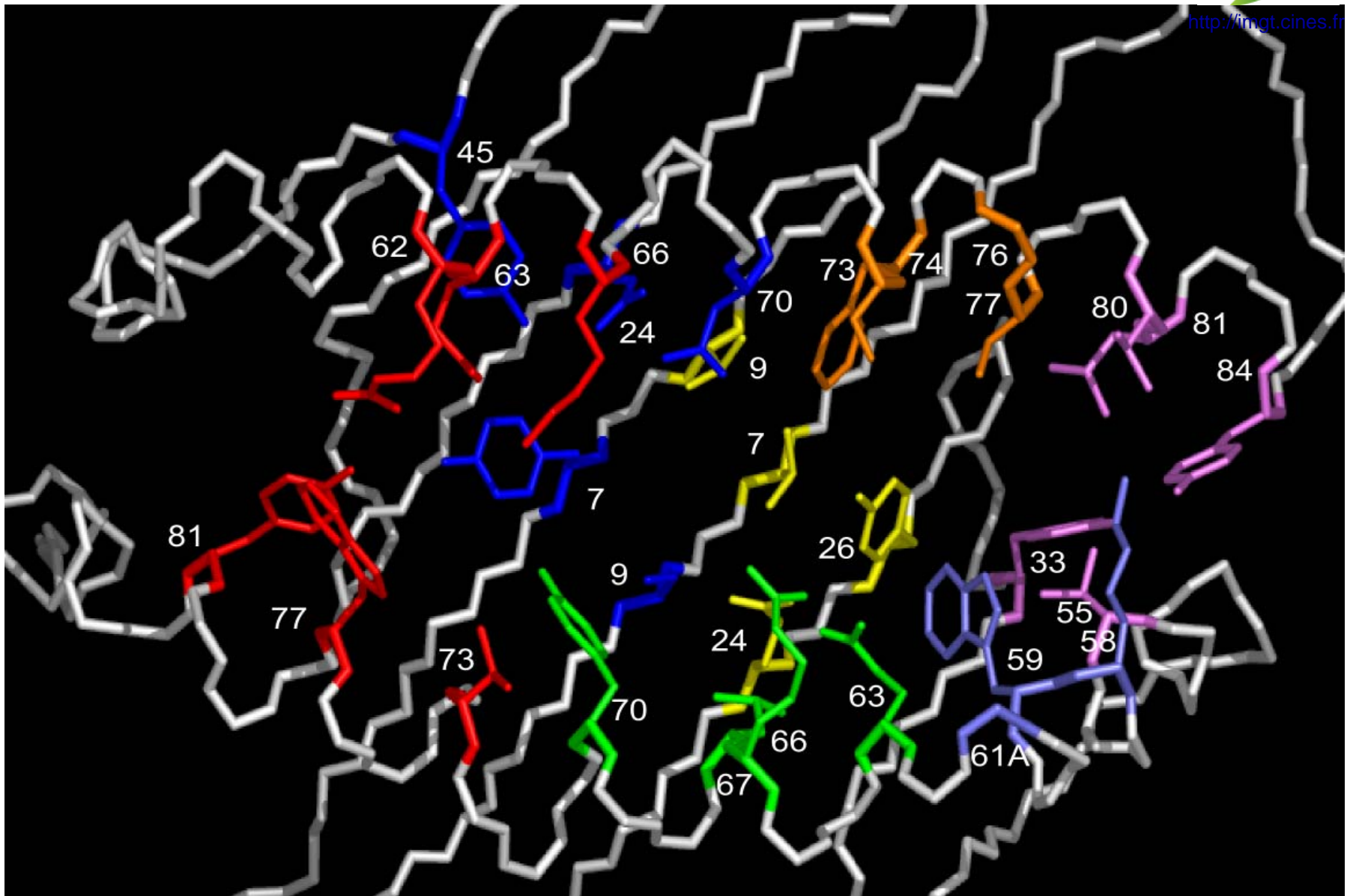
[D1] C-LIKE-DOMAIN

IGHG1
(FC-GAMMA1)

