

IMGT, the Informatics Framework to Support Effective Collaboration in Complex Environments

Marie-Paule Lefranc

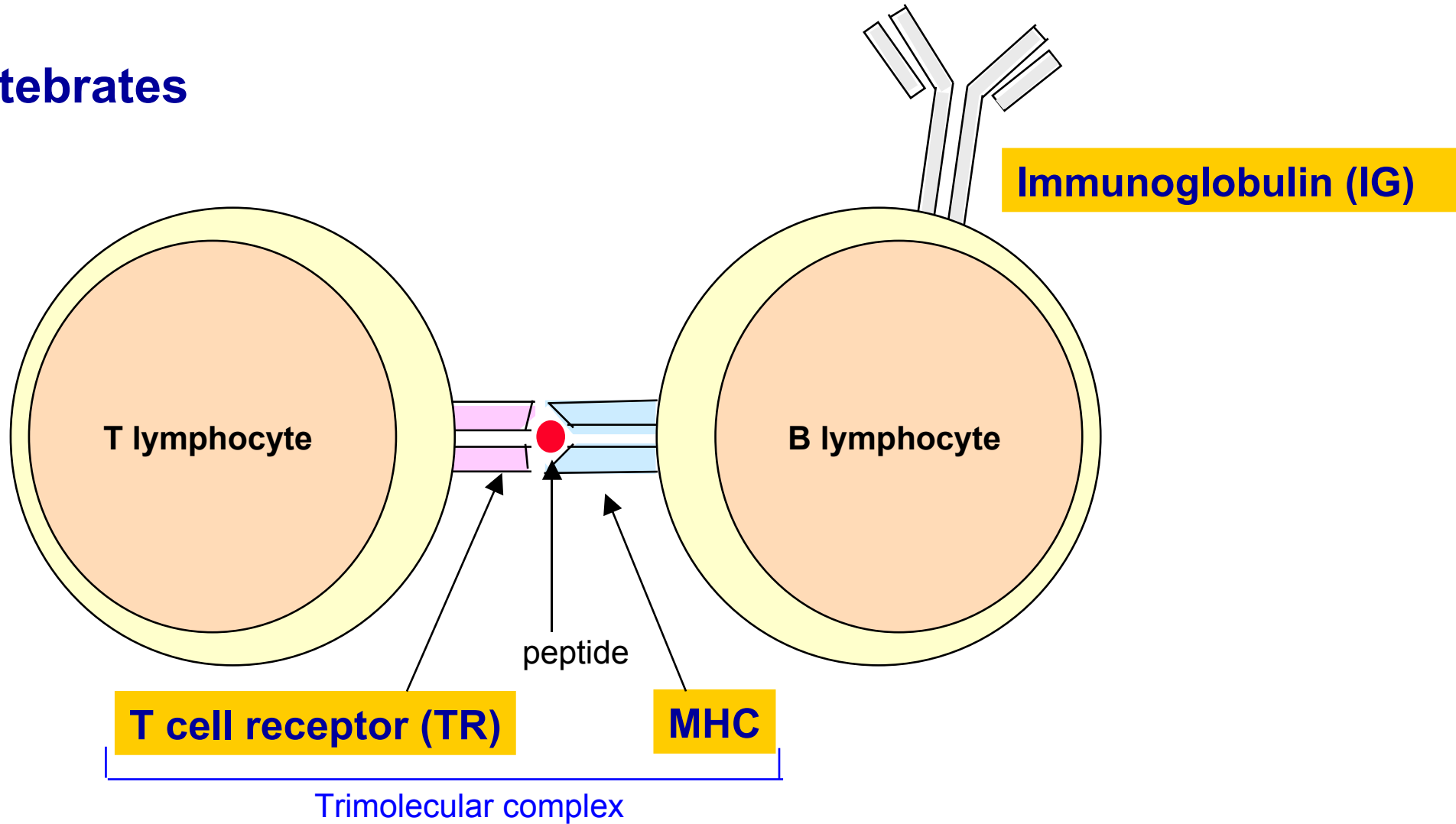
Professor, Montpellier University, CNRS
Senior Member, The University Institute of France
Group leader, Institute of Human Genetics, Montpellier

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

Outline

- IMGT® domain of expertise
- IMGT-ONTOLOGY axioms and concepts
- IMGT/V-QUEST, IMGT/JunctionAnalysis, IMGT/3Dstructure-DB
- Conclusions and Perspectives

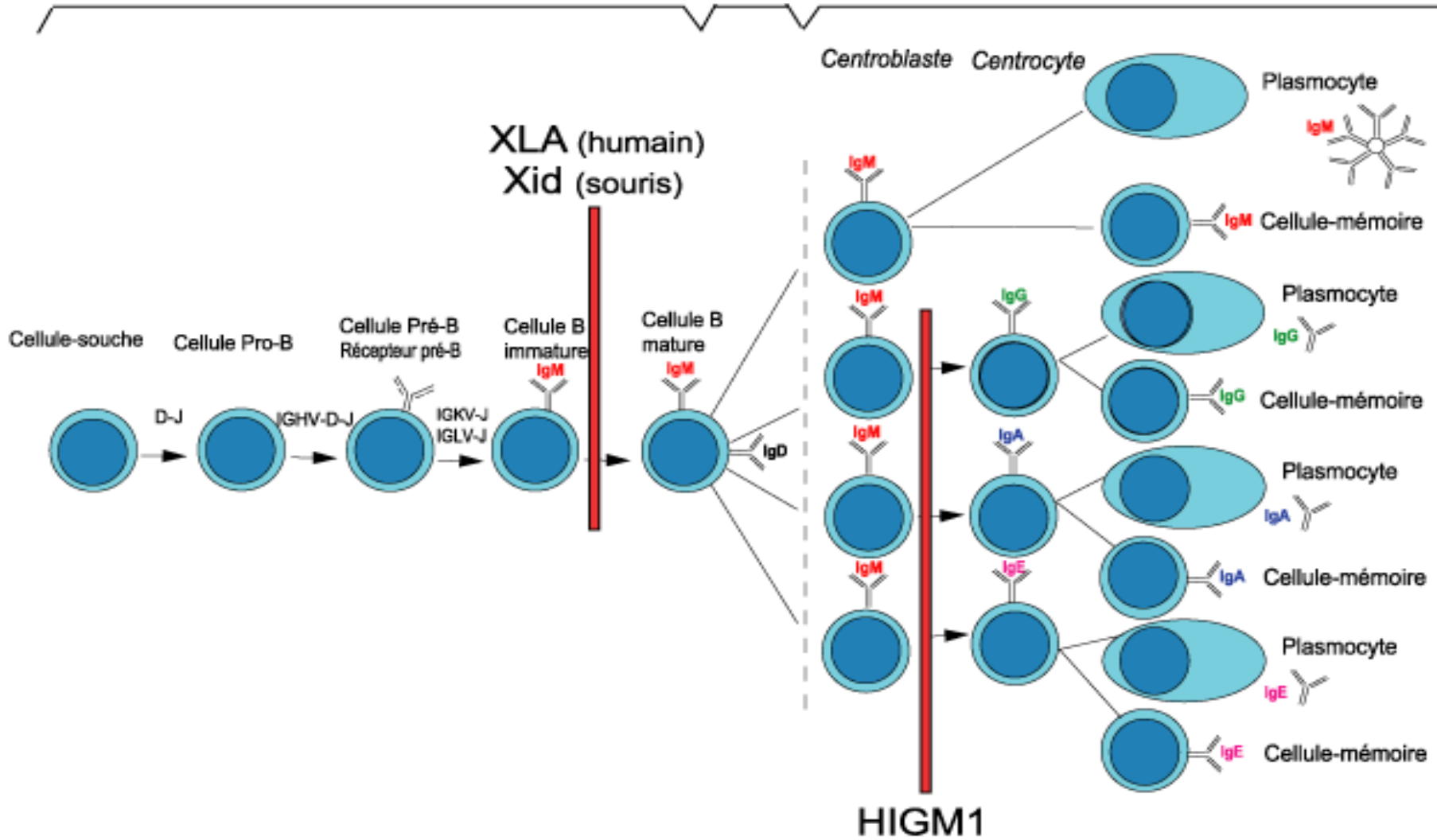
Vertebrates



Bone marrow

Blood

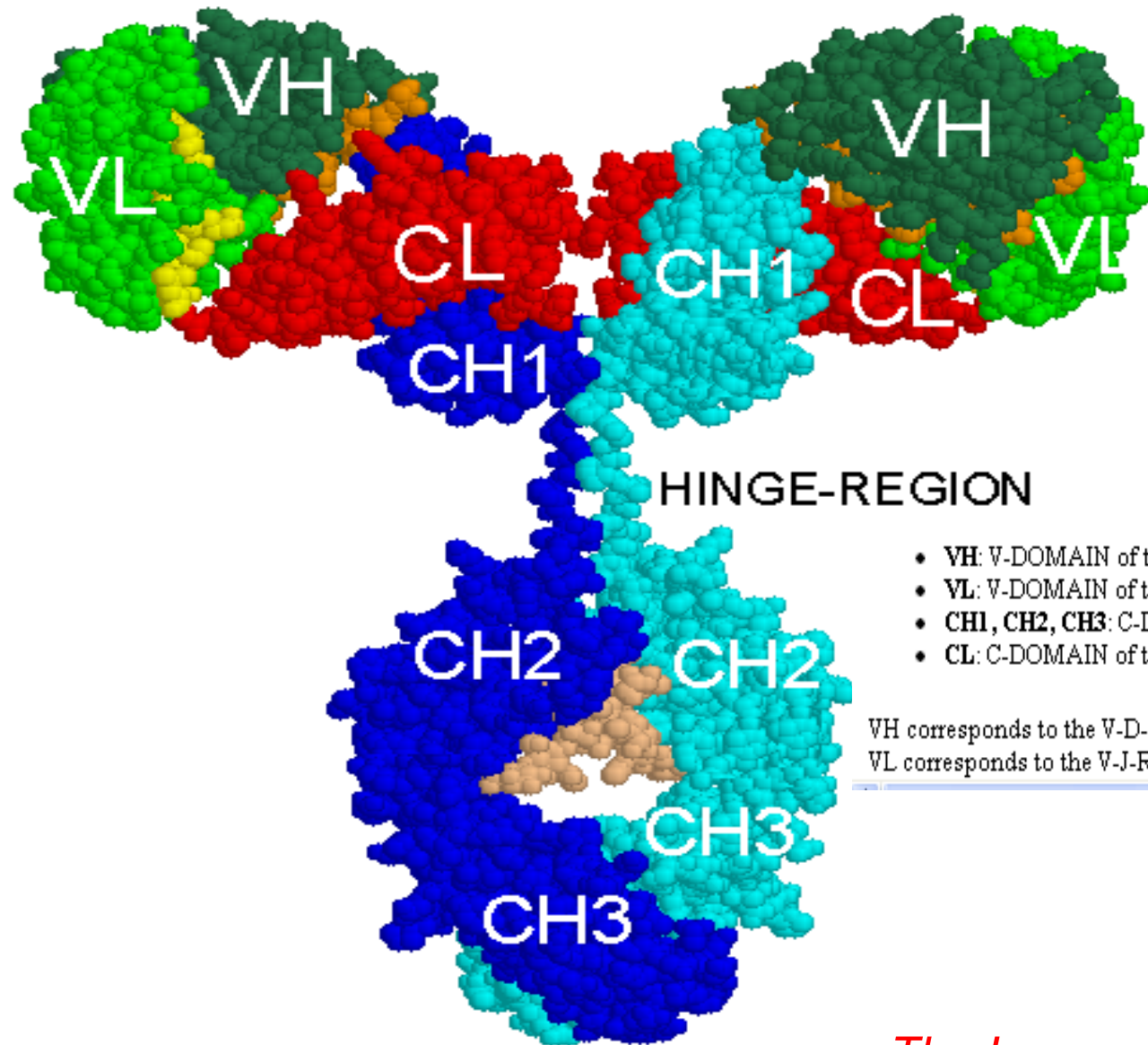
Lymph nodes, spleen



V-D-J and V-J
rearrangements

Hypermutations,
selection

Spacefill 3D representation of an IgG



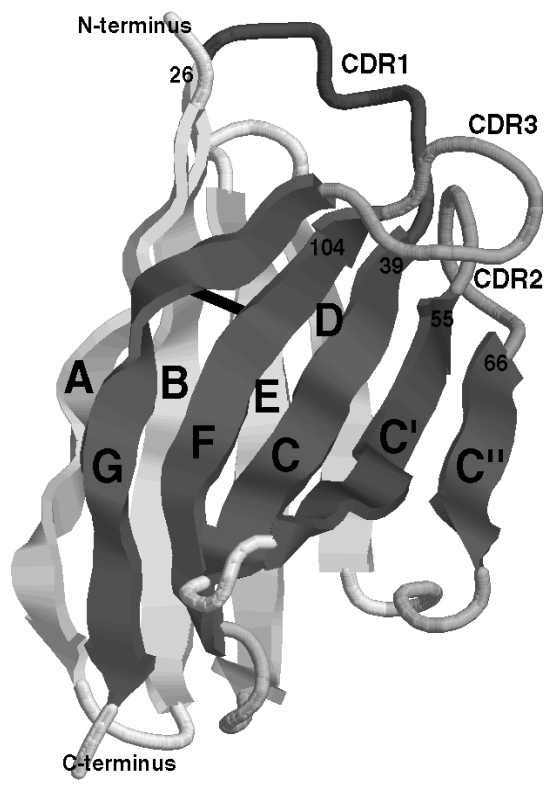
- **VH**: V-DOMAIN of the immunoglobulin heavy chain
- **VL**: V-DOMAIN of the immunoglobulin light chain
- **CH1, CH2, CH3**: C-DOMAIN of the immunoglobulin heavy chain
- **CL**: C-DOMAIN of the immunoglobulin light chain

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.

