

Nomenclature of the Human Immunoglobulin Kappa (IGK) Genes

Marie-Paule Lefranc

IMGT Nomenclature Committee, CNRS, Université Montpellier II, Montpellier, France

Key Words

Human · IMGT · Immunoglobulin · Kappa chain genes · Orphans

Abstract

'Nomenclature of the Human Immunoglobulin Kappa (IGK) Genes', the 17th report of the 'IMGT Locus in Focus' section, provides the first complete list of all the human IGK genes. In the most frequent haplotypes, the human IGK locus spans 1,820 kb and the IGKV genes are organized in two clusters separated by 800 kb. In those haplotypes where both the proximal and distal IGKV clusters are present, the total number of human IGK genes per haploid genome is 82 (107 genes if the orphans are included) of which 37–41 are functional. If only the proximal IGKV cluster is present, which is found in a rare haplotype, the total number of genes per haploid genome is 46 (71 genes if the orphans are included) of which 23–25 genes are functional. IMGT/HUGO gene names and definitions

of the human IGK genes on chromosome 2p11.2, and IGK orphans on chromosomes 1, 2, 15, and 22 are provided with the gene functionality and the number of alleles according to the rules of the IMGT Scientific chart, with the accession numbers of the IMGT reference sequences, and the accession ID of the Genome Database GDB and NCBI LocusLink databases in which all the IMGT human IGK genes have been entered. The tables are available at the IMGT Marie-Paule page from **IMGT**, the international ImMunoGeneTics database (<http://imgt.cines.fr:8104>) created by Marie-Paule Lefranc, Université Montpellier II, CNRS, France.

Copyright © 2001 S. Karger AG, Basel

Introduction

'Nomenclature of the Human Immunoglobulin Kappa (IGK) Genes' is the 17th report of the 'IMGT Locus in Focus' section launched in the April 1998 issue of *Experi-*

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2001 S. Karger AG, Basel
0254-9670/01/0183-0161\$17.50/0

Accessible online at:
www.karger.com/journals/eci

Prof. Marie-Paule Lefranc, IMGT, Laboratoire d'ImmunoGénétique Moléculaire, LIGM, UPR CNRS 1142, IGH, 141, rue de la Cardonille F-34396 Montpellier Cedex 5 (France)
Tel. +33 4 99 61 99 65, Fax +33 4 99 61 99 01
E-Mail lefranc@ligm.igh.cnrs.fr, IMGT: <http://imgt.cines.fr:8104>

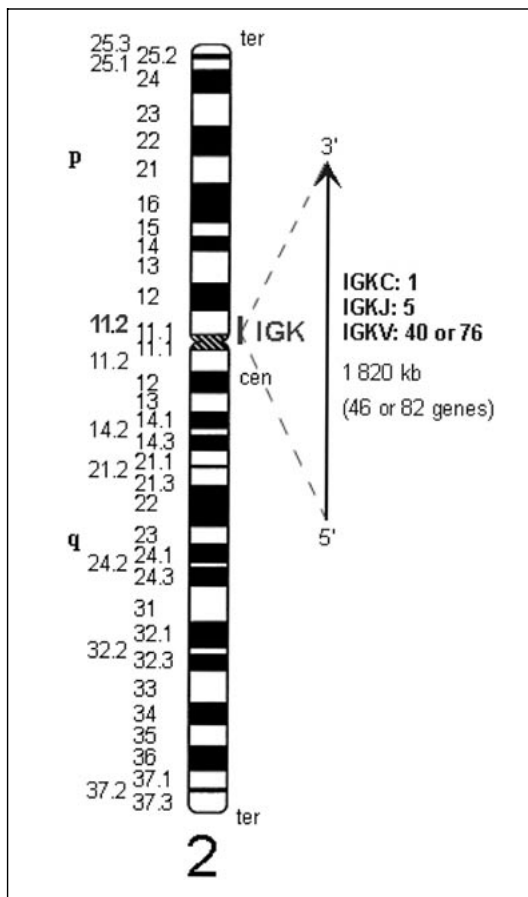


Fig. 1. Chromosomal localization of the human IGK locus at 2p11.2. A vertical line indicates the localization of the IGK locus at 2p11.2. The arrow indicates the orientation 5' → 3' of the locus, and the gene group order in the locus. The arrow is proportional to the size of the locus, indicated in kilobases. The total number of genes in the locus is shown between parentheses; the lower number (46 genes) corresponds to the rare haplotype which only comprises the proximal V-CLUSTER of 40 IGKV genes, whereas the higher number (82 genes) corresponds to the common haplotype which comprises both the distal and the proximal V-CLUSTERS and 76 IGKV genes. The number of functional genes defines the potential IGK repertoire which comprises 37–41 genes (31–35 IGKV, 5 IGKJ, and 1 IGKC) per haploid genome if both the proximal and distal V-CLUSTER are present. In the rare haplotype with only the proximal V-CLUSTER, the potential IGK repertoire comprises 23–25 genes (17–19 IGKV, 5 IGKJ, and 1 IGKC) per haploid genome [18].

mental and Clinical Immunogenetics [1–17]. This report comprises three tables and three figures entitled, respectively: (1) 'Complete list of the human IGK genes on chromosome 2 at 2p11.2'; (2) 'Human IGK orphans on chromosomes 1, 2, 15, and 22'; (3) 'Correspondence between the human IGKV gene nomenclatures'; (4) 'Chromosomal localization of the human IGK locus at 2p11.2'; (5) 'Representation of the human IGK locus at 2p11.2', and (6) 'The CLASSIFICATION concept of the IMGT-ONTOLOGY exemplified for the IGKV genes'. The tables provide the first complete list of all the human IGK genes. In the most frequent haplotypes, the human IGK locus spans 1,820 kb and the IGKV genes are organized in two clusters separated by 800 kb. In those haplotypes where both the proximal and distal IGKV clusters are present the total number of human IGK genes per haploid genome is 82 (107 genes if the orphans are included) of which 37–41 are functional. If only the proximal IGKV cluster is present, which is found in a rare haplotype, the total number of genes per haploid genome is 46 (71 genes if the orphans are included) of which 23–25 genes are functional. IMGT/HUGO gene names and definitions of the human IGK genes on chromosome 2p11.2, and IGK orphans on chromosomes 1, 2, 15, and 22 are provided with the gene functionality and the number of alleles, according to the rules of the IMGT Scientific chart, with the accession numbers of the IMGT reference sequences, and the accession ID of the Genome Database GDB and NCBI LocusLink databases in which all the IMGT human IGK genes have been entered. Detailed references for individual IGKV, IGKJ, and IGKC genes are available in references 3, 7, and 18. These tables and figures are available at the IMGT Marie-Paule page from **IMGT**, the international ImMunoGeneTics database (<http://imgt.cines.fr:8104>) created by Marie-

Paule Lefranc, Université Montpellier II, CNRS, Montpellier, France [19–21].

Human IGK Locus at 2p11.2

The human IGK locus is located at band 2p11.2 [22] of the short arm of chromosome 2 [23] (fig. 1). The orientation of the locus has been determined by the analysis of translocations, involving the IGK locus, in leukemia and lymphoma.

The human IGK locus at 2p11.2 spans 1,820 kb (fig. 2). It consists of 76 IGKV genes [3, 7, 24–29] belonging to 7 subgroups, 5 IGKJ segments [30], and a unique IGKC gene [31] (table 1). In the most frequent haplotypes, the 76 IGKV genes are organized in two clusters separated by 800 kb. The IGKV distal cluster (in 5' of the locus and in the most centromeric position) spans 400 kb and comprises 36 genes. The IGKV proximal cluster (in 3' of the locus, closer to IGKC, and in the most telomeric position) spans 600 kb and comprises 40 genes. The potential genomic IGK repertoire comprises 31–35 functional IGKV genes belonging to 5 subgroups, the 5 IGKJ segments, and the unique IGKC gene. One rare IGKV haplotype has been described which contains only the proximal cluster. This haplotype comprises the 40 proximal IGKV genes belonging to 7 subgroups, of which 17–19 are functional and belong to 5 subgroups. If both the proximal and distal IGKV clusters are present, the total number of human IGK genes per haploid genome is 82 (107 genes if the orphans are included) of which 37–41 are functional [18]. If only the proximal IGKV cluster is present, the total number of genes per haploid genome is 46 (71 genes if the orphans are included) of which 23–25 genes are functional [18].

Orphans

Twenty-five IGKV orphans have been identified and sequenced: one on chromosome 1, 2 on the short arm of chromosome 2 but outside of the main IGK locus, 12 on the long arm of chromosome 2, one on chromosome 15, 5 on chromosome 22, and 4 outside of chromosome 2 [3] (table 2).

IGK Gene Nomenclature and IMGT Scientific Chart

Gene Names

Gene names (tables 1, 2) are according to the IMGT gene name nomenclature for Ig and TcR of all vertebrates based on the 'CLASSIFICATION' concept of the IMGT-ONTOLOGY [32] (Appendix), and according to the rules of the IMGT Scientific chart [19, 21] available at <http://imgt.cines.fr:8104>. IMGT gene names and IMGT gene definitions for the human Ig [33] and TcR genes [34] have been approved by the Human Genome Organization (HUGO) Nomenclature Committee in 1999. Note that, in the HUGO symbols (<http://www.gene.ucl.ac.uk/nomenclature>), slashes and parentheses are omitted, and capital letters replace the lowercase letters found in some provisional IMGT gene names. Otherwise the gene symbols and all the full names (including slashes and parentheses) are identical in IMGT and HUGO nomenclatures.

Functionality

Criteria of functionality (F: Functional, P: Pseudogene, ORF: Open Reading Frame) (tables 1, 2) have been described in the IMGT Scientific chart [1]. The definition of functionality is based on sequence analysis. As examples, the instances functional (for germline V, D, J, and for C sequences) mean that

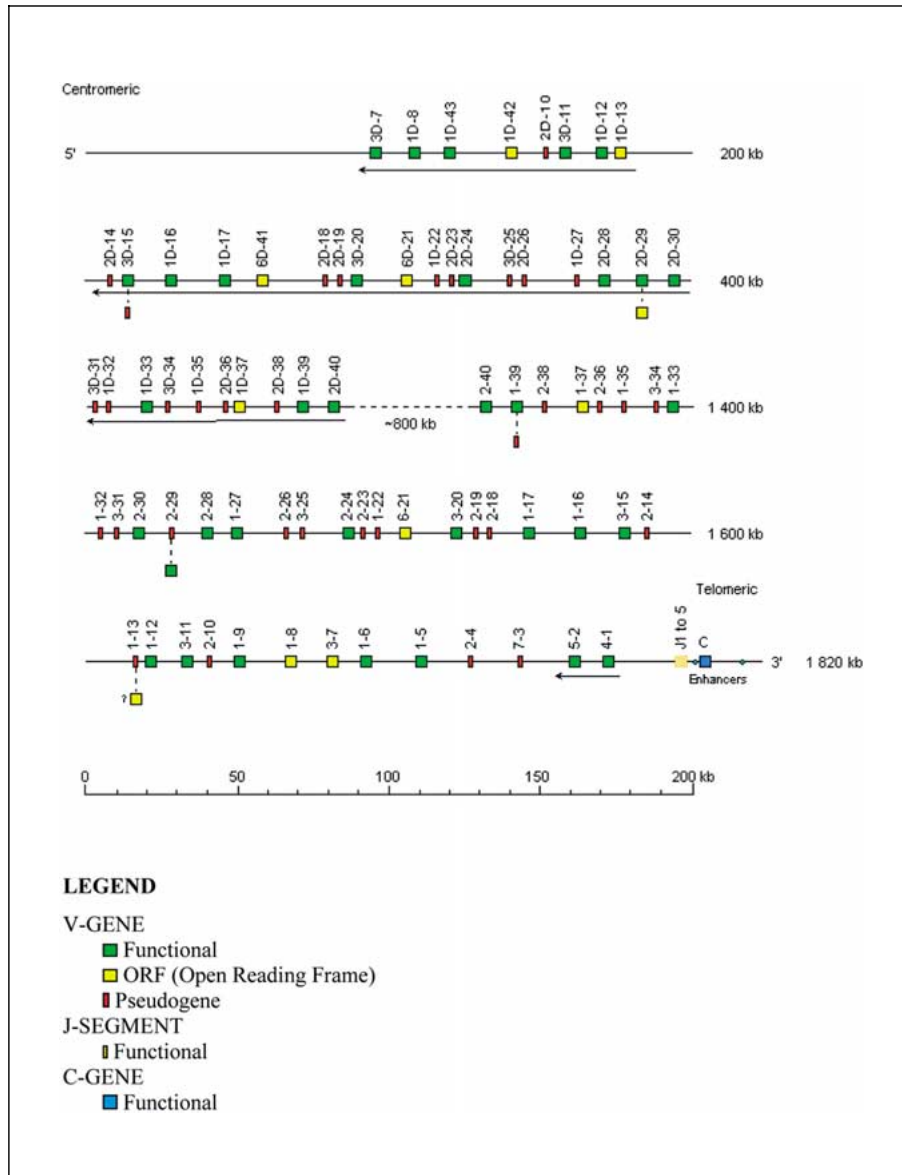
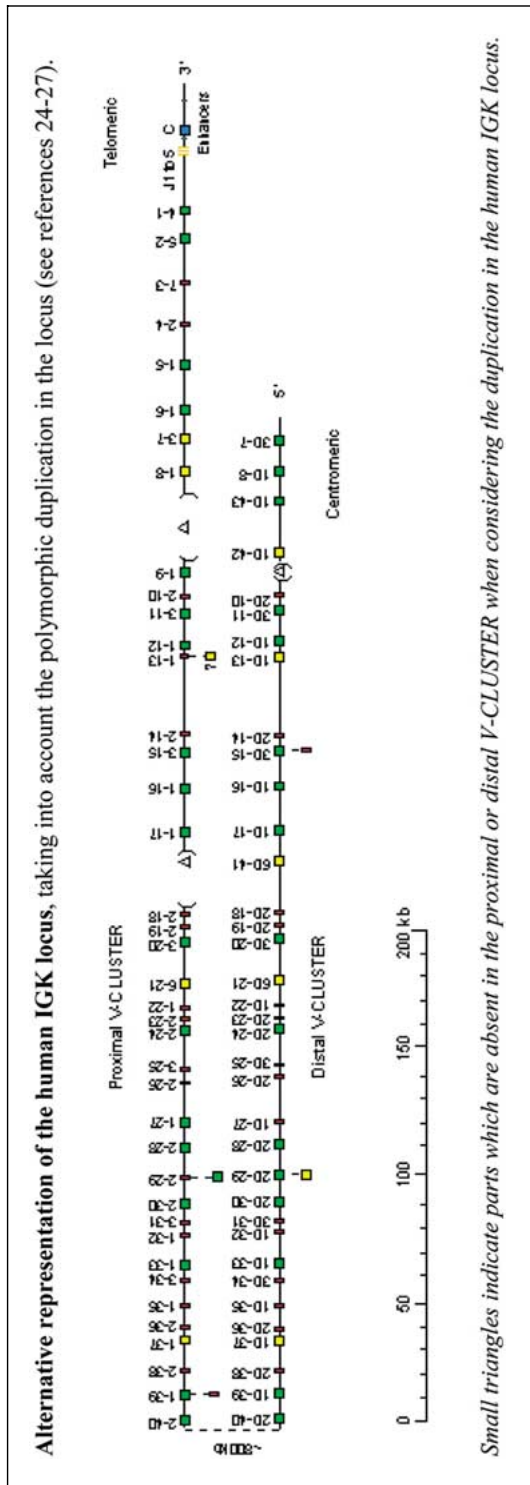


Fig. 2. Representation of the human IGK locus at 2p11.2. The boxes representing the genes are not to scale. Exons are not shown. The IGKV genes of the proximal V-CLUSTER are designated by a number for the subgroup, followed by an hyphen and a number for the localization from 3' to 5' in the locus. The IGKV

genes of the distal duplicated V-CLUSTER are designated by the same numbers as the corresponding genes in the proximal V-CLUSTER, with the letter D added. Arrows show the IGKV gene polarity of which is opposite to that of the J-C-CLUSTER.



the coding regions have an open reading frame without a stop codon, and that there is no described defect in the splicing sites, and/or recombination signals, and/or regulatory elements. According to the gravity of the identified defects, functionality can be defined as ORF, pseudogene or vestigial (for germline V, D, J, and for C genes) [1]. Complete definitions are available in the IMGT Scientific chart at the IMGT Marie-Paule page. Information on gene rearrangement, DNA transcription into mRNA, and RNA translation into a polypeptide chain is provided in the IMGT 'Germline gene tables' in the IMGT Repertoire (columns designated as R, T, and Pr, respectively), and has been published in a previous 'IMGT Locus in Focus' report [3]. This information is extracted from the literature and through IMGT/LIGM-DB sequence database search [20, 21]. The IMGT/V-QUEST tool, available at the IMGT Home page at <http://imgt.cines.fr:8104>, allows the identification of the germline IGKV, and IGKJ genes from IGKV-J genomic rearrangements and transcripts, and provides translation and a two-dimensional representation (Collier de Perles) of the variable regions [20, 21, 35].

Reference Sequences

For each gene, an IMGT reference sequence accession number is given (tables 1, 2). For the functional or ORF genes, the IMGT reference sequence accession number is that corresponding to allele*01. Note that the number *01 does not necessarily mean that other alleles are already known, but it signifies that any new polymorphic sequence will be described by comparison to that allele *01. Although the IMGT accession numbers are the same as those from the EMBL/GenBank/DDBJ general databases, the content of the IMGT/LIGM-DB flat files differs by the expert annotations, added by IMGT.

Table 1. Complete list of the human IGK genes on chromosome 2 at 2p11.2

IGKV gene nomenclature: IGKV genes are designated by a number for the subgroup, followed by an hyphen and a number for the localization from 3' to 5' in the locus. The IGKV genes of the distal duplicated V-CLUSTER are designated by the same number as the corresponding genes in the proximal V-CLUSTER, with the letter D added.

a Proximal cluster

IMGT gene groups	IMGT gene names (1)	IMGT Functionality	IMGT reference sequence accession numbers	Number of alleles	IMGT gene definition (2)	GDB accession ID (3)	LocusLink accession ID (3)
IGKC	IGKJ						
	IGKJ1	F	J00241/V00557	4	Immunoglobulin kappa constant	GDB:120088	3514
	IGKJ2	F	J00242	1	Immunoglobulin kappa joining 1	GDB:9953169	28950
	IGKJ3	F	J00242	1	Immunoglobulin kappa joining 2	GDB:9953424	28949
	IGKJ4	F	J00242	1	Immunoglobulin kappa joining 3	GDB:9953426	28948
IGKJ5	IGKJ5	F	J00242	1	Immunoglobulin kappa joining 4	GDB:9953428	28947
	IGKJ5	F	J00242	1	Immunoglobulin kappa joining 5	GDB:9953430	28946
	IGKV1-5	F	Z00001	3	Immunoglobulin kappa variable 1-5	GDB:9953432	28944
	IGKV1-6	F	M64858	1	Immunoglobulin kappa variable 1-6	GDB:9953434	28943
	IGKV1-8	ORF	Z00014	1	Immunoglobulin kappa variable 1-8	GDB:9953436	28942
IGKV2	IGKV1-9	F	Z00013	1	Immunoglobulin kappa variable 1-9	GDB:9953438	28941
	IGKV1-12	F	V01577	1 (+1?)*	Immunoglobulin kappa variable 1-12	GDB:9953440	28940
	IGKV1-13	P (+ORF?)	Z00010	1 (+1?)*	Immunoglobulin kappa variable 1-13	GDB:9953442	28939
	IGKV1-16	F	J00248	1	Immunoglobulin kappa variable 1-16	GDB:9953444	28938
	IGKV1-17	F	X72808	1	Immunoglobulin kappa variable 1-17	GDB:9953446	28937
	IGKV1-22	P	X71885	-	Immunoglobulin kappa variable 1-22	GDB:9953448	28936
	IGKV1-27	F	X63398	1	Immunoglobulin kappa variable 1-27	GDB:9953450	28935
	IGKV1-32	P	X71883	-	Immunoglobulin kappa variable 1-32	GDB:9953452	28934
	IGKV1-33	F	M64856	1	Immunoglobulin kappa variable 1-33	GDB:9953454	28933
	IGKV1-35	P	X71890	-	Immunoglobulin kappa variable 1-35	GDB:9953456	28932
	IGKV1-37	ORF	X59316	1	Immunoglobulin kappa variable 1-37	GDB:9953458	28931
	IGKV1-39	F, P	X59315	2	Immunoglobulin kappa variable 1-39	GDB:9953460	28930
	IGKV2-4	P	X72814	-	Immunoglobulin kappa variable 2-4	GDB:9953462	28929
	IGKV2-10	P	Z00012	-	Immunoglobulin kappa variable 2-10	GDB:9953464	28928
	IGKV2-14	P	X72810	-	Immunoglobulin kappa variable 2-14	GDB:9953466	28927
IGKV2-18	P	X63400	-	Immunoglobulin kappa variable 2-18	GDB:9953468	28926	
IGKV2-19	P	X12692	-	Immunoglobulin kappa variable 2-19	GDB:9953470	28925	
IGKV2-23	P	X71885	-	Immunoglobulin kappa variable 2-23	GDB:9953472	28924	
IGKV2-24	F	X12684	1	Immunoglobulin kappa variable 2-24	GDB:9953474	28923	
IGKV2-26	P	X71884	-	Immunoglobulin kappa variable 2-26	GDB:9953476	28922	
IGKV2-28	F	X63397	1	Immunoglobulin kappa variable 2-28	GDB:9953478	28921	
IGKV2-29	P, F	X63396	2	Immunoglobulin kappa variable 2-29	GDB:9953480	28920	

IGKV2-30	F	X63403	1	Immunoglobulin kappa variable 2-30	GDB:9953482	28919
IGKV2-36	P	X71889	-	Immunoglobulin kappa variable 2-36	GDB:9953484	28918
IGKV2-38	P	X71888	-	Immunoglobulin kappa variable 2-38	GDB:9953486	28917
IGKV2-40	F	X59314	2	Immunoglobulin kappa variable 2-40	GDB:9953488	28916
IGKV3-7	ORF	X02725	3	Immunoglobulin kappa variable 3-7	GDB:9953490	28915
IGKV3-11	F	X01668	2	Immunoglobulin kappa variable 3-11	GDB:9953492	28914
IGKV3-15	F	M23090	1	Immunoglobulin kappa variable 3-15	GDB:9953494	28913
IGKV3-20	F	X12686	2	Immunoglobulin kappa variable 3-20	GDB:9953496	28912
IGKV3-25	P	X06583	-	Immunoglobulin kappa variable 3-25	GDB:9953498	28911
IGKV3-31	P	X71883	-	Immunoglobulin kappa variable 3-31	GDB:9953500	28910
IGKV3-34	P	X71891	1	Immunoglobulin kappa variable 3-34	GDB:9953502	28909
IGKV4-1	F	Z00023	1	Immunoglobulin kappa variable 4-1	GDB:9953504	28908
IGKV5-2	F	X02485	1	Immunoglobulin kappa variable 5-2	GDB:9953506	28907
IGKV6-21	ORF	X63399	1	Immunoglobulin kappa variable 6-21	GDB:9953508	28906
IGKV7-3	P	X12682	-	Immunoglobulin kappa variable 7-3	GDB:9953510	28905

b Distal cluster

IMGT gene groups	IMGT gene names (1)	IMGT functionality	IMGT reference sequence accession numbers	Number of alleles	IMGT gene definition (2)	GDB accession ID (3)	LocusLink accession ID (3)
IGKV	IGKV1D-8	F	Z00008	1	Immunoglobulin kappa variable ID-8	GDB:9953512	28904
	IGKV1D-12	F	X17263	1 (+1?)*	Immunoglobulin kappa variable ID-12	GDB:9953514	28903
	IGKV1D-13	ORF	X17262	1 (+1?)*	Immunoglobulin kappa variable ID-13	GDB:9953516	28902
	IGKV1D-16	F	K01323	2	Immunoglobulin kappa variable ID-16	GDB:9953518	28901
	IGKV1D-17	F	X63392	1	Immunoglobulin kappa variable ID-17	GDB:9953520	28900
	IGKV1D-22	P	X71887	-	Immunoglobulin kappa variable ID-22	GDB:9953522	28899
	IGKV1D-27	P	Z00004	-	Immunoglobulin kappa variable ID-27	GDB:9953524	28898
	IGKV1D-32	P	X71896	-	Immunoglobulin kappa variable ID-32	GDB:9953526	28897
	IGKV1D-33	F	M64855	1	Immunoglobulin kappa variable ID-33	GDB:9953528	28896
	IGKV1D-35	P	X71894	-	Immunoglobulin kappa variable ID-35	GDB:9953530	28895
	IGKV1D-37	ORF	X71893	1	Immunoglobulin kappa variable ID-37	GDB:9953532	28894
	IGKV1D-39	F	X59312	1	Immunoglobulin kappa variable ID-39	GDB:9953534	28893
	IGKV1D-42	ORF	X72816	1	Immunoglobulin kappa variable ID-42	GDB:9953536	28892
	IGKV1D-43	F	X72817	1	Immunoglobulin kappa variable ID-43	GDB:9953538	28891
	IGKV2D-10	P	X17265	-	Immunoglobulin kappa variable 2D-10	GDB:9953540	28890
	IGKV2D-14	P	X72811	-	Immunoglobulin kappa variable 2D-14	GDB:9953542	28889
	IGKV2D-18	P	X63395	-	Immunoglobulin kappa variable 2D-18	GDB:9953544	28888
	IGKV2D-19	P	X71882	-	Immunoglobulin kappa variable 2D-19	GDB:9953546	28887
	IGKV2D-23	P	X71887	-	Immunoglobulin kappa variable 2D-23	GDB:9953548	28886
	IGKV2D-24	F	X63401	1	Immunoglobulin kappa variable 2D-24	GDB:9953550	28885

Table 1 (continued)

b Distal cluster

IMGT gene groups	IMGT gene names (1)	IMGT Functionality	IMGT reference sequence accession numbers	Number of alleles	IMGT gene definition (2)	GDB accession ID (3)	LocusLink accession ID (3)
	IGKV2D-26	P	X12689	-	Immunoglobulin kappa variable 2D-26	GDB:9953552	28884
	IGKV2D-28	F	X12691	1	Immunoglobulin kappa variable 2D-28	GDB:9953554	28883
	IGKV2D-29	F, ORF	M31952	2	Immunoglobulin kappa variable 2D-29	GDB:9953556	28882
	IGKV2D-30	F	X63402	1	Immunoglobulin kappa variable 2D-30	GDB:9953558	28881
	IGKV2D-36	P	X71893	-	Immunoglobulin kappa variable 2D-36	GDB:9953560	28880
	IGKV2D-38	P	X71892	-	Immunoglobulin kappa variable 2D-38	GDB:9953562	28879
	IGKV2D-40	F	X59311	1	Immunoglobulin kappa variable 2D-40	GDB:9953564	28878
	IGKV3D-7	F	X72820	1	Immunoglobulin kappa variable 3D-7	GDB:9953566	28877
	IGKV3D-11	F	X17264	1	Immunoglobulin kappa variable 3D-11	GDB:9953568	28876
	IGKV3D-15	F, P	X72815	2	Immunoglobulin kappa variable 3D-15	GDB:9953570	28875
	IGKV3D-20	F	X12687	1	Immunoglobulin kappa variable 3D-20	GDB:9953572	28874
	IGKV3D-25	P	X71886	-	Immunoglobulin kappa variable 3D-25	GDB:9953574	28873
	IGKV3D-31	P	X71896	-	Immunoglobulin kappa variable 3D-31	GDB:9953576	28872
	IGKV3D-34	P	X71895	-	Immunoglobulin kappa variable 3D-34	GDB:9953578	28871
	IGKV6D-21	ORF	X12683	1	Immunoglobulin kappa variable 6D-21	GDB:9953580	28870
	IGKV6D-41	ORF	X12688	1	Immunoglobulin kappa variable 6D-41	GDB:9953582	28869

(*) An asterisk indicates alleles which could not be assigned to the proximal or distal V-CLUSTER gene.

(1) IMGT gene names have been approved by the Human Genome Organization (HUGO) Nomenclature Committee in 1999.

All the IGK gene names (gene symbols) are identical in IMGT and HUGO nomenclatures.

(2) IGK gene definitions (full names) are identical in IMGT and HUGO nomenclatures.

(3) Other entries concerning the IGK locus and groups, in the OMIM, GDB, and LocusLink genome databases, and in HUGO:

IMGT designation	IMGT definition	OMIM	GDB	LocusLink	HUGO
IGK locus	Immunoglobulin kappa locus	-	GDB:9991242	-	IGK@
IGKJ group	Immunoglobulin kappa joining group	146970	GDB:9955668	7842	IGKJ@
IGKV group	Immunoglobulin kappa variable group	146980	GDB:119341	3519	IGKV@

Table 2. Human IGK orphans on chromosomes 1, 2, 15, and 22

a On chromosome 1

IMGT gene groups (1)	IMGT gene names	IMGT functionality	IMGT reference sequence accession numbers	Number of alleles	IMGT gene definition (2)	Chromosomal localisation	GDB accession ID (3)	LocusLink Accession ID (3)
IGKV	IGKV1/OR1-1	P	M20809	-	Immunoglobulin kappa variable 1/OR1-1	Ipter-Iqter	GDB:9953584	3525

b On chromosome 2

IGKV	IGKV1/OR2-0	ORF	Y08392	-	Immunoglobulin kappa variable 1/OR2-0	2p11.2	GDB:9953586	28867
	IGKV1/OR2-3	P	X05102	-	Immunoglobulin kappa variable 1/OR2-3	2cen-2q11	GDB:9953588	28866
	IGKV1/OR2-6	P	X05103	-	Immunoglobulin kappa variable 1/OR2-6	2cen-2q11	GDB:9953590	28865
	IGKV1/OR2-9	P	X51879	-	Immunoglobulin kappa variable 1/OR2-9	2cen-2q11	GDB:9953592	28864
	IGKV1/OR2-11	P	X51885	-	Immunoglobulin kappa variable 1/OR2-11	2cen-2q11	GDB:9953594	28863
	IGKV1/OR2-108	ORF	X51887	-	Immunoglobulin kappa variable 1/OR2-108	2q12-2q14	GDB:9953596	28862
	IGKV2/OR2-1	P	X05101	-	Immunoglobulin kappa variable 2/OR2-1	2cen-2q11	GDB:9953600	28861
	IGKV2/OR2-2	P	X51884	-	Immunoglobulin kappa variable 2/OR2-2	2cen-2q11	GDB:9953604	28859
	IGKV2/OR2-4	P	X51883	-	Immunoglobulin kappa variable 2/OR2-4	2cen-2q11	GDB:9953606	28858
	IGKV2/OR2-7	P	X51881	-	Immunoglobulin kappa variable 2/OR2-7	2cen-2q11	GDB:9953608	28857
	IGKV2/OR2-8	P	X51880	-	Immunoglobulin kappa variable 2/OR2-8	2cen-2q11	GDB:9953610	28856
	IGKV2/OR2-10	P	X51886	-	Immunoglobulin kappa variable 2/OR2-10	2cen-2q11	GDB:9953612	28855
	IGKV3/OR2-5	P	X51882	-	Immunoglobulin kappa variable 3/OR2-5	2cen-2q11	GDB:9953614	28854
	IGKV3/OR2-268	ORF	X74459	-	Immunoglobulin kappa variable 3/OR2-268	2p12	GDB:9953616	3523

c On chromosome 15

IGKV	IGKV1/OR15-118	P	M20812	-	Immunoglobulin kappa variable 1/OR15-118	15pter-15qter	GDB:9953598	3526
------	----------------	---	--------	---	------------------------------------------	---------------	-------------	------

d On chromosome 22

IGKV	IGKV1/OR22-1	P	Z00040	-	Immunoglobulin kappa variable 1/OR22-1	22q11	GDB:9953620	3530
	IGKV1/OR22-5	P	Z00003	-	Immunoglobulin kappa variable 1/OR22-5	22q11	GDB:9953622	28850
	IGKV2/OR22-3	P	Z00041	-	Immunoglobulin kappa variable 2/OR22-3	22q11	GDB:9953626	3529
	IGKV2/OR22-4	P	M20707	-	Immunoglobulin kappa variable 2/OR22-4	22q11	GDB:9953628	28847
	IGKV3/OR22-2	P	Z00042	-	Immunoglobulin kappa variable 3/OR22-2	22q11	GDB:9953630	3527



Table 2 (continued)**e** Outside chromosome 2, not localized (NL)

IGMT gene groups	IGMT gene names (1)	IGMT Functionality	IGMT reference sequence accession numbers	Number of alleles	IGMT gene definition (2)	Chromosomal localisation	GDB accession ID (3)	LocusLink Accession ID (3)
IGKV	IGKV1/OR-1	P	M23653	-	Immunoglobulin kappa variable 1/OR-1	NL	GDB:9953632	3531
	IGKV1/OR-2	P	X64640	-	Immunoglobulin kappa variable 1/OR-2	NL	GDB:9953633	3532
	IGKV1/OR-3	P	X64641	-	Immunoglobulin kappa variable 1/OR-3	NL	GDB:9953634	3533
	IGKV1/OR-4	P	X64642	-	Immunoglobulin kappa variable 1/OR-4	NL	GDB:9953635	3534

(1) Note that in the HUGO symbols, slashes of the orthon names are omitted. Otherwise the gene names (gene symbols) are identical in IMGT and HUGO nomenclatures.
(2) Gene definitions (full names) are identical (including slashes) in IMGT and HUGO nomenclatures.
(3) Other entries concerning the IGK orphans in the OMIM, GDB, and LocusLink genome databases, and in HUGO:

IGMT designation	IGMT definition	OMIM	GDB	LocusLink	HUGO
IGKV/OR2	Immunoglobulin kappa variable orphans on chromosome 2	147185	GDB:128324	3524	IGKVOR2@

Alleles

The number of alleles of the human IGKV, IGKJ and IGKC genes (tables 1, 2) is according to *Tables of alleles* and *Alignments of alleles*, in the IMGT Repertoire, at <http://imgt.cines.fr:8104>. A dash (-) indicates that allele polymorphism of the pseudogenes has not been studied. Alignments of all known germline functional and ORF sequences assigned to the different alleles, by comparison to the allele *01, are displayed in reference [18]. Human IGK entries in reference 18 include 51 genes and 66 alleles, with a total of 142 sequences.

Genome Database Accession Numbers

All IMGT/HUGO human IGK gene symbols, full names and reference sequence accession numbers have been entered into Genome Database GDB, Toronto, Canada (<http://www.gdb.org>), and into LocusLink at NCBI (National Center for Biotechnology Information), Bethesda, USA (<http://www.ncbi.nlm.nih.gov/LocusLink>). Accession ID to these genome databases are provided in tables 1 and 2. Links to OMIM (Online Mendelian Inheritance in Man, MIM) (<http://www.ncbi.nlm.nih.gov/Omim>) are cited when there are entries in OMIM. Links to the individual IMGT, GDB and LocusLink gene entries are available from <http://imgt.cines.fr:8104> from IMGT Repertoire > List of human Ig and TcR genes > Immunoglobulins.

Correspondences between Nomenclatures and Numberings

Correspondence between the human IGKV gene nomenclatures is reported in table 3.

In order to easily compare sequences of immunoglobulins and T cell receptors, a unique numbering has been defined for the

Table 3. Correspondence between the different human IGKV nomenclatures

Proximal V-CLUSTER from 3' (top of left column) to 5' (bottom of left column)		Distal V-CLUSTER from 3' (bottom of right column) to 5' (top of right column)	
IMGT IGKV gene name [3]	Zachau's group [25]	IMGT IGKV gene name [3]	Zachau's group [25]
4-1	B3		
5-2	B2		
7-3	B1		
2-4	L13		
1-5	L12, L12a		
1-6	L11		
3-7	L10, L10a	3D-7	L25
1-8	L9	1D-8	L24, L24a
		1D-43	L23, L23a
		1D-42	L22
1-9	L8		
2-10	L7	2D-10	L21
3-11	L6, L6a	3D-11	L20
1-12	L5	1D-12	L19
1-13	L4	1D-13	L18
2-14	L3	2D-14	L17
3-15	L2	3D-15	L16, L16a, L16b, L16c
1-16	L1	1D-16	L15, L15a
1-17	A30	1D-17	L14
		6D-41	A14
2-18	A29	2D-18	A13
2-19	A28	2D-19	A12
3-20	A27, A27a	3D-20	A11, A11a
6-21	A26	6D-21	A10
1-22	A25	1D-22	A9
2-23	A24	2D-23	A8
2-24	A23	2D-24	A7
3-25	A22	3D-25	A6
2-26	A21	2D-26	A5
1-27	A20	1D-27	A4, A4a
2-28	A19	2D-28	A3
2-29	A18a, A18b	2D-29	A2a, A2c
2-30	A17	2D-30	A1
3-31	A16, A16a	3D-31	O10
1-32	A15, A15a	1D-32	O9
1-33	O18, O18a	1D-33	O8
3-34	O17	3D-34	O7
1-35	O16	1D-35	O6
2-36	O15	2D-36	O5
1-37	O14	1D-37	O4
2-38	O13	2D-38	O3
1-39	O12, O12a	1D-39	O2
2-40	O11, O11a	2D-40	O1

The IGKV genes of the proximal V-CLUSTER are designated by a number for the subgroup, followed by an hyphen and a number for the localization from 3' to 5' in the locus. The IGKV genes of the distal duplicated V-CLUSTER are designated by the same numbers as the corresponding genes in the proximal V-CLUSTER, with the letter D added. Homologous genes in the proximal and distal V-CLUSTER are on the same line. Boxes are empty when no homologous gene is found in the other cluster.

variable regions [35, 36]. Correspondence between the IMGT unique numbering and other numberings for the human IGKV genes is available from the IMGT Scientific chart and reference 35. The IMGT unique numbering relies on the high conservation of the structure of the variable region. This numbering takes into account and combines the definition of the framework (FR) and complementarity determining regions (CDR) [37], structural data from X-ray diffraction studies [38], and the characterization of the hypervariable loops [39]. The unique numbering has allowed the redefinition of the limits of the FR and CDR [35]. The FR-IMGT and CDR-IMGT lengths themselves become crucial information characterizing the variable regions belonging to a group, a subgroup, and/or a gene. For example, for a germline gene of the human IGKV1 subgroup, the lengths of the 3 CDR-IMGT, in number of amino acids is designated as [6.3.7] (IMGT Repertoire > 2D and 3D structures) [7, 35]. The unique numbering is used as the output of the IMGT/V-QUEST alignment tool, and in the *Alignments of alleles* (IMGT Repertoire > Proteins and alleles) [18].

Acknowledgments

I thank Gérard Lefranc, Ruth Lovering, Donna Maglott, Chris Porter, Sue Povey, Conover Talbot and all the immunologist colleagues for their contribution to the IMGT Nomenclature Committee. I am grateful to Valérie Contet for the editorial work. IMGT is funded by the European Union's BIOTECH (BIO4CT96-0037) and 5th PCRDT (QLG2-2000-01287) programs, the Centre National de la Recherche Scientifique (CNRS), Ministère de la Recherche, and Ministère de l'Éducation Nationale. Subventions have been received from the Association pour la Recherche sur le Cancer (ARC), and Région Languedoc-Roussillon.

Appendix

The 'CLASSIFICATION' Concept of the IMGT-ONTOLOGY

The 'CLASSIFICATION' concept of the IMGT-ONTOLOGY (fig. 3) organizes immunogenetics knowledge useful to name and classify the immunoglobulin genes [32].

Locus: A locus is a group of immunoglobulin genes that are ordered and localized in the same chromosomal location in a given species. 'Locus' IGK (2p12) is one of the three main immunoglobulin loci in the human genome. Immunoglobulin genes have also been identified in other chromosomal locations outside the main loci which represent new instances of the locus concept. However, the genes they contain, designated as orphans, are not functional.

Group: A group is a set of genes which share the same 'gene type' (V, D, J or C) and participate potentially in the synthesis of a polypeptide of the same 'chain type'. By extension, a group includes the related pseudogenes and orphans. A 4-letter root designates the 'group': for example, IGKV, IGKJ, and IGKC for the immunoglobulin kappa genes.

Subgroup: A subgroup is a set of genes which belong to the same group in a given species and share at least 75% identity at the nucleotide level (in the germline configuration for V, D, and J). For example, the IGKV genes belong to 7 subgroups.

Gene: A gene is defined as a DNA sequence that can be potentially transcribed and/or translated (this definition includes the regulatory elements in 5' and 3', and the introns, if present). Instances of the 'gene' concept are gene names. By extension, orphans and pseu-

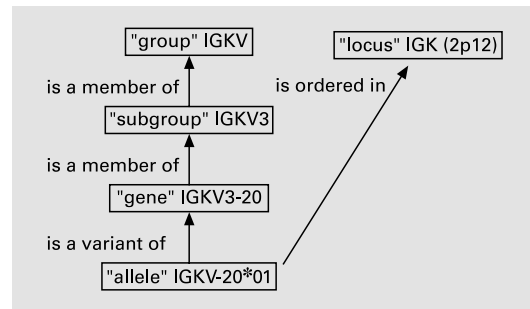


Fig. 3. The 'CLASSIFICATION' concept of the IMGT-ONTOLOGY, exemplified for the IGKV genes.

dogenes are also instances of the 'gene' concept. For each gene, IMGT has defined a reference sequence [19]. For the V, D, and J genes, the reference sequence corresponds to a germline entity. The rules for the choice of the reference sequences are described at <http://imgt.cines.fr:8104> in the IMGT Scientific chart.

Allele: An allele is a polymorphic variant of a gene. Alleles are described, exhaustively and in a standardized way, for the four 'core' coding regions, that

is the germline V-REGIONS, D-REGIONS, and J-REGIONS, and for the C-REGIONS, from immunoglobulin genes. These alleles refer to sequence polymorphisms, with mutations described at the sequence level [1]. Their sequences are compared to the reference sequence designated as *01 (see IMGT Scientific chart at <http://imgt.cines.fr:8104> for IMGT description of mutations, and IMGT allele nomenclature for sequence polymorphisms).

References

- Lefranc MP: IMGT (ImMunoGeneTics) Locus on Focus. A new section of Experimental and Clinical Immunogenetics. *Exp Clin Immunogenet* 1998;15:1-7.
- Pallarès N, Fripiat JP, Giudicelli V, Lefranc MP: The human immunoglobulin lambda variable (IGLV) genes and joining (IGLJ) segments. *Exp Clin Immunogenet* 1998;15:8-18.
- Barbié V, Lefranc MP: The human immunoglobulin kappa variable (IGKV) genes and joining (IGKJ) segments. *Exp Clin Immunogenet* 1998;15:171-183.
- Martinez C, Lefranc MP: The mouse (*Mus musculus*) immunoglobulin kappa variable (IGKV) genes and joining (IGKJ) segments. *Exp Clin Immunogenet* 1998;15:184-193.
- Pallarès N, Lefebvre S, Contet V, Matsuda F, Lefranc MP: The human immunoglobulin heavy variable (IGHV) genes. *Exp Clin Immunogenet* 1999;16:36-60.
- Ruiz M, Pallarès N, Contet V, Barbié V, Lefranc MP: The human immunoglobulin heavy diversity (IGHD) and joining (IGHJ) segments. *Exp Clin Immunogenet* 1999;16:173-184.
- Scaviner D, Barbié V, Ruiz M, Lefranc MP: Protein displays of the human immunoglobulin heavy, kappa and lambda variable and joining regions. *Exp Clin Immunogenet* 1999;16:234-240.
- Folch G, Lefranc MP: The human T cell receptor beta variable (TRBV) genes. *Exp Clin Immunogenet* 2000;17:42-54.
- Scaviner D, Lefranc MP: The human T cell receptor alpha variable (TRAV) genes. *Exp Clin Immunogenet* 2000;17:83-96.
- Scaviner D, Lefranc MP: The human T cell receptor alpha joining (TRAJ) genes. *Exp Clin Immunogenet* 2000;17:97-106.
- Folch G, Lefranc MP: The human T cell receptor beta diversity (TRBD) and beta joining (TRBJ) genes. *Exp Clin Immunogenet* 2000;17:107-114.
- Artero S, Lefranc MP: The Teleostei immunoglobulin heavy IGH genes. *Exp Clin Immunogenet*, 2000;17:148-161.
- Artero S, Lefranc MP: The Teleostei immunoglobulin light IGL1 and IGL2 V, J and C genes. *Exp Clin Immunogenet*, 2000;17:162-172.
- Folch G, Scaviner D, Contet V, Lefranc MP: Protein displays of the human T cell receptor alpha, beta, gamma and delta variable and joining regions. *Exp Clin Immunogenet* 2000;17:205-215.
- Bosc N, Lefranc MP: The mouse (*Mus musculus*) T cell receptor beta variable (TRBV), diversity (TRBD), and joining (TRBJ) genes. *Exp Clin Immunogenet* 2000;17:216-228.
- Bosc N, Contet V, Lefranc MP: The mouse (*Mus musculus*) T cell receptor delta variable (TRDV), diversity (TRDD), and joining (TRDJ) genes. *Exp Clin Immunogenet* 2001;18:51-58.
- Lefranc MP: Nomenclature of the human immunoglobulin heavy (IGH) genes. *Exp Clin Immunogenet* 2001;18:100-116.
- Lefranc MP, Lefranc G: The Immunoglobulin Facts Book. New York, Academic Press, Hartcourt, in press.
- Lefranc MP, Giudicelli V, Ginestoux C, Bodmer J, Müller W, Bontrop R, Lemaître M, Malik A, Barbié V, Chaume D: IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 1999;27:209-212.
- Ruiz M, Giudicelli V, Ginestoux C, Stoehr P, Robinson J, Bödmer J, Marsh S, Bontrop R, Lemaître M, Lefranc G, Chaume D, Lefranc MP: IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 2000;28:219-221.
- Lefranc MP: IMGT ImMunoGeneTics database. *Int Bioforum* 2000;4:98-100.
- Emanuel BS, Selden JR, Chaganti RS, Jhanwar S, Nowell PC, Croce CM: The 2p breakpoint of a 2;8 translocation in Burkitt lymphoma interrupts the V kappa locus. *Proc Natl Acad Sci USA* 1984;81:2444-2446.
- Malcolm S, Barton P, Murphy C, Ferguson-Smith MA, Bentley DL, Rabbitts TH: Localization of human immunoglobulin kappa light chain variable region genes to the short arm of chromosome 2 by in situ hybridization. *Proc Natl Acad Sci USA* 1982;79:4957-4961.
- Zachau HG: The immunoglobulin kappa locus - or - what has been learned from looking closely at one-tenth of a percent of the human genome. *Gene* 1993;135:167-173.

- 25 Zachau HG: The immunoglobulin K genes. *Immunologist* 1996;4:49–54.
- 26 Huber C, Huber E, Lautner-Rieske A, Schable KF, Zachau HG: The human immunoglobulin kappa locus. Characterization of the partially duplicated L regions. *Eur J Immunol* 1993;23:2868–2875.
- 27 Schäble KF, Zachau HG: The variable genes of the human immunoglobulin kappa locus. *Biol Chem Hoppe-Seyler* 1993;374:1001–1022.
- 28 Cox JP, Tomlinson IM, Winter G: A directory of human germ-line V kappa segments reveals a strong bias in their usage. *Eur J Immunol* 1994;24:827–836.
- 29 Schäble KF, Thiebe R, Flugel A, Meindl A, Zachau HG: The human immunoglobulin kappa locus: Pseudogenes, unique and repetitive sequences. *Biol Chem Hoppe-Seyler* 1994;375:189–199.
- 30 Hieter PA, Maizel JV Jr, Leder P: Evolution of human immunoglobulin kappa J region genes. *J Biol Chem* 1982;257:1516–1522.
- 31 Hieter PA, Max EE, Seidman JG, Maizel JV Jr, Leder P: Cloned human and mouse kappa immunoglobulin constant and J region genes conserve homology in functional segments. *Cell* 1980;22:197–207.
- 32 Giudicelli V, Lefranc MP: Ontology for immunogenetics: IMGT-ONTOLOGY. *Bioinformatics* 1999;12:1047–1054.
- 33 Lefranc MP: Nomenclature of the human immunoglobulin genes. *Curr Protocols Immunol* 2000;A.1P.1–A.1P.37.
- 34 Lefranc MP: Nomenclature of the human T cell receptor genes. *Curr Protocols Immunol* 2000;A-1O.1–A.1O.23.
- 35 Lefranc MP: The IMGT unique numbering for immunoglobulins, T cell receptors and Ig-like domains. *Immunologist* 1999;7:132–136.
- 36 Lefranc MP: Unique database numbering system for immunogenetic analysis. *Immunol Today* 1997;8:509.
- 37 Kabat EA, et al: Sequences of Proteins of Immunological Interest. Washington, Public Health Service, NIH, 1991.
- 38 Satow Y, Cohen GH, Padlan EA, Davies DR: Phosphocholine binding immunoglobulin Fab McPC603. An X-ray diffraction study at 2.7 Å. *J Mol Biol* 1986;190:593–604.
- 39 Chothia C, Lesk AM: Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* 1987;196:901–917.