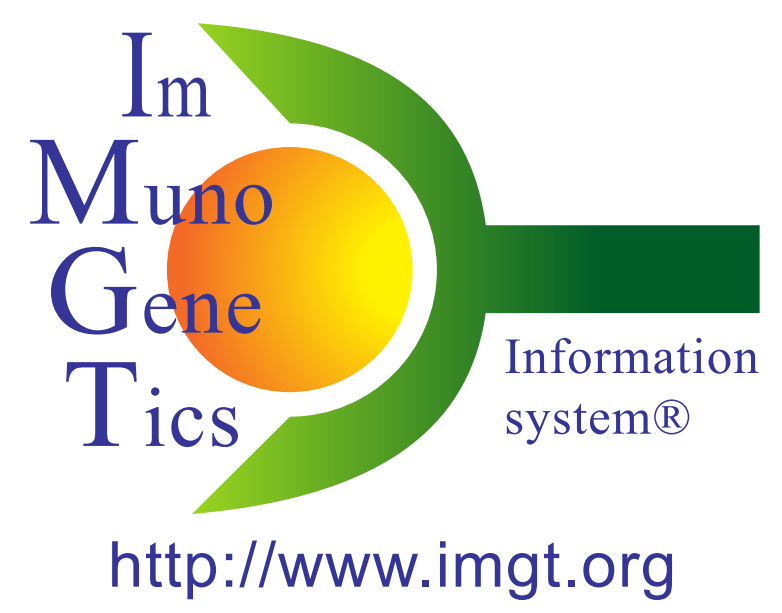


IMGT[®] genomic annotation of the dog (*Canis lupus familiaris*) seven immunoglobulin (IG) or antibody and T cell receptor (TR) loci

Imène Chentli*, Perrine Pégorier*, Saida Saljoqi*, Géraldine Folch, Joumana Jabado-Michaloud, Véronique Giudicelli, Patrice Duroux, Sofia Kossida and Marie-Paule Lefranc

*Equal contribution

IMGT[®], the international ImMunoGeneTics information system[®], Laboratoire d'ImmunoGénétique Moléculaire (LIGM), Institut de Génétique Humaine (IGH), UMR 9002 CNRS-UM, Université de Montpellier (UM), Montpellier (France)



IMGT[®], the international ImMunoGeneTics information system[®], <http://www.imgt.org> [1], is the global reference in immunogenetics and immunoinformatics [2], founded in 1989 by Marie-Paule Lefranc at Montpellier (Université de Montpellier and CNRS). IMGT[®] is a high-quality integrated knowledge resource specialized in the immunoglobulins (IG) or antibodies, T cell receptors (TR), major histocompatibility (MH) of human and other vertebrate species, and in the immunoglobulin superfamily (IgSF), MH superfamily (MhSF) and related proteins of the immune system (RPI) of vertebrates and invertebrates.

The genome of the vertebrates with jaws (*Gnathostomata*), which appeared in the evolution about 450 million years ago, includes the IG, TR and MH genes characteristic of the adaptive immune repertoires. Currently, there are 244 annotated vertebrate genomes including 112 from mammals at NCBI.

In humans and other mammals, there are seven main loci for IG and TR: three for IG (IGH, IGK and IGL) and four for TR (TRA, TRB, TRD and TRG). IMGT[®] genomic annotated data are classically displayed in IMGT Repertoire Web Resources (Locus description, Locus representation, Gene tables, Alignments of alleles). So far the number of species present in the IMGT Web Resources reaches 40, however only two species, *Homo sapiens* and *Mus musculus*, have been fully annotated for their seven antigen receptor loci. The seven IG and TR loci of the dog (*Canis lupus familiaris*) were recently described [3]. The dog represents the first species for which the seven antigen receptor (IG and TR) loci are annotated simultaneously in IMGT[®]. The biocuration was performed on the loci extracted from genome assembly. Six hundred thirty-three genes (413 IG and 220 TR genes) were identified in the seven loci spanning more than 6.15 megabases (Mb) and were fully annotated. This biocuration has been performed manually and the standardized annotation has allowed data entry in IMGT[®] databases and tools (IMGT/LIGM-DB [4], IMGT/GENE-DB [5], IMGT/V-QUEST [6], IMGT/HighV-QUEST [7] and IMGT/DomainGapAlign [8]).