CH2
C-DOMAIN

MHC class I IMGT contact sites

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

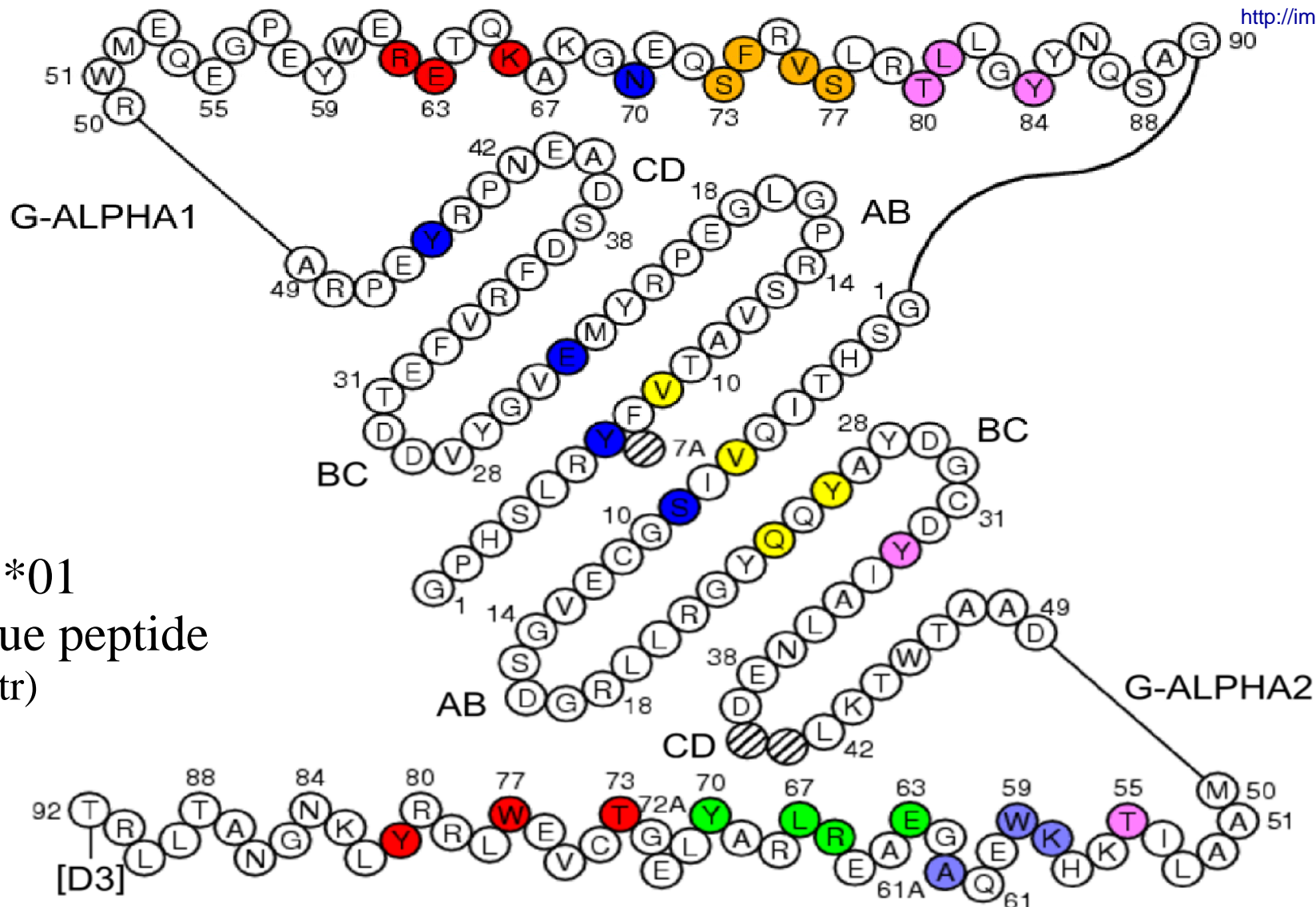
(code 1jtr)

8 residue peptide

(Kaas et al. In press)

MHC class I IMGT contact sites

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)

IMGT, the international ImMunoGeneTics information system®
<http://imgt.cines.fr>



<http://imgt.cines.fr>



Im
Muno
Gene
Tics

Information
system®

<http://imgt.cines.fr>