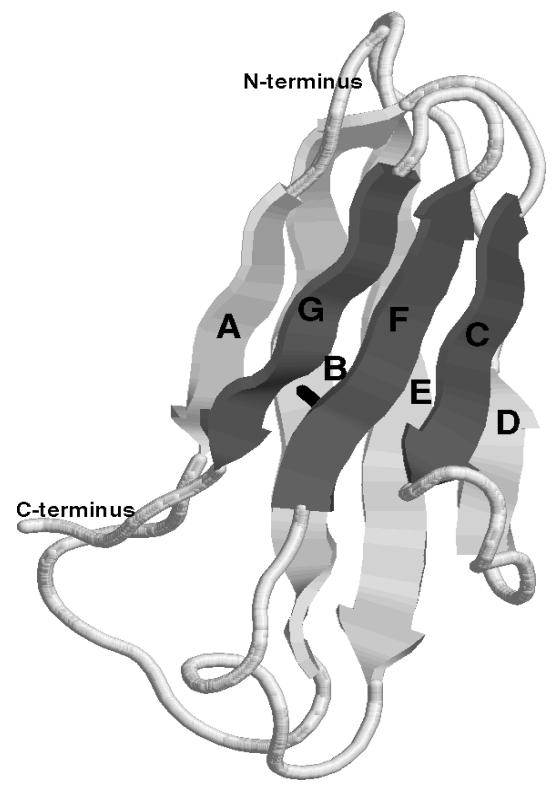
Structural domains

IG and TR

V-DOMAIN



C-DOMAIN



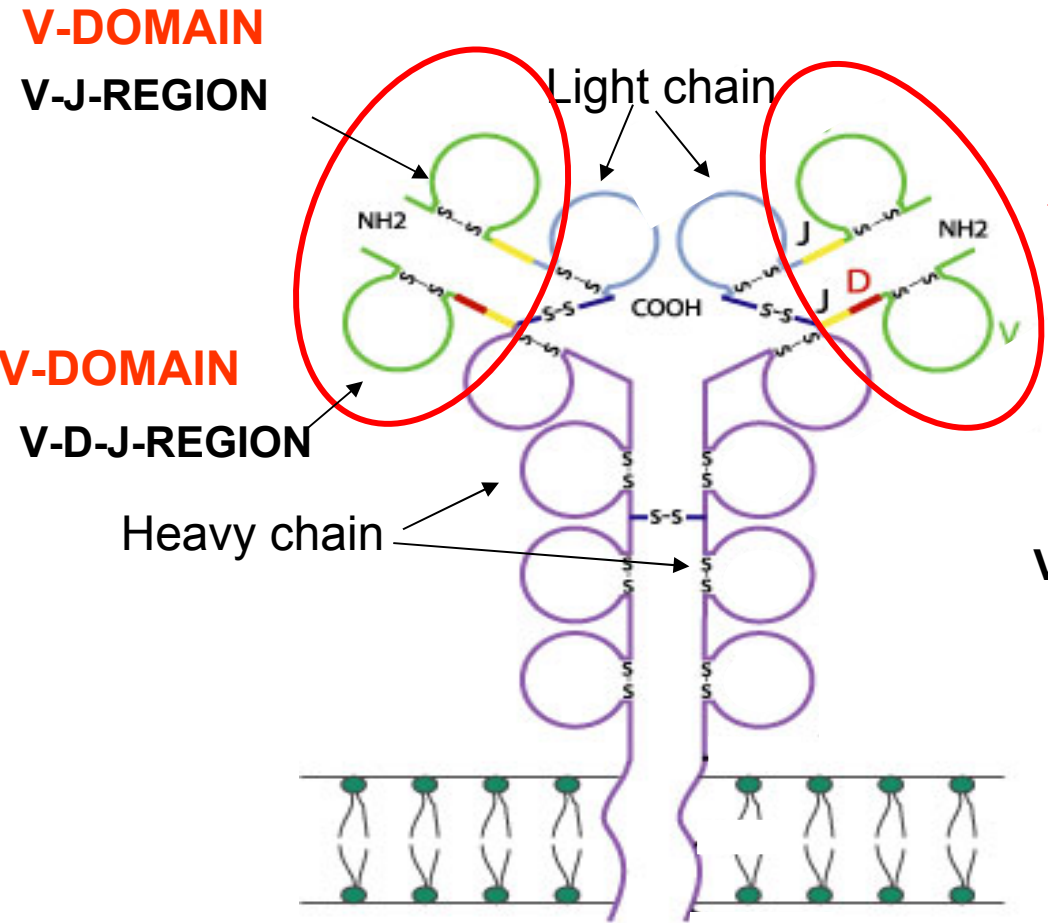
MHC

G-DOMAINS

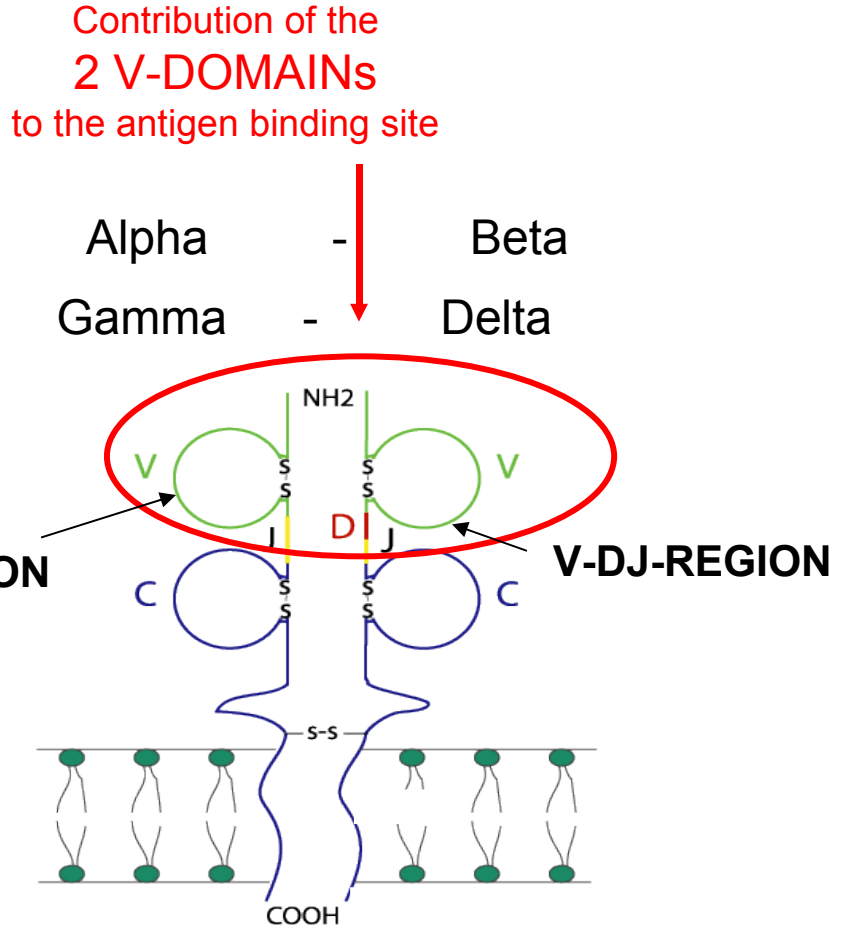


Immunoglobulin (IG)

T cell receptor (TR)



Membrane IgM

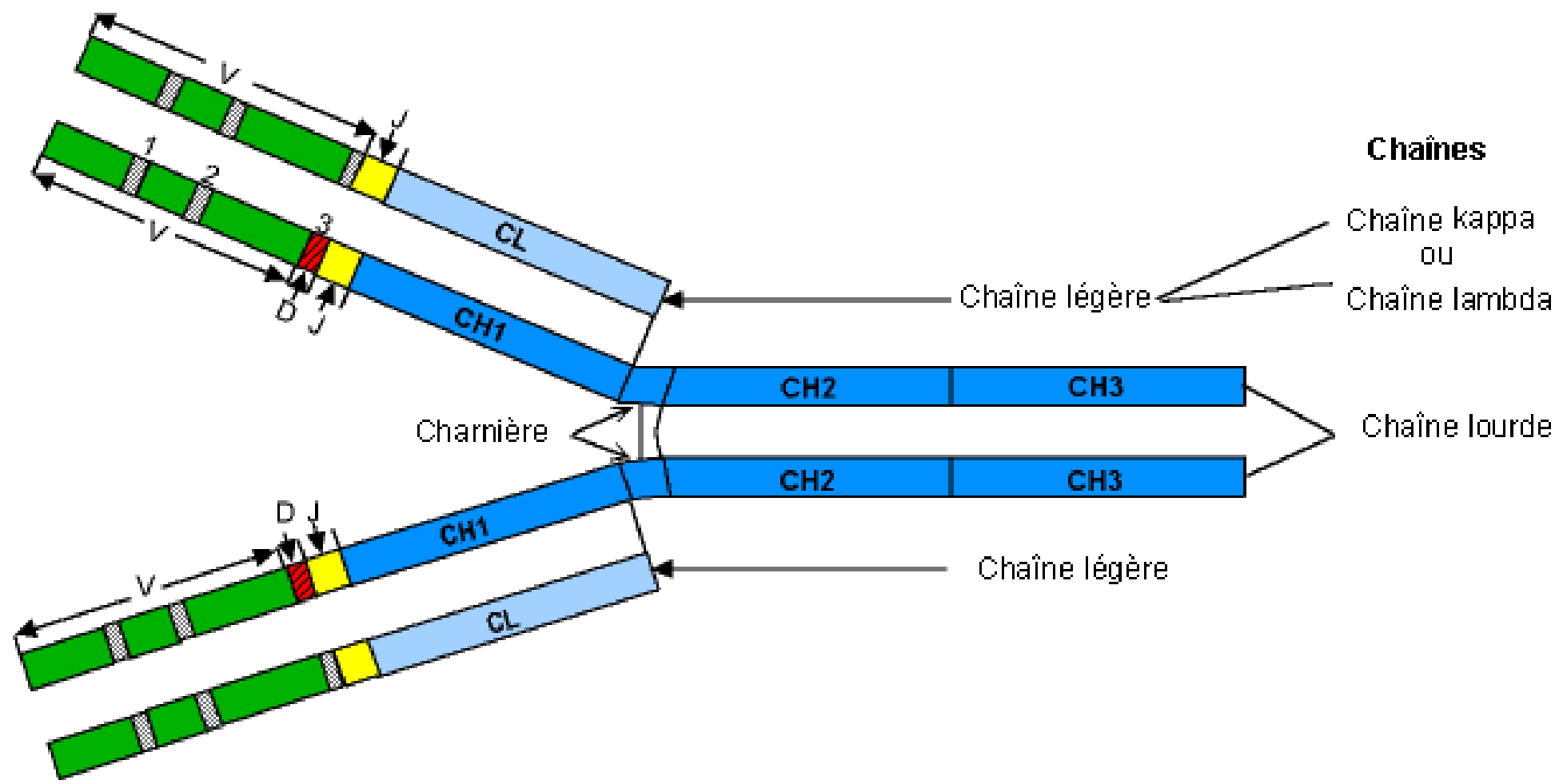


T cell receptor

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

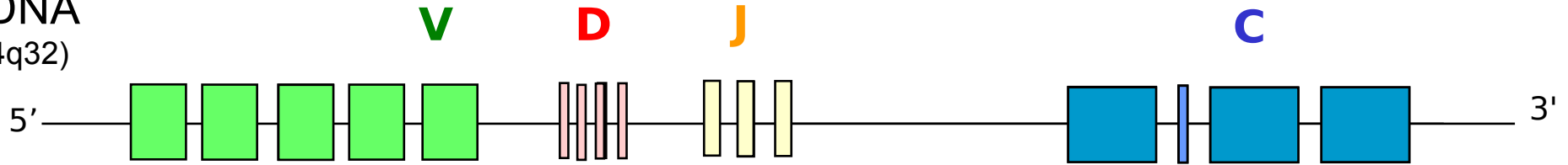
Alpha - Beta
Gamma - Delta

Immunoglobulin IgG



Immunoglobulin (IG) synthesis

genomic DNA
(IGH Locus 14q32)

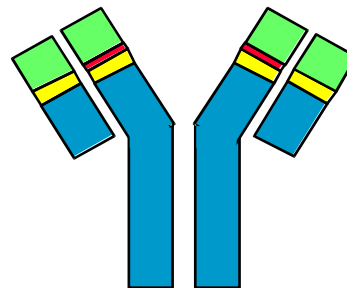


rearranged
DNA



mRNA

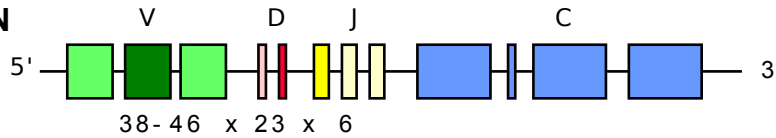
2×10^{12} different IG
per individual



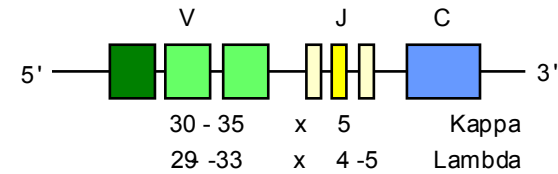
Immunoglobulin (IG) synthesis

150 FUNCTIONAL IG GENES

HEAVY CHAIN



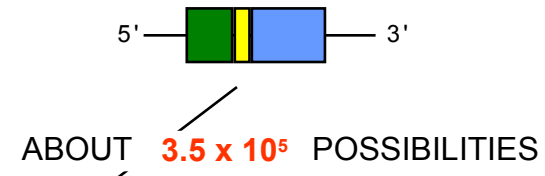
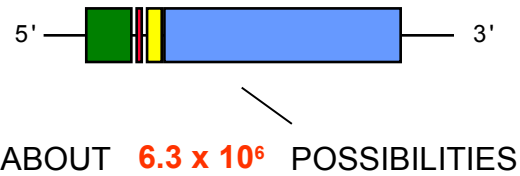
LIGHT CHAIN



