

Immunoinformatique

Marie-Paule Lefranc

Laboratoire d'ImmunoGénétique Moléculaire
Université Montpellier 2, UPR CNRS 1142, IGH
Institut Universitaire de France

Master recherche 2 Bio-Med, TC3,
Universités Montpellier 1 et 2, ENSCM, 27 Septembre 2007

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?

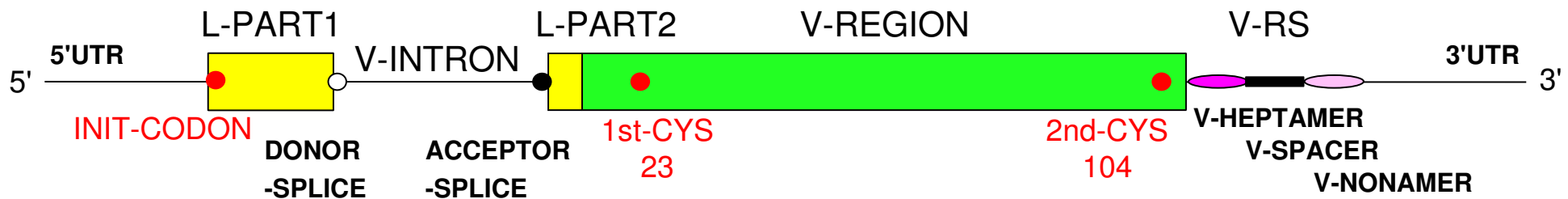
1. Unique structure of the immunoglobulin (IG) and T cell receptor (TR) genes.

Genomic DNA in germline configuration

V-GENE

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

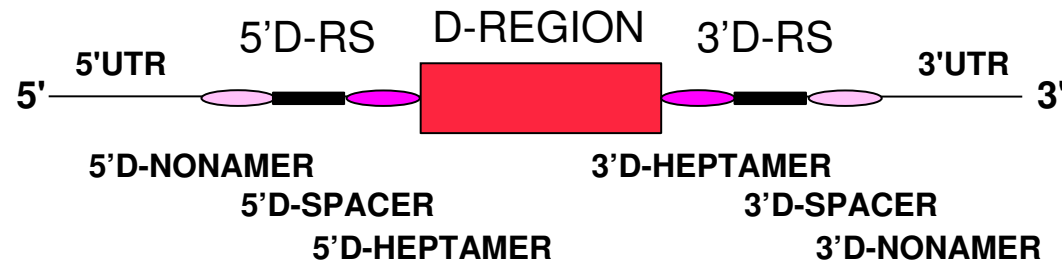
tgagagctcc	gttcctcacc	atggactgga	cctggaggat	cctcttcttg	gtggcagcag	60
ccacaggtaa	gaggctccct	agtcccagtg	atgagaaaga	gattgagtc	agtccaggga	120
gatctcatcc	acttctgtgt	tctctcca	gaaggagccact	ccaggtgca	gctggtgcag	180
tctggggctg	aggtgaagaa	gcctggggcc	tcagtgaagg	tctcctgcaa	ggcttctgga	240
tacaccttca	ccggctacta	tatgcactgg	gtgcgacagg	cccctggaca	agggcttgag	300
tggatgggat	ggatcaacc	taacagtgg	ggcaciaaact	atgcacagaa	gtttcagggc	360
agggtcacca	tgaccaggga	cacgtccatc	agcacagcct	acatggagct	gagcaggctg	420
agatctgacg	acacggccgt	gtattactgt	gcgagagaca	cagtggtgaaa	acccacatcc	480
tgagggtgtc	agaaacc	caa	gggaggaggc	ag		



Genomic DNA in germline configuration

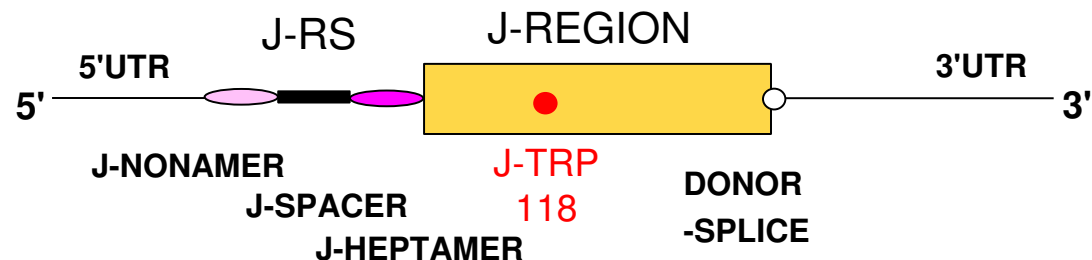
D-GENE

ggcagctcta c ggtttttggc c tgagctgaga ac cactgtgc taactgggga cacagtcatt
ggcagctcta c ggtttttggc c tgagctgaga ac cactgtgc taactgggga cacagtcatt



J-GENE

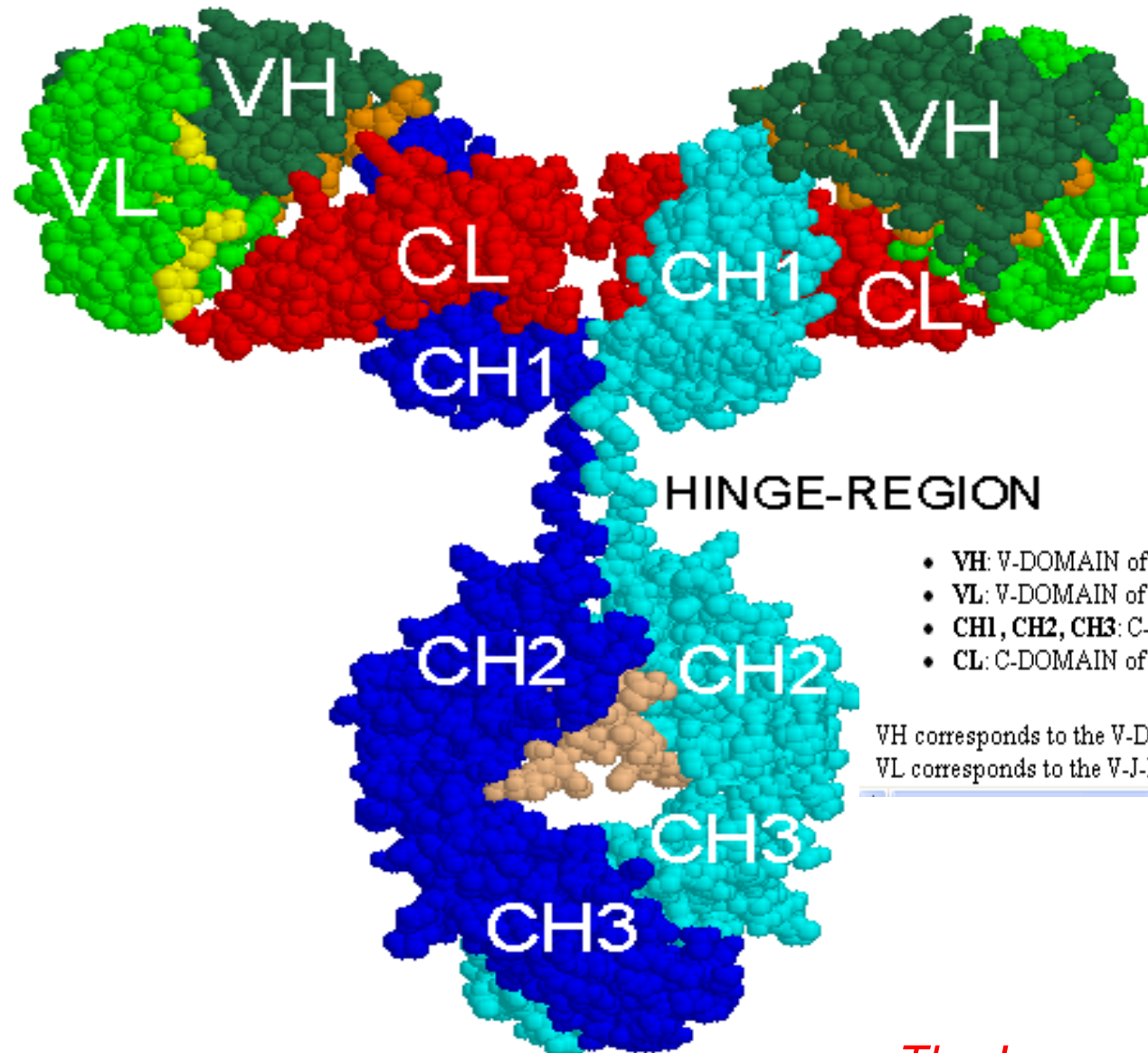
gggtttct gt gcccctgg ctcagggtg act caccgtg gctgaatact
tccagcactg gggccagggc accctgggtca ccgtctctc ag



Why was IMGT necessary?

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

Spacefill 3D representation of an IgG

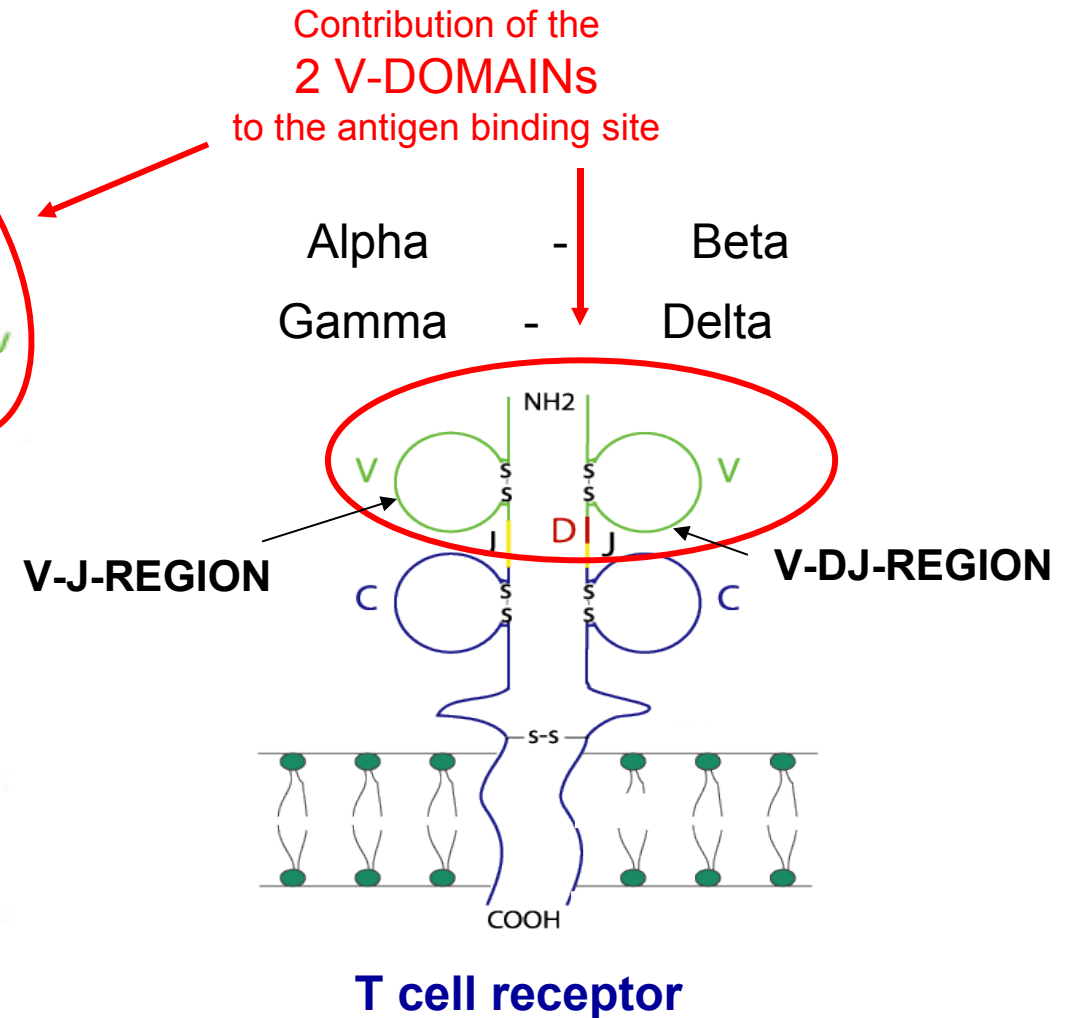
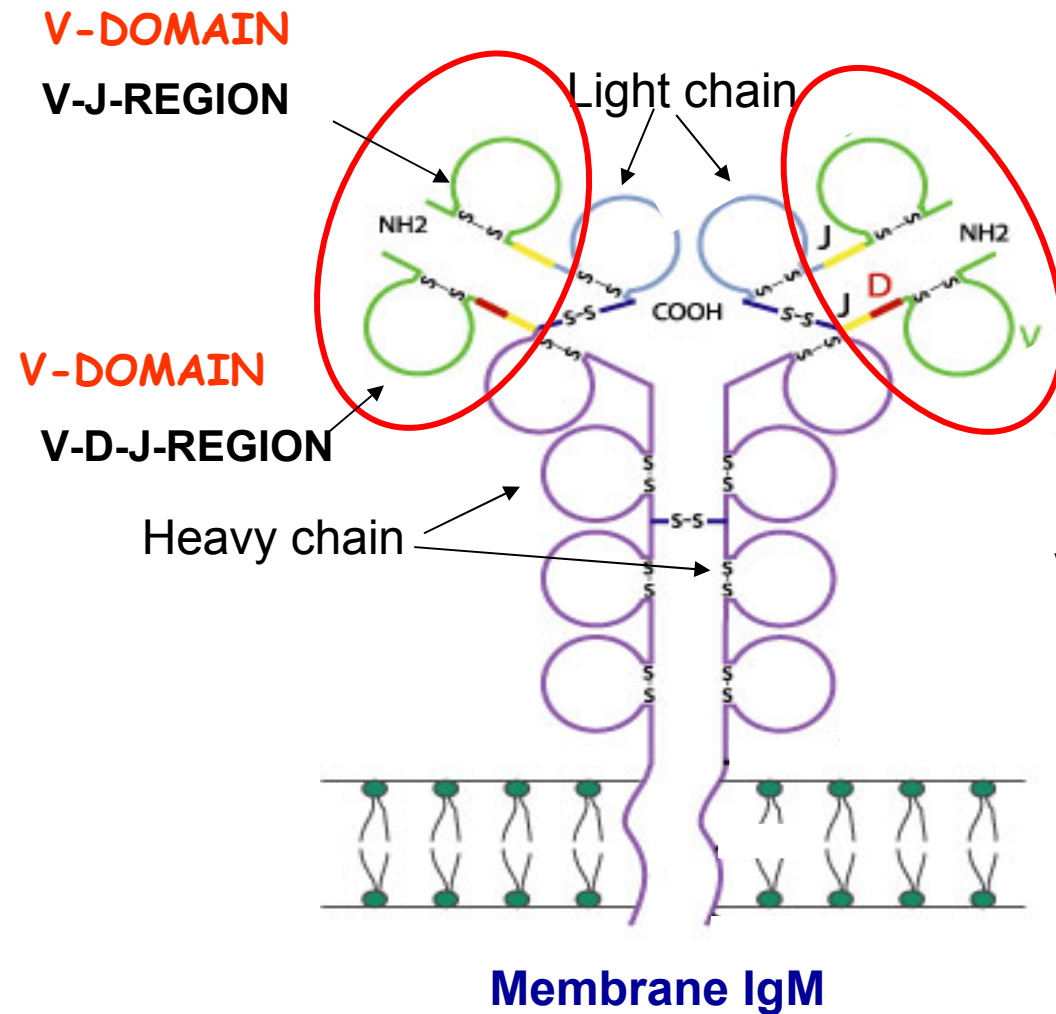


