Hum Genet (1999) 104 : 361–369

REVIEW ARTICLE

Marie-Paule Lefranc · Nathalie Pallarès
Jean-Pol Frippiat

Allelic polymorphisms and RFLP in the human immunoglobulin lambda light chain locus

Received: 23 December 1998 / Accepted: 1 March 1999

Abstract The organization of the human immunoglobulin lambda light chain locus (IGL) was recently described. This locus has been entirely sequenced. To evaluate the extent of the genomic variability existing inside that locus, we compiled all the available sequences of germline IGLV genes to find variants of Vλ sequences. We also looked for RFLP polymorphisms in a reputedly highly polymorphic human population from eastern Senegal, and compiled all RFLP data previously published. Analysis of these data indicates that IGLV alleles are frequent and increase the diversity of the lambda light chain repertoire in the human population. In contrast, RFLP and polymorphism by insertion and/or deletion are limited in that locus. This observation reinforces our hypothesis that the human IGL locus has undergone less evolutionary shuffling than the human kappa or heavy-chain loci.

Introduction

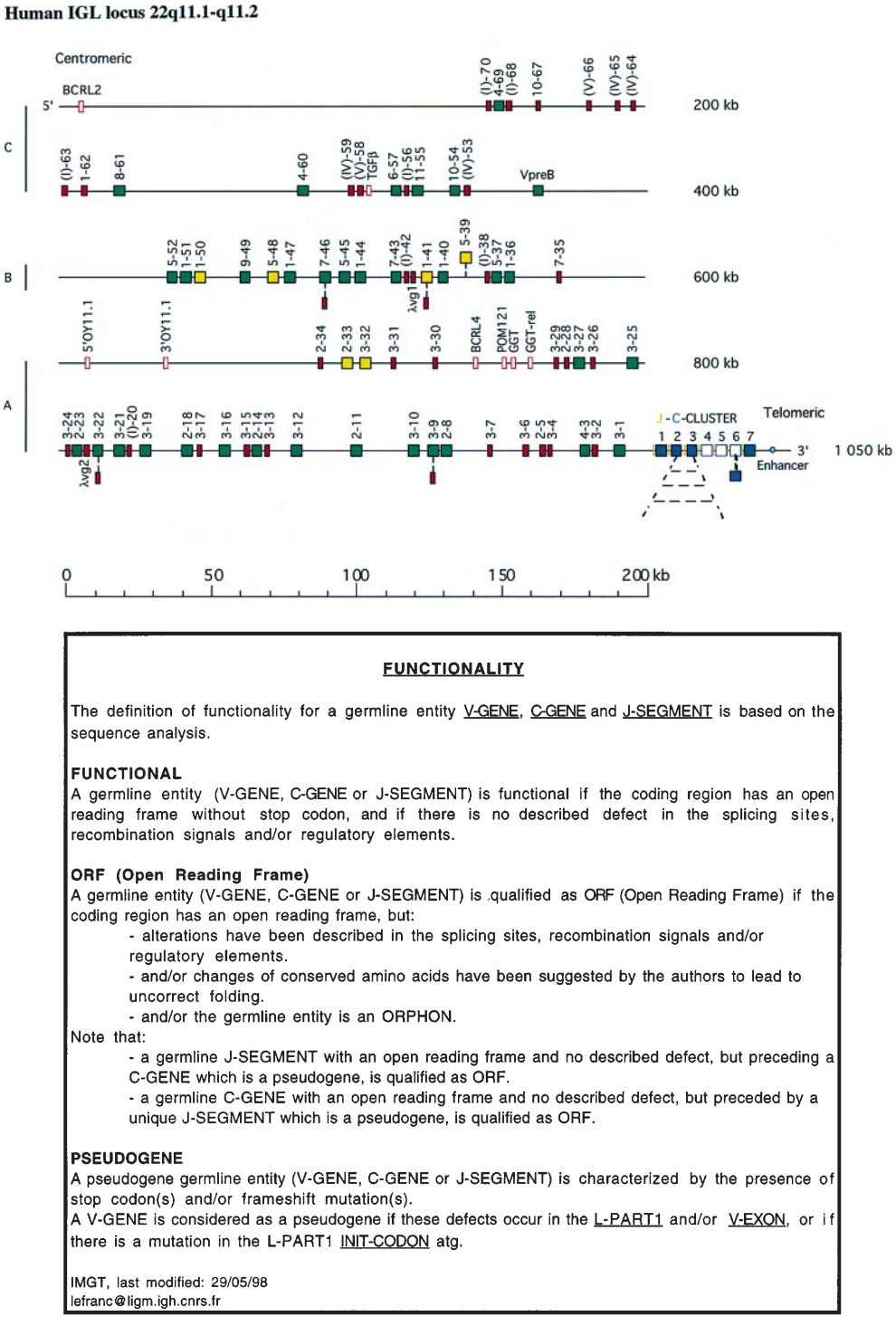
Immunoglobulin lambda light chains are encoded by V, J and C genes, which define the major human immunoglobulin lambda light chain (IGL) locus, located on chromosome 22 at band q11.2 (Erikson et al. 1981; de la Chapelle et al. 1983). Recently, a complete physical map of the locus was produced (Frippiat et al. 1995) and all functional IGLV genes were sequenced (Williams et al. 1996). One megabase of DNA covering the major IGL locus was also sequenced (Kawasaki et al. 1997).