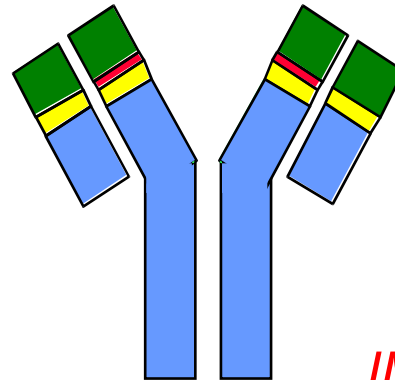
6300 POTENTIAL RECOMBINATIONS

185 + 165 POTENTIAL RECOMBINATIONS

N-DIVERSITY
SOMATIC MUTATIONS
x 1000



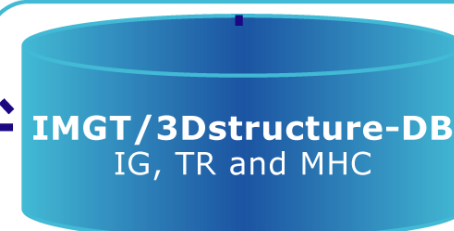
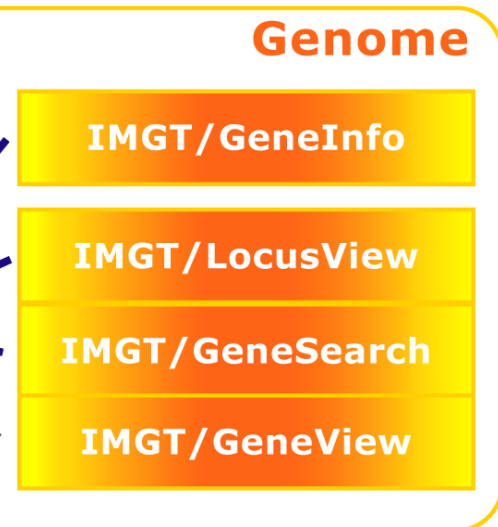
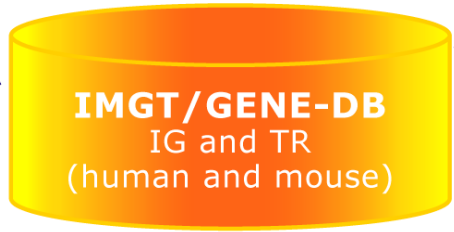
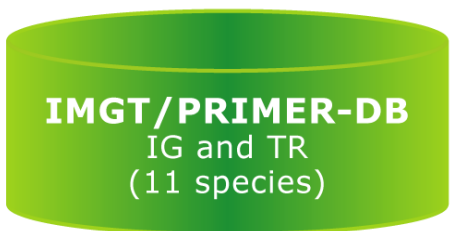
2×10^{12}
DIFFERENT ANTIBODIES



Sequences



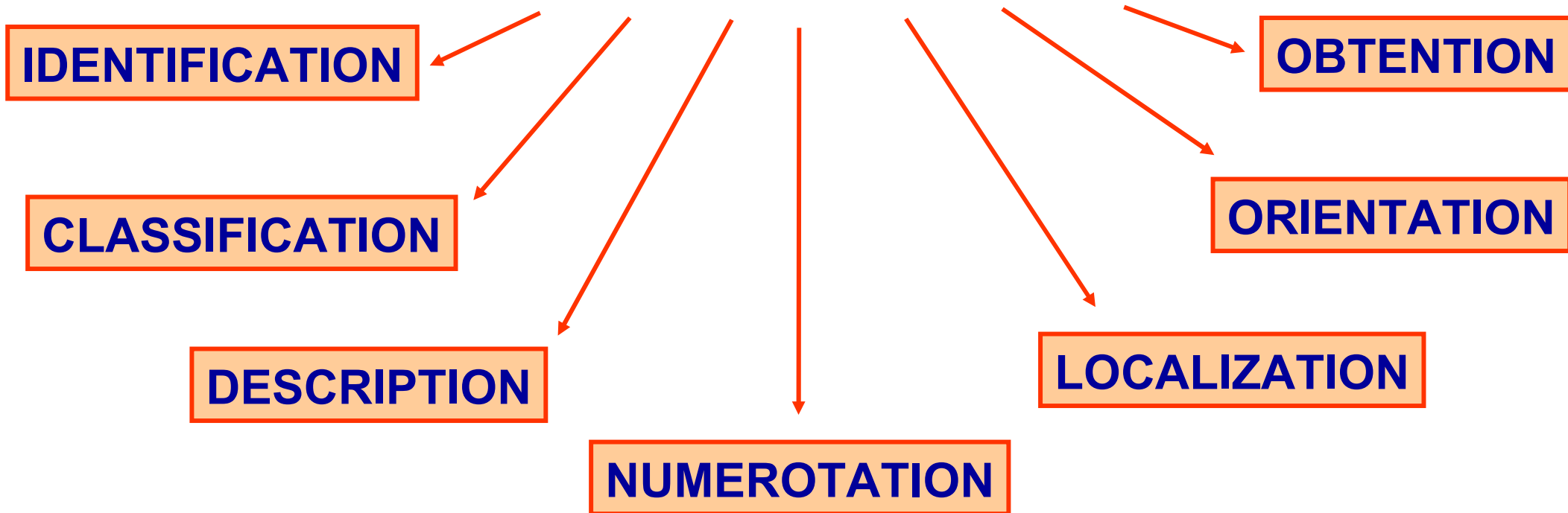
<http://imgt.cines.fr>



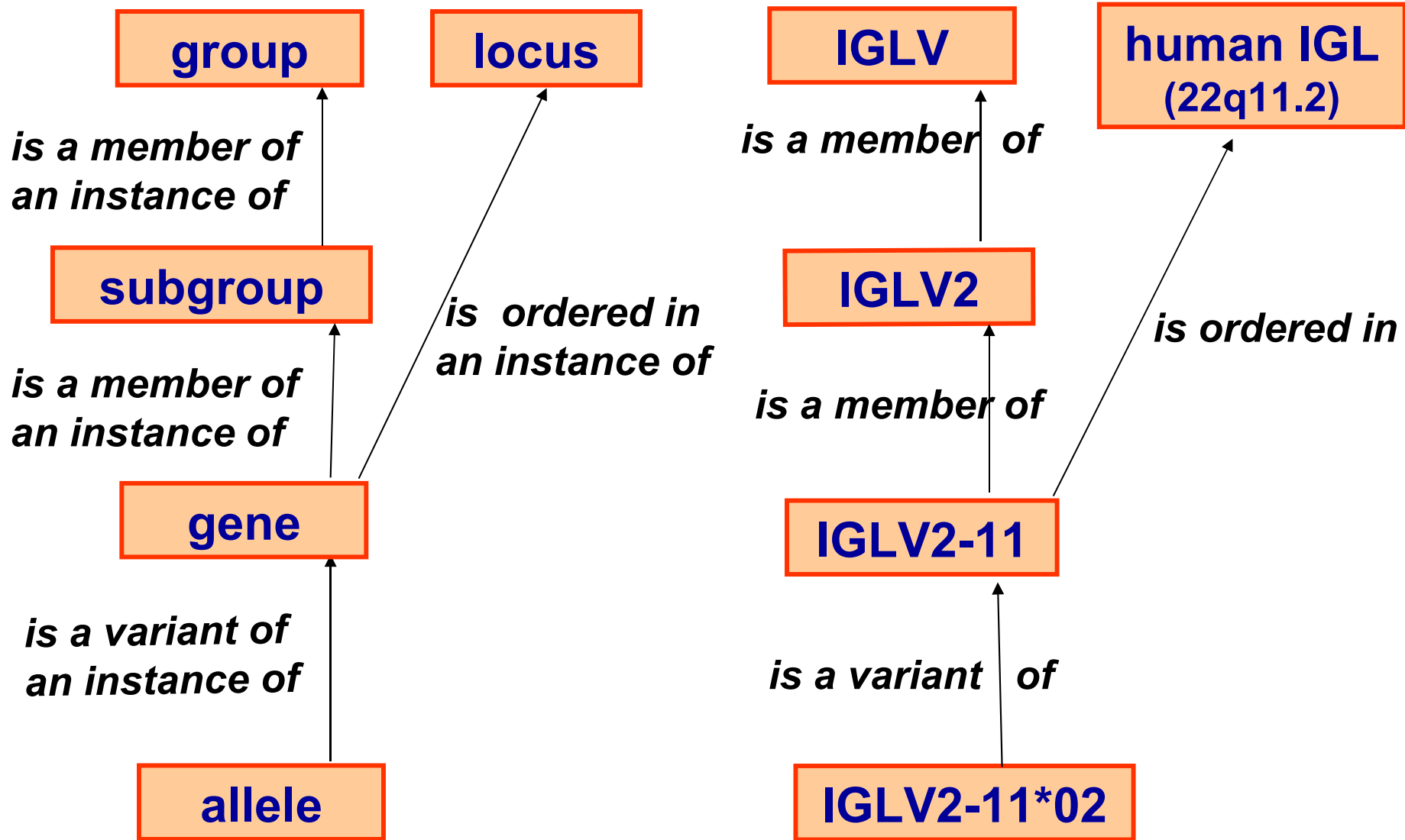
2D and 3D structures

IMGT-ONTOLOGY seven axioms:

To share, reuse and represent knowledge
in Immunogenetics and Life Sciences



CLASSIFICATION axiom

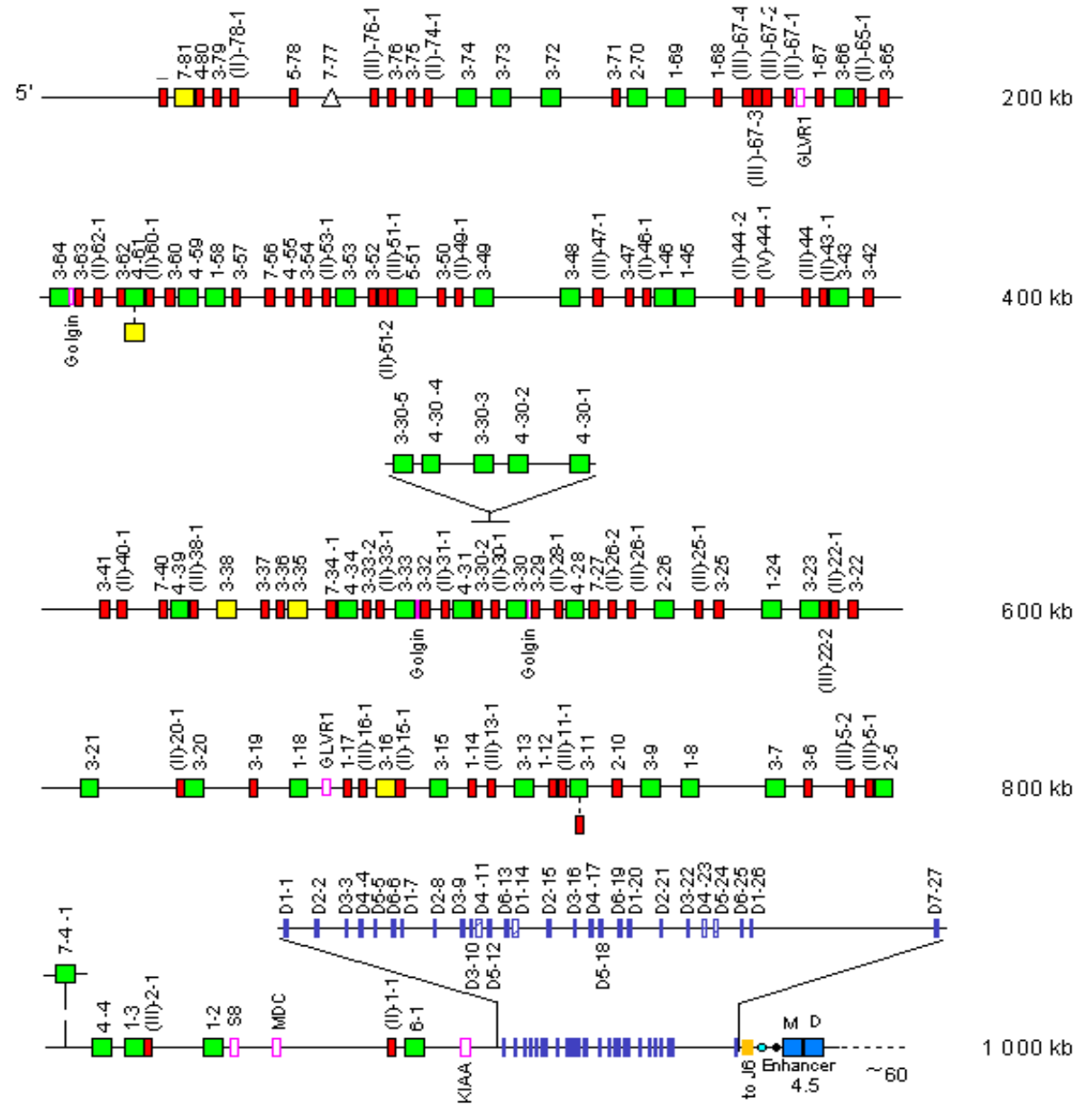


« Concepts »

« Instances »