- VH: V-DOMAIN of the immunoglobulin heavy chain
- VL: V-DOMAIN of the immunoglobulin light chain
- CH1, CH2, CH3: C-DOMAIN of the immunoglobulin heavy chain
- CL: C-DOMAIN of the immunoglobulin light chain

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.

Immunoglobulin (IG)

T cell receptor (TR)

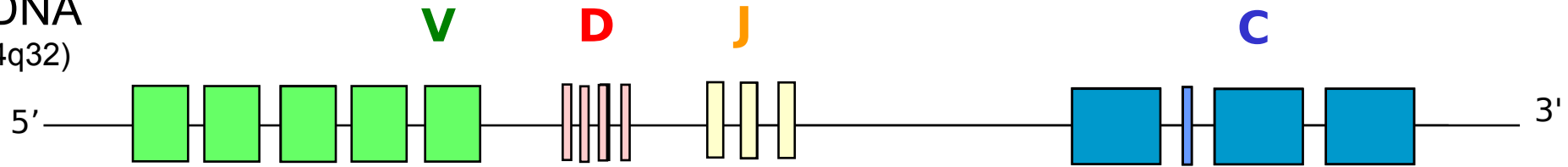


ImmunoGeneTics
Information system®

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

Immunoglobulin (IG) synthesis

genomic DNA
(IGH Locus 14q32)

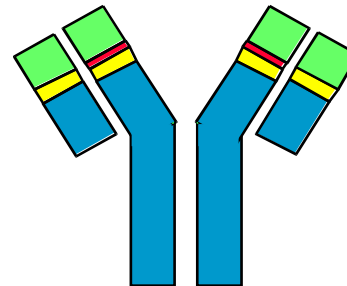


rearranged
DNA



mRNA

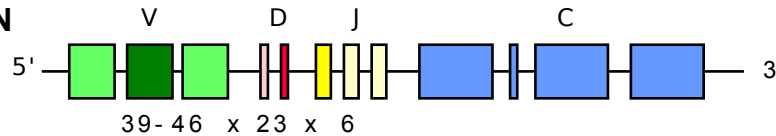
2×10^{12} different IG
per individual



Immunoglobulin (IG) synthesis

150 FUNCTIONAL IG GENES

HEAVY CHAIN



6300

POTENTIAL RECOMBINATIONS

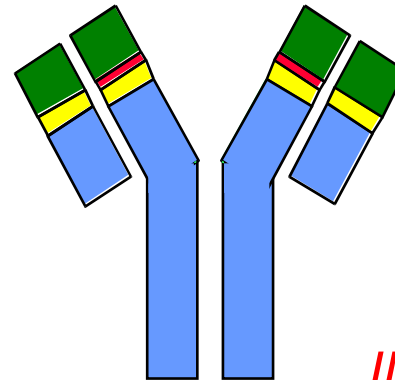


ABOUT 6.3×10^6 POSSIBILITIES

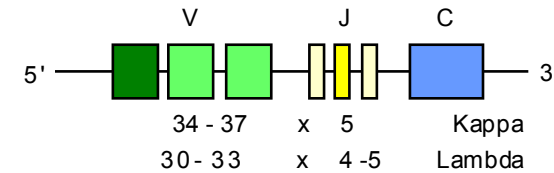
N-DIVERSITY
SOMATIC MUTATIONS

$\times 1000$

2×10^{12}
DIFFERENT ANTIBODIES



LIGHT CHAIN



185 + 165

POTENTIAL RECOMBINATIONS



ABOUT 3.5×10^5 POSSIBILITIES

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

[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,](#)

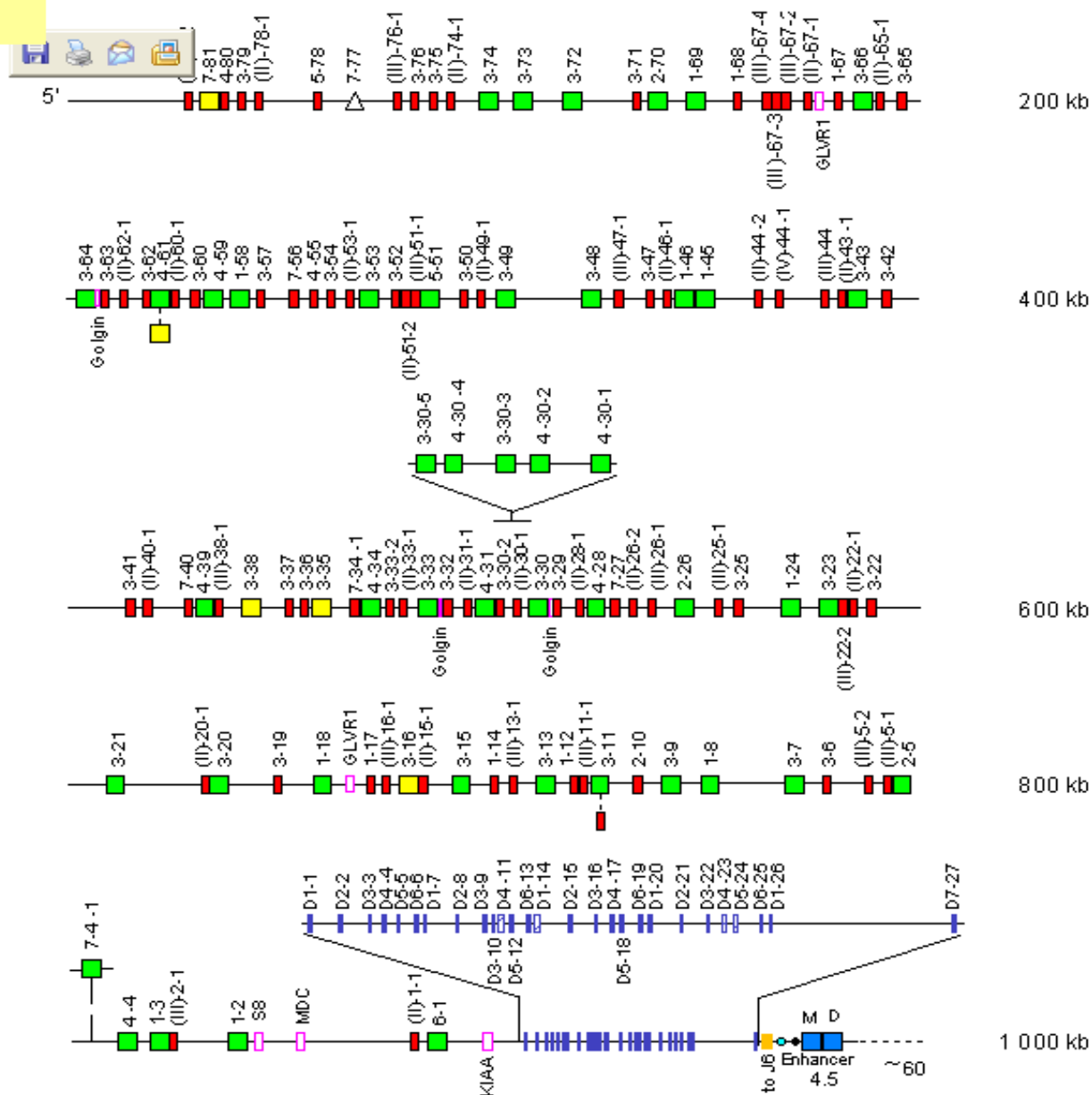
[Meeting announcements](#)

[IMGT Immunoinformatics page,...](#)

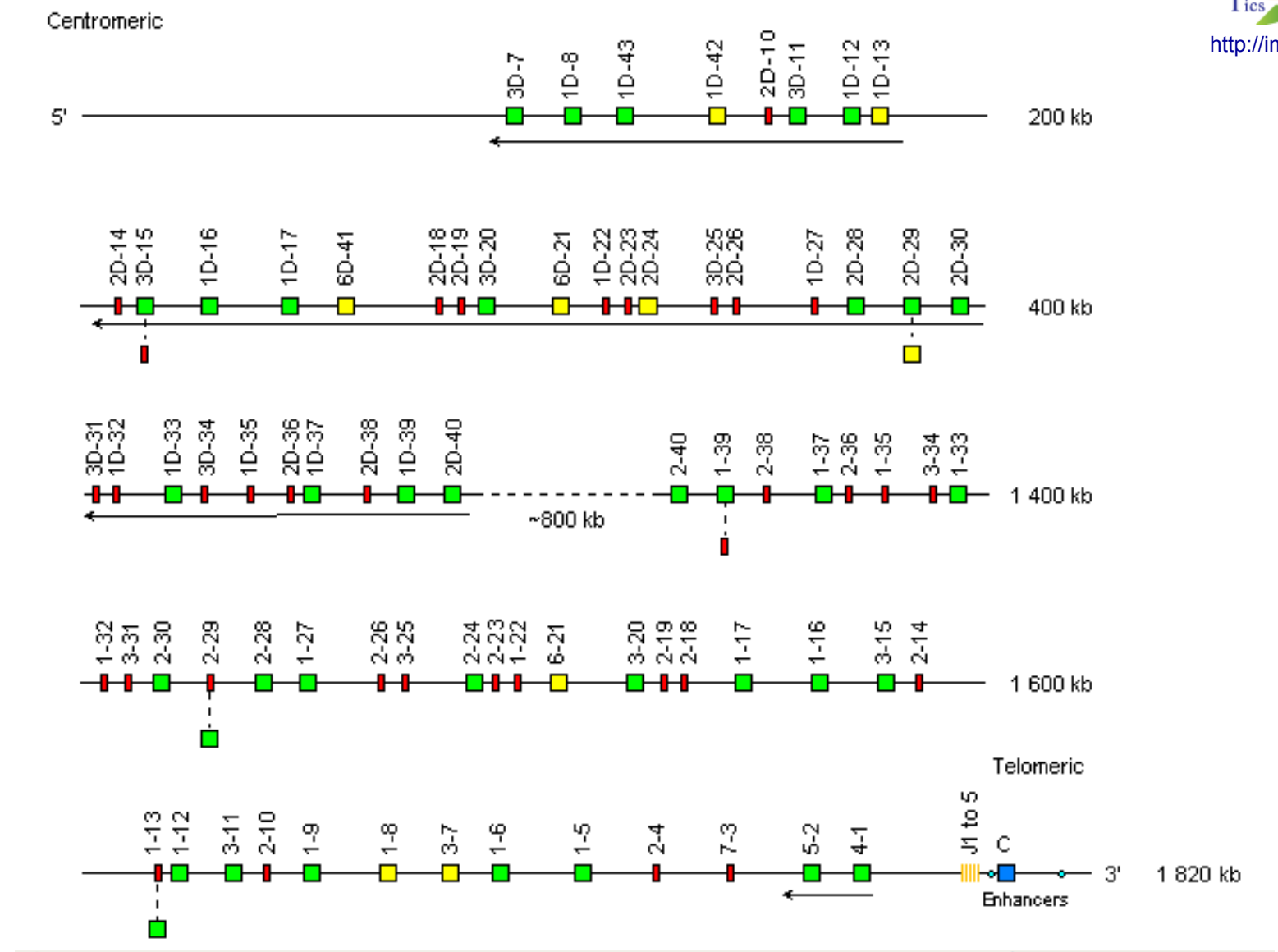
IMGT Education

[Tutorials](#) [IMGT Lexique](#)

Chromosome 14q32.33



Chromosome 2p11.2



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(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

X02850 , TRAV8-6*01

AE000659, TRAV8-6*02

M86361 , TRAV8-6*02

X02850 , TRAV8-6*01

AE000659, TRAV8-6*02

M86361 , TRAV8-6*02

X02850 , TRAV8-6*01

AE000659, TRAV8-6*02

M86361 , TRAV8-6*02

X02850 , TRAV8-6*01

AE000659, TRAV8-6*02

M86361 , TRAV8-6*02

X02850 , TRAV8-6*01

AE000659, TRAV8-6*02

M86361 , TRAV8-6*02

X02850 , TRAV8-6*01

AE000659, TRAV8-6*02

M86361 , TRAV8-6*02

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	GTG	CCT	GTG	TTT	GAA	GAA	GCC	CCT	GTG	GAG

																			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											
CTG	AGG	TGC	AAC	TAC	TCA	TGG	TCT	GTT	TCA	GTG	TAT	...	...	...	...	...	...	CTC	TTC											

																			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											
TGG	TAT	GTG	CAA	TAC	CCC	AAC	CAA	GGA	CTC	CAG	CTT	CTC	CTG	AGA	TAT	TTA	TCA	GGA	TCC											

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										
ACC	CTG	...	...	...	GTT	GAA	AGC	ATC	AAC	...	...	...	...	...	...	...	GGT	TTT	GAG										

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	GTG	CAT	ATA	AGC	GAC	ACG	GCT

										CDR3-IMGT							
101	102	103	104	105	106	107	108										
E	Y	F	C	A	V	S											
GAG	TAC	TTC	TGT	GCT	GTG	AGT	GA										

										R	#c
										--G	

IMGT Protein display

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	PSEVTAVEGAIVQINCTYQ	TSGFYG	LSWYQQHDGGAPTFLSY	NALDG	LEETG	RFSSFLSRSDSYGYLLLQELQMKDSASYFC	AVR					
AE000658, TRAV1-2	GQNIDQ	PTEMTATEGAIVQINCTYQ	TSGFNG	LFWYQQHAGEAPTFLSY	NVLDG	LEEKQ	RFSSFLSRSGYGYLLLQELQMKDSASYLC	AVR					
AE000658, TRAV2	KDQVFQ	PSTVASSEGAVVEIFCNHS	VSNAYN	FFWYLFHPGCGAPRLLVK	GSK	PSQQG	RYNMTYER	FSSSLILQVREADAAVYYC	AVE				
AE000658, TRAV3	AQSVAQPEDQVNVAEGNPLTVKCTYS	VSGNPY	LFWYVQYPNRLQFLLLK	YITGDNL	VKGSY	GFEAEFNKSTSFHLKKPSALVSDSALYFC	AVRD						
AE000658, TRAV4	LAKTTQ	PISMDSYEGQEVNITCSHN	NIATNDY	ITWYQQFPSPQGRFIIQ	GYKT	KVTNE	VASLFIPADRKSSLSLPRVSLSDTAVYYC	LVGD					
AE000659, TRAV5	GEDVEQS	LFLSVREGDSSVINCTYT	DSSSTY	LYWYKQEPGAGLQLLTY	IFSNMD	MKQDQ	RLTVLLNKKDKHLSLRIADTQTGDSAIYFC	AES					
AE000659, TRAV6	SQKIEQNSEALNIQEGKTATLTCNYT	NYSPAY	LQWYRQDPGRGPVFLLL	IRENEK	EKRKE	RLKVTFTDTT	KQSLF	ITASQPADSATYLC	ALD				
AE000659, TRAV7	ENQVEHSPHFLGPQQGDVASMSCTYS	VSREFNN	LQWYRQNTGMGPKHLLS	MYSAGY	EKQKG	RLNATLLK	NGSSLYITAVQPEDSATYFC	AVD					
AE000659, TRAV8-1	AQSVSQHNHHVILSEAASLELGCNYS	YGGTVN	LFWYVQYPGQHLQLLLK	YFSGDPL	VKGIK	FEAEFIKSKFSFNLRKPSVQWSDTAEYFC	AVN						
AE000659, TRAV8-2	AQSVTQ	QSHVSVSEGTPVLLRCNYS	SSYSPS	LFWYVQHPNKLQLLLK	YTSATL	VKGIN	GFEAEFPKSETSFHLTKPSAHMSDAAEYFC	VVS					
AE000659, TRAV8-3	AQSVTPQDIHITVSEGASLELRCNYS	YGATPY	LFWYVQSPGQGLQLLLK	YFSGDTL	VQGIK	GFEAEFPKRSQSSFNLRKPSVHWSDAAEYFC	AVG						
AE000659, TRAV8-4	AQSVTQLGSHVSV	IGALVLLRCNYS	SSVTPY	LFWYVQYPNQGLQLLLK	YTSATL	VKGIN	GFEAEFPKSETSFHLTKPSAHMSDAAEYFC	AVS					
X02850, TRAV8-6	AQSVTQLDSQVPVFEEAPVELRCNYS	SSVSVY	LFWYVQYPNQGLQLLLK	YLSGSTL	VEIN	GFEAEFNKSTSFHLRKPSVHISDTAEYFC	AVS						
AE000660, TRAV8-7	TQSVTQLDGHITVSEEAPELELCNYS	YSGVPS	LFWYVQYSSQSLQLLLK	DLTEATQ	VKGIR	GFEAEFPKSETSFYLRKPSTHVSDDAAEYFC	AVGDR						
AE000659, TRAV9-1	GDSSVWOTEGQWLPSEGDSLIVNCSYF	TTQVPS	LEWYVQYPPGEGPOLHLK	AMKAND	KGRNK	GFEAMVREKETTSEHLFKDSVQFSDSAVYFC	ALS						

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

IMGT databases and tools

Sequences



<http://imgt.cines.fr>

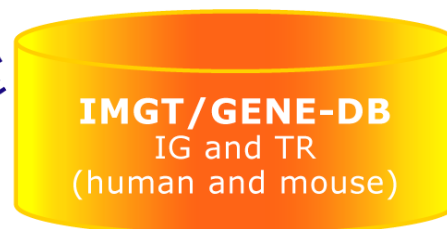
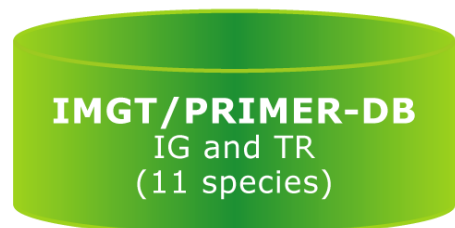


IMGT/V-QUEST

IMGT/JunctionAnalysis

IMGT/Allele-Align

IMGT/PhyloGene



Genome

IMGT/GeneInfo

IMGT/LocusView

IMGT/GeneSearch

IMGT/GeneView

IMGT/3Dstructure-DB
IG, TR and MHC

IMGT/StructuralQuery

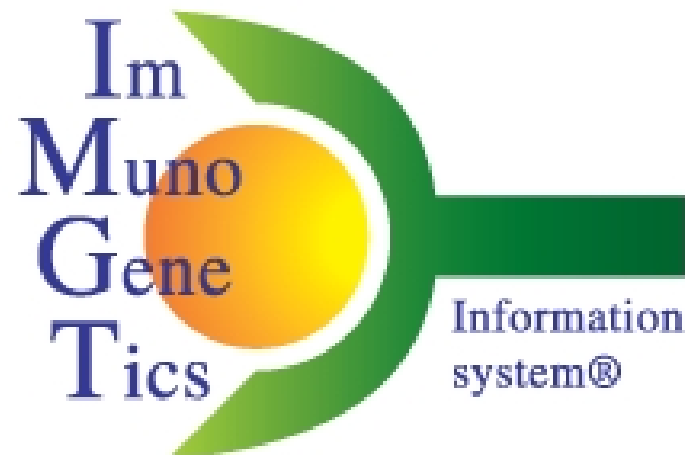
2D and 3D structures

Locus representation: Human IGL

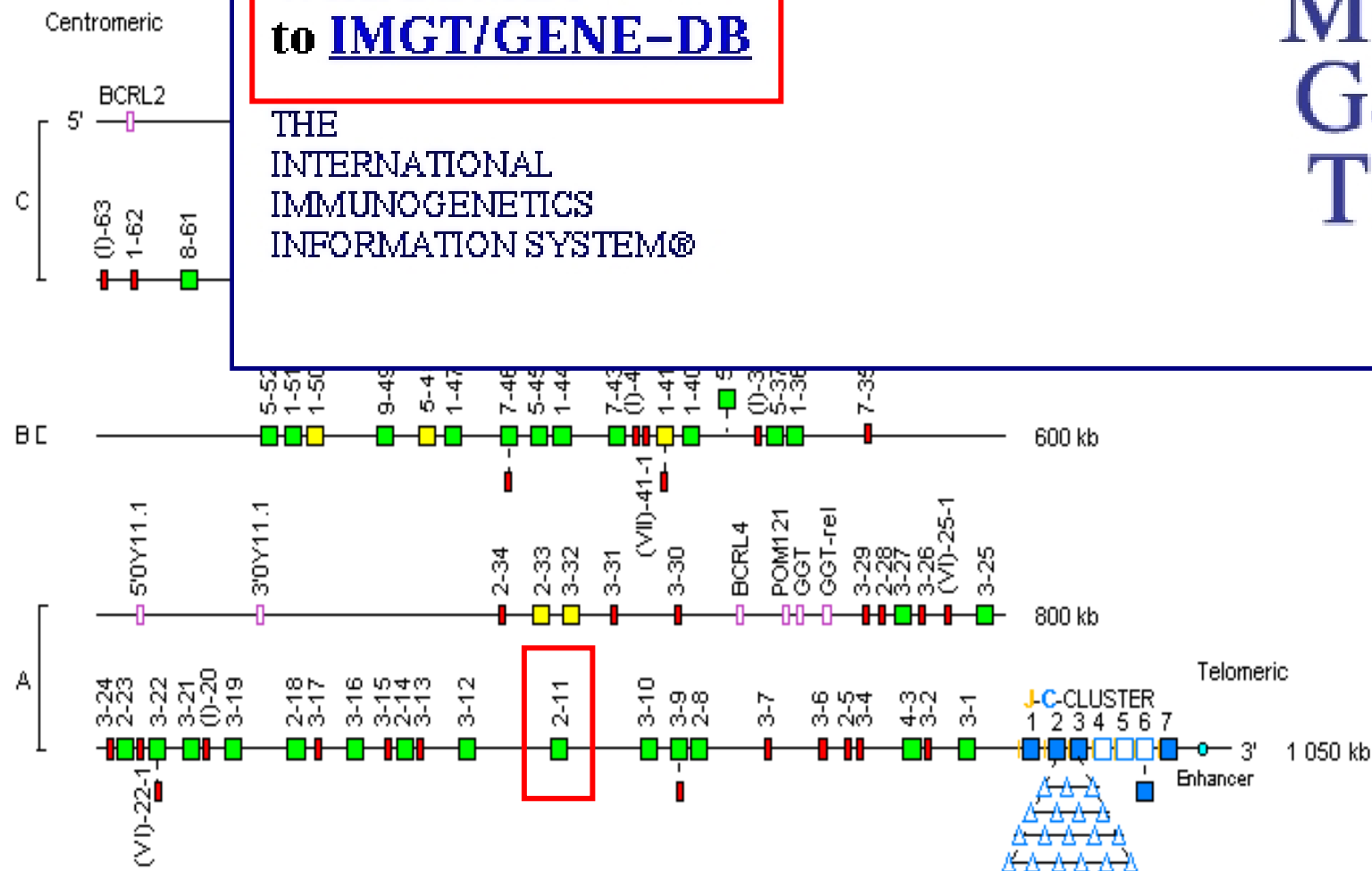
Human IGL 22q11.2

**WELCOME !
to IMGT/GENE-DB**

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>



FT V-GENE

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

FT

FT

FT FR2-IMGT

FT

FT

FT CONSERVED-TRP

FT

FT CDR2-IMGT

FT

FT

FT FR3-IMGT

FT

FT

FT 2nd-CYS

FT CDR3-IMGT

FT

FT

XX

SQ Sequence 297 BP; 60 A; 93 C; 71 G; 73 T; 0 other;

cagtctgcc tgactcagcc tcgctcagtg tccgggtctc ctggacagtc agtcaccatc 60

tctctacta gaaccacac tatattaat attataact atctctctc ataccaacac 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](#)



cDNA (in databases: mRNA!)

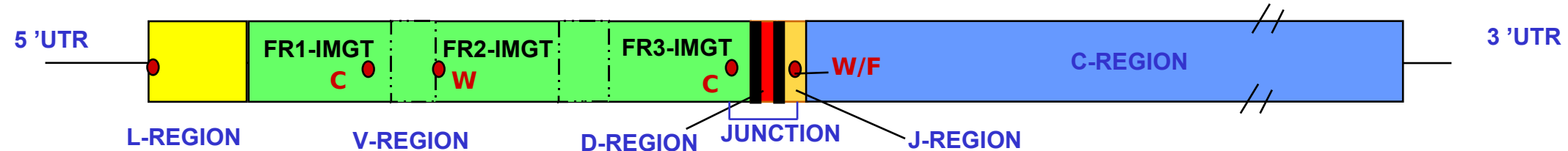
```
..... .gagga ttcacccatgg aactgggggct ccgctggggtt ttccttggtg ctttttttaga 120
aggtgtccag tgtgaggtgc aactgggtgga gtctgggggga ggcctgggtca agccggggggg 180
gtccctgaga ctctcctgtg cagcctcttg attagcttc agtacctatg ccatgaactg 240
ggtccgccag gctccaggga aggggctgga atgggtctca agtattagta gtagaagtga 300
ttacatatac tatagagact cagtgaaggg ccgattcacc atctccagag acaacgccaa 360
gaattcactg tatctgcaaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgcgagagat tcttgtaatg gtgctatatg ttatggtttc agtccttggg gccagggaac 480
cctgggtcacc gtctcctcag catccccgac cagccccaag gtcttcccgc tgagcctctg 540
cagcacccag ccagatggga acgtgggtcat cgcctgcctg gtccagggtt 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 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 ccttcaactg cactgctgcc taccgccagt ccaagacccc 1140
gctaaccgcc acccttctaa aatccgga aaacacattccgg cccgagggtcc acctgctgcc 1200
gccgccgtcg gaggagctgg ccctgaacga gctgggtgacg ctgacgtgcc tggcacgtgg 1260
cttcagcccc aaggacgtgc tggttcgctg gctgcagggg tcacaggagc tgccccgcga 1320
gaagtacctg acttgggcat cccggcagga gccagccag ggcaccacca ccttcgctgt 1380
gaccagcata ctgcgcgtgg cagccgagga ctggaagaag ggggacacct tctcctgcat 1440
ggtggggccac gaggccctgc cgtgggcctt cacacagaag accatcgacc gcttggcggg 1500
taaaccacc catgtcaatg tgtctgttgt catggcggag gtggacggca cctgctactga 1560 ....
```

cDNA

```

..... .gagga ttcacccatgg aactggggct ccgctggggtt ttccttggtg cttttttaga 120
aggtgtccag tgtgaggtgc aactgggtgga gtctggggga ggctgggtca agccgggggg 180
gtccctgaga ctctcctgtg cagcctctgg attaatgcttc agtacctatg ccatgaactg 240
gggccgccag gctccaggga aggggctgga atgggtctca agtattagta gtagaagtga 300
ttacatatatc tatagagact cagtgaaggg ccgattcacc atctccagag acaacgccaa 360
gaattcactg tatctgcaaa tgaatagcct gagagtcgac gacacggctg tctattactg 420
tgcgagaga ttccttgtaatg gtgctatatg ttatggtttc agtccctggg gccagggaac 480
cctggtcacc gtctcctcag catccccgac cagccccaag gtcttcccgc tgagcctctg 540
cagcaccacag ccagatggga acgtgggtcat cgcttgcttg gtccagggtt tcttccccc 600
ggagccactc agtgtgacct ggagcgaaag cggacagggc gtgaccgcca gaaacttccc 660
accagccag gatgcctccg gggacctgta caccacgagc agccagctga ccctgccggc 720
cacacagtgc ctagccggca agtccgtgac atgccacgtg aagcactaca cgaatcccag 780
ccaggatgtg actgtgcctt gccagttcc ctcaactcca cctaccccat ctccctcaac 840
tccacctacc ccatctccct catgctgcca ccccgcactg tcaactgcacc gaccggccct 900
cgaggacctg ctcttagggt cagaagcgaa cctcacgtgc aactgaccg gcctgagaga 960
tgcttcgggt gtcaccttca cctggacgcc ctcaagtggg aagagcgctg ttcaaggacc 1020
acctgagcgt gacctctgtg gctgctacag cgtgtccagt gtcttgccgg gctgtgccga 1080
gccatggaac catgggaaga ccttcacttg cactgctgcc taccgccagt ccaagacccc 1140
gctaaccgcc accctctcaa aatccggaaa cacattccgg cccgagggtc acctgctgcc 1200
gccgccgtcg gaggagctgg ccctgaacga gctgggtgacg ctgacgtgcc tggcacgtgg 1260
cttcagcccc aaggacgtgc tggttcgctg gctgcagggg tcacaggagc tgccccgcga 1320
gaagtacctg acttgggcat cccggcagga gccagccag ggcaccacca ccttcgctgt 1380
gaccagcata ctgcgcgtgg cagccgagga ctggaagaag ggggacacct tctcctgcat 1440
ggtggggccac gaggccttgc cgctggcctt cacacagaag accatcgacc gcttggcggg 1500
taaaccaccc catgtcaatg tgtctgttgt catggcggag gtggacggca cctgctactga 1560 ....

```



The IMG-T-ONTOLOGY concepts

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

IMG-T/LIGM-DB Consultation module v3 - Netscape

Fichier Edition Afficher Aller Communicator Aide

ID HSCOMFRAA IMG-T/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

DESCRIPTION

IMGT/LIGM-DB Consultation module v3 - Mozilla Firefox		
File	Edit	View
Go	Bookmarks	Tools
Help		
FH	Key	Location/Qualifiers
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_notes="automatically annotated with IMGT tools"
FT		/organism="Homo sapiens"
FT	V-D-J-REGION	1..375
FT		/translation="QVHLVESGGAVFHPGRSLRLSRAASGETESSYGMHWVRQAPAKGLEWVA
FT		VIWYDGSNKYYADSVKGRFTISRDN SKNTLYLAKHVTIAAAAGRRGAGMDVWGQGT
FT		TVTVSS"
FT	V-REGION	1..296
FT		/allele="IGHV3-33*01, putative"
FT		/gene="IGHV3-33"
FT		/CDR_lengths="[8,8,18]"
FT		/putative_limit="3' side"
FT		/translation="QVHLVESGGAVFHPGRSLRLSRAASGETESSYGMHWVRQAPAKGLEWVA
FT		VIWYDGSNKYYADSVKGRFTISRDN SKNTLYLAKHVTIAAAAGRRGAGMDVWGQGT
FT		TVTVSS"
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="GFTTFSSYG"
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="YYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYC"

OBTENTION

CLASSIFICATION

NUMEROTATION

IMG T/V-QUEST

IMGT/V-QUEST

Alignment for V-GENE

<u>AF402940</u>		score	GTGCAGCTGCTCGAGCAGTCTGGGGCT__GAGGTGAGCAAGCCTGGGGCCTCAGTAAAGGTTTCCTGCA
<u>X62109</u>	IGHV1-3*01	1146	CA.GTC.A...T.T.....AG.....G.....
<u>X62107</u>	IGHV1-3*02	1110	CA.GTT.A...G.T.....AG.....G.....
<u>M99637</u>	IGHV1-8*01	957	CA.GT..A...G.T.....AG.....G....C.....
<u>L06612</u>	IGHV1-46*03	948	CA.GT..A...G.T.....AG.....G.....
<u>X92343</u>	IGHV1-46*01	948	CA.GT..A...G.T.....AG.....G.....

Alignment for J-GENE

<u>AF402940</u>		score	CTTCACGGGGCGGGACGCTTTGGACGTCTGGGGCCAAGGGACCACGGTCACCGTCTCCTCA
<u>J00256</u>	IGHJ3*01	181	_____.T.....T..T.....A.T.....T...G
<u>X86355</u>	IGHJ6*02	179	T.A.TACTACTACT...G.A.....
<u>X86355</u>	IGHJ3*02	172	_____.T.....T..TA.....A.T.....T...G



V-GENE

JUNCTION

J-GENE

V-QUEST Search Page - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

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

WELCOME !
to the **IMGT/V-QUEST** 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

Terminé

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

Analyse your Immunoglobulin sequences

Your selection: Human

Your sequences are compared to the **Human IG set** from the [IMGT/V-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 :

```
gaggtgcagctgttggagtcgggggggaggttggtagagcctgggggggtccctgagactc
tctgtgcagcctctggattcacctttagcagctttgccatgagctgggtccgccaggt
cctgggaaggggctggactgggtctcagaaattagtggtagtggcggtaccacatattac
gcaggctccgtgaagggccggttcaccatctccagagacaactccaagaatacgtgttt
ctgcaaatgaatagcctgacagccgaagacacggccgtatattactgtgcgaagaggtt
```

- ☐ 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...GGCTTGGTACAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACCTTTAGCAGCTTTGCC...
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	GTTTTGGAGTGGTTATCTGTGGTTTCGACCCCTGGGGCCAGGGAACCCTGGTCACCGTCTCCTCAGCTTCCACCAA
X86355	IGHJ5*02	219	-CAAC-----
J00256	IGHJ5*01	201	-CAAC-----T-----A-----
X86355	IGHJ4*02	177	AC-AC--T---TA-----

IMGT/V-QUEST

IMGT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

http://imgt.cines.fr/cgi-bin/IMGtdnap.jv OK

Translation of the JUNCTION

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

Applet input [1}

IMG/TV-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 ----->

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

----->

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

<----- F R 2 - I M G T ----->

	40	45	50
input	... ATG AGC TGG GTC CGC CAG GCT CCT GGG AAG GGG CTG GAC TGG GTC TCA		
M99660 IGHV3-23*01	---	---	---
J00236 IGHV3-23*02	---	---	---
U29481 IGHV3-23*03	---	---	---
Z12358 IGHV3-48*03	---	---	---
M99675 IGHV3-48*01	---	---	---

Applet input [1]

IMGT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

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

Translation

input

M99660 IGHV3-23*01

```

<----- F R 1 - I M G T -----
1           5           10           15
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

--- --T --- ... ---

----->
20           25           30           CDR1 - IMGT           35
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
--- --A- --- ... ..

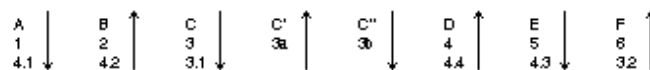
<----- F R 2 - I M G T -----
40           45           50
M S W V R Q A P G K G L D W V S
... .. ATG AGC TGG GTC CGC CAG GCT CCT GGG AAG GGG CTG GAC TGG GTC TCA
E
... .. --- --A- --- --G- ---

-->
CDR2 - IMGT
55           60           65           70
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- ---

```

Applet input [1]

input



IMGT/V-QUEST

Recent improvements

IMGIT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

WELCOME ! to IMGIT/V-QUEST

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



(none)

Citing IMGIT/V-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 [IMGIT/V-QUEST reference directory sets](#)

Nucleotide sequences

Enter your sequence(s) in FASTA format :

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

```
>AY867144
gatgttgtgatgactcagtcctccaggccacctgtctttgtctccagggaagagccacc
ctctcctgcaggggccagtcagactgttagcagcagttacttagcctggtagcagcagaaa
cttggccaggctcccaggctcctgactctatggtgcacccagtagggctagtggtatctca
gacagggttcagcggcagtggtctgtggacagacttcacactcaacatcagcagactggag
cctgaagattttgcagtgattactgtcagcagcatgatacctcactctcactttcggc
ggagggaccagggtggtgatcaad
>test
acgttttatctccagcttggttccttgcccaaaagtgtacgggttaagtgt
tataatgttggcagtaataagctgcaaaatcatcaggctgcaggctgctg
atggtgagagtgaaatctgtcccgatccactgccgctgaaccttgatgg
```

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

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

IMGT/V-QUEST - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

Case of complementary reverse sequences

Sequence number 1: test

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

```
>test (complementary reverse)
caggtgcagctggtgcagctctgggggagggcgtagtccagcctgggaggtccttgagactc
tctgtgcagagggcggattcactttcagttcctattctatgcactgggtccgccaggcg
ccaggcaaggcctggaatgggtggcggttatgtggtttgatggaactgaaaaatactct
gcagagtcctgtgaagggccgattcaccatctccagagacaattccaagaacacattgttt
ttgcaaatgaacagcctgagagccgacgacacggctgtgtattactgtgcgagagagcct
gattggttattatggggggacaggggtgctctggacgtctggggccaagggaaccaggctc
accgtctcctcagacatcgtgatgaccagtcctcctccaccctgtctgcacatctgtagga
gacagagtcaccatcacttgccggggcagtcagggcattagtagctgggtggcctgggtat
cagcagaaaccagggaagccctaaagctcctgatgtatgagggcgtctagtttagaaagt
gggggtcccatcaaggttcagcggcagtggtacgggacagaattcactctcaccatcagc
agcctgcagcctgatgattttgcagcttattactgccaaacattataaacacttaccggtac
acttttggccagggaaccaagctggagataaaaacgt
```

Terminé

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

```
>AY867144
gatgttgtgatgactcagtcctccaggcaccctgtcttctgtctccaggggaagagccacc
ctctcctgcagggccagtcagactgttagcagcagttacttagcctggtaccagcagaaa
cttgccaggtctcccaggctcctgatctatggtgcatccagtagggctagtggcatctca
gacaggttcagcggcagtggtctgggacagacttcacactcaacatcagcagactggag
cctgaagattttgcagtgattactgtcagcagcatgatacctcacctctcactttcggc
ggagggaccagggtggtgatcaaa
>test
acgttttatctccagcttggttccttgcccaaaagtgtacgggtaagtgt
tataatggtggcagtaataaagctgcaaaatcatcaggctgcaggctgctg
atggtgagagtgaattctgtcccgatccactgccgctgaaccttgatgg
```

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

Parcourir...

Start Clear the form

Selection of output
parameters

Output parameters

Display type :	HTML												
Includes JCTA :	yes												
Results :	1. Details Select the fields to be displayed : <table border="0"> <tr> <td>☒ Alignment for V-GENE</td> <td>☒ Alignment for D-GENE</td> </tr> <tr> <td>☒ Alignment for J-GENE</td> <td>☒ JUNCTION</td> </tr> <tr> <td>☒ Alignment with FR-IMGT and CDR-IMGT delimitations</td> <td>☒ Translation</td> </tr> <tr> <td>☒ Mutation statistics</td> <td>☒ Annotations</td> </tr> <tr> <td>☒ "input" V-REGION for IMGT/Phylogene</td> <td>☒ Collier de Perles</td> </tr> <tr> <td>☒ Mutation table</td> <td></td> </tr> </table>	☒ Alignment for V-GENE	☒ Alignment for D-GENE	☒ Alignment for J-GENE	☒ JUNCTION	☒ Alignment with FR-IMGT and CDR-IMGT delimitations	☒ Translation	☒ Mutation statistics	☒ Annotations	☒ "input" V-REGION for IMGT/Phylogene	☒ Collier de Perles	☒ Mutation table	
☒ Alignment for V-GENE	☒ Alignment for D-GENE												
☒ Alignment for J-GENE	☒ JUNCTION												
☒ Alignment with FR-IMGT and CDR-IMGT delimitations	☒ Translation												
☒ Mutation statistics	☒ Annotations												
☒ "input" V-REGION for IMGT/Phylogene	☒ Collier de Perles												
☒ 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 :													

Terminé

Alignment for V-GENE

Closest V-GENEs (evaluated from the 1st nt to the 2nd-CYS codon plus 15 nt of the CDR3-IMGT)

	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)

Score and nucleotide identity

Alignment

```

AY867144      <----- FR1-IMGT ----->
gatgttgtgatgactcagtcctccaggcaccctgtctttgtctccagggaaagagccaccctctcctgcagggccagtcagactgttagc
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-----a-----
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...gacaggttcagcggcagtgagg.....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---c-
X74460 IGKV3/OR2-268*02 .....-cc---c-c---c-...-c-----t---c-
X72820 IGKV3D-7*01  .....-cc---c-c---c-...-c-----t---c-
    
```

IMGT/V-QUEST

IMGTV-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          cactttcggcggaggggaccagggtggtgatcaaa
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

ImMunoGeneTics
Information system®

View from above

Generation of the JUNCTION diversity

3'V-REGION

N-REGION

D-REGION

N-REGION

5'J-REGION

tgtgcgaaaga



tacc



agcatattgtg

gtggtgactgctat

tcc



gat



acaactgggtcg actcctgg

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	tgtgcgagaga	aga	.tagcaatggctacaa....	aata
#2 Z47269	IGHV1-69*06	tgtgcgagag.	gggggggctaagg	tcgaatttttgagtggtt.....	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	...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	112.3	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	V	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 - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

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

OK

Citing IMGT/JunctionAnalysis: Yousfi Monod, M. et al. Bioinformatics 2004, 20 Supplement 1:I379-I385 [PMID: 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
>247269, 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

ImMunoGeneTics Information system®

http://imgt.cines.fr

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

IMGT/JunctionAnalysis - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils ?

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

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....	aata	tttgactactgg	IGHJ4*02
#2	Z47269	IGHV1-69*06	tgtgcgaga	gggggggctaagg	tcgaatttttgagtggtt.....	tcattgggt	...actggttcgaccctgg	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							
#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 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:

	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	tgtgt <u>ac</u> g.....		tgttgt	gcagcgcctggtac	ccaaatatc		...actttgacc <u>act</u> gg	IGHJ4*02	IGHD6-13*01	1	2	1	5/15	
#2	Z70257	IGHV3-7*02	tgtgcgag.		ggatgg	cagctc <u>t</u> tatgcc	cgccc	ctactgggtacttcgatctctgg	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	acgatttttgagtggttatt....	ccccggggga	..atgcttttgatatctgg	IGHJ3*02	IGHD3-3*01		0	0	0	12/17	
#5	Z70610	IGHV4-34*09	tgtgcgagag.		tcgggag	cgatttttgagtggttatt....	cccga	ca	tgatgcttttgatatctgg	IGHJ3*02	IGHD3-3*01		0	0	0	9/12
#6	Z70611	IGHV4-59*01	tgtgcgaga..		ca	tggt <u>a</u> actataa.	tgccggcggtg	...actggttcgaccctgg	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.....	cccttc	gactactgg	IGHJ4*02	IGHD3-16*01		0	2	0	7/14	
#9	Z70615	IGHV4-59*01	tgtgcgagag.		ggctg	gtaaagaggg.....	tttcggaa	.tactgggtacttcgatctctgg	IGHJ2*01	IGHD5-24*01		0	2	0	7/13	
#10	Z70616	IGHV4-34*01	tgtgcgagag.		cgg	gtttggg.....	ttecc	...actggttcgaccctgg	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	tgccccgctcctgccaaat	gtattactatggttcgggga.....	tatgtacg	tttgactactgg	IGHJ4*03	IGHD3-10*01		0	0	0	15/28	

The eleven IMGT 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			K R
Medium	138-154	V			H
Small	108-117		P	T	E D Q
Very small	60-90	A	G	S	
		Aliphatic	Sulfur	Hydroxyl	Basic
					Acidic
					Amide
		Nonpolar		Uncharged	Charged
				Polar	

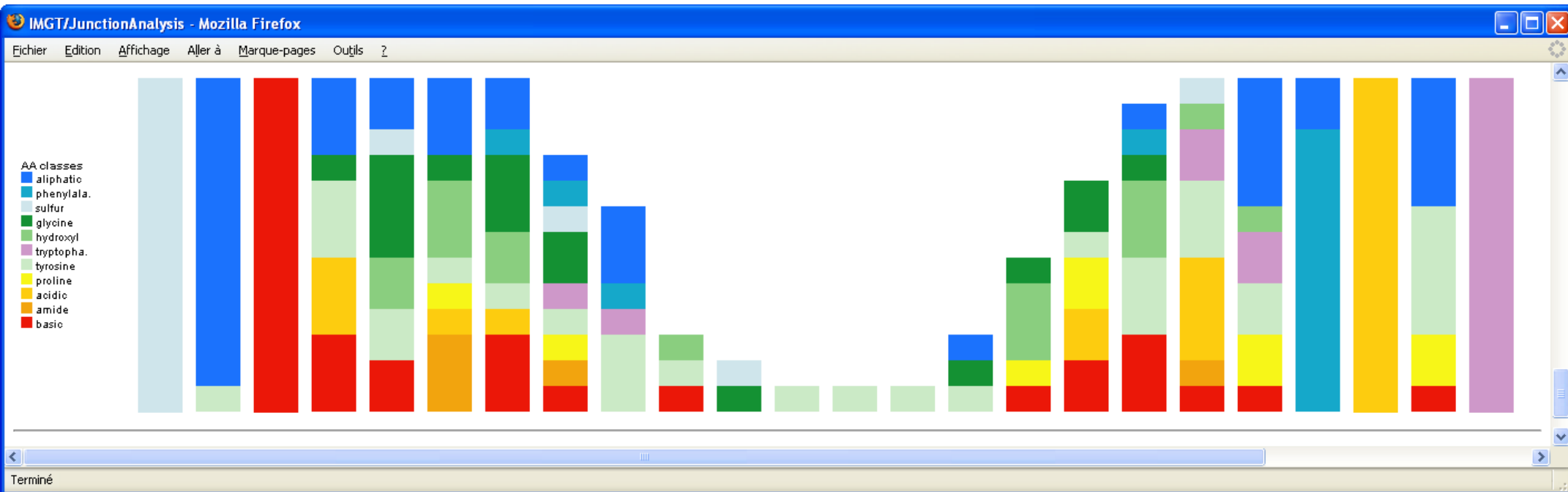
JUNCTION alignments

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	
	tgt	gta	cgt	gtt	gtg	cag	cgc	ctg	gta	...	...	...	...	...	...	...	ccc	aaa	tat	cac	ttt	gac	cac	tgg					
#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	
	tgt	gcg	agg	gat	ggc	agc	tct	tat	gcc	...	...	...	...	...	...	cgc	ccc	tac	tgg	tac	ttc	gat	ctc	tgg					
#3	Z70606	C	A	R	A	T	T	H												Y	A	F	D	V	W	+	11	1,604.77	
	tgt	gcg	aga	gcg	act	acg	cac	...	...	...	...	...	...	...	...	...	...	...	tat	gct	ttt	gat	gtc	tgg					
#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
	tgt	gcc	aga	gta	acg	att	ttt	gga	gtg	gtt	...	...	...	...	att	ccc	cgg	ggg	aat	gct	ttt	gat	atc	tgg					
#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
	tgt	gcg	aga	gtc	ggg	agc	gat	ttt	tgg	agt	ggt	...	...	...	tat	tcc	cga	cat	gat	gct	ttt	gat	atc	tgg					
#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	
	tgt	gcg	aga	cat	ggt	aac	tat	aat	gcc	...	...	...	...	...	...	...	ggc	ggt	gac	tgg	ttc	gac	ccc	tgg					
#7	Z70613	C	A	R	A	A	A	G											T	S	L	F	D	Y	W	+	12	1,531.71	
	tgt	gcg	aga	gca	gca	gct	ggt	...	...	...	...	...	...	...	...	...	...	acc	tcc	ctc	ttt	gac	tac	tgg					
#8	Z70614	C	A	R	H	Y	N	<u>S</u>	G										<u>T</u>	Y	P	L	D	Y	W	+	13	1,846.01	
	tgt	gcg	aga	cac	tat	aat	tcg	ggg	...	...	...	...	...	...	...	...	...	act	tat	ccc	ctc	gac	tac	tgg					
#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	
	tgt	gcg	aga	ggg	ctg	gta	aag	agg	ggt	...	...	...	...	...	...	tcg	gaa	tac	tgg	tac	ttc	gat	ctc	tgg					
#10	Z70616	C	A	R	A	G	L	G											S	H	W	F	D	P	W	+	12	1,602.8	
	tgt	gcg	aga	gcg	ggt	ttg	ggt	...	...	...	...	...	...	...	...	...	...	tcc	cac	tgg	ttc	gac	ccc	tgg					
#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	
	tgt	gcg	aga	gac	cgg	ggc	ggg	atg	...	...	...	...	...	...	...	...	ggt	cgg	gat	gct	ttt	gat	atc	tgg					
#12	Z70621	C	A	R	H	H	D	L	W	F								G	E	F	D	P	L	D	Y	W	+	16	2,307.53
	tgt	gcg	aga	cac	cac	gat	tta	tgg	ttc	...	...	...	...	...	...	ggg	gag	ttt	gac	ccc	ctt	gac	tac	tgg					
		C	A	R	D	C	P	A	P	A	K	M	Y	Y	Y	G	S	G	I	C	T	F	D	Y	W	+			

#13 Z70622

IMGT/JunctionAnalysis

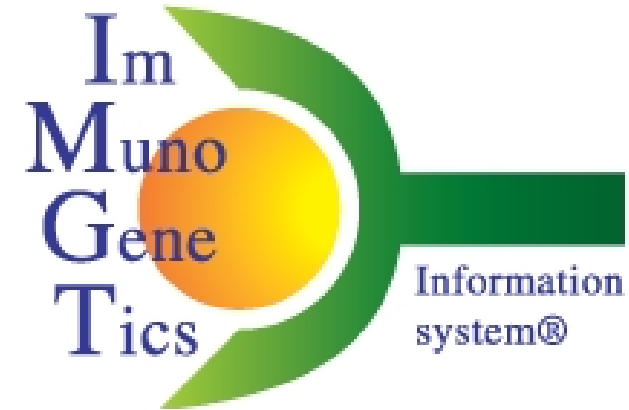


IMGT/3Dstructure-DB and IMGT/StructuralQuery

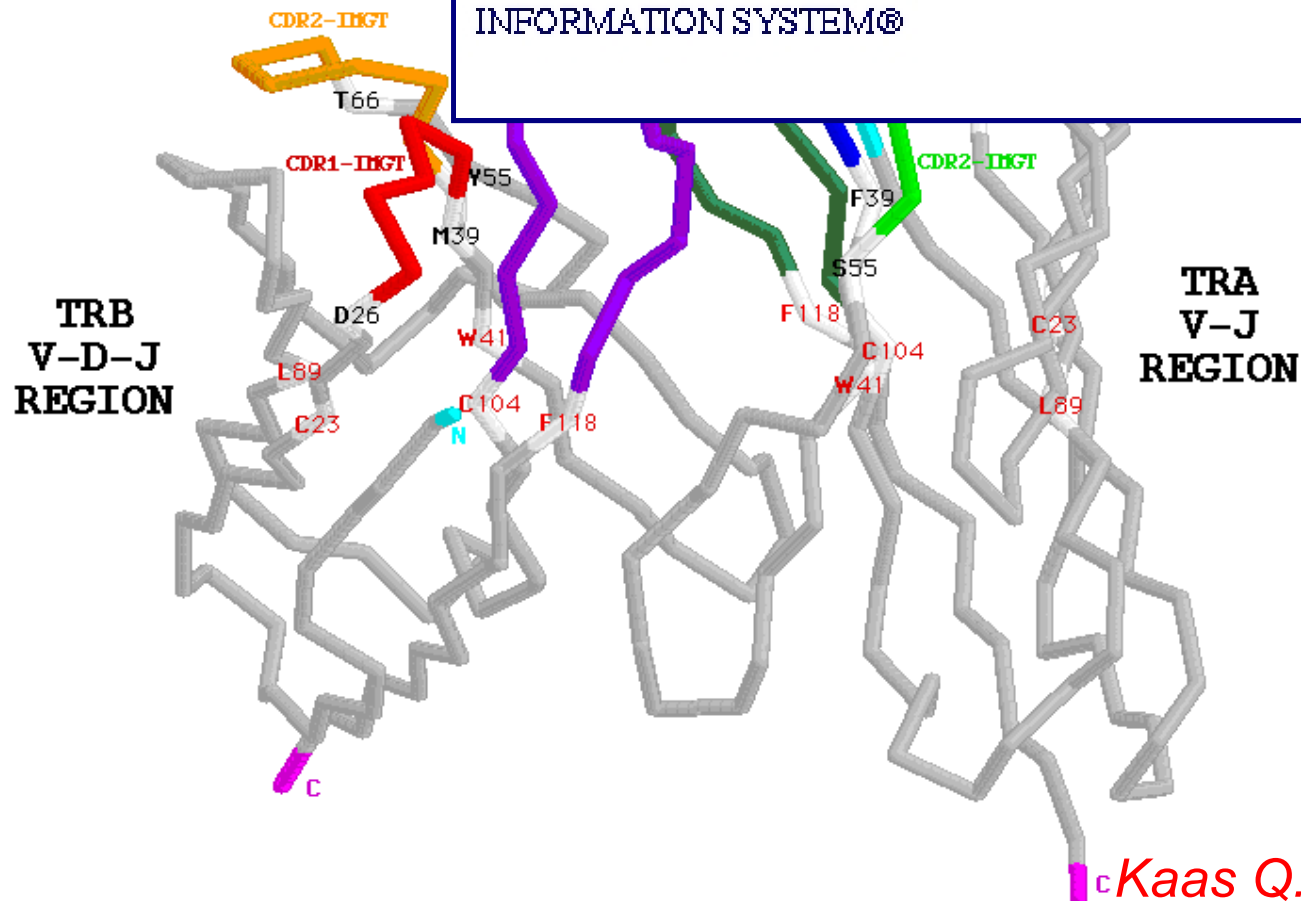
WELCOME !

to IMGT/3Dstructure-DB

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



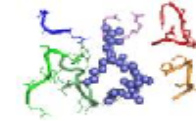
<http://imgt.cines.fr>



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

IMGT/3Dstructure-DB

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>[RI TLKESGPPLVKPTQTLTLTCSFSFGSLSDFGVGVWIRQPPGKALEWLAII YSDDDKRYSPSINTRLTI TKDTSRNQVVLVMTRVSPV] [D] [IGHV_1] [IG DTATYFCAHRRGPTTLFGVPIARGPUNAMDWQGGITVTI SSTS TKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1 GVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYTQNVNHNKPKNTKVDKRVPEKS]</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]
C-DOMAIN	IMGT domain description	CH1
	IGHG 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]
		STKGPSVFPLAPCSKSTA...GGTAALGCLVKDYFP..EPVTVSWNSGALTS...GVHTFPAVLQSS.....GLYSLSSVTVTPSSSL...GTQTYTQNVNHNK...SNTKVDKRV

Chain ID	1u8k_A	
IMGT chain description	L-KAPPA	
Chain amino acid sequence	<pre>[ALQLTQSPSSLSASVGDRTITITCRASQGVTSALAWYRQKPGSPQQLLIYDASSLESQVPSRFSGSGSGTEFTLTITSLRPEDFATYYQQQ] [IGKJ_1] [IGK_1] [IGKC_1 LHFYPTFGGGTRVDVRRITVAAPSVPFPPSDRLQKSGTASVQCLINNFYPREAKVQWVKVINALQSGNSQESVTEQDSKDSYSLSSLT] LSKADYEKKHVEGEVTHQGLSSPVTKSFNRGEC]</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
	CDR-IMGT lengths	[6.3.9]
	Sheet composition	[A B D E] [C C' C' F G]
		<pre>[CDR1] [CDR2] ALQLTQSPSSLSASVGDRTITITCRASQGVTSALAWYRQKPGSPQQLLIYDASSLESQVPSRFSGSGSGTEFTLTITSLRPEDFATYYQQQ [CDR3] ISTSLRPEDFATYYQQQLHFLHF...YPHTFGGGTRVDVRRITVAAPSVPFPPSDRLQKSGTASVQCLINNFYPREAKVQWVKVINALQSGNSQESVTEQDSKDSYSLSSLT</pre>
	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	

IMGT/3Dstructure-DB

Chain ID	1u8k_B	
IMGT chain description	VH-CH1	Region identification and delimitation
Chain amino acid sequence	<pre>[RI TLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAI I YSDDDKRYSPSINTRLTI TKDTSKNQVVLVMTRVSPV][D][IGHJ_1] [DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI SSTSTKGPSVFPLAPCS KSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1 GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTCNVNHKPSNTKVDKRVPEPKS]</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 . PLVKPTQTLTLTCSFS GFSLSDFGVG . . VGWIRQPPGKALEWLAI YSDDDK . . . RYSPSIN . TRLTI TKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI</pre>	
	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%),

IMGT/3Dstructure-DB

Chain ID	1u8k_B	
IMGT chain description	VH-CH1	
Chain amino acid sequence	Identification of the IMGT gene and the IMGT closest allele(s) <pre>[RI TLKESGPPLVKPTQTLTLTCSFS] [D DTATYFCAHRRGPTTLFGVPIAR HC_1 GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYTQNVNHKPSNTRVDKRVPEPKS MTRVSPV IG ISGALTS</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 . PLVKPTQTLTLTCSFS GFSLSDFGVG . . VGWIRQPPGKALEWLAI YSDDDK . . RYSPSIN . TRLTI TKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIAR GPVNAMDVWGQGI 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

IMGT/3Dstructure-DB

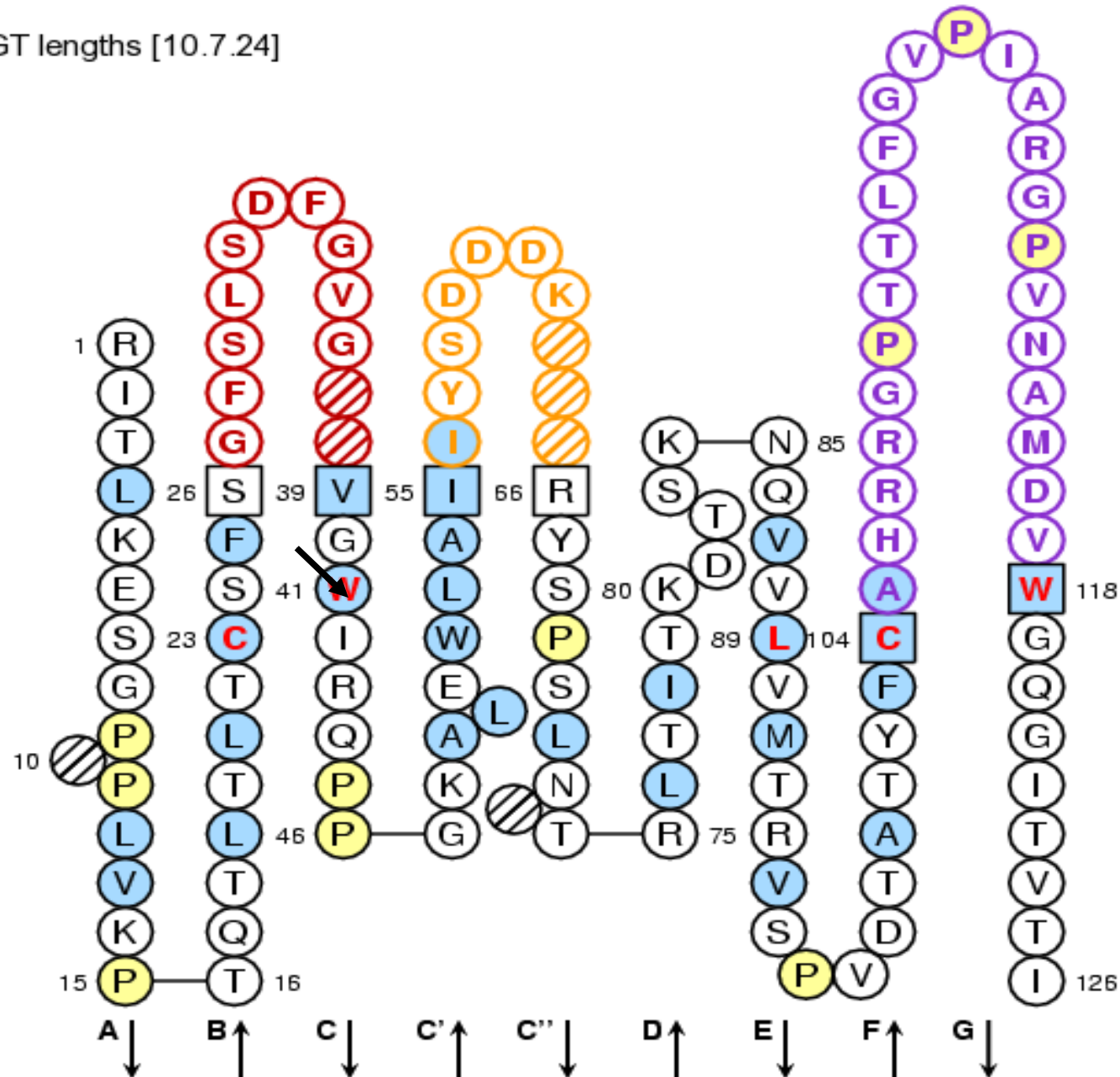
Chain ID	1u8k_B	
IMGT chain description	VH-CH1	
Chain amino acid sequence	<pre>[RI TLKESGPPLVKPTQTLTLTCSFSGFSLSDFGVGVGWIRQPPGKALEWLAI] [D] [IGHJ_1] [DTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI SSTSTKGPSVFPLAPCSKSTAGGTAALGCLVKDYFPEPVTVSWNSGALTS HC_1] GVHTFPAVLQSSGLYSLSVVTVPSSSLGTQTYTQNVNHNKPSNTKVDKRVPEPKS</pre>	
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]
<pre> [CDR1] RI TLKESGP . PLVKPTQTLTLTCSFSGFSLSDFGVG . . VGWIRQPPGKALEWLAI IYSDDDK . . RYSPSIN . TRLTI TKDTSKNQVVLV [CDR3] MTRVSPVDTATYFCAHRRGPTTLFGVPIARGPVNAMDVWGQGI TVTI</pre>		
	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

CDR-IMGT lengths [10.7.24]



Contact analysis

41V - TRP (W)

chain : 1u8k_B

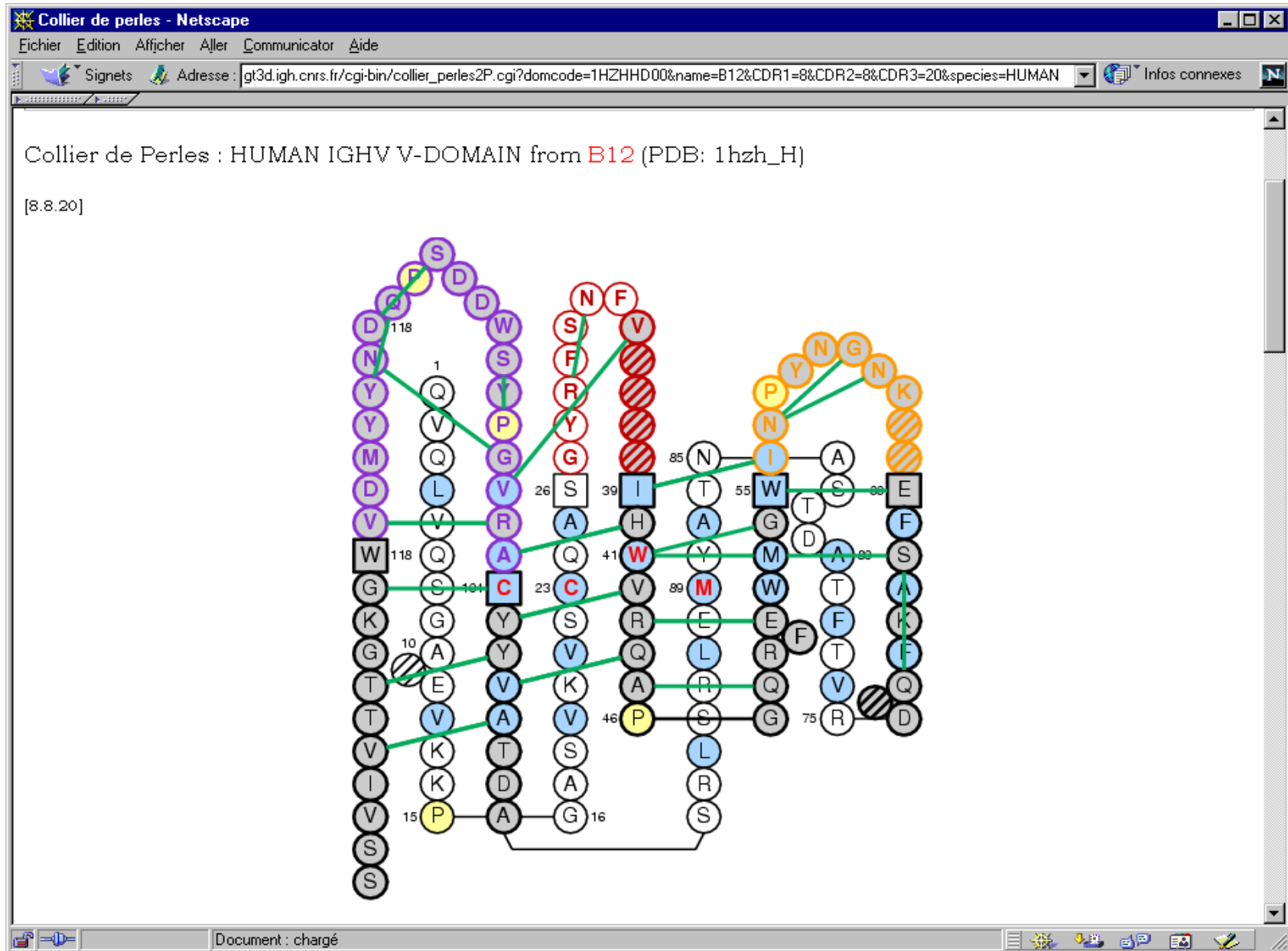
Secondary structure	Extended conformation
Phi	-122.64
Psi	137.12
ASA	0.0

Residue local structure

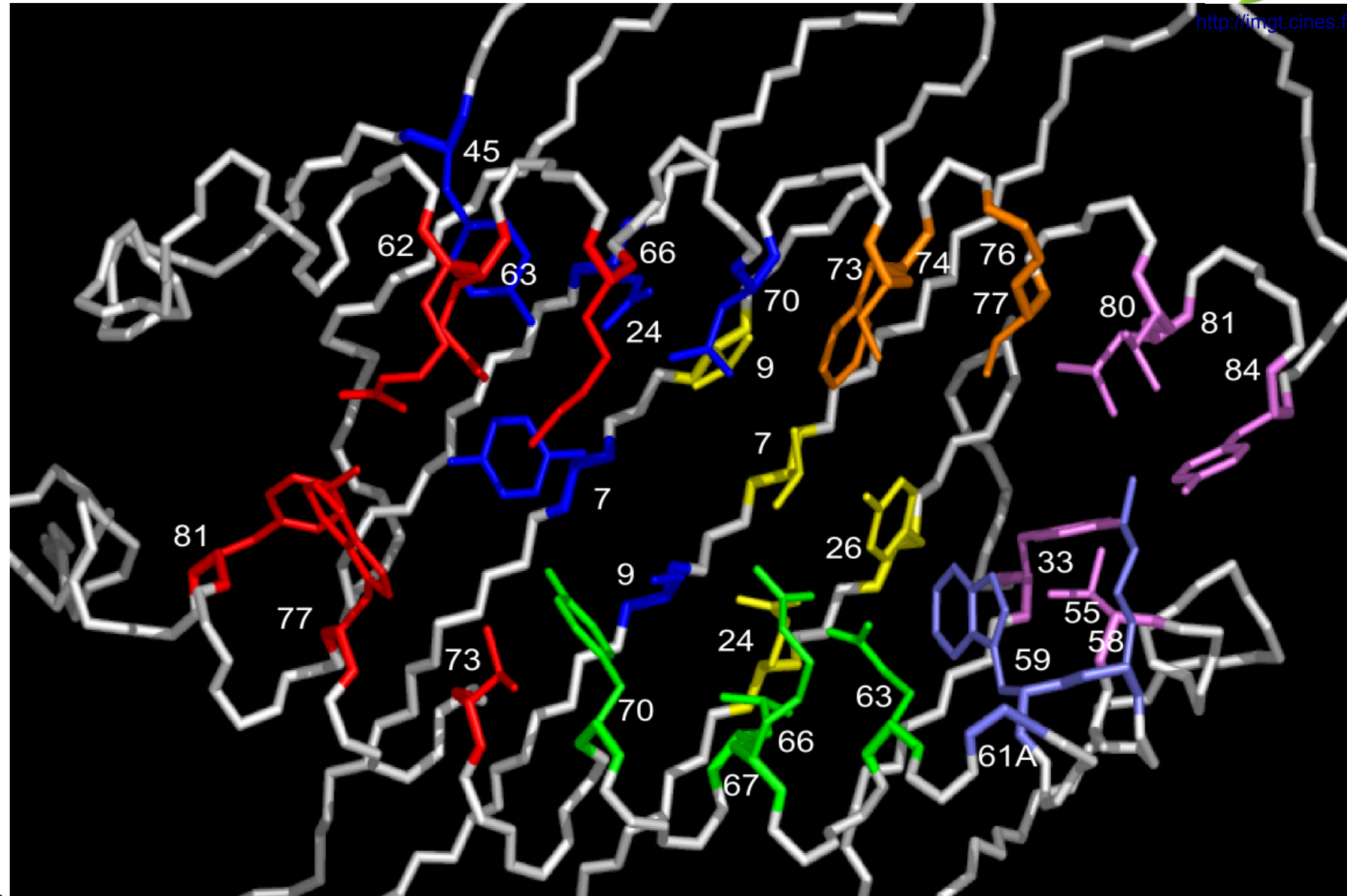
	IMGT Num	Code 1C	Code 3C	Chain	Domain	Tot	NCov	Pol	HB	NPol	Cov	SS
Info	6V1	E	GLU	1u8k_B	VH	6	6	0	0	6	0	0
Info	21V1	L	LEU	1u8k_B	VH	17	17	0	0	17	0	0
Info	22V1	T	THR	1u8k_B	VH	8	8	0	0	8	0	0
Info	23V1	C	CYS	1u8k_B	VH	10	10	0	0	10	0	0
Info	39V1	V	VAL	1u8k_B	VH	2	2	1	0	1	0	0
Info	43V1	R	ARG	1u8k_B	VH	2	2	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



MHC class I IMGT contact sites



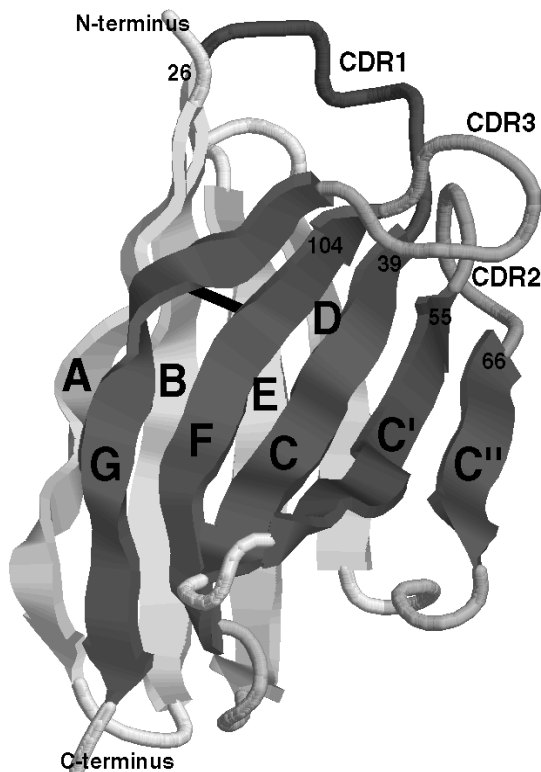
C1	1 E
C2	--
C3	2 Q
C4	3 Y
C5	(4 K)
C6	5 F
C7	--
C8	--
C9	6 Y
C10	7 S
C11	8 V

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

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

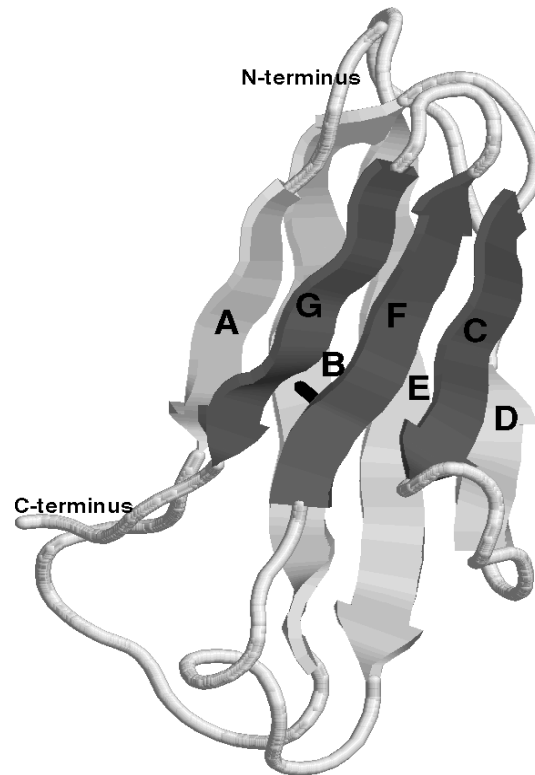
IMGT unique numbering

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

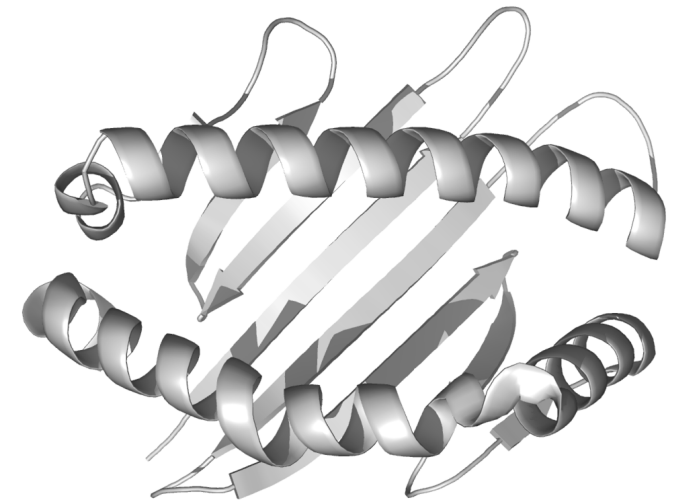


Immunoglobulin superfamily (IgSF)

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



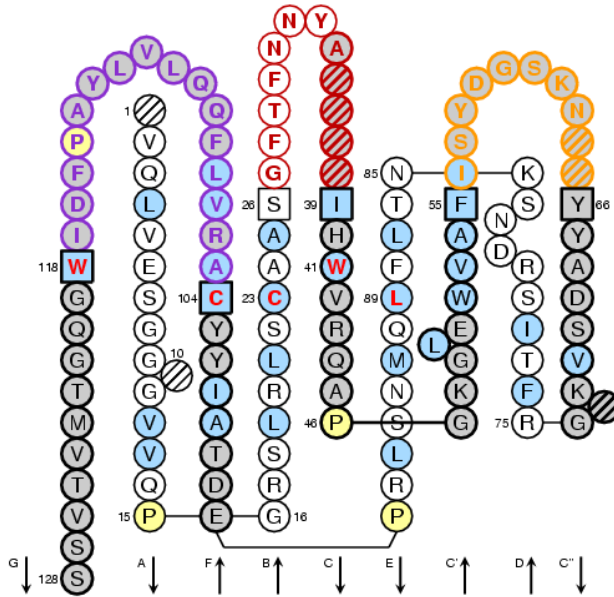
G-DOMAIN (MHC)
AND
G-LIKE-DOMAIN
(other than MHC)



MHC superfamily (MhcSF)

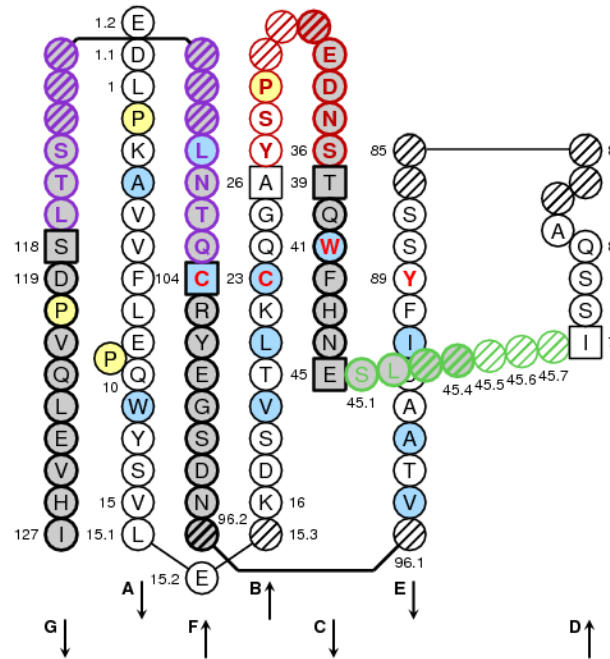
V-LIKE-DOMAIN

MOG



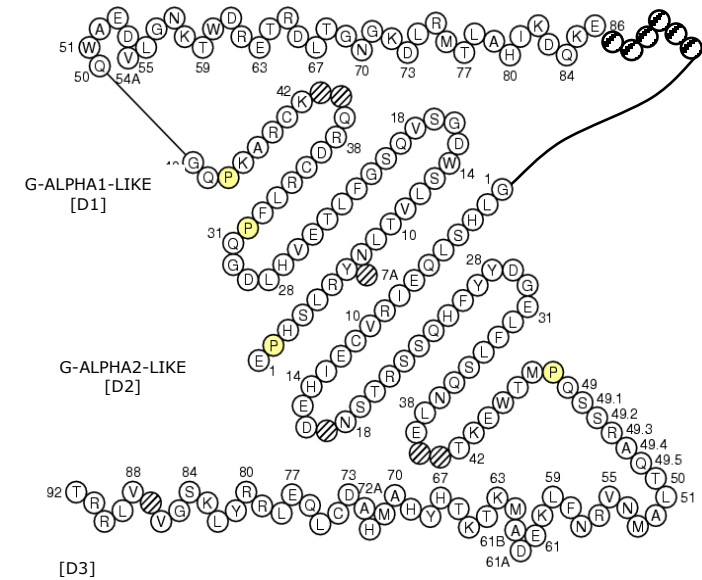
C-LIKE-DOMAIN

FCGR3B



G-LIKE-DOMAIN

MICA



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

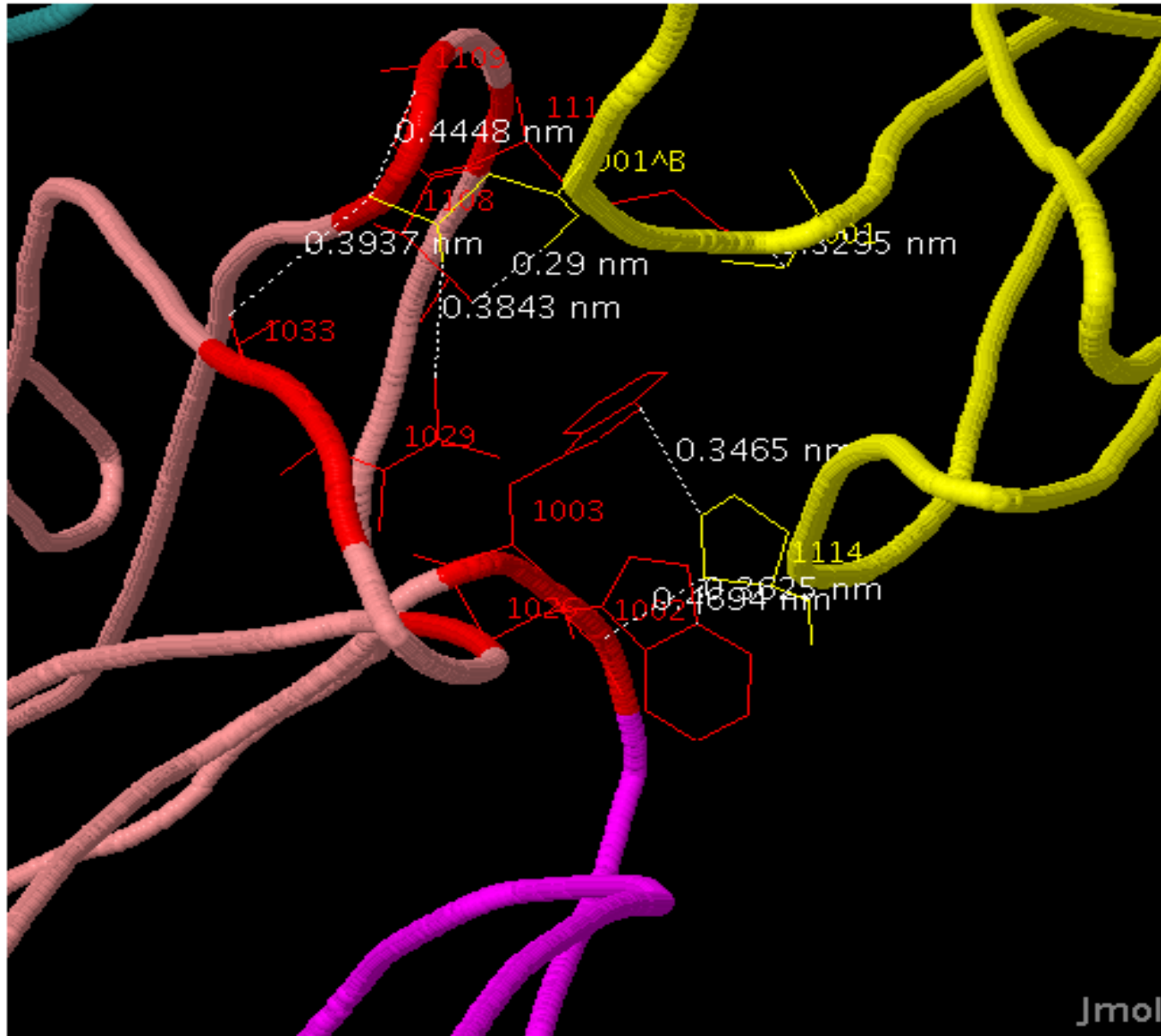
Bertrand, G. et al., Tissue Antigens, 64, 119-131 (2004)

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

Interactions between domains

FCGR3B

[D2]
C-LIKE-
DOMAIN



[D1] C-LIKE-DOMAIN

IGHG1
(FC-GAMMA1)

CH2
C-DOMAIN