

IMGT-ONTOLOGY and IMGT databases, tools and Web resources for immunoinformatics

Marie-Paule Lefranc
Université Montpellier, CNRS

First international Immunoinformatics Symposium
Yokohama, Japan, 26-27 February 2004
RIKEN Yokohama Institute

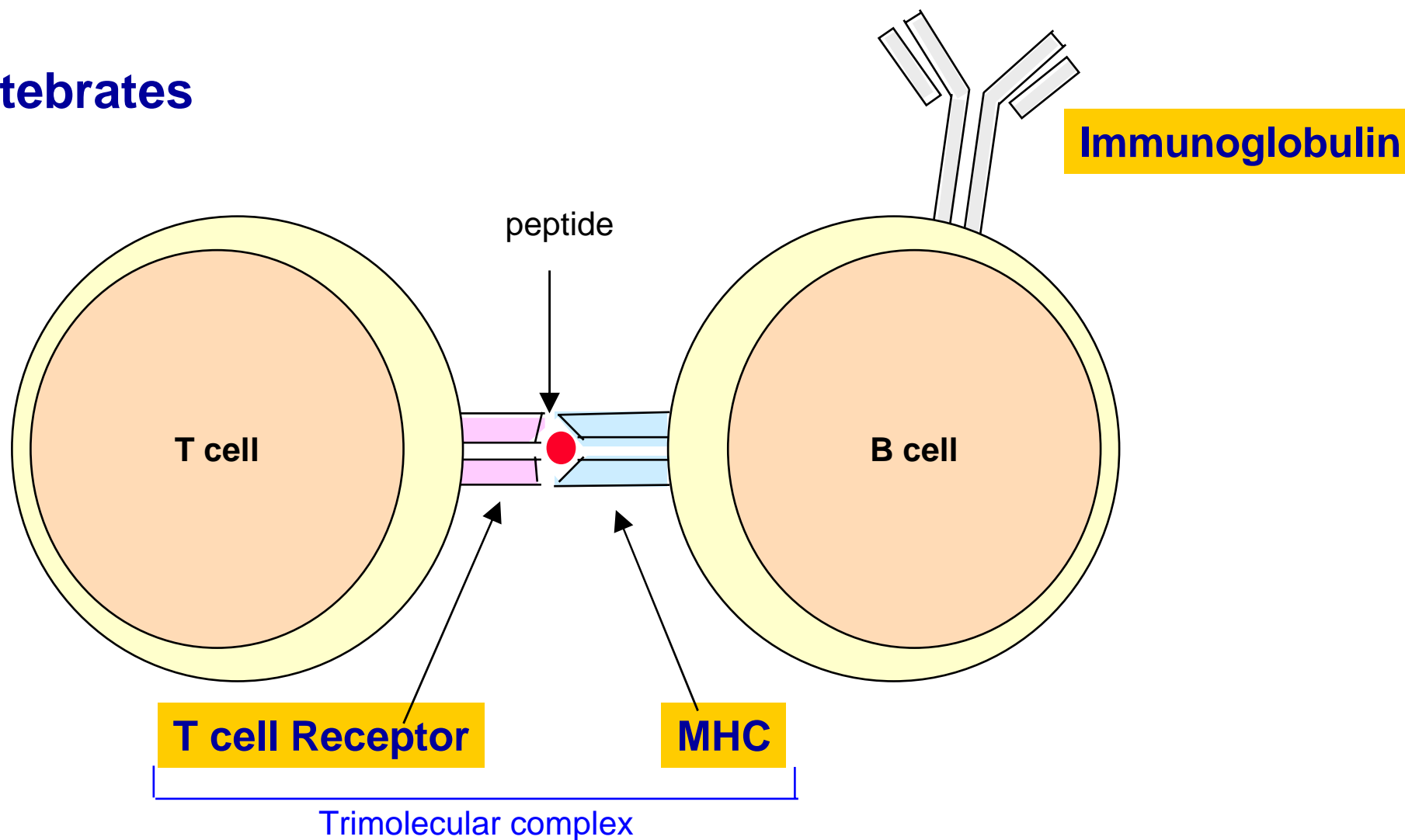


The international ImMunoGeneTics information system®
Coordinator: M.-P. Lefranc, Montpellier, France <http://imgt.cines.fr>



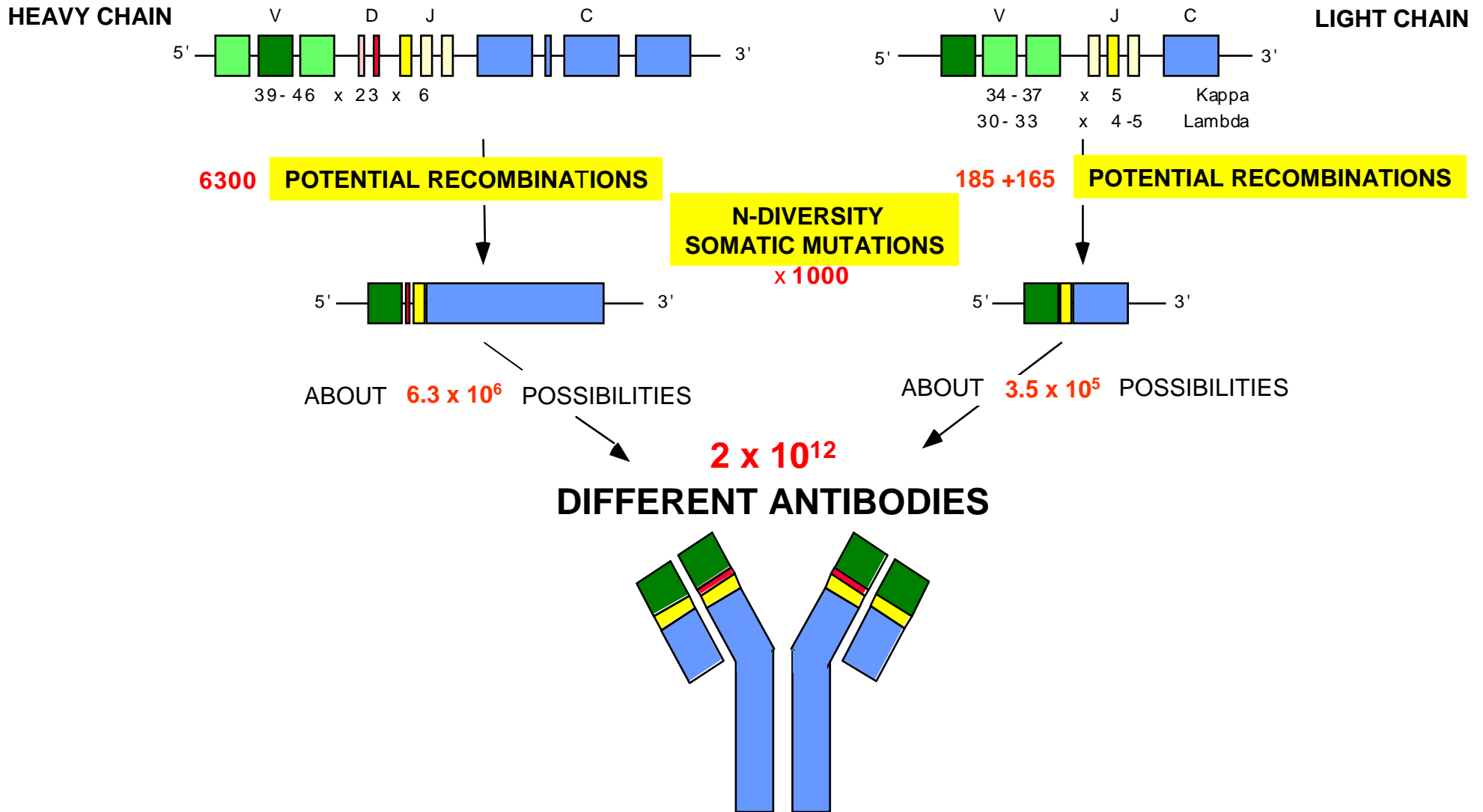
IMGT domain of research: the adaptive immune system

Vertebrates



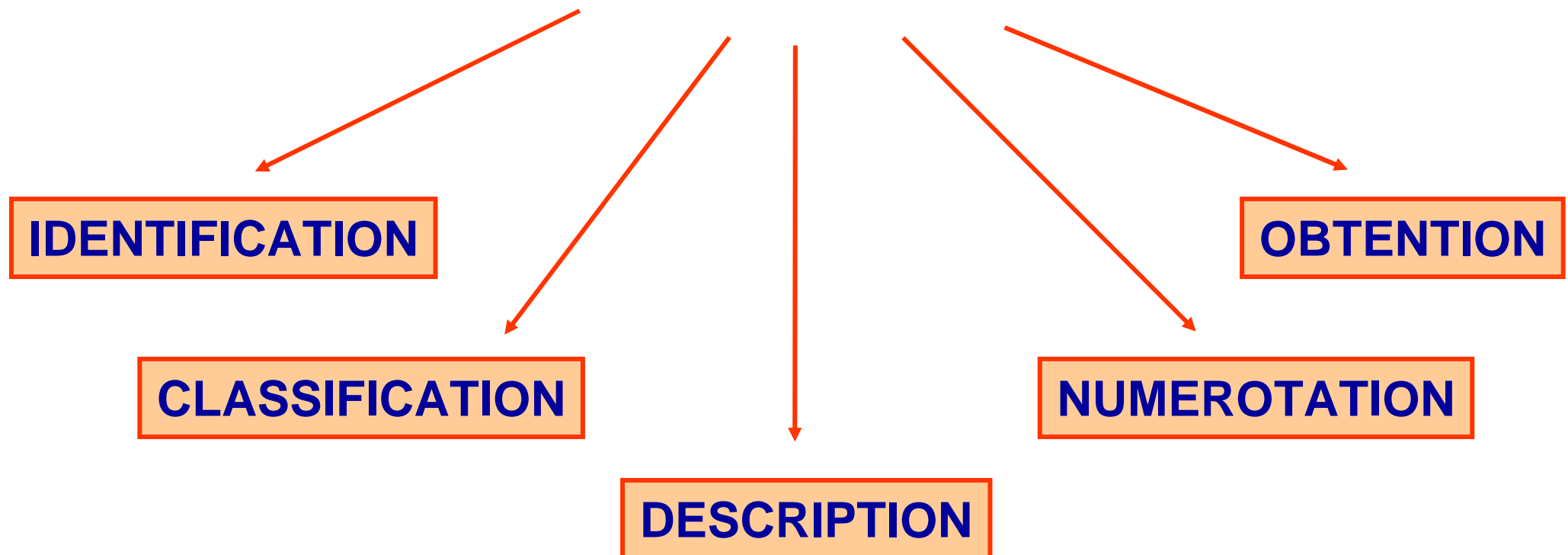
Immunoglobulin (IG) and T cell receptor (TR) synthesis

150 FUNCTIONAL IG GENES

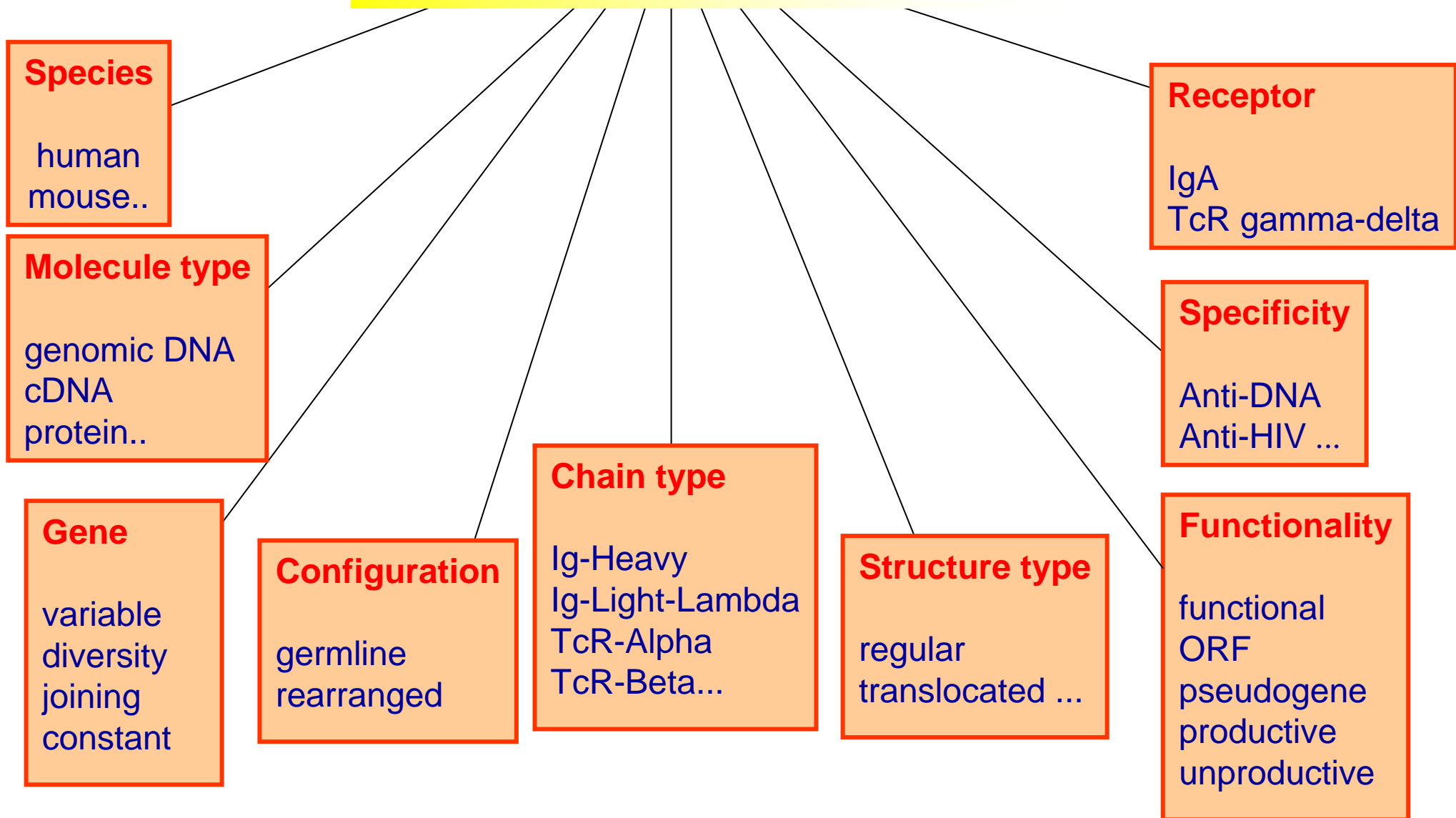


IMGT-ONTOLOGY five main concepts

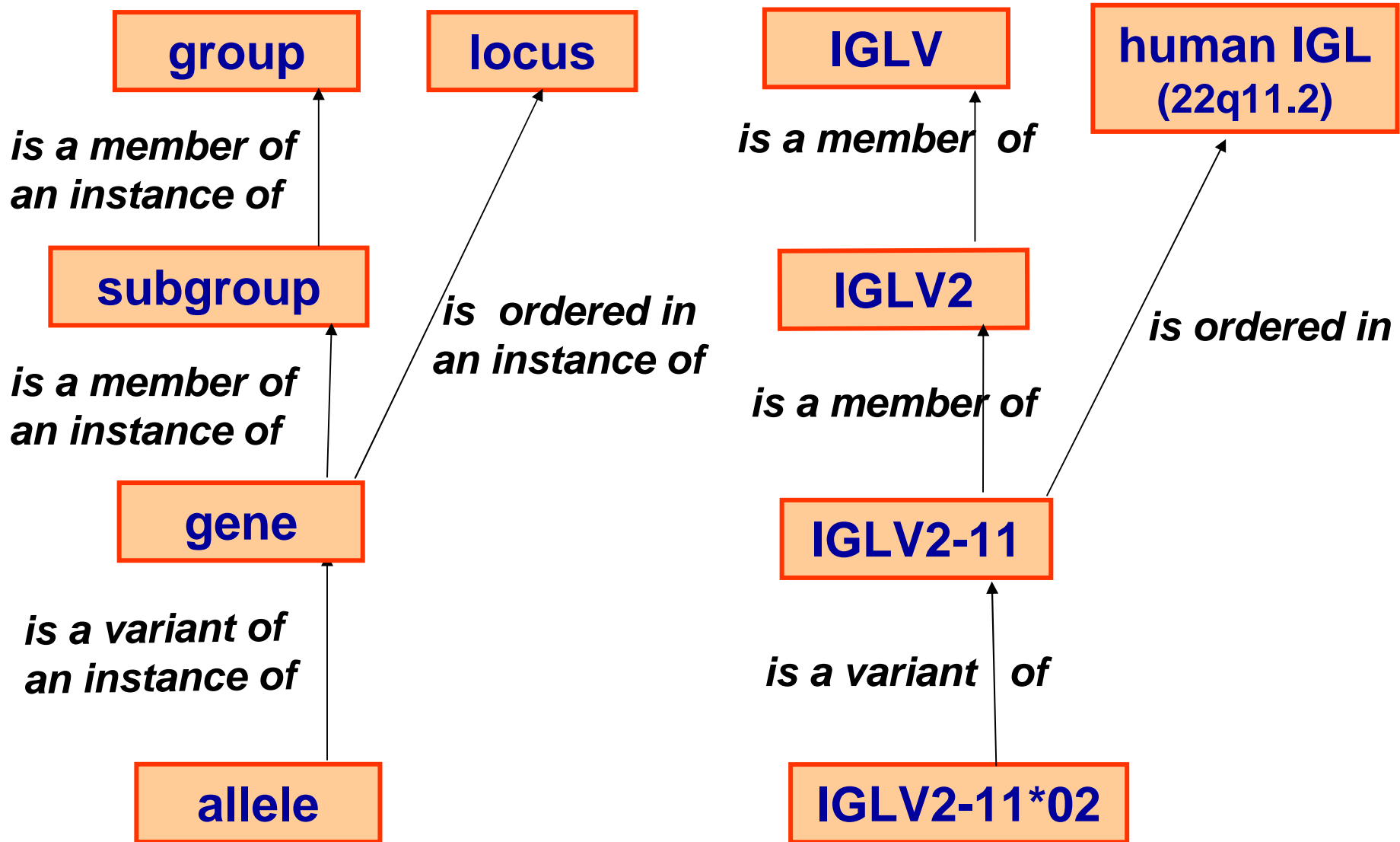
to share, reuse and represent knowledge
in immunogenetics



"IDENTIFICATION" concept



"CLASSIFICATION" concept

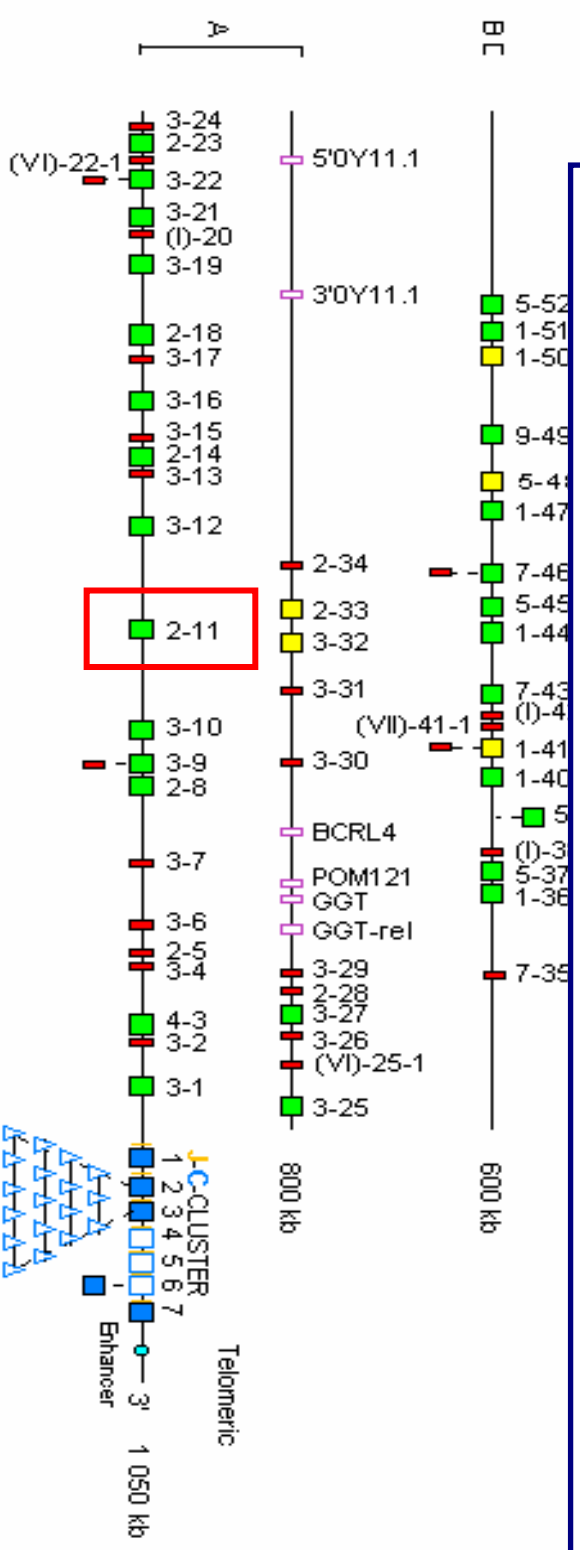
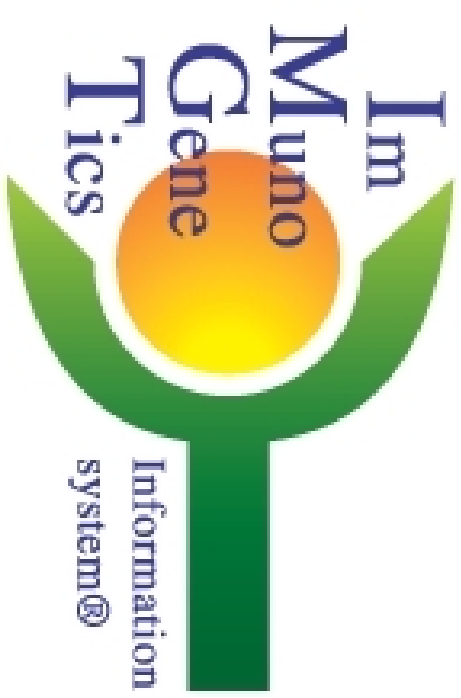


Locus representation: Human IGL

Human IGL 2

WELCOME!
to IMGT/GENE-DB

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>



● **IG: 339–354 genes in the major loci**

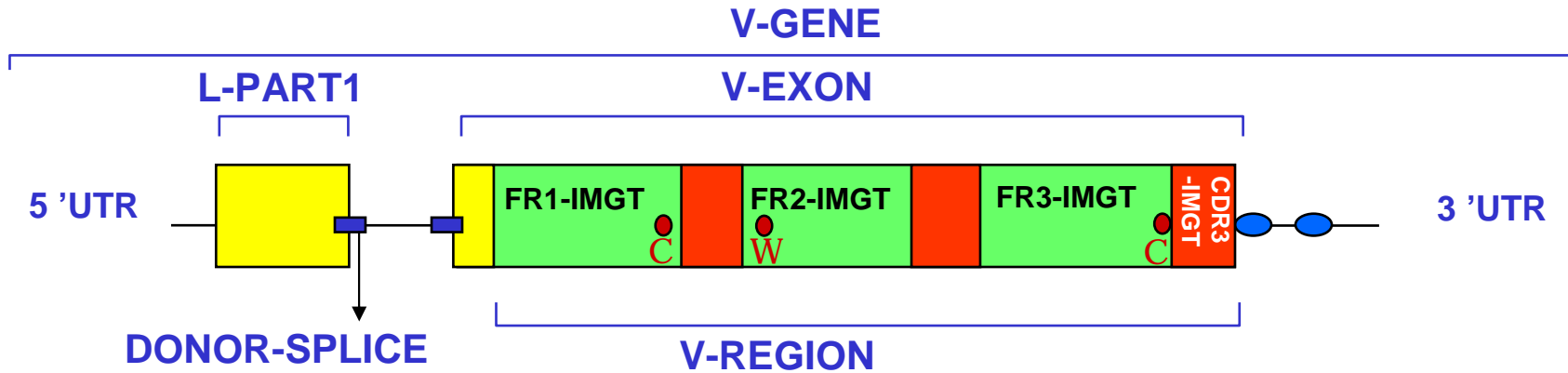
Locus and chromosomal localization	V	D	J	C	Total number of genes in the major locus[1]
IGH at 14q32.33	123–129	27	9	11	170–176
IGK at 2p11.2	(40 or) 76	0	5	1	(46 or) 82
IGL at 22q11.2	73–74	0	7–11	7–11	87–96
Total number of genes for the 3 IG loci	236–279	27	21–25	19–23	303–354

^a included one processed gene

● **TR: 228–234 genes in the major loci**

Locus and chromosomal localization	V	D	J	C	Total number of genes in the major locus[2]
TRA at 14q11.2	54 ^a	0	61	1	116 ^a
TRB at a 7q34	64–67	2	14	2	82–85
TRG at 7p14	12–15	0	5	2	19–22
TRD at 14q11.2	3 (8) ^a	3	4	1	11 (16) ^a
Total number of genes for the 4 TR loci	133–139	5	84	6	228–234

"DESCRIPTION" concept



Label 1

Label 2

V-GENE

V-EXON

FR3-IMGT

CDR3-IMGT

L-PART1

DONOR-SPLICE

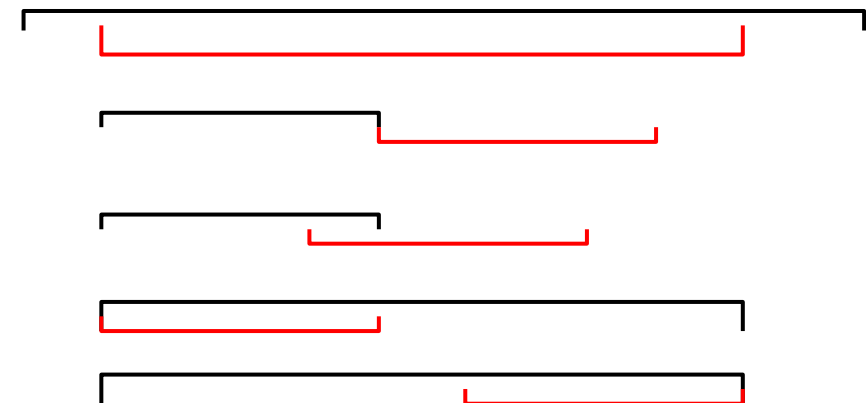
V-REGION

FR1-IMGT

V-REGION

CDR3-IMGT

Relations entre Labels



IMGT/LIGM-DB Consultation module v3 - Netscape
 Fichier Edition Afficher Aller Communicator Aide
 Signets Adresse : <http://ligm.igh.cnrs.fr:8104/cgi-bin/IMGTlect.jv> Infos connexes

FT	V-GENE	<1..297>
FT		(partial)
FT		
FT		
FT		
FT		
FT		
FT		
FT	V-REGION	
FT		
FT		
FT		
FT		
FT	FR1-IMGT	
FT		
FT		
FT	1st-CYS	
FT	CDR1-IMGT	
FT		
FT		
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="KRPSGVPPDRFSGSKSGNTASLTISGLQAEDEADYYC"
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 atatctccta 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](#)



<http://imgt.cines.fr>



Document : chargé

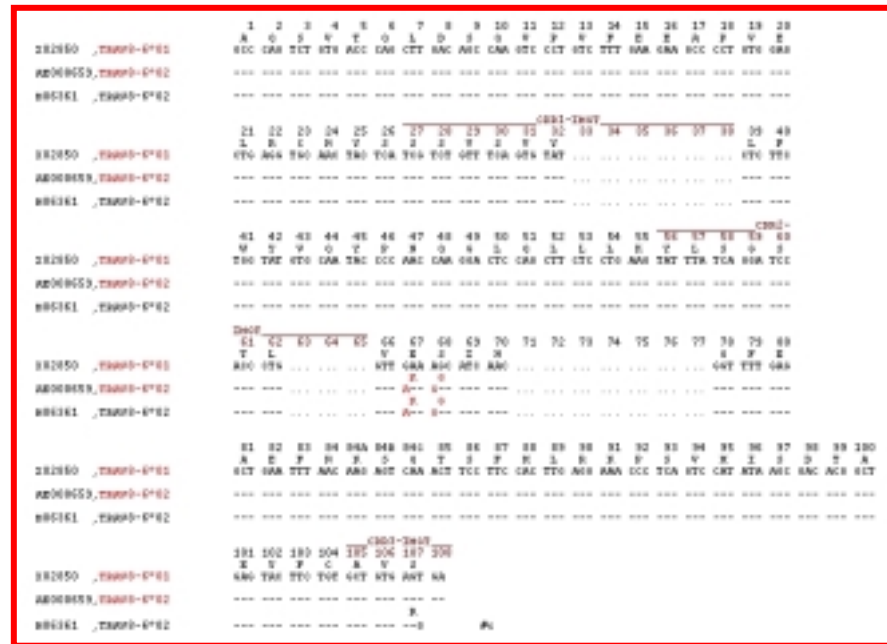
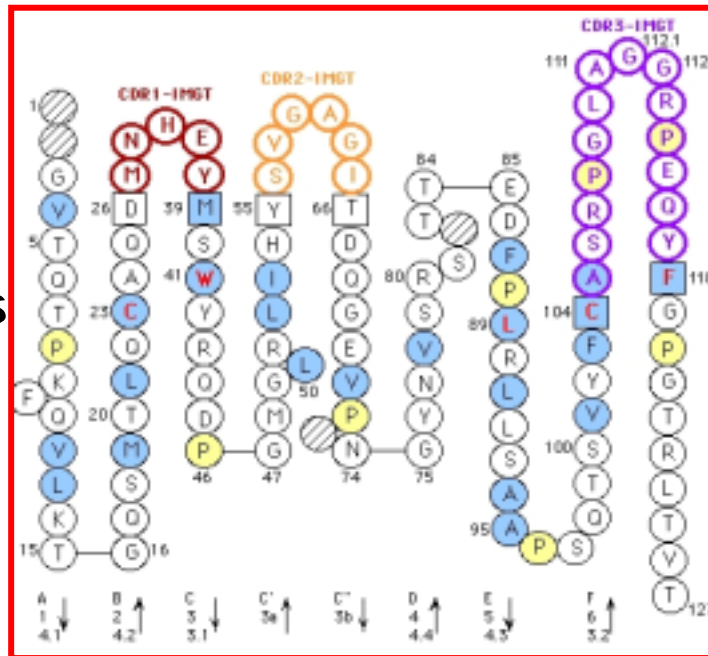


<http://imgt.cines.fr>



"NUMEROTATION" concept

Collier de Perles



Alignment of alleles

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	PSEVTAVEGAI	VQINCTYQ	TSGFYG	LSWYQQHDGGAPTFLSY	NALDGL	LEETG	RFSSFLSRSDSYGYLLQLQMKDSASYFC	AVR				
AE000658, TRAV1-2	GQNIQ	PTEMTATEGAI	VQINCTYQ	TSGFNG	LFWYQQHAGEAPTFLSY	NVLDG	LEEKQ	RFSSFLSRSKGYSYLLKELQMKDSASYLC	AVR				
AE000658, TRAV2	KDQVFQ	PSTVASSEGAVVE	IFCNHS	VSNAYN	FFWYLFHPGCA	PRLLVK	GSK	PSQQG	RYNMTYER	FSSLLILQVREADAAVYYC	AVE		
AE000658, TRAV3	AQSVAQPEDQV	NVAEGNPLT	VKCTYS	VSGNPY	LFWYVQYPN	RGLQLLK	YITGDNL	VKGSY	GFEAEFNK	QTSFHLKPSALVSDSALYFC	AVRD		
AE000658, TRAV4	LAKITQ	PISMSDYE	QGEVNI	TCSHN	NIATNDY	ITWYQQF	PSQGRFIIQ	GYKT	KVINE	VASLFI	PADRSSTLSLPRVLSDTAVYYC	LVGD	
AE000659, TRAV5	GEDVEQS	LFLSVREGD	SSVINGTYT	DSSSTY	LYWYKQEP	GAGLQLLTY	IFSNMD	MKQDQ	RLTVLL	NKDKHL	SLRIADTQTGDSAIYFC	AES	
AE000659, TRAV6	SQKIEQNSEAL	NIQEGK	TATLTCNYT	NYSPAY	LQWYRQDP	GRGPVFLLL	IRENEK	EKRKE	RLKVT	FDITL	KQSLFHITASQPADSATYLC	ALD	
AE000659, TRAV7	ENQVEHSPH	FLGPQQGD	VASMSCTYS	VSRFNN	LQWYRQNT	GMGPKHLLS	MYSAGY	EKQKG	RLNAT	LLK	NGSSLYITAVQPEDSATYFC	AVD	
AE000659, TRAV8-1	AQSVSQNH	HVILSEAA	SELEGCNYS	YGQTVN	LFWYVQYP	GGQLQLLK	YFSGDPL	VKGIK	GFEAEF	IKSKF	SPNLRKPSVQWSDAEYFC	AVN	
AE000659, TRAV8-2	AQSVTQLD	SHSVSE	QTPVLLRCNYS	SSYSPS	LFWYVQHP	NKGLQLLLK	YTSAATL	VKGIN	GFEAEF	KKSET	SFHLTKPSAHMSDAAEYFC	VVS	
AE000659, TRAV8-3	AQSVTQPD	IHIITV	SEGASLELRCNYS	YGATPY	LFWYVQSP	GGQLQLLK	YFSGDTL	VQGIK	GFEAEF	KRSQSS	PNLRKPSVHWSDAEYFC	AVG	
AE000659, TRAV8-4	AQSVTQLG	SHSVSE	GALVLLRCNYS	SSVPPY	LFWYVQYP	NQGLQLLLK	YTSAATL	VKGIN	GFEAEF	KKSET	SFHLTKPSAHMSDAAEYFC	AVS	
XO2850, TRAV8-6	AQSVTQLD	SQVPV	EEAPVELRCNYS	SSVSVY	LFWYVQYP	NQGLQLLLK	YLSGSTL	VESIN	GFEAEF	NK	QTSFHLRKP	SVHISDTAEYFC	AVG
AE000660, TRAV8-7	TQSVTQLD	GHIITV	SEEAPELRCNYS	YSGVPS	LFWYVQY	SSQLQLLK	DLTEATQ	VKGIR	GFEAEF	KKSET	SFYLKPSVHWSDAEYFC	AVGDR	
AE000659, TRAV9-1	QDSVWQTE	GGVLR	PSERD	SLIVNCSYF	TTQVPS	LFWYVQV	PGEGPOLHLK	AMKAMD	KGRMK	GFEAMV	RKETTSEHLEKDSVQESDSAVYFC	ALS	

V-DOMAIN 3D representation (TR A6, 1ao7)

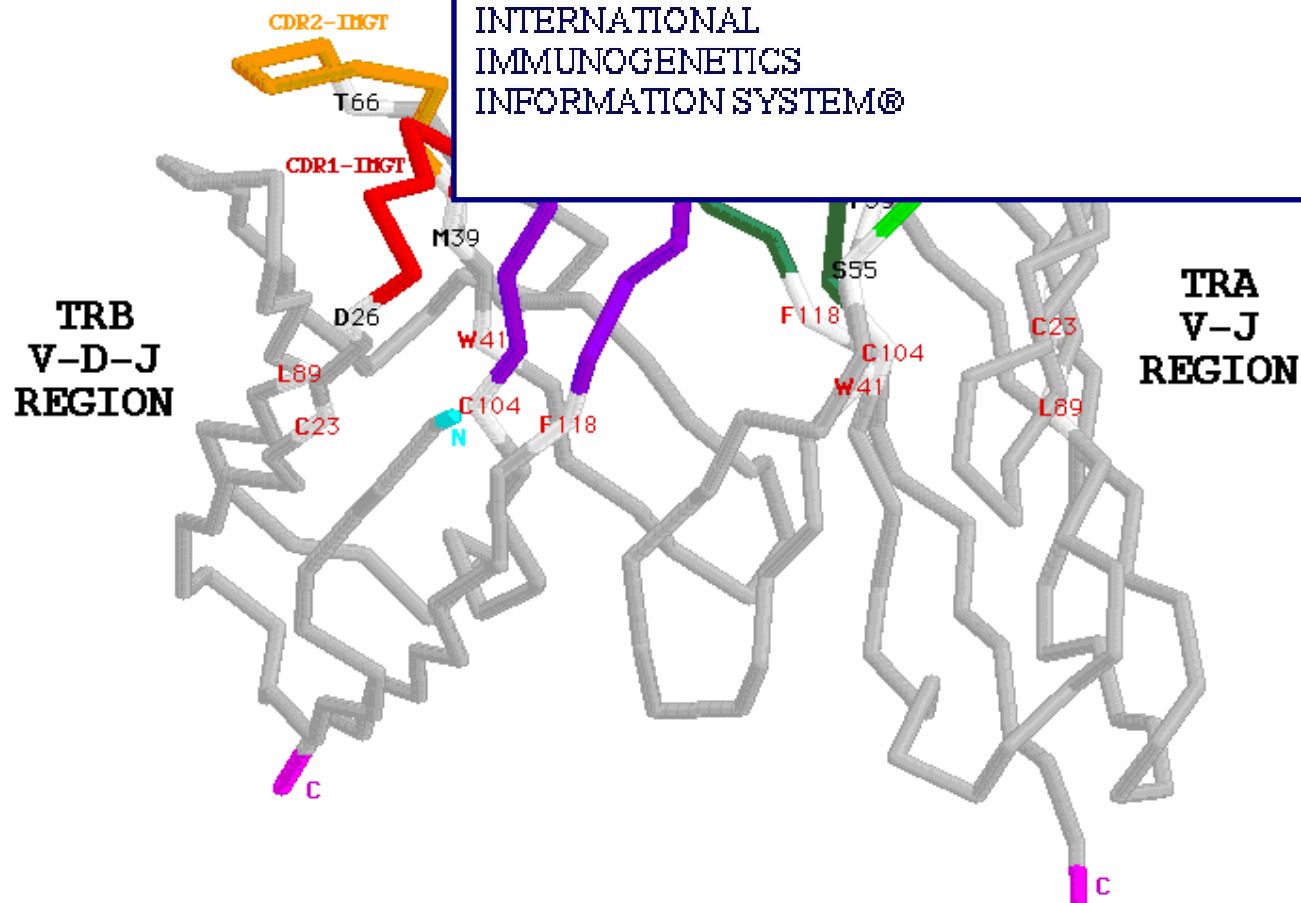
WELCOME !

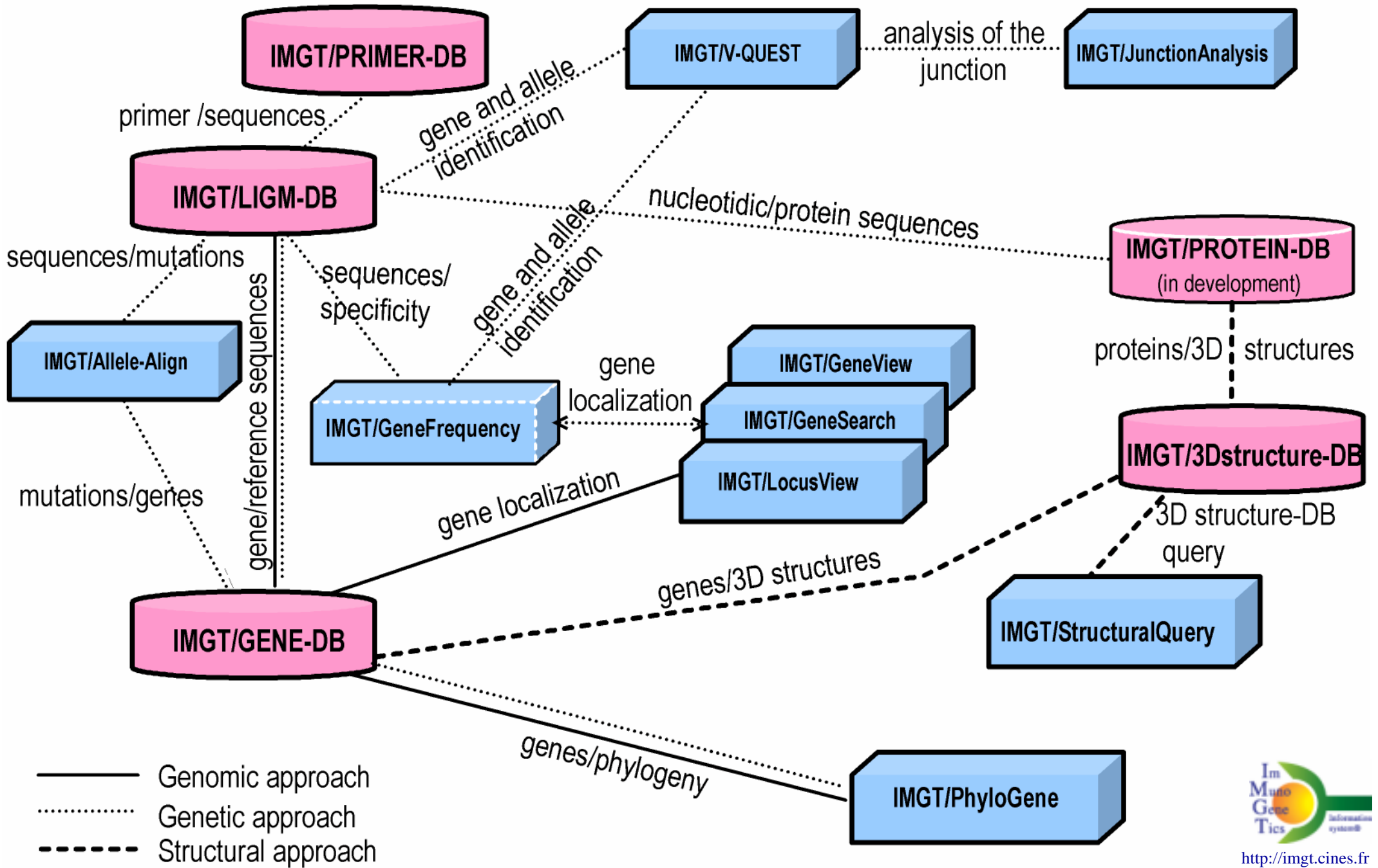
to IMGT/3Dstructure-DB

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®

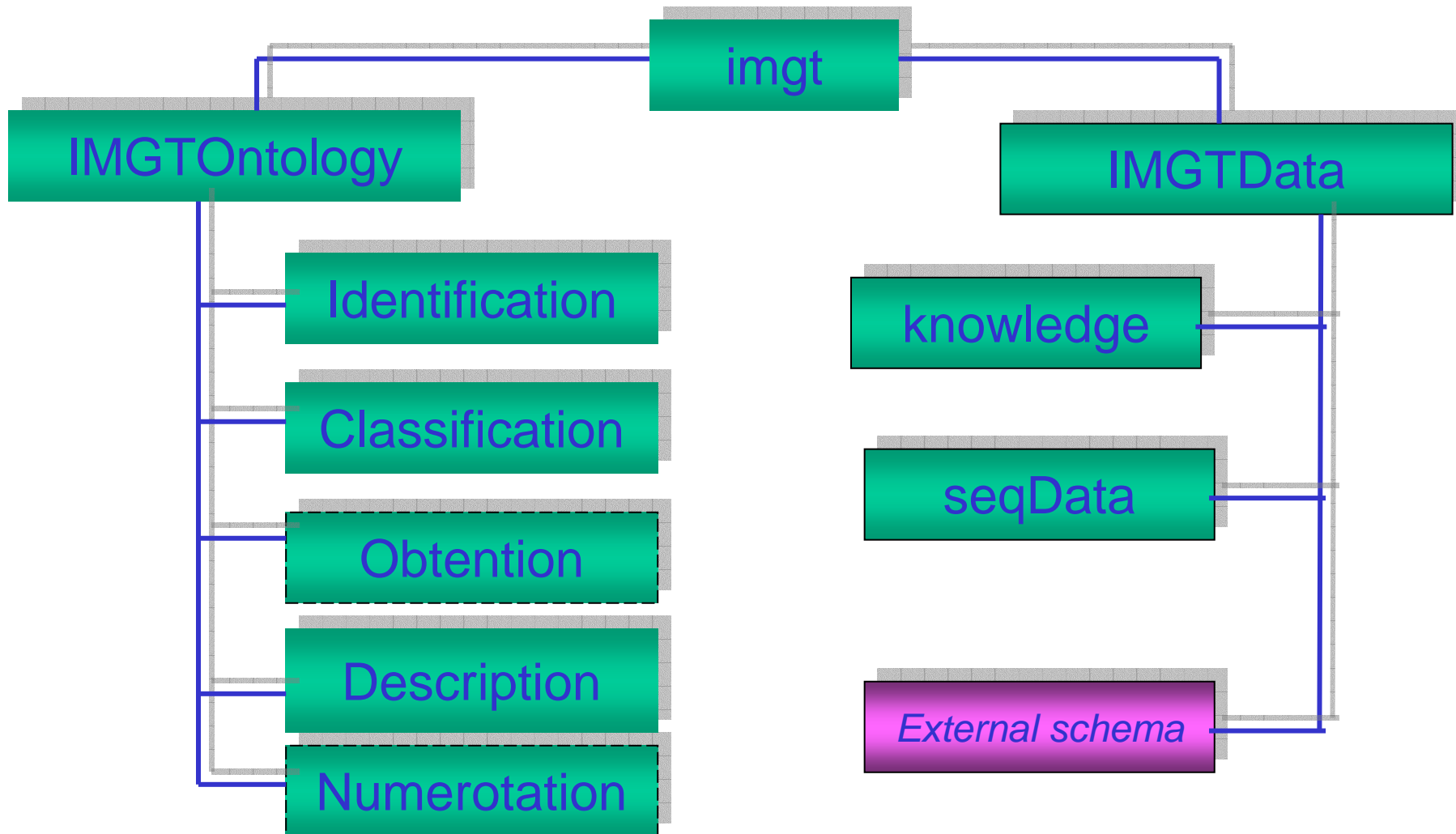


<http://imgt.cines.fr>

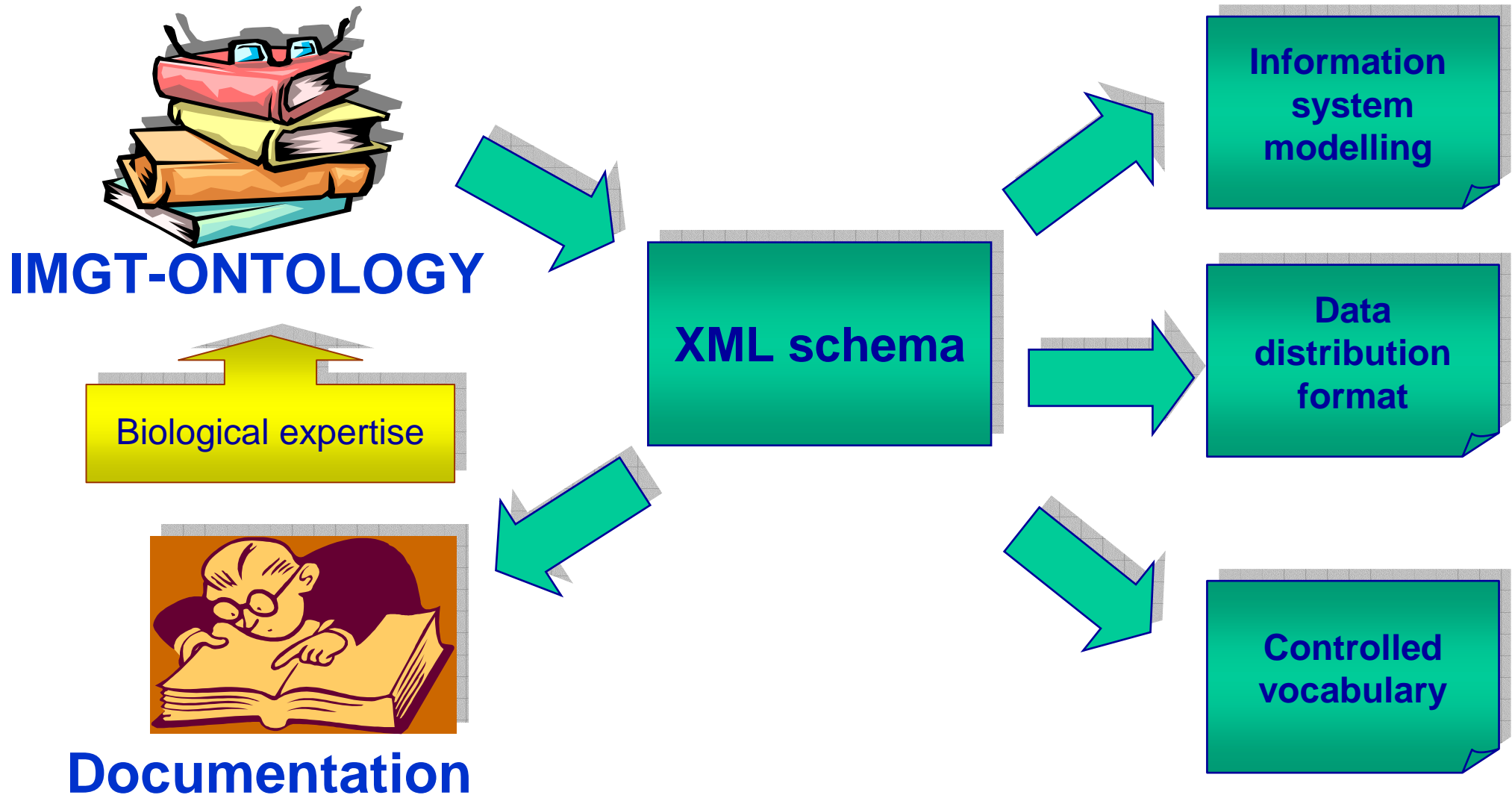




IMGT-ML schema



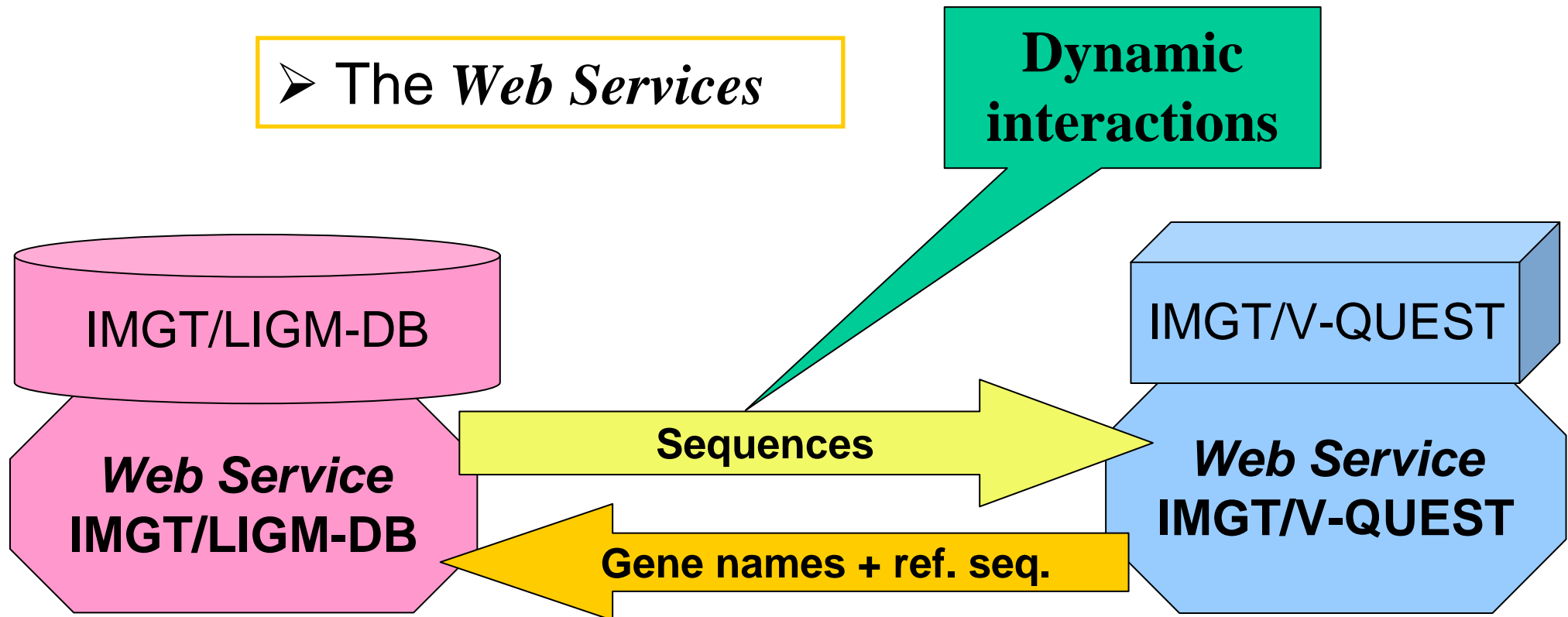
IMG-T-ML architecture



Informatic answers to the biological problems

- ❑ Use IMGT-ONTOLOGY (and IMGT-ML)
- ❑ Allow IMGT components to dynamically interact

➤ The *Web Services*



Example of IMGTV-QUEST results

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	181	_____T.....T..T.....A.T.....T...G
X86355	IGHJ6*02	179	T.A.TACTACTACT...G.A.....
X86355	IGHJ3*02	172	_____T.....T..TA.....A.T.....T...G

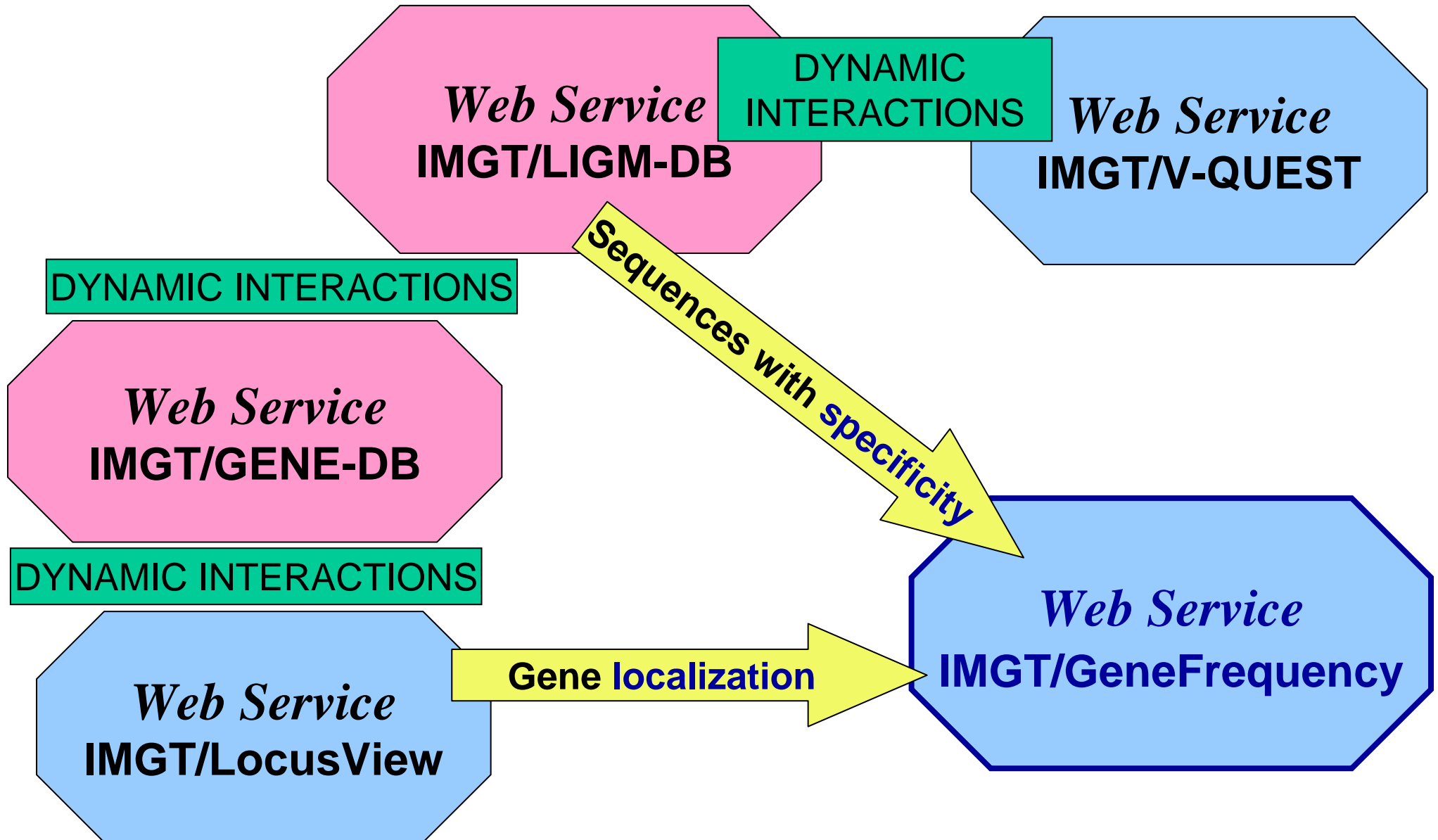


V-GENE

JUNCTION

J-GENE

Diagram of collaboration: Analyse de repertoires



Example of IMGT/GeneFrequency results

Your Selection :

Human IGH, IGK and IGL Locus Specificity anti-thyroid peroxidase (TPO)

For the D and J genes, the number of genes is shown between parentheses when genes names could not be indicated for a click on the zoom for the D and J genes names.

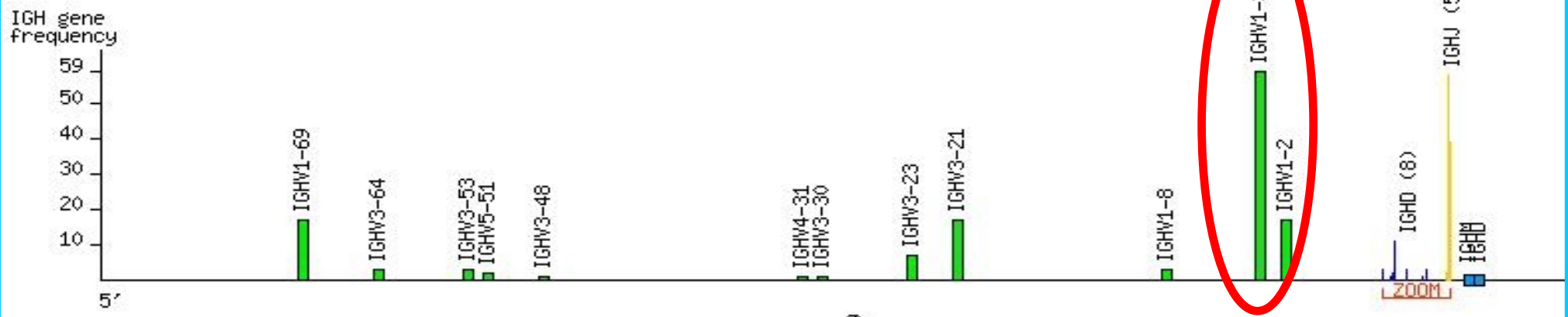
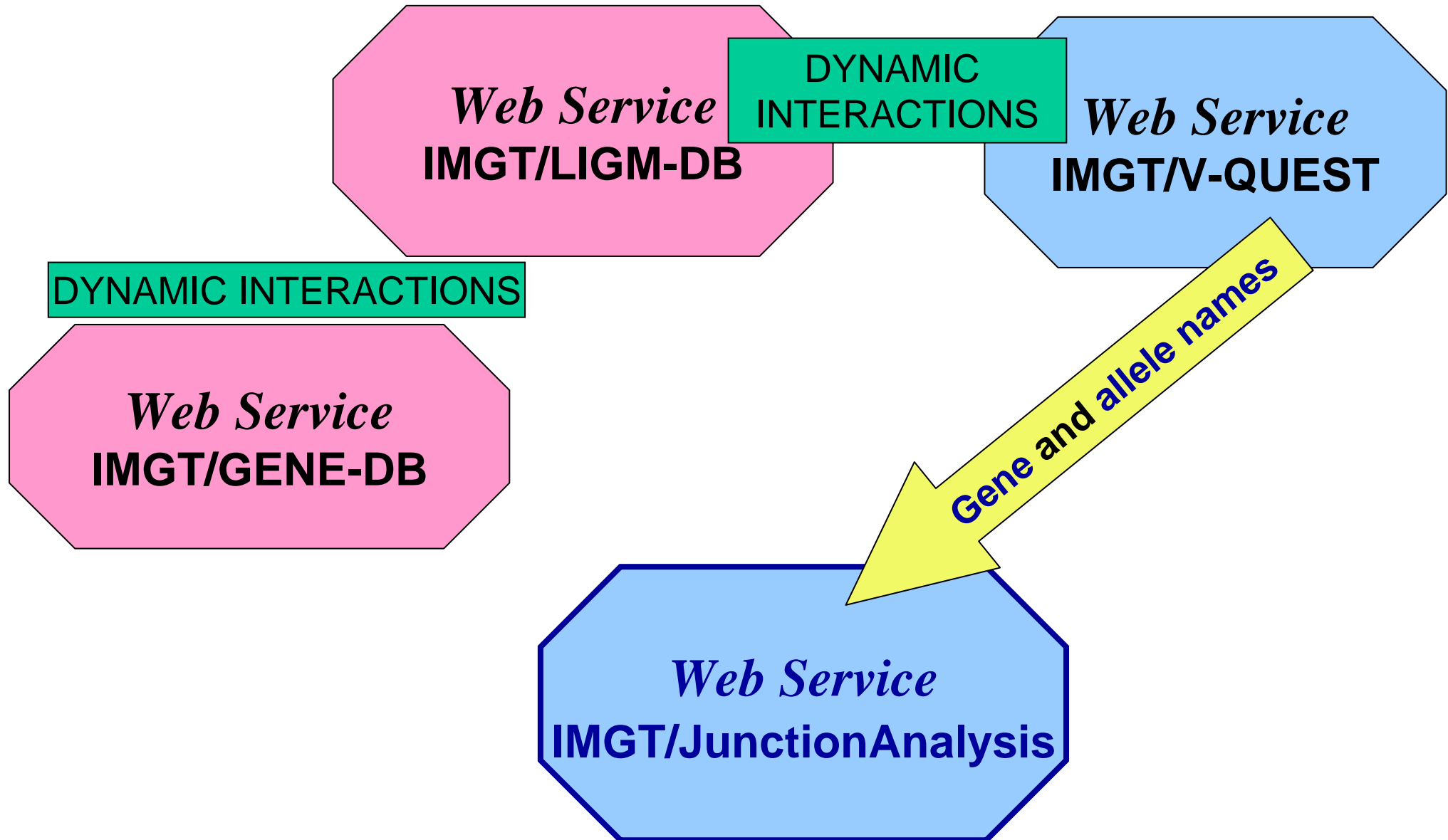


Diagram of collaboration: Analyse des jonctions



Example of IMGT/JunctionAnalysis results



<http://imgt.cines.fr>

Analysis of the JUNCTIONS

Input	V name	V-REGION	D-REGION	N2	J-REGION
#1 AF402940	IGHV1-3*01	tgtgcgagag.gcttcacgggg.....	cgggacgctttggacgtctgg

Input	J name	D name	Vmut	Dmut	Jmut	Ngc
#1 AF402940	IGHJ3*01	IGHD3-10*01	0	4	2	5/6

Translation of the JUNCTIONS

	105	107	109	112	114	116	118	CDR3-IMGT								
	104	106	108	110	113	115	117	frame length								
	C	A	R	G	F	T	G	R	D	A	L	D	V	W		
#1 AF402940	tgt	gcg	aga	ggc	ttc	acg	ggg	cgg	gac	gct	ttg	gac	gtc	tgg	+	12

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

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>

Analysis of the JUNCTIONS

Input	V name	V-REGION	M1	D-REGION	M2		
#1 M62724	IGHV7-4-1+02	tatacaagaga	aga	.taaccaatgactacaa.....	aata		
#2 Z47269	IGHV1-69*06	tgtagagag.	gggggggctaaggtogaattttggagtggt.....	tcatgggt		
Input J name J-REGION J name D name Vmut Dmut Jmut Ngc							
#1 M62724tttaactactag	IGHJ4*02	IGHD5-24*01	0	2	0	1/7
#2 Z47269	...actggttcgacccctgg	IGHJ5*02	IGHD3-3*02	0	2	0	13/20

Translation of the JUNCTIONS

	104	105	106	107	108	109	110	111	111.1	111.2	111.3	112.4	112.2	112.1	112	113	114	115	116	117	118	+	CDR3-IMGT length
#1 M62724	C	A	R	E	D	S	N	G							Y	K	I	F	D	Y	W	+	13
	gtg	gcg	aga	gaa	gat	agc	aat	ggc							tac	aaa	ata	ttt	gac	tac	tggtgg		
	C	A	R	G	G	A	K	V							G	Y	W	F	D	P	W		
#2 Z47269				gag	gag	gag	ggt	gac	gaa	ttt	ttg	gag	tggttt	ttt	cat	gggtac	tggttc	gac	ccc	tggtgg		+	20

IMGT-Choreography: Expressed IG and TR repertoires

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

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>

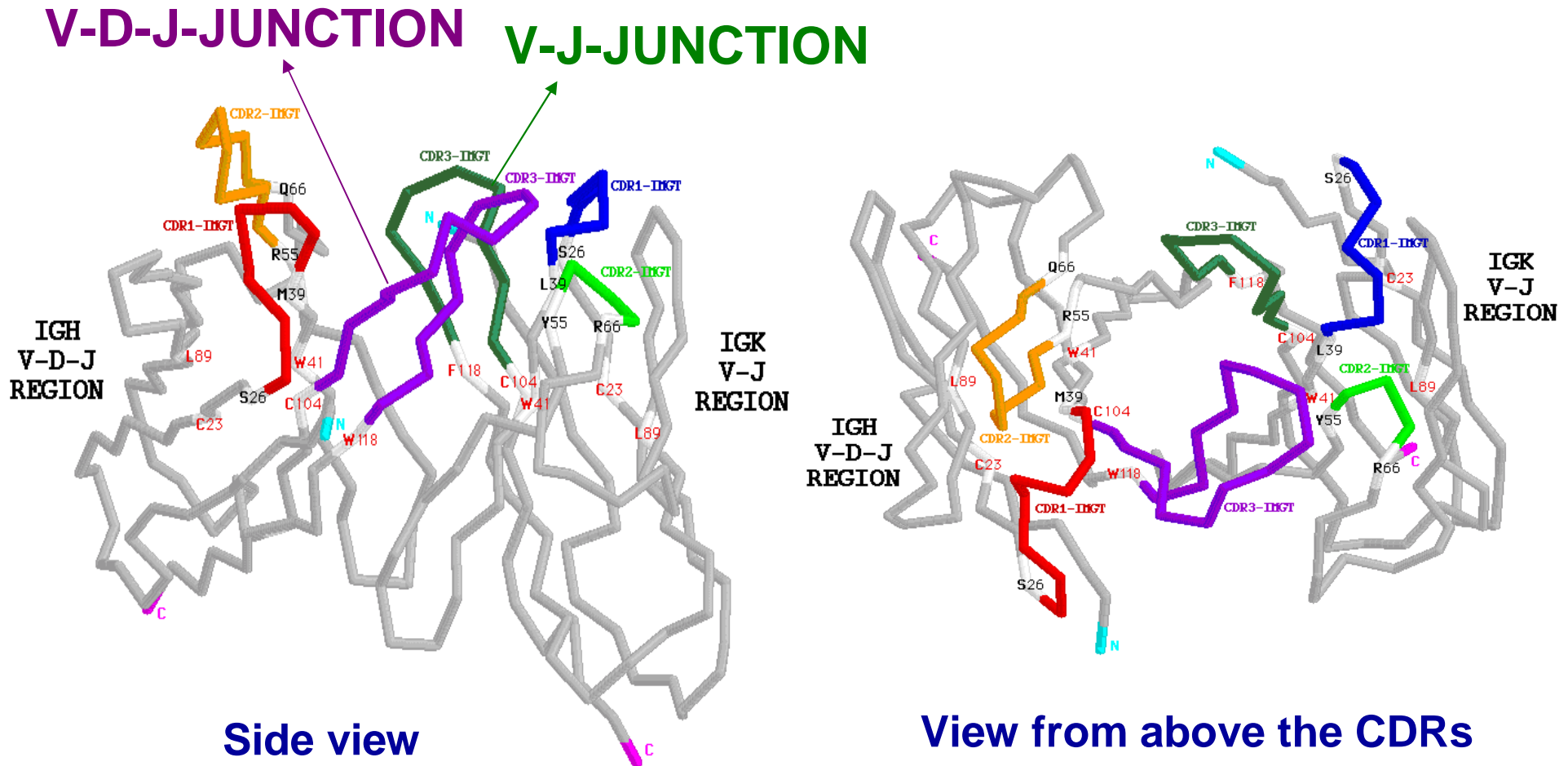
Analysis of the JUNCTIONS

	Input	V name	V-REGION	N	J-REGION	J name	Vmut	Jmut	Ngc
#1	AF490920	IGKV1-33*01	tgtcaacactatgatgatttccc...		attcactttc	IGKJ3*01	3	0	0/0
#2	AF490935	IGKV4-1*01	tgtcagcaatattatagtactctc.		..tcactttc	IGKJ4*01	0	0	0/0
#3	AF490937	IGKV4-1*01	tgtcagcaatattatagtggctctcc		.gtacactttt	IGKJ2*01	2	0	0/0
#4	AF490932	IGKV3-15*01	tgtcagcactataataactggcctcc	cc	tgtacactttt	IGKJ2*01	1	0	2/2

Translation of the JUNCTIONS

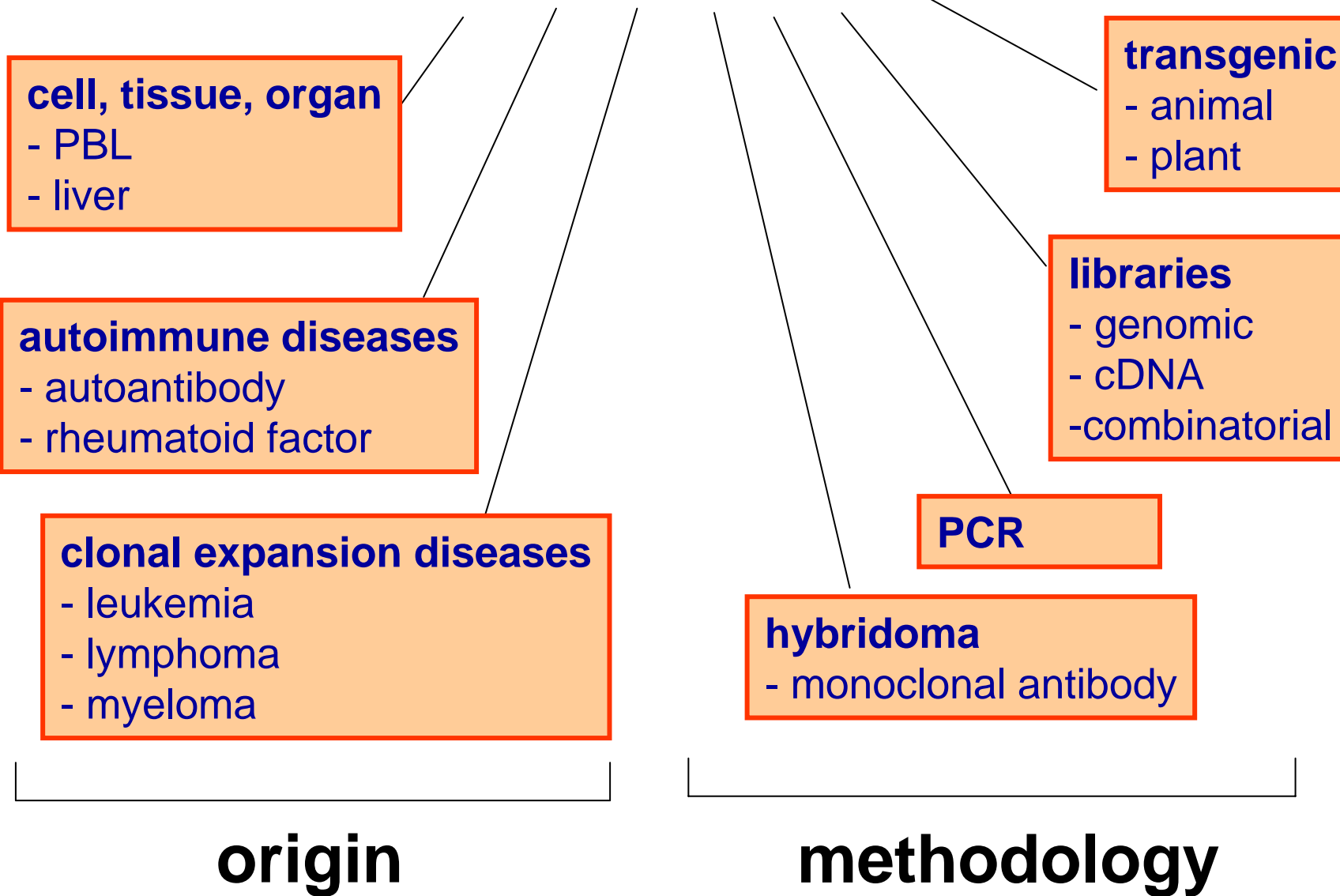
		104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	frame	CDR3-IMGT length
#1	AF490920	tgt	caa	cac	tat	gat	gat					ttc	cca	ttc	act	ttc	+	9
		C	Q	Q	Y	Y	S					T	P	L	T	F		
#2	AF490935	tgt	cag	caa	tat	tat	agt					act	cct	ctc	act	ttc	+	9
		C	Q	Q	Y	Y	S				G	P	P	Y	T	F		
#3	AF490937	tgt	cag	caa	tat	tat	agt				ggt	cct	ccg	tac	act	ttt	+	10
		C	Q	H	Y	N	N	W			P	P	L	Y	T	F		
#4	AF490932	tgt	cag	cac	tat	aat	aac	tgg	cct	ccc	ctg	tac	act	ttt			+	11

IMGT-Choreography: 3D structures/specificities



V-DOMAINS (*Mus musculus* E5.2 Fv)

"OBTENTION" concept



Immunoinformatics



<http://imgt.cines.fr>

Data integration specific to Immunology

- *interactions host-pathogens
- *vaccinology
- *immunomodulation...

Gene

Transcript

Microarrays

Protein

3D

Organelle

Gene regulation
Pathways
Networks

Cell

Bioinformatics
databases, and tools

Population

Organism

Organ

Tissue

Mathematical and
computational models

Collection of
clinical data

Who is using IMGT?

Medical research:

repertoire in autoimmune diseases, AIDS, leukemias, lymphomas, myelomas, translocations, detection of residual diseases

Therapeutic approaches:

immunotherapy, grafts, immunomodulation, immunosuppression

Veterinary research:

IG and TR repertoire of domestic and farm species

Biotechnology related to antibody engineering:

chimeric, humanized, human antibodies, scFv, combinatorial libraries, intrabodies

Genome diversity:

comparative and developmental immunology, evolution of the adaptive immune system

IMGT, the international ImMunoGeneTics information system®

<http://imgt.cines.fr>



The IMGT team at Montpellier, France



<http://imgt.cines.fr>