M.-P. Lefranc · N. Pallarès · J.-P. Frippiat
Laboratoire d'ImmunoGénétique Moléculaire,
LIGM, Université Montpellier II,
UPR CNRS 1142, IGH, 141 rue de la Cardonille,
F-34396 Montpellier cedex 5, France

J.-P. Frippiat (✉)
Laboratoire de Biologie Expérimentale et d'Immunologie,
Université Henri Poincaré, Nancy 1, Boulevard des Aiguillettes,
BP 239, F-54506 Vandoeuvre-lès-Nancy cedex, France,
e-mail: frippiat@scbiol.u-nancy.fr,
Fax: +33-3-83912446

The IGL locus is contained in 900 kb of DNA. It comprises 70 IGLV genes, 7 IGLJ segments, 7 IGLC genes and 34 IGLV vestigial sequences (genes with long deletions or insertions) (Frippiat et al. 1995; Kawasaki et al. 1997) (Fig. 1). All IGLV genes and IGLJ segments are in the same transcriptional orientation and rearrange by deletion. A total of 57 IGLV genes (39 of them are functional or have an ORF) define 11 IGLV subgroups based on nucleotide homology (Fig. 2) and conserved amino acids (Chuchana et al. 1990a). The 13 remaining genes are pseudogenes which could not be assigned to subgroups with functional genes (a subgroup is only defined if it contains at least one functional or ORF IGLV gene). Thirty IGLV genes, four IGLJ segments and four IGLC genes are expressed. The expressed Vλ repertoire is mainly due to five IGLV genes: genes 1-40, 1-44, 2-8, 2-14 and 3-21 which encode 60% of the repertoire (Ignatovich et al. 1997). The IGLV2 subgroup, which comprises five functional genes, is the most expressed (36% of the repertoire), followed by the IGLV1 and IGLV3 subgroups (comprising five and ten functional genes, respectively). By contrast, the other eight subgroups represent less than 3.6% of the expressed repertoire (Ignatovich et al. 1997).

Human IGLV genes encode six different combinations of canonical structures for the complementarity-determining regions, CDR1 and CDR2 (combination 6.3 for subgroup IGLV3; 7.7 for subgroup IGLV4; 7.8 for subgroup IGLV9; 8.3 for subgroups IGLV1, 6, 10; 9.3 for subgroups IGLV1, 2, 7, 8 and combination 9.7 for subroup IGLV5; available in IMGT repertoire at http://imgt.cnusc.fr:8104). Canonical structures are determined by the loop length and the identity of key structural determining residues involved in the packing of the loop (Chothia and Lesk 1987; Wu and Cygler 1993). Human IGLV genes belonging to the same subgroup express the same combination of canonical structures for CDR1 and CDR2, except in the IGLV1 subgroup where two types of combinations are observed. IGLV genes from sheep, which express predominantly lambda light chains (Reynaud et al. 1991), are most closely related to members of the human IGLV1, 2 and 3 subgroups and appear to en-



FUNCTIONALITY

The definition of functionality for a germline entity V-GENE, C-GENE and J-SEGMENT is based on the sequence analysis.