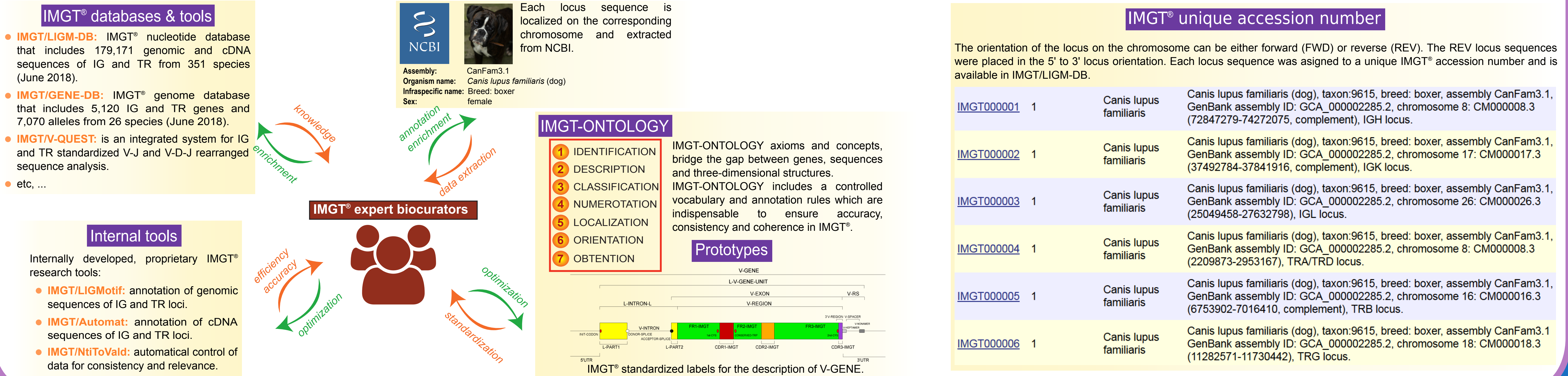
[1] Lefranc M.-P. et al. Nucl. Acids Res. 43:D413-422 (2015) PMID: 25378316
[5] Giudicelli V. et al. Nucl. Acids Res. 33:D256-261 (2005) PMID: 15608191

[2] Lefranc M.-P. Front. Immunol. 5:22 (2014) PMID: 24600447
[6] Brochet X. et al. Nucl. Acids Res. 36:W503-508 (2008) PMID: 18503082

[3] Martin J. et al. Immunogenetics. 70(4):223-336 (2018) PMID: 28924718
[7] Alamyar E. et al. Immunome Res. 8:1-2 (2012) PMID: 22647994

[4] Giudicelli V. et al. Nucl. Acids Res. 34:D781-D784 (2006) PMID: 16381979
[8] Ehrenmann F. et al. Nucl. Acids Res. 38:D301-307 (2010) PMID: 19900967

IMGT[®] Expert Biocuration Pipeline



Immunoglobulins (IG): IGH, IGK & IGL

Locus description

The **413** dog IG genes are distributed on 3 loci, among 3 chromosomes:

- IGH locus in REV orientation on chromosome 8
- IGK locus in REV orientation on chromosome 17
- IGL locus in FWD orientation on chromosome 26

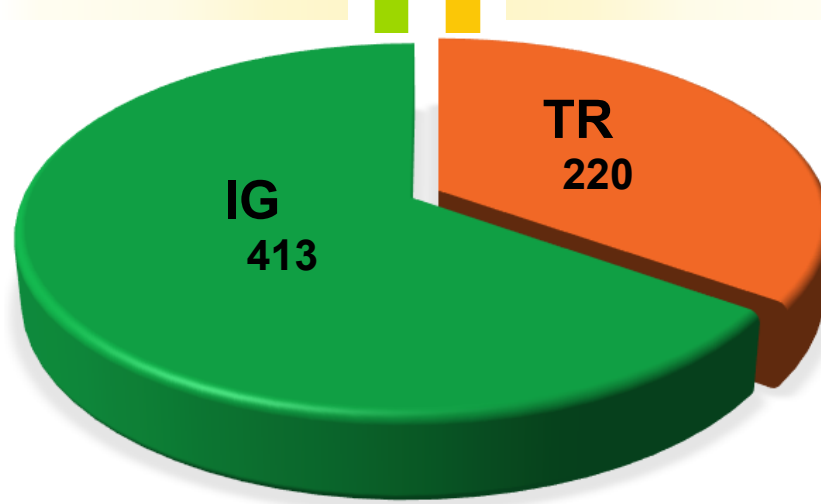
89 IGHV genes belonging to 4 IGHV subgroups, 6 IGHJ genes, 6 IGHJ genes and 5 IGHJ genes on 1,425 kilobases (kb).

22 IGKV genes belonging to 5 IGKV subgroups, 5 IGKJ genes and 1 IGKC gene on 349 kb.

261 IGLV genes belonging to 7 IGLV subgroups, 9 IGLJ genes and 9 IGLC genes on 2,583 kb.

	F	ORF	P	Total
IGH	36	2	51	89
D	5	1	0	6
J	5	1	0	6
C	4	1	0	5
IGK	13	1	8	22
J	4	1	0	5
C	1	0	0	1
IGL	69	12	180	261
J	9	0	0	9
C	9	0	0	9
Total	413			

Genes: 413
Alleles: 416



T cell receptor (TR): TRA, TRB, TRD & TRG

Locus description

The **220** dog TR genes are distributed on 4 loci, among 3 chromosomes:

- TRA/TRD loci in FWD orientation on chromosome 8
- TRB locus in REV orientation on chromosome 16
- TRG locus in FWD orientation on chromosome 18

56 TRAV genes belonging to 30 TRAV subgroups, 59 TRAJ genes and 1 TRAC gene on 743 kilobases (kb).

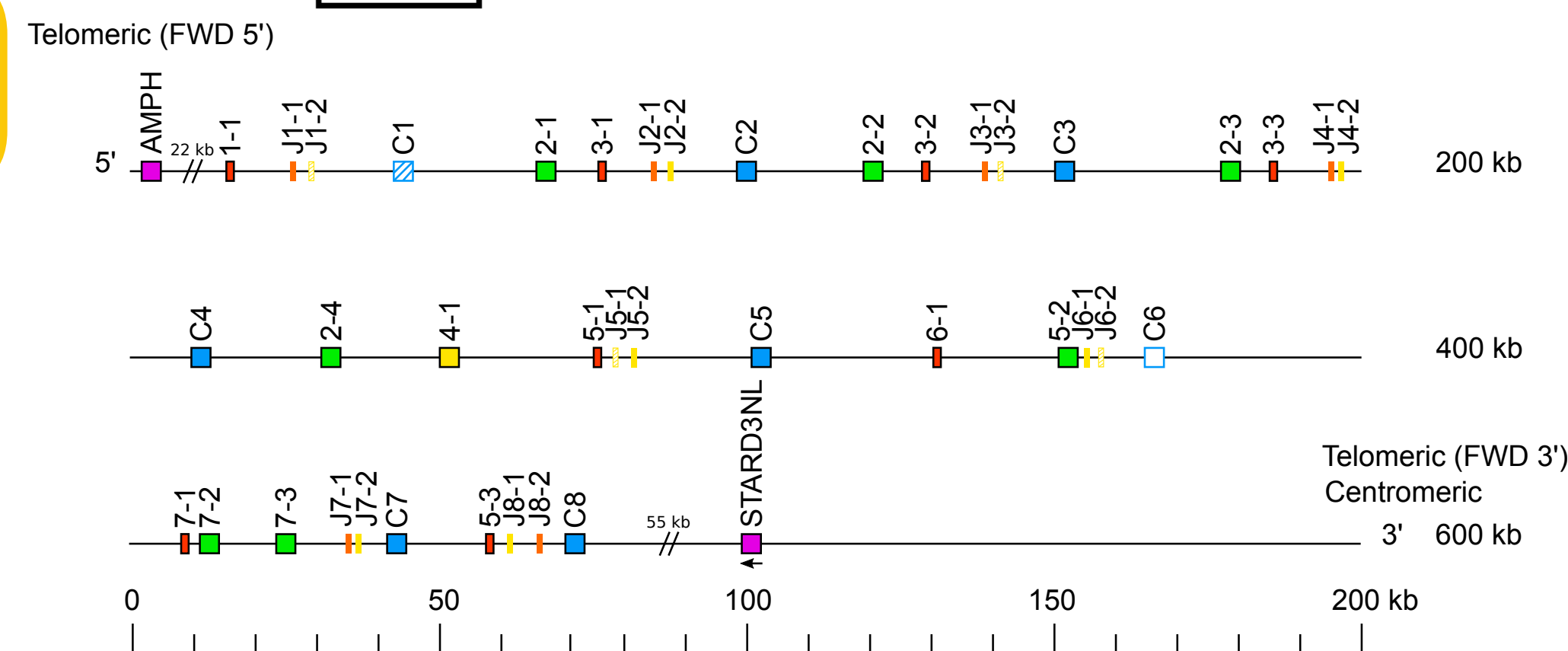
36 TRBV genes belonging to 25 TRBV subgroups, 2 TRBD genes, 12 TRBJ genes and 2 TRBC genes on 271 kb.

5 TRDV genes belonging to 5 TRDV subgroups, 2 TRDD genes, 4 TRDJ genes and 1 TRDC gene on 344 kb.

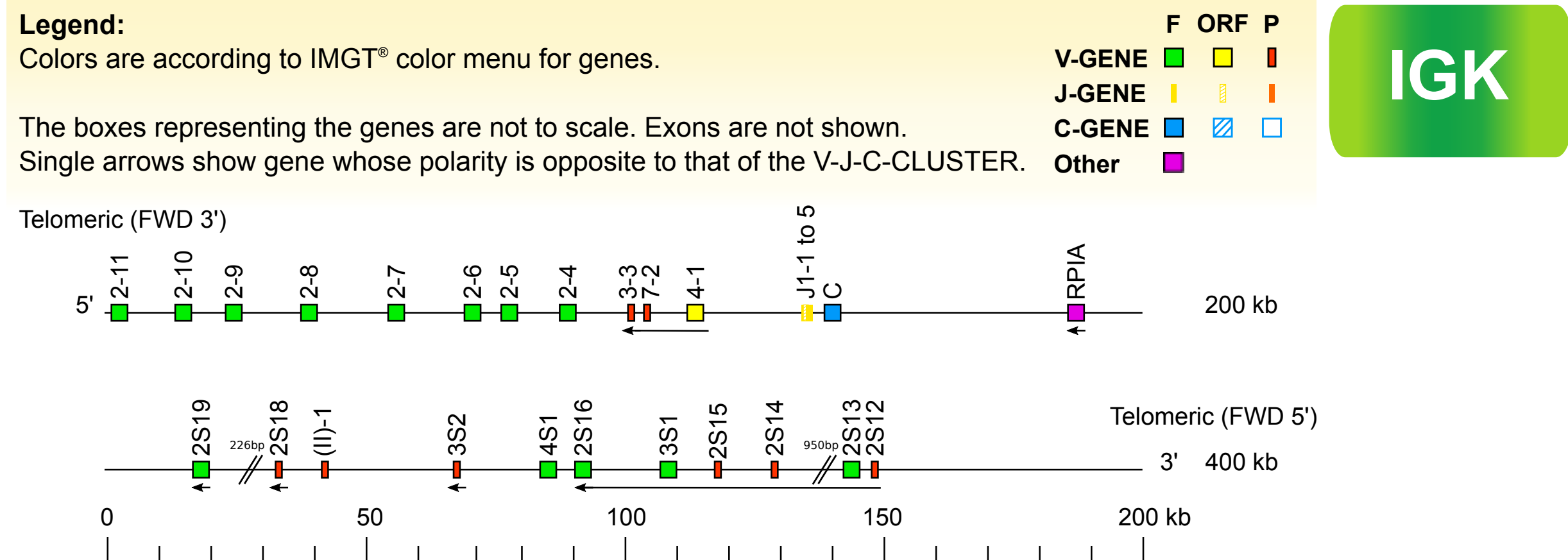
16 TRGV genes belonging to 7 TRGV subgroups, 16 TRGJ genes and 8 TRGC genes on 447 kb.

	F	ORF	P	Total
TRA	34	0	22	56
J	40	12	7	59
C	1	0	0	1
TRB	22	1	13	36
D	2	0	0	2
J	9	2	1	12
C	2	0	0	2
TRD	3	1	1	5
D	2	0	0	2
J	3	1	0	4
C	1	0	0	1
TRG	8	0	8	16
J	7	3	6	16
C	6	1	1	8
Total	220			

Locus representation



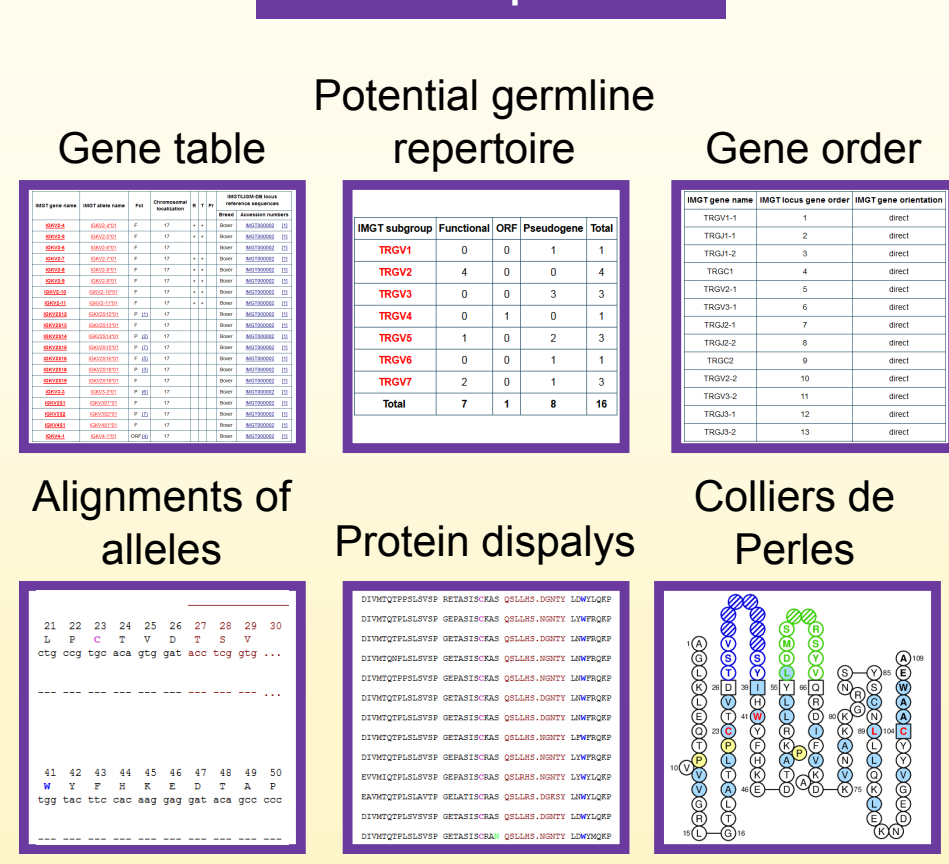
Locus representation



Locus in Genome Assembly

IMGT locus	IGK locus
Taxonomy	IMGT locus ID: Canlupfam_IGK_1
Genome assembly	CanFam3.1
GenBank assembly ID	GCA_000002285.2
RefSeq assembly ID	GCF_000002285.3
BAC library	
Chromosome	17
Chromosome sequence ID and locus positions (GenBank assembly)	CM000017.3 (37492784-37841916, complement)
Chromosome sequence ID and locus positions (RefSeq assembly)	NC_006599.3 (37492784-37841916, complement)
IMGT locus orientation on the chromosome	REV
IMGT/LIGM-DB locus reference sequence (ID)	IMGT000002
IMGT/LIGM-DB locus reference sequence length (bp)	349133

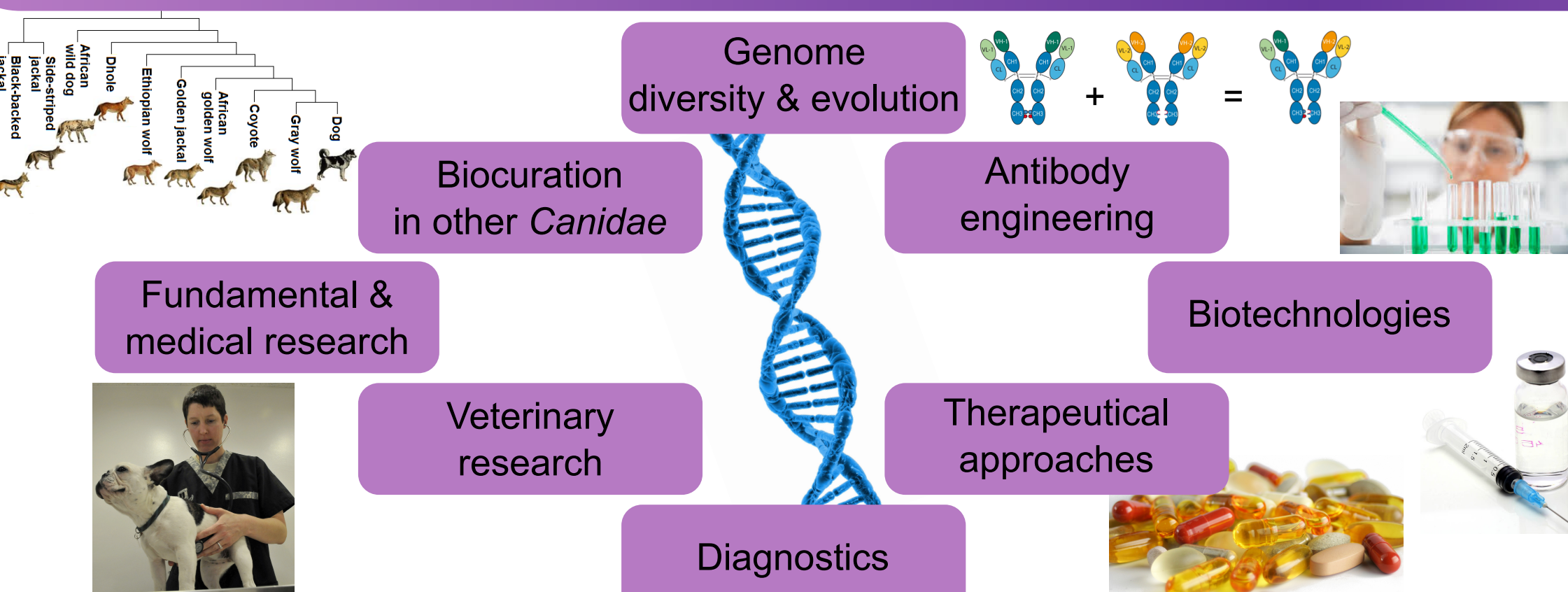
IMGT Repertoire



Locus in Genome Assembly

IMGT locus	TRG locus
Taxonomy	IMGT locus ID: Canlupfam_TRG_1
Genome assembly	CanFam3.1
GenBank assembly ID	GCA_000002285.2
RefSeq assembly ID	GCF_000002285.3
BAC library	
Chromosome	18
Chromosome sequence ID and locus positions (GenBank assembly)	CM000018.3 (11282571-11730442)
Chromosome sequence ID and locus positions (RefSeq assembly)	NC_006600.3 (11282571-11730442)
IMGT locus orientation on the chromosome	FWD
IMGT/LIGM-DB locus reference sequence (ID)	IMGT000006
IMGT/LIGM-DB locus reference sequence length (bp)	447872

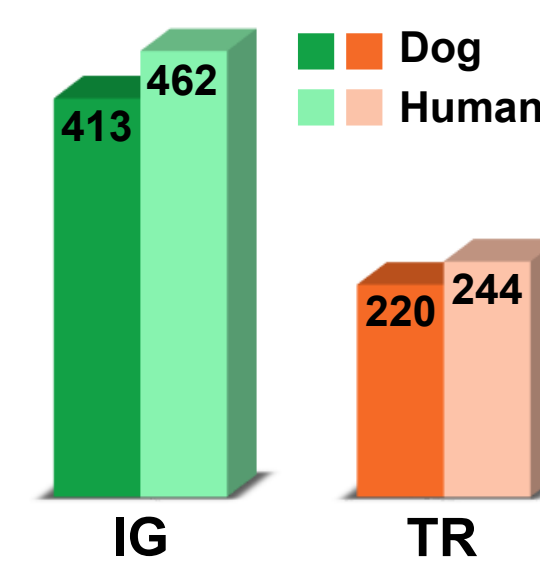
Dog repertoire: a model for veterinary and human medicine



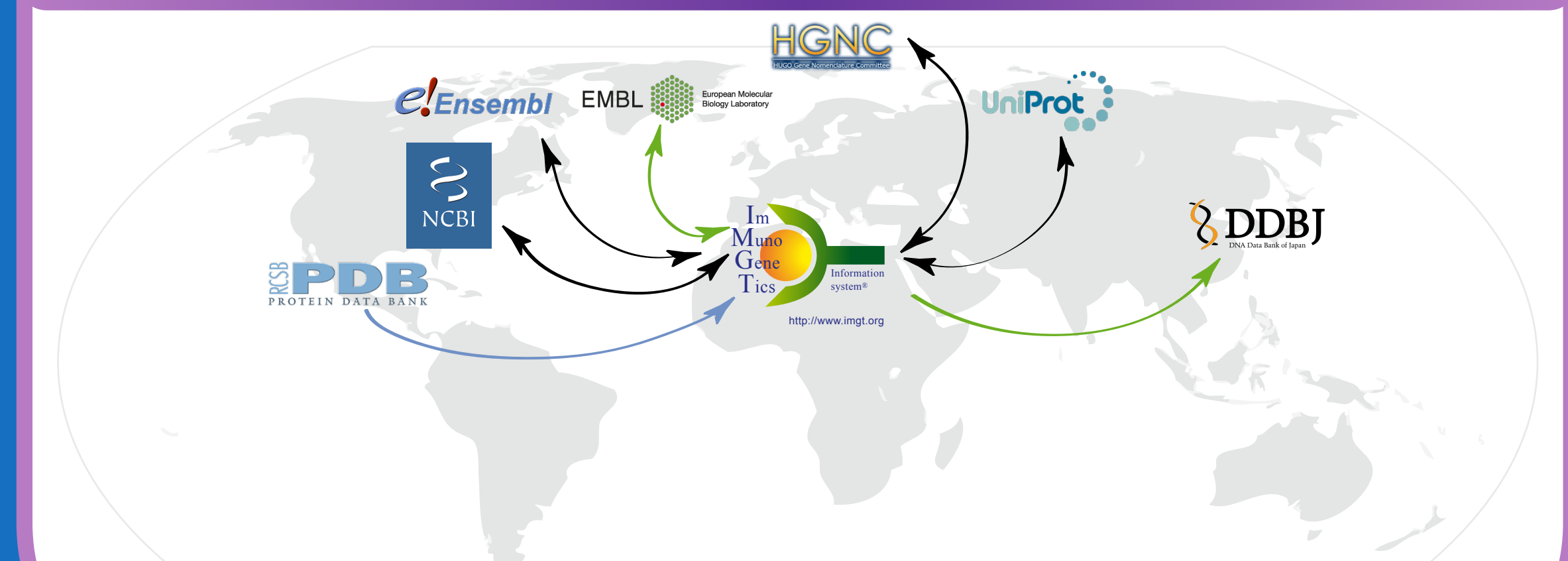
The annotation of the seven dog loci gives access to the study and the comparison of the expressed adaptive immune repertoires in veterinary normal and pathological situations.

The dog offers a unique opportunity for potential applications in veterinary and human medicine.

For example, the treatment of canine lymphoma has been predictive of the human response to that treatment.



External links



IMGT[®] founder and executive director emeritus: Marie-Paule Lefranc (Marie-Paule.Lefranc@igh.cnrs.fr)
IMGT[®] director: Sofia Kossida (Sofia.Kossida@igh.cnrs.fr)
Bioinformatics manager: Véronique Giudicelli (Veronique.Gudicelli@igh.cnrs.fr)
Computer manager: Patrice Duroux (Patrice.Duroux@igh.cnrs.fr)
Corresponding authors: Imène Chentli (Imene.Chentli@igh.cnrs.fr), Perrine Pégorier (Perrine.Pegorier@igh.cnrs.fr) & Saida Saljoqi (Saida.Saljoqi@igh.cnrs.fr)