IG and TR genes: V, D, J and C

Homo sapiens
 IGH locus on
 chromosome 14
 at 14q32.33



Locus representation: Human IGL

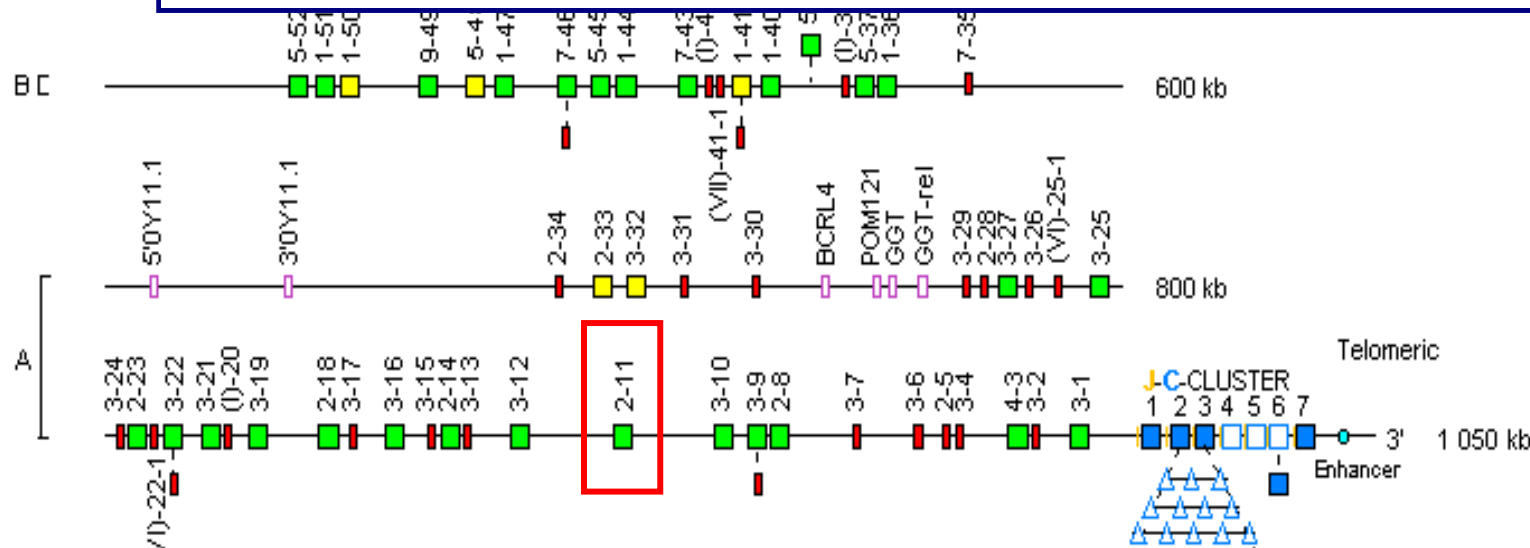
Human IGL 2

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

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>



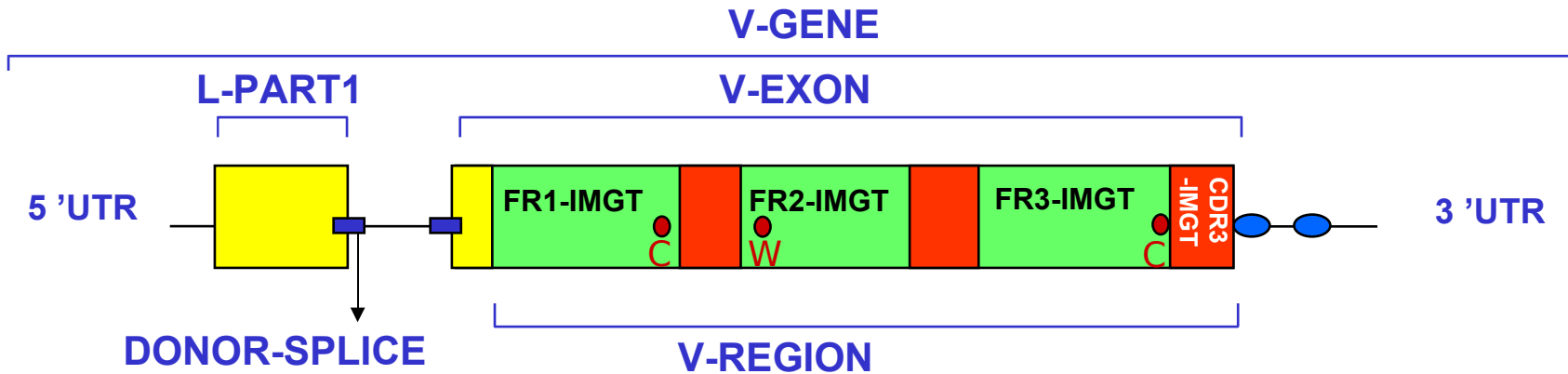
IG and TR: 1538 genes and 2523 alleles (human, mouse)

CLASSIFICATION axiom

- The IMGT-ONTOLOGY main concepts of classification include 'group', 'subgroup', 'gene', 'allele'.
- They allowed to set up the nomenclature for IG and TR genes (V, D, J, C genes).
- IMGT gene names were approved by HGNC in 1999 and entered in GDB, LocusLink and Entrez Gene (NCBI).
- IMGT/GENE-DB is the international reference database for IG and TR genes (direct links from Entrez Gene NCBI).
- WHO-IUIS/IMGT 2007 report (*Dev. Comp. Immunol., Immunogenetics*).

DESCRIPTION axiom

PROTOTYPE for a V-GENE



Label 1

Label 2

Relations entre Labels

V-GENE

V-EXON



FR3-IMGT

CDR3-IMGT



L-PART1

DONOR-SPLICE



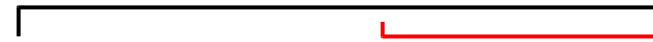
V-REGION

FR1-IMGT



V-REGION

CDR3-IMGT



IMGT/LIGM-DB

IMGT/LIGM-DB Consultation module v3 - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

FH Key Location/Qualifiers

FH	Key	Location/Qualifiers
FT	L-V-D-J-C-SEQUENCE	<1..375> 114,673 sequences from 210 species
FT		/partial
FT		/db_xref="taxon:9606"
FT		/cell_type="B-cell hybridoma 2F7"
FT		/IMGT_note="automatically annotated with IMGT tools"
FT		/organism="Homo sapiens"
FT	V-D-J-REGION	1..375
FT		/translation="QVHLVESGGAVFHPGRSLRLSRAASGFTFSSYGMHWVRQAPAKGLEWVAVIWIYDGSNKYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHVTIAAAGRRGAGMDVWGQTTVTVSS"
FT	V-REGION	1..296
FT		/allele="IGHV3-33*01, putative"
FT		/gene="IGHV3-33"
FT		/CDR_length="[8.8.18]"
FT		/putative_limit="3' side"
FT		/translation="QVHLVESGGAVFHPGRSLRLSRAASGFTFSSYGMHWVRQAPAKGLEWVAVIWIYDGSNKYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAK"
FT	FR1-IMGT	1..75
FT		/AA_IMGT="1 to 26, AA 10 is missing"
FT		/translation="QVHLVESGGAVFHPGRSLRLSRAAS"
FT	CDR1-IMGT	76..99
FT		/AA_IMGT="27 to 34"
FT		/translation="GFTFSSYG" IMGT-ONTOLOGY: 277 IMGT labels for sequences
FT	FR2-IMGT	100..150
FT		/AA_IMGT="39 to 55" 285 IMGT labels for 3D structures
FT		/translation="MHWVRQAPAKGLEWVAV"
FT	CONSERVED-TRP	106..108
FT	CDR2-IMGT	151..174
FT		/AA_IMGT="56 to 63"
FT		/translation="IWIYDGSNK" SO (Sequence ontology): 67 IMGT labels
FT	FR3-IMGT	175..288
FT		/AA_IMGT="66 to 104, AA 73 is missing"
FT		/translation="YYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCA"

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

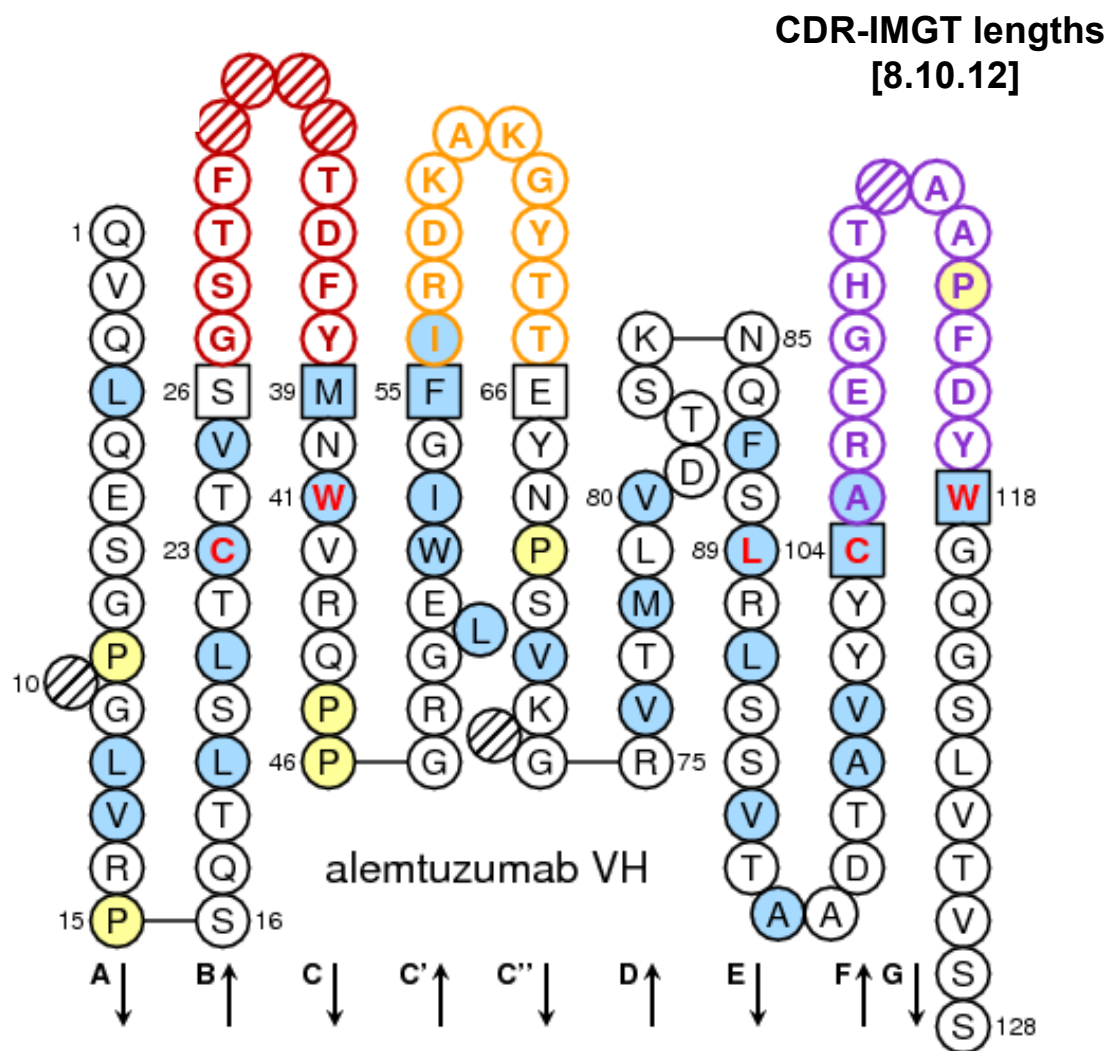
Done

DESCRIPTION axiom

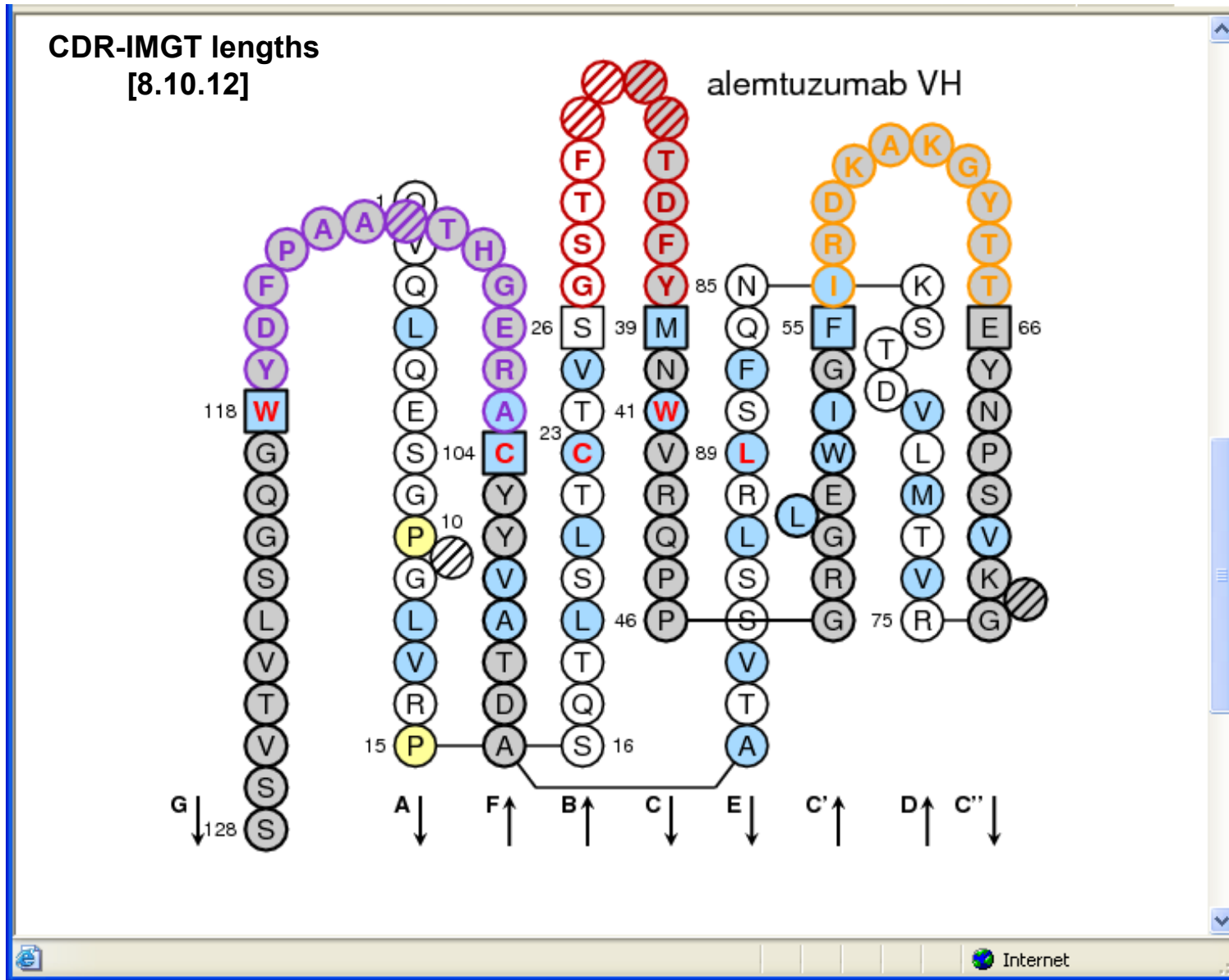
- The IMGT-ONTOLOGY concepts of description comprise the standardized IMGT labels and relations.
- They have allowed to describe the IG, TR and MHC sequences and 3D structures, **whatever the receptor type, the chain type, or the species.**
- They are particularly useful to describe IG, TR, and MHC and their complexes (**IG/antigen, TR/pMHC**).
- It is possible to query the IMGT® databases (**IMGT/LIGM-DB for sequences, IMGT/3Dstructure-DB for 3D structures**) with IMGT labels.
- Sequence Ontology (SO) includes IMGT labels.

NUMEROTATION axiom

IMGT
Collier
de
Perles

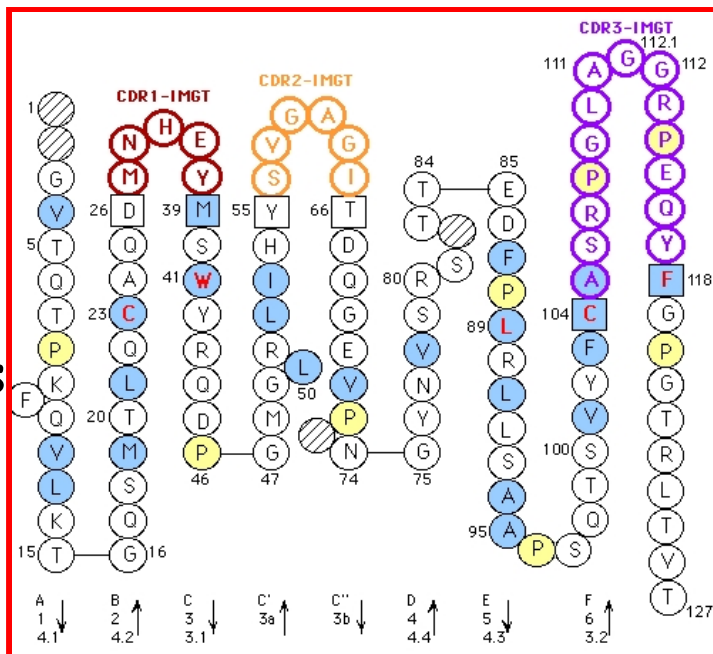


IMGT Collier de Perles on two layers



IMGT Web resources: 10 000 pages HTML

IMGT
Collier
de Perles



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
X02850 ,TRAV8-6*01	A	Q	S	V	T	Q	L	D	S	Q	V	P	V	F	E	E	A	P	V	E
AE000659,TRAV8-6*02
M86361 ,TRAV8-6*02

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
X02850 ,TRAV8-6*01	L	R	C	N	V	S	S	V	S	V	V	V	V	V	L	F				
AE000659,TRAV8-6*02
M86361 ,TRAV8-6*02

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
X02850 ,TRAV8-6*01	T	G	T	A	T	C	A	R	C	C	C	A	R	G	A	G	A	G	A	G
AE000659,TRAV8-6*02
M86361 ,TRAV8-6*02

	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
X02850 ,TRAV8-6*01	T	L	L	V	V	E	S	I	N											
AE000659,TRAV8-6*02
M86361 ,TRAV8-6*02

	81	82	83	84	84A	84B	84c	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
X02850 ,TRAV8-6*01	A	E	F	H	K	S	Q	T	S	F	H	L	R	K	P	S	V	H	I	S	D	T	A
AE000659,TRAV8-6*02
M86361 ,TRAV8-6*02

	101	102	103	104	105	106	107	108
X02850 ,TRAV8-6*01	E	Y	F	C	A	V	S	
AE000659,TRAV8-6*02
M86361 ,TRAV8-6*02

IMGT
Alignment
of alleles

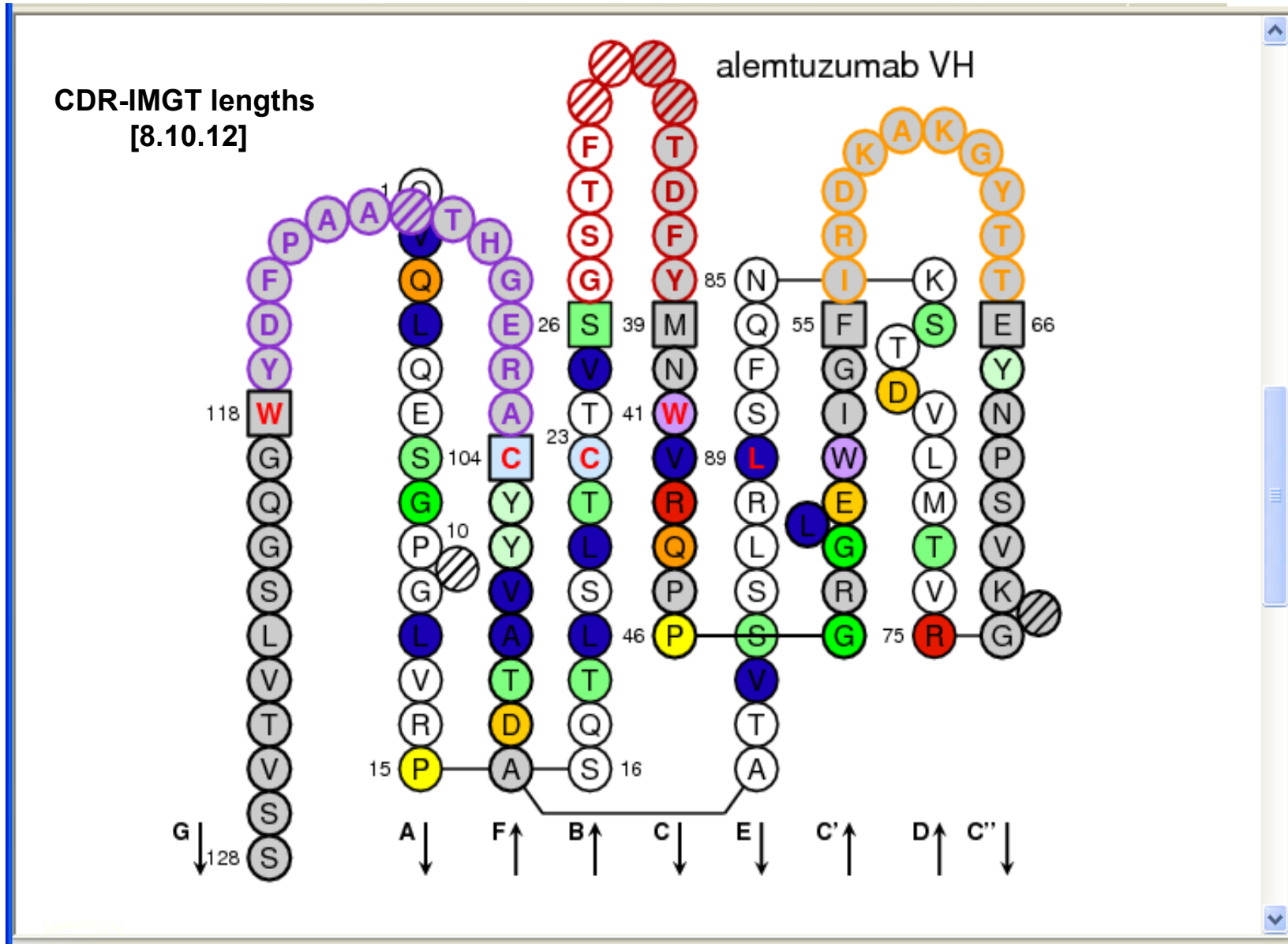
IMGT
Protein
Display

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)
AE000658, TRAV1-1	GQSLEQ.PSEVTAVEGAIQINCTYQ	TSGFYG.....	LSWYQQHGGAPTFLSY	NALDG.....	LEETG.....	RFSSFLSRSDSYGYLLLQELQMKDSASYFC
AE000658, TRAV1-2	GQNIDQ.PTEMTATEGAIQINCTYQ	TSGFNG.....	LFWYQQHAGEAPTFLSY	NVLDT.....	LEEKG.....	RFSSFLSRSGYGYLLLQELQMKDSASYLC
AE000658, TRAV2	KDQVFQ.PSTVASSEGAIVVEIFCNHS	VSNAYN.....	PFWYLFHPGCAPIRLLVK	GSK.....	PSQQG.....	RYNMTYER.FSSSLLILQVREADAAVYFC
AE000658, TRAV3	AQSVAQPEDQVNAEAGNPLTVKCTYS	VSGNPY.....	LFWYVQYPNRLQLFLK	YITGDNL.....	VKGSY.....	GFEAEFNKSTSPHLKPKSALVSDSALYFC
AE000658, TRAV4	LAKITQ.PISMDSYEGQEVNITOSH	NIATNDY.....	ITWYQQFSPQGRFIIQ	GYKT.....	KVINE.....	VASFLIPADRKSFTLSLPRVSLSDTAVYFC
AE000659, TRAV5	GEDVEQS.LFLSVREGDSSVINCTYT	DSSSTY.....	LYWYKQEPGAGLQLLTY	IFSNMD.....	MKQDQ.....	RLTLLNKKDKHLSLRIADTQTGDSAIYFC
AE000659, TRAV6	SQKIEQNSALNIQEGKTATLTCNYT	NYSPAY.....	LQWYRQDPGRGPVFLLL	IRENEK.....	EKRKE.....	RLKVTFTDITLQKSLFHITASQPADSATYFC
AE000659, TRAV7	ENQVEHSPHFLGPQQGDVASMCTYS	VSRFNN.....	LQWYRQNTGMGPKHLLS	MYSAGY.....	EKQKG.....	RLNATLLK.NGSSLYITAVQPEDSATYFC
AE000659, TRAV8-1	AQSVSQHNHHVILSEASLELGCNYS	YGGTVM.....	LFWYVQYPGQHLQLLLK	YFSGDPL.....	VKGIK.....	GFEAEFIKSKFSPNLRKPSVQVSDTAEYFC
AE000659, TRAV8-2	AQSVTQLDSHVSVSEGTVPVLLRCNYS	SSYSPS.....	LFWYVQHPNKGQLQLLLK	YTSAAITL.....	VKGIN.....	GFEAEFKKSETSPHLTKPSAHMSDAAEYFC
AE000659, TRAV8-3	AQSVTQPDIIHITVSEGAFLVLLRCNYS	YGATPY.....	LFWYVQSPGQGLQLLLK	YFSGDITL.....	VQGIK.....	GFEAEFKKSSPFLRKP SVHWSDAAEYFC
AE000659, TRAV8-4	AQSVTQLGSHVSVSEGALVLLRCNYS	SSVPPY.....	LFWYVQYPNQLQLLLK	YTSAAITL.....	VKGIN.....	GFEAEFKKSETSPHLTKPSAHMSDAAEYFC
X02850 ,TRAV8-6	AQSVTQLDSQVPVFEAAPVELRCNYS	SSVSVY.....	LFWYVQYPNQLQLLLK	YLSGSL.....	VESIN.....	GFEAEFNKSTSPHLRKP SVHISDTAEYFC
AE000660, TRAV8-7	TQSVTQLDGHITVSEEAPELRCNYS	YSGVPS.....	LFWYVQYSSQSLQLLLK	DLTEATQ.....	VKGIK.....	GFEAEFKKSETSPFLRKPSTHVSDAAEYFC
AE000659, TRAV9-1	GDSSVQVTEGQVLPSEGDSLIYVCSYF	TTQVPS.....	LFWYVQVPPGEGPOLH.K	AMKAND.....	KGRNK.....	GFEAMYRKFSTSPHLKPKSVQFSDSAVYFC

The eleven IMGT amino acid classes according to the physicochemical properties

'Volume' classes		'Hydropathy' classes								
	in Å ³	Hydrophobic		Neutral		Hydrophilic				
Very large	189-228	F		W		Y				
Large	162-174	I	L	M			K	R		
Medium	138-154	V				H	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	

IMGT Collier de Perles amino acid profile



Comparison with human germline genes:

- **Percentage of identity of the V-REGION**
- **Percentage of identity of the 4 FR-IMGT**
(91 VH positions, 89 V-KAPPA positions)
- **Number of amino acids with IMGT class change in the 4 FR-IMGT**
(*'class' refers to the 11 IMGT physicochemical properties aa classes*)

alemtuzumab	IGHV4-59*01	(73 %)	84.61 % (77/91)	2/91
	IGKV1-33*01	(86.32 %)	97.75 % (87/89)	1/89
bevacizumab	IGHV7-4-1*02	(72.40 %)	74.72 % (68/91)	1/91
	IGKV1-33*01	(87.40 %)	92.13 % (82/89)	2/89
trastuzumab	IGHV3-66*01	(81.63 %)	90.10 % (82/91)	0/91
	IGKV1-39*01	(86.32 %)	93.25 % (83/89)	1/89

NUMEROTATION axiom

- The IMGT-ONTOLOGY concepts of numerotation include IMGT unique numbering and IMGT Collier de Perles for V-DOMAIN (IG and TR).
- They have been extended to the C-DOMAIN (IG and TR) and G-DOMAIN (MHC).
- They have allowed to bridge the gap between sequences and 3D structures in IMGT/3Dstructure-DB.
- They are used for mutations, polymorphisms, CDR-IMGT lengths, contact analysis, potential immunogenicity evaluation and paratope definition.
- WHO-INN programme requires the CDR-IMGT lengths for antibody.

WELCOME ! to the IMGTV-QUEST Search page

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>

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

You are in the new IMGTV-QUEST, upgraded for multiple sequences and with new functionalities. **NEW!**

Analyse your Immunoglobulin nucleotide sequences

- [Human](#)
- [Mouse](#)
- [Chondrichthyes](#)
- [Teleostei](#)
 - [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](#)

Analysis by batches of up to 50 sequences in a single run

Nucleotide sequences

Enter your sequence(s) in [FASTA format](#) (FASTA format is required):

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

```
>AY393054
gtgggttttccctgttgctattttaaaagggtgtccaatgtgaggtgcagctggtggagctctgggggaggcttggtagcagccagggcgg
tcctgagactctcctgtgcagcttctggattgacctttggtgattactttatgagctggttccgccaggctccaggggaagggactgga
gtgggtaggtttcattaagagcgaacttatggtgggacaacagaatacggccgctctgtgaaaggcagattcatcatctcgagagatg
atccaaaagcatcgctatttgcaaatgaacagcctggaaaccgaggacacagccatatactgtagtcccgggggtagtgtctat
taccagaaacttccagcagtggggcccgggcaccatggtcacctctcctcagcctccaccaagggcccacgggtctccccctggc
accctctccaagagcacctctgggggcacagcggccctgggctgctggtcaaggactacttcccc
>AY393055
gtgggttttccctgttgctatttcaaaagggtgtccagtgtgaggtgcagctggtggagactggaggaggcttgatccagcctgggggg
tcctgagactctcctgtgcagcctctgggttcaccgtcagtagcaactacatgagctgggtccgccaggctccaggggaaggggctgga
```

Or give the path access to a local file containing your sequence(s) in [FASTA format](#) (FASTA format is required):

Selection of parameters for the results

Display type :

Nb of nucleotides per line in alignment:

A. Detailed view

1. <input checked="" type="checkbox"/> Alignment for V-GENE	5. <input type="checkbox"/> Sequence of the JUNCTION ('nt' and 'AA')	11. IMGTV Collier de Perles
2. <input checked="" type="checkbox"/> Alignment for D-GENE	6. <input checked="" type="checkbox"/> V-REGION alignment according to the IMGTV numbering	<input checked="" type="radio"/> links to IMGTV Collier de Perles
3. <input checked="" type="checkbox"/> Alignment for J-GENE	7. <input checked="" type="checkbox"/> V-REGION translation	<input type="radio"/> IMGTV Collier de Perles (PNG format, slow)
4. <input checked="" type="checkbox"/> Results of IMGTVJunctionAnalysis	8. <input type="checkbox"/> V-REGION mutation table	<input type="radio"/> no IMGTV Collier de Perles
<input type="radio"/> with full list of eligible D-GENES	9. <input type="checkbox"/> V-REGION mutation statistics	
<input checked="" type="radio"/> without list of eligible D-GENES	10. <input type="checkbox"/> V-REGION mutation hot spots	
12. <input type="checkbox"/> Sequences of V-, V-J- or V-D-J- REGION ('nt' and 'AA') with gaps in FASTA <small>Access to IMGTV/PhyloGene for V-REGION ('nt')</small>		13. <input type="checkbox"/> Annotations by IMGTV/Automat

B. Synthesis view

1. <input checked="" type="checkbox"/> Alignment for V-GENES	5. <input checked="" type="checkbox"/> V-REGION protein display (with color)
2. <input checked="" type="checkbox"/> V-REGION alignment according to the IMGTV numbering	6. <input checked="" type="checkbox"/> V-REGION protein display (mutations displayed)
3. <input checked="" type="checkbox"/> V-REGION translation	7. <input checked="" type="checkbox"/> V-REGION most frequently occurring AA
4. <input checked="" type="checkbox"/> V-REGION protein display	8. <input checked="" type="checkbox"/> Results of IMGTVJunctionAnalysis

Advanced parameters

Selection of IMGTV reference directory set	<input type="text" value="F+ORF+ in frame P"/>	<input checked="" type="radio"/> With all alleles <input type="radio"/> With allele *01 only
Selection of parameters for IMGTVJunctionAnalysis	Nb of D-GENES in IGH JUNCTIONs (default is 1) <input type="text" value="default"/>	Number of accepted mutations: <input type="text" value="default"/> in 3'V-REGION
		<input type="text" value="default"/> in D-REGION
		<input type="text" value="default"/> in 5'J-REGION
More options for Detailed view	Nb of nucleotides to exclude in 5' of the V-REGION for the evaluation of the nb of mutations (in results 8 and 9) <input type="text" value=""/>	Nb of nucleotides to add (or exclude) in 3' of the V-REGION for the evaluation of the alignment score (in result 1) <input type="text" value=""/>

3. V-REGION translation

```

<-----FR1 - IMGT
1           5           10           15
E V Q L V Q S G A ... E V K K P
X56368 IGHV5-51*03 gag gtg cag ctg gtg cag tct gga qca ... gag gtg aaa aag ccg
seq1          R
seq2          R
seq3          G           E
seq4          G           E
seq5          R
seq6          R
----->
                20           25           30
G E S L K I S C K G S G Y S F
X56368 IGHV5-51*03 ggg gag tct ctg aag atc tcc tgt aag ggt tct gga tac agc ttt
seq1          A
seq2          R A
seq3          A           T
seq4          A           T
seq5          A           T
seq6          A           T

```

Selection of parameters for the results

Display type:

Nb of nucleotides per line in alignment:

A. Detailed view

- | | | |
|------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|
| 1. <input checked="" type="checkbox"/> Alignment for V-GENE | 5. <input type="checkbox"/> Sequence of the JUNCTION ('nt' and 'AA') | 11. IMG T Collier de Perles |
| 2. <input checked="" type="checkbox"/> Alignment for D-GENE | 6. <input checked="" type="checkbox"/> V-REGION alignment according to the IMGT numbering | <input checked="" type="radio"/> links to IMG T Collier de Perles |
| 3. <input checked="" type="checkbox"/> Alignment for J-GENE | 7. <input checked="" type="checkbox"/> V-REGION translation | <input type="radio"/> IMG T Collier de Perles (PNG format, slow) |
| 4. <input checked="" type="checkbox"/> Results of IMG T/JunctionAnalysis | 8. <input type="checkbox"/> V-REGION mutation table | <input type="radio"/> no IMG T Collier de Perles |
| <input type="radio"/> with full list of eligible D-GENEs | 9. <input type="checkbox"/> V-REGION mutation statistics | |
| <input checked="" type="radio"/> without list of eligible D-GENEs | 10. <input type="checkbox"/> V-REGION mutation hot spots | |
| 12. <input type="checkbox"/> Sequences of V-, V-J- or V-D-J- REGION ('nt' and 'AA') with gaps in FASTA | | 13. <input type="checkbox"/> Annotations by IMG T/Automat |
| Access to IMG T/PhyloGene for V-REGION ('nt') | | |

B. Synthesis view

- | | |
|-----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
| 1. <input checked="" type="checkbox"/> Alignment for V-GENEs | 5. <input checked="" type="checkbox"/> V-REGION protein display (with color) |
| 2. <input checked="" type="checkbox"/> V-REGION alignment according to the IMGT numbering | 6. <input checked="" type="checkbox"/> V-REGION protein display (mutations displayed) |
| 3. <input checked="" type="checkbox"/> V-REGION translation | 7. <input checked="" type="checkbox"/> V-REGION most frequently occurring AA |
| 4. <input checked="" type="checkbox"/> V-REGION protein display | 8. <input checked="" type="checkbox"/> Results of IMG T/JunctionAnalysis |

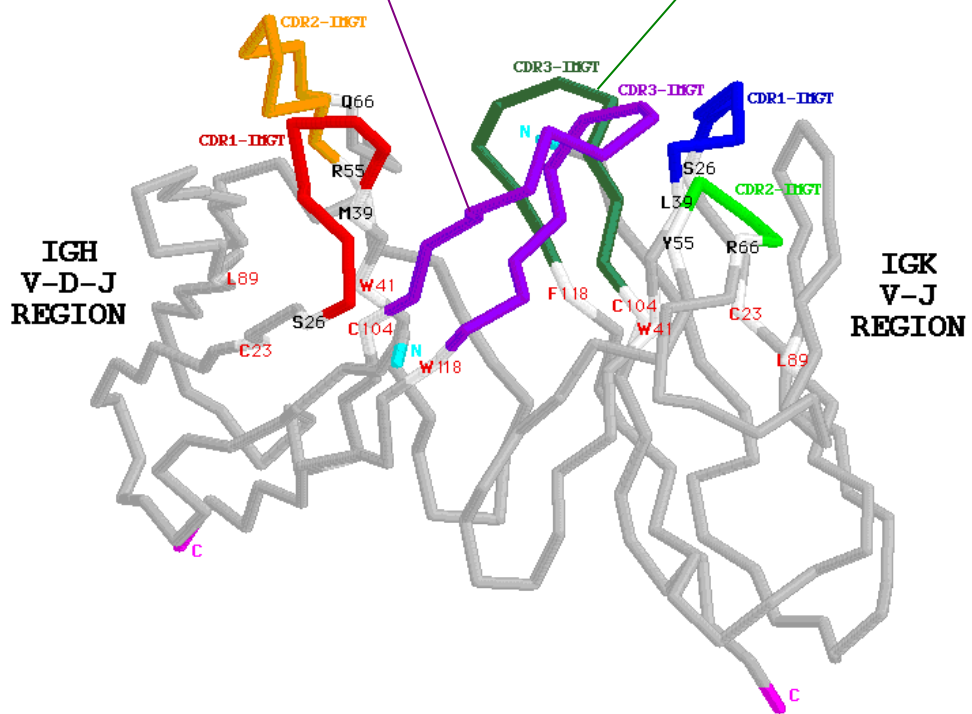
Selection of advanced parameters

Advanced parameters

Selection of IMGT reference directory set	<input type="text" value="F+ORF+ in frame P"/>	<input checked="" type="radio"/> With all alleles <input type="radio"/> With allele *01 only
Selection of parameters for IMG T/JunctionAnalysis	Nb of D-GENEs in IGH JUNCTIONs (default is 1) <input type="text" value="default"/>	Number of accepted mutations: <input type="text" value="default"/> in 3'-REGION <input type="text" value="default"/> in D-REGION <input type="text" value="default"/> in 5'-REGION
More options for Detailed view	Nb of nucleotides to exclude in 5' of the V-REGION for the evaluation of the nb of mutations (in results 8 and 9) <input type="text" value=""/>	Nb of nucleotides to add (or exclude) in 3' of the V-REGION for the evaluation of the alignment score (in result 1) <input type="text" value=""/>

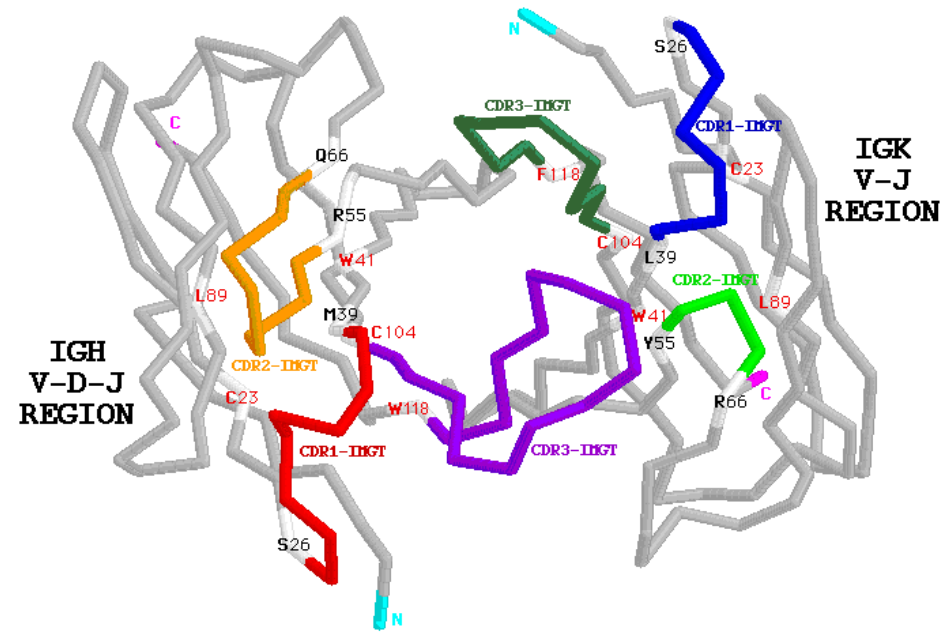
V-DOMAINS: VH and V-KAPPA

VH **V-KAPPA**
V-D-J junction **V-J junction**



Side view of the V-DOMAINS

VH **V-KAPPA**



View from above the CDRs

Mouse (Mus musculus) E5.2Fv

CDR3-IMG1= Complementarity determining region (105-117)
V-J junction (104-118)
V-D-J junction (104-118)

Immunoglobulin V-D-J generation of sequence 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

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		tgttgtgcagc <u>g</u> cctggtac	ccaaatac		...actttgaccactgg	IGHJ4*02	IGHD6-13*01	1	2	1	5/15
#2	Z70257	IGHV3-7*02		ggatggcagctcttatgcc	cgccc		ctactggtaactcgatctctgg	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		..atgcttttgatatctgg	IGHJ3*02	IGHD3-3*01	0	0	0	12/17
#5	Z70610	IGHV4-34*09		tcgggagcgatttttgagtggttatt....	cccga	ca	tgatgcttttgatatctgg	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		cagcagctggtac	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		.tactggtaactcgatctctgg	IGHJ2*01	IGHD5-24*01	0	2	0	7/13
#10	Z70616	IGHV4-34*01		cgggttggg.....	ttccc		...actggttcgaccctgg	IGHJ5*02	IGHD3-16*01	0	0	0	6/8
#11	Z70620	IGHV4-30-4*01		ccggggcgggatggtt.....	cgg		.gatgcttttgatatctgg	IGHJ3*02	IGHD3-16*01	1	4	0	5/5
#12	Z70621	IGHV4-39*01		ccacgatttatggttcggggagtt.....	tgaccccc	ttgactactgg	IGHJ4*02	IGHD3-16*01	0	1	0	12/21
#13	Z70622	IGHV4-39*06	t	tgccccgctcctgccaaaat	gtattactatggttcgggga.....	tatgtacg	tttgactactgg	IGHJ4*03	IGHD3-10*01	0	0	0	15/28

IMGT/JunctionAnalysis: analysis of the IG and TR junctions

JUNCTION alignments with translation and IMGT AA classes

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

	104	105	106	107	108	109	110	111	111.1	111.2	111.3	112.3	112.2	112.1	112	113	114	115	116	117	118
#1 seq1	C	S	P	G	G	S	<u>A</u>	Y						Y	<u>H</u>	E	<u>H</u>	F	Q	Q	W
	tgt	agt	ccc	ggg	ggt	agt	g <u>ct</u>	tat	tac	<u>c</u> ac	gaa	<u>c</u> ac	ttc	cag	cag	tgg
#2 seq2	C	<u>V</u>	K	P	T	D	D	D	G				<u>H</u>	R	A	E	Y	F	Q	<u>Y</u>	W
	tgt	g <u>tg</u>	aaa	ccc	acg	gat	gat	gat	ggc	<u>c</u> ac	cgg	gct	gaa	tac	ttc	cag	<u>t</u> ac	tgg
#3 seq3	C	S	P	G	G	S	<u>A</u>	Y						Y	<u>H</u>	E	<u>D</u>	F	Q	Q	W
	tgt	agt	ccc	ggg	ggt	ag <u>c</u>	g <u>ct</u>	tat	tac	<u>c</u> ac	gaa	g <u>a</u> c	ttc	cag	cag	tgg
#4 seq4	C	S	P	G	G	S	<u>A</u>	Y						Y	<u>H</u>	E	<u>H</u>	F	Q	Q	W
	tgt	agt	ccc	ggg	ggt	agt	g <u>ct</u>	tat	tac	<u>c</u> ac	gaa	<u>c</u> ac	ttc	cag	cag	tgg
#5 seq5	C	A	R	Q	N	P	P	E	Y	S	G	A	Y	<u>H</u>	<u>D</u>	G	W	F	D	P	W
	tgt	gcg	aga	caa	aac	ccc	ccc	gag	tat	agt	ggc	gca	tat	<u>c</u> at	g <u>a</u> t	ggg	tgg	ttc	gac	ccc	tgg
#6 seq6	C	A	R	E	M	L	Y	G	S	G	<u>G</u>	Y	Y	P	P	D	A	F	E	<u>L</u>	W
	tgt	gcg	aga	gag	atg	ctc	tat	ggt	tcg	ggg	g <u>gt</u>	tat	tac	ccc	cct	gat	gca	ttt	gag	<u>g</u> tc	tgg
#7 seq7	C	A	R	Q	N	P	P	E	Y	S	G	A	Y	<u>H</u>	<u>D</u>	G	W	F	D	P	W
	tgt	gcg	aga	cag	aat	ccc	ccc	gag	tat	agt	ggc	gca	tat	<u>c</u> at	g <u>a</u> t	ggg	tgg	ttc	gac	ccc	tgg
#8 seq8	C	A	R	E	M	L	Y	G	S	G	<u>G</u>	Y	Y	P	P	D	A	F	E	V	W
	tgt	gcg	aga	gag	atg	ctc	tat	ggt	tcg	ggg	g <u>gt</u>	tat	tac	ccc	cct	gat	g <u>c</u> a	ttt	gag	gtc	tgg
#9 seq9	C	A	R	Q	N	P	P	E	Y	S	G	A	Y	<u>H</u>	<u>D</u>	G	W	F	D	P	W
	tgt	gcg	aga	cag	aac	ccc	ccc	gag	tat	agt	ggc	gca	tat	<u>c</u> at	g <u>a</u> t	ggg	tgg	ttc	gac	ccc	tgg

Analysis of the 3D structures



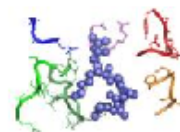
<http://imgt.cines.fr>

THANK YOU

for using [IMGT/3Dstructure-DB](#)

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

IMGT/3Dstructure-DB card for : **1ce1**



Entry code Search

[Query page](#)

IMGT protein name	IMGT receptor type	IMGT receptor description	Ligand(s)	Species	CC	Chain ID
CAMPATH-1H, alemruzumab , MABCAMPATH®	IG	FAB-GAMMA-1_KAPPA		Humanized	1	[1ce1_H 1ce1_L]
		Peptide	CD52 (synthetic peptide)	Synthetic	1	[1ce1_P]

Experimental technique **X-ray diffraction**

Resolution (in angstrom) **1.90**

PDB release date **25-JUN-99**

[Epitope and Chain details](#)

[Contact analysis](#)

[3D visualization](#)
Jmol or QuickPDB

[Renumbered IMGT file](#)

[IMGT numbering comparison](#)

[References and links](#)

[Printable card](#)

Contact analysis:

[IMGT/3Dstructure-DB Domain pair contacts](#) ([overview](#)) of **1ce1**

Atom contact types

Non covalent

Covalent

Atom contact categories

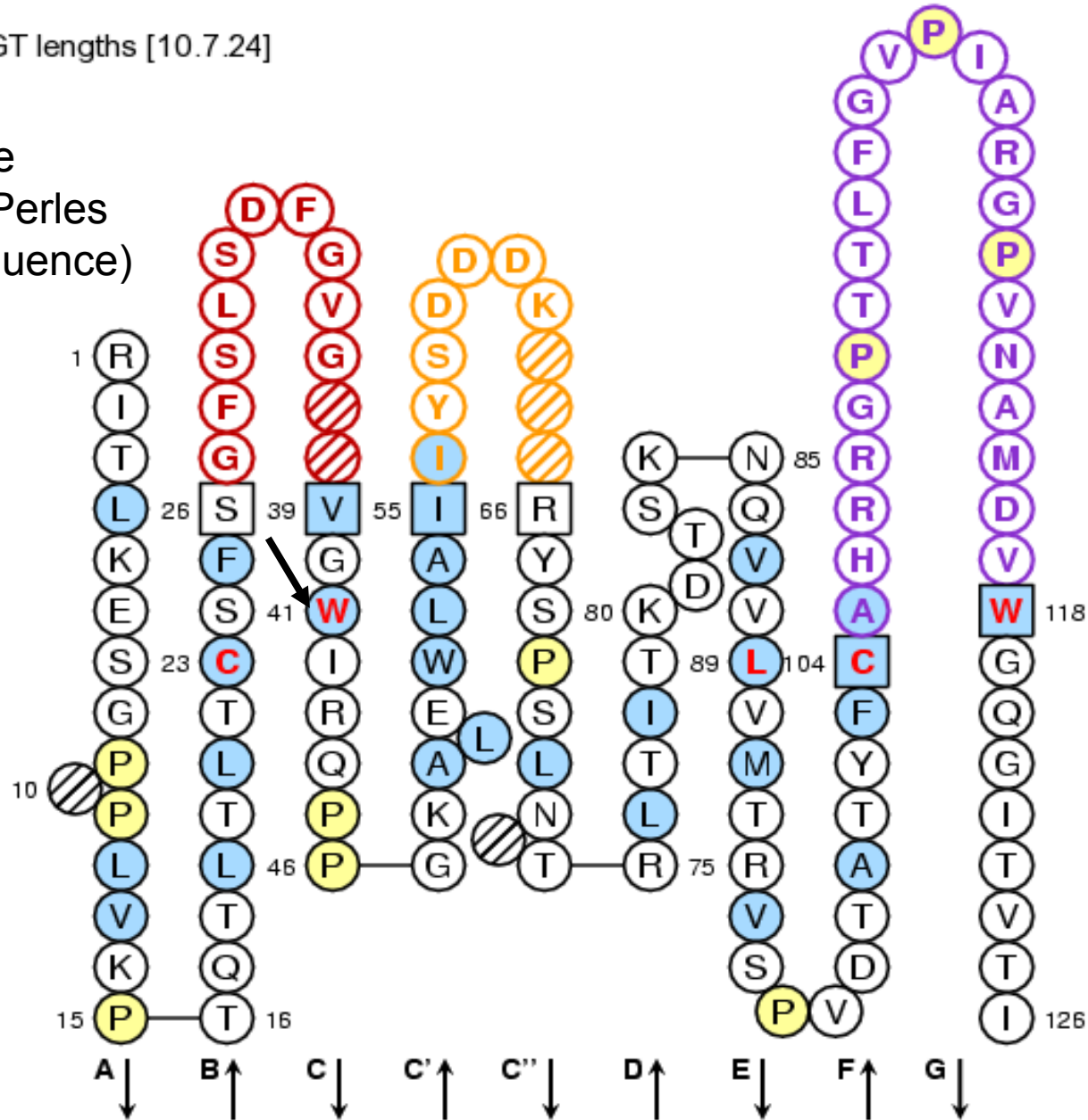
(BB) Backbone/backbone

Kaas Q. et al.

Access to atomic pair contacts in IMG/3Dstructure-DB

CDR-IMG/ lengths [10.7.24]

Click on residue
in IMG/ Collier de Perles
(or in amino acid sequence)



41V - TRP (W)

chain : 1u8k_B

Secondary structure	Extended conformation
Phi	-122.64
Psi	137.12
ASA	0.0

Residue local structure

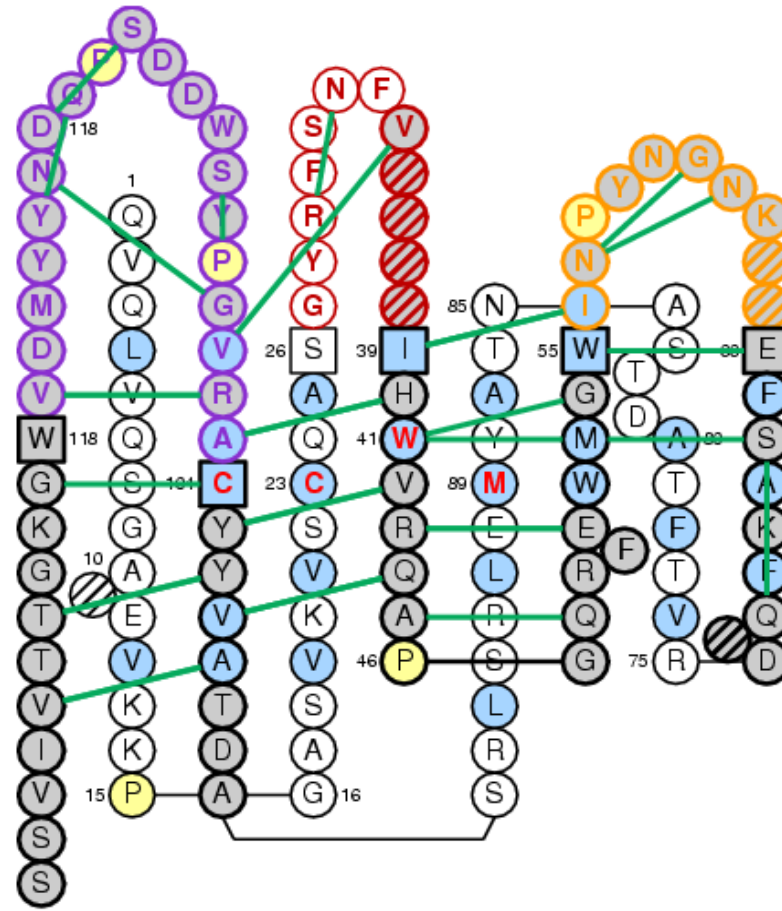
	IMG_T Num	Code 1C	Code 3C	Chain	Domain	Tot	NCov	Pol	HB	NPol	Cov	SS
Info	6V1	E	GLU	1u8k_B	VH	6	6	0	0	6	0	0
Info	21V1	L	LEU	1u8k_B	VH	17	17	0	0	17	0	0
Info	22V1	T	THR	1u8k_B	VH	8	8	0	0	8	0	0
Info	23V1	C	CYS	1u8k_B	VH	10	10	0	0	10	0	0
Info	39V1	V	VAL	1u8k_B	VH	2	2	1	0	1	0	0
Info	43V1	R	ARG	1u8k_B	VH	2	2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

Hydrogen bonds (IMGT Collier de Perles on 2 layers)

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

[8.8.20]



Contacts VH-(Ligand), V-KAPPA-(Ligand)

IMGT molecule name	IMGT description	Chain ID	IMGT chain description	Domain number	IMGT domain description
CAMPATH-1H, alemtozumab , MABCAMPATH@	FAB-GAMMA-1_KAPPA	1ce1_H	VH-CH1	[D1]	VH
				[D2]	CH1
		1ce1_L	L-KAPPA	[D1]	V-KAPPA
				[D2]	C-KAPPA
CD52 (synthetic peptide)	Peptide	1ce1_P	Peptide		

	Unit 1		Unit 2		Residue contacts	Number of residues			Atom contact types		
	Domain	Chain	Domain	Chain		Total	From 1	From 2	Total	Polar	Hydrogen
DomPair	VH	1ce1_H	CH1	1ce1_H	19	17	8	9	125	9	1
DomPair			V-KAPPA	1ce1_L	63	45	24	21	532	61	6
DomPair			(Ligand)	1ce1_P	25	19	12	7	216	40	9
DomPair	CH1	1ce1_H	VH	1ce1_H	19	17	9	8	125	9	1
DomPair			C-KAPPA	1ce1_L	68	58	28	30	498	40	6
DomPair	V-KAPPA	1ce1_L	VH	1ce1_H	63	45	21	24	532	61	6
DomPair			C-KAPPA	1ce1_L	18	18	8	10	137	19	2
DomPair			(Ligand)	1ce1_P	16	14	7	7	171	37	5
DomPair	C-KAPPA	1ce1_L	CH1	1ce1_H	68	58	30	28	498	40	6
DomPair			V-KAPPA	1ce1_L	18	18	10	8	137	19	2

Contacts VH-(Ligand)

IMGT/3Dstructure-DB Domain pair contacts

Contacts of

Domain Chain
VH 1ce1_H

with

Domain Chain
(Ligand) 1ce1_P

Atom contact types

- Non covalent
 Polar
 Hydrogen bond
 Non polar
 Covalent
 Disulfide

[Check all](#)
[Uncheck all](#)

Atom contact categories

- (BB) Backbone/backbone
 (SS) Side chain/side chain
 (BS) Backbone/side chain
 (SB) Side chain/backbone

[Check all](#)
[Uncheck all](#)

Show

Summary:

Residue contacts	Number of residues			Atom contact types		
	Total	From 1	From 2	Total	Polar	Hydrogen
25	19	12	7	216	40	9

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

	Order				R@P	Order				Atom contacts				
	IMGT Num	Residue	Domain	Chain		IMGT Num	Residue	Domain	Chain	Total	Polar	Hydrogen		
R@P	38	TYR	Y	VH	1ce1_H	R@P	2	THR	T		1ce1_P	4	0	0
R@P	38	TYR	Y	VH	1ce1_H	R@P	7	ALA	A		1ce1_P	13	1	0
R@P	38	TYR	Y	VH	1ce1_H	R@P	8	ASP	D		1ce1_P	14	2	2

Contacts VH-(Ligand)

	IMGT Num	Residue	Domain	Chain		IMGT Num	Residue	Domain	Chain	Total	Polar	Hydrogen	
R@P	38	TYR	Y	VH	1ce1_H	R@P	2	THR	T	1ce1_P	4	0	0
R@P	38	TYR	Y	VH	1ce1_H	R@P	7	ALA	A	1ce1_P	13	1	0
R@P	38	TYR	Y	VH	1ce1_H	R@P	8	ASP	D	1ce1_P	14	2	2
R@P	55	PHE	F	VH	1ce1_H	R@P	6	SER	S	1ce1_P	5	0	0
R@P	55	PHE	F	VH	1ce1_H	R@P	7	ALA	A	1ce1_P	16	0	0
R@P	55	PHE	F	VH	1ce1_H	R@P	8	ASP	D	1ce1_P	1	0	0
R@P	57	ARG	R	VH	1ce1_H	R@P	7	ALA	A	1ce1_P	9	3	2
R@P	57	ARG	R	VH	1ce1_H	R@P	8	ASP	D	1ce1_P	20	6	1
R@P	61	LYS	K	VH	1ce1_H	R@P	8	ASP	D	1ce1_P	11	2	1
R@P	66	GLU	E	VH	1ce1_H	R@P	7	ALA	A	1ce1_P	1	0	0
R@P	107	GLU	E	VH	1ce1_H	R@P	2	THR	T	1ce1_P	13	2	1
R@P	107	GLU	E	VH	1ce1_H	R@P	4	SER	S	1ce1_P	5	2	0
R@P	107	GLU	E	VH	1ce1_H	R@P	7	ALA	A	1ce1_P	5	0	0
R@P	108	GLY	G	VH	1ce1_H	R@P	1	GLY	G	1ce1_P	2	1	0
R@P	108	GLY	G	VH	1ce1_H	R@P	2	THR	T	1ce1_P	9	2	0
R@P	109	HIS	H	VH	1ce1_H	R@P	1	GLY	G	1ce1_P	24	4	0
R@P	109	HIS	H	VH	1ce1_H	R@P	2	THR	T	1ce1_P	21	5	0
R@P	109	HIS	H	VH	1ce1_H	R@P	3	SER	S	1ce1_P	9	2	1
R@P	110	THR	T	VH	1ce1_H	R@P	1	GLY	G	1ce1_P	1	1	0
R@P	110	THR	T	VH	1ce1_H	R@P	3	SER	S	1ce1_P	11	4	1
R@P	112	ALA	A	VH	1ce1_H	R@P	3	SER	S	1ce1_P	3	1	0
R@P	113	ALA	A	VH	1ce1_H	R@P	2	THR	T	1ce1_P	3	0	0
R@P	113	ALA	A	VH	1ce1_H	R@P	3	SER	S	1ce1_P	7	2	0
R@P	113	ALA	A	VH	1ce1_H	R@P	4	SER	S	1ce1_P	4	0	0
R@P	114	PRO	P	VH	1ce1_H	R@P	4	SER	S	1ce1_P	5	0	0

Contacts V-KAPPA-(Ligand)

Residue contacts	Number of residues			Atom contact types		
	Total	From 1	From 2	Total	Polar	Hydrogen
16	14	7	7	171	37	5

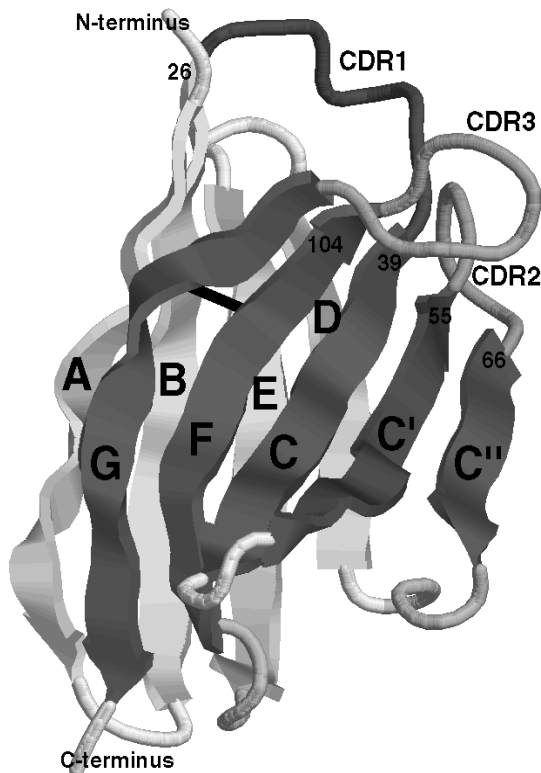
List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

R@P	Order					R@P	Order				Atom contacts			
	IMGT Num	Residue	Domain	Chain	IMGT Num		Residue	Domain	Chain	Total	Polar	Hydrogen		
R@P	38	TYR	Y	V-KAPPA	1ce1_L	R@P	3	SER	S		1ce1_P	1	0	0
R@P	38	TYR	Y	V-KAPPA	1ce1_L	R@P	5	PRO	P		1ce1_P	21	0	0
R@P	56	ASN	N	V-KAPPA	1ce1_L	R@P	3	SER	S		1ce1_P	3	2	0
R@P	107	HIS	H	V-KAPPA	1ce1_L	R@P	4	SER	S		1ce1_P	20	4	1
R@P	107	HIS	H	V-KAPPA	1ce1_L	R@P	5	PRO	P		1ce1_P	12	2	0
R@P	107	HIS	H	V-KAPPA	1ce1_L	R@P	6	SER	S		1ce1_P	14	3	1
R@P	108	ILE	I	V-KAPPA	1ce1_L	R@P	5	PRO	P		1ce1_P	12	1	0
R@P	108	ILE	I	V-KAPPA	1ce1_L	R@P	6	SER	S		1ce1_P	12	3	0
R@P	109	SER	S	V-KAPPA	1ce1_L	R@P	6	SER	S		1ce1_P	11	2	0
R@P	114	ARG	R	V-KAPPA	1ce1_L	R@P	6	SER	S		1ce1_P	18	3	1
R@P	114	ARG	R	V-KAPPA	1ce1_L	R@P	7	ALA	A		1ce1_P	4	2	0
R@P	114	ARG	R	V-KAPPA	1ce1_L	R@P	8	ASP	D		1ce1_P	6	2	0
R@P	116	ARG	R	V-KAPPA	1ce1_L	R@P	2	THR	T		1ce1_P	1	1	0
R@P	116	ARG	R	V-KAPPA	1ce1_L	R@P	4	SER	S		1ce1_P	9	4	1
R@P	116	ARG	R	V-KAPPA	1ce1_L	R@P	6	SER	S		1ce1_P	20	6	1
R@P	116	ARG	R	V-KAPPA	1ce1_L	R@P	7	ALA	A		1ce1_P	7	2	0

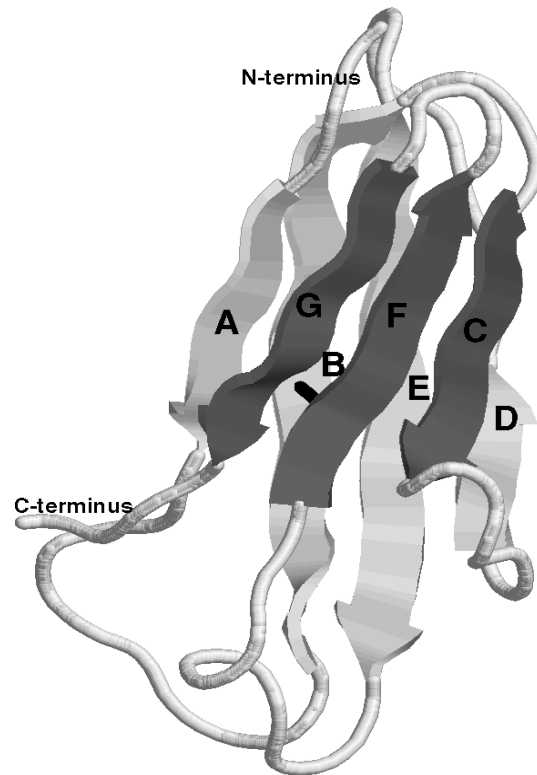
IMGT unique numbering

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

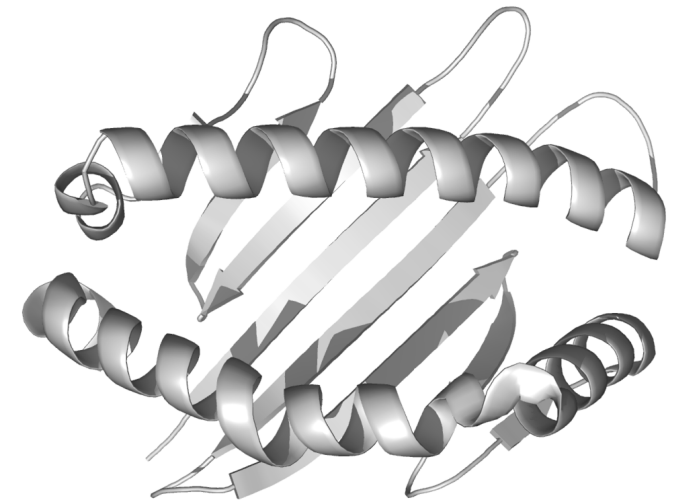


Immunoglobulin superfamily (IgSF)

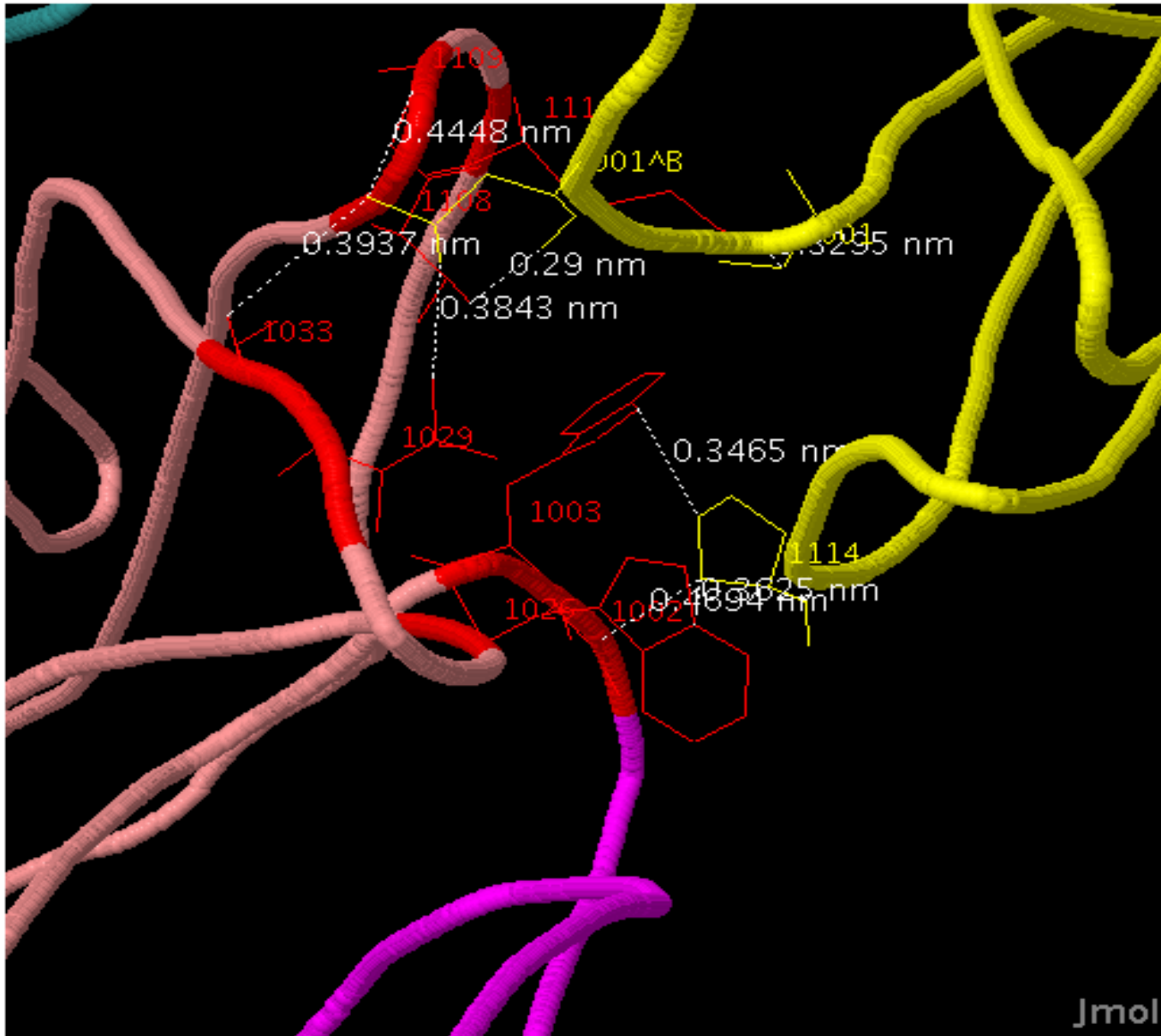
C-DOMAIN (IG,TR)
AND
C-LIKE-DOMAIN
(other than IG,TR)



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



MHC superfamily (MhcSF)



IGHG1
(FC-GAMMA1)

CH2

(1e4k)

[D2]

FCGR3B

[D1]

CONCLUSIONS and PERSPECTIVES

Three IMGT® biological approaches: genetic, genomic, structural.
Knowledge is based on the IMGT-ONTOLOGY axioms and concepts.

On the informatics side: building of IMGT-Choreography on interactions between IMGT® components to answer biological questions.

The same axioms and concepts are valid for a multiscale (molecule, cell, organ, organism) and systemic approach (system immunobiology).



Many thanks to the IMGT® team at Montpellier, France