FUNCTIONAL
A germline entity (V-GENE, C-GENE or J-SEGMENT) is functional if the coding region has an open reading frame without stop codon, and if there is no described defect in the splicing sites, recombination signals and/or regulatory elements.

ORF (Open Reading Frame)
A germline entity (V-GENE, C-GENE or J-SEGMENT) is qualified as ORF (Open Reading Frame) if the coding region has an open reading frame, but:
- alterations have been described in the splicing sites, recombination signals and/or regulatory elements.
- and/or changes of conserved amino acids have been suggested by the authors to lead to uncorrect folding.
- and/or the germline entity is an ORPHON.

Note that:
- a germline J-SEGMENT with an open reading frame and no described defect, but preceding a C-GENE which is a pseudogene, is qualified as ORF.
- a germline C-GENE with an open reading frame and no described defect, but preceded by a unique J-SEGMENT which is a pseudogene, is qualified as ORF.

PSEUDOGENE
A pseudogene germline entity (V-GENE, C-GENE or J-SEGMENT) is characterized by the presence of stop codon(s) and/or frameshift mutation(s).
A V-GENE is considered as a pseudogene if these defects occur in the L-PART1 and/or V-EXON, or if there is a mutation in the L-PART1 INIT-CODON atg.

IMGT, last modified: 29/05/98
lefranc@ligm.igh.cnrs.fr

Fig. 1 Schematic representation of the human major IGL locus at chromosome 22q11.2. *A, B, C* refer to three distinct clusters based on the IGLV gene subgroup content (Frippiat et al. 1995). Gene functionality is according to IMGT (*box*) (Lefranc 1998). Functional IGLV genes are drawn as *green squares*, IGLV ORF as *yellow squares* and IGLV pseudogenes as *red rectangles*. Most of the functional IGLV genes have been found expressed as mRNA, whereas none of the ORF has yet been observed as mRNA. Functional IGLC genes are shown as *blue squares* and IGLC pseudogenes as *white squares*. *Dashes* indicate IGLC genes not yet sequenced. Note that *boxes* representing the genes are not to scale. IGLV names are made up of a number for the subgroup (Frippiat et al. 1995; Williams et al. 1996), followed by a dash and a number for the localization from 3′ to 5′ in the locus. Pseudogenes which could not be assigned to subgroups with functional genes are designated by a roman number between parentheses, corresponding to the clans as defined by Kawasaki et al. (1997), followed by a dash and a number for the localization from 3′ to 5′ in the locus

code the same loop structures as these subgroups (Williams et al. 1996). IGLV1, 2 and 3 subgroups contain numerous genes. These genes may have been intensively duplicated during the course of evolution because they encode canonical structures which are particularly useful to our immune system. Furthermore, the location of IGLV2 and IGLV3 genes close to the J-C cluster could facilitate their expression, for example by an enhanced recombination efficiency.

Analysis of the map of the IGL locus (Fig. 1) showed that IGLV genes which belong to the same subgroup (except subgroup 4) are located in the same area of the locus. Three clusters, corresponding to three clans of subgroups, were defined to show that characteristic: cluster A (proximal to IGLC1) contains all the genes of the IGLV2 and IGLV3 subgroups and one gene of the IGLV4 subgroup; cluster B (in the middle of the locus) contains the genes of the IGLV1, 5, 7, and 9 subgroups; and cluster C (distal to

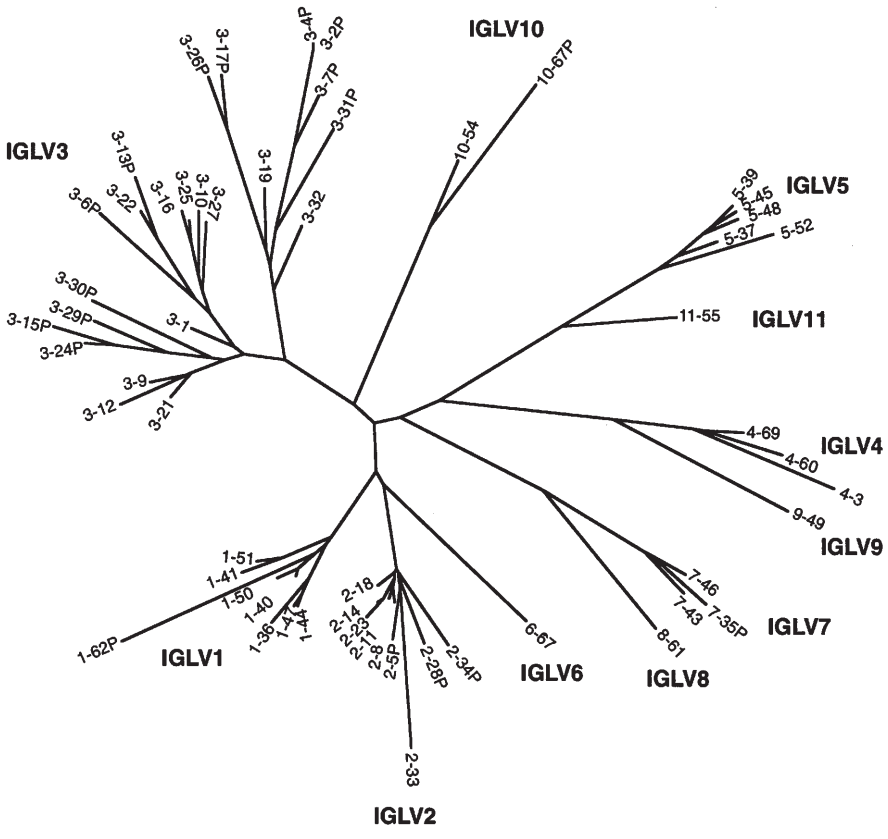


Fig. 2 IGLV family tree. The nucleotide sequences of the 57 IGLV genes (39 of them are functional or have an ORF) were used with the software PHYLIP (Felsenstein 1989) to construct an unrooted family tree. IGLV names are made of a number for the subgroup followed by a dash and a number for the localization from 3′ to 5′ in the locus. *P* at the end of the IGLV name indicates a pseudogene. This tree shows the relationships between the different IGLV genes and their classification into 11 subgroups (indicated in *bold*)

IGLC1) contains the two other members of the IGLV4 subgroup plus the genes of the IGLV6, 8, 10 and 11 subgroups (Frippiat et al. 1995; Kawasaki et al. 1997). This clustered organization differs from that of the human immunoglobulin heavy-chain (IGH) (Cook and Tomlinson 1995) and kappa light chain (IGK) (Zachau 1993) loci where members of the different families are interspersed. However, mouse VH genes are also organized in family clusters (Sheehan et al. 1993).

A minor IGL locus, located on human chromosome 8q11.2, was recently described (Frippiat et al. 1997). It contains only one member of the IGLV8 subgroup and one IGLV pseudogene. This minor locus has likely been generated by a duplication and translocation event which occurred before the evolutionary divergence of humans and gorillas. By comparison, minor loci associated with the human major IGH and IGK loci contain more genes. Eight VH genes and a cluster of DH segments have been reported on chromosome 15q11.2, and 16 VH genes were discovered on chromosome 16p11.2 (Nagaoka et al. 1994; Tomlinson et al. 1994). Twelve Vκ genes are located outside the major IGK locus on the long arm of chromosome 2, and 11 Vκ genes are dispersed on other chromosomes (Zachau 1993). These V genes and D segments are known as orphons.

The precise dimension of the major IGL locus may vary from one individual to another. Duplications of IGLC2 and/or IGLC3 can increase the number of IGLC genes by up to 11 genes per haploid genome (Taub et al. 1983; Ghanem et al. 1988; Kay et al. 1992). Several RFLP polymorphisms affecting IGLV genes (Blancher et al. 1990; Chuchana et al. 1990b, 1993; Frippiat et al. 1995) were also reported. Associations between IGLC or IGLV allele frequencies and susceptibility to Graves' disease, rheumatoid arthritis, silicosis or dilated cardiomyopathy have been described (Sidebottom et al. 1987; Williams et al. 1988; Nishi et al. 1992; Honda et al. 1993). There is also a preferential association between proteins of the IGLV6 subgroup and amyloidosis AL (Solomon et al. 1982). Different sequences, corresponding to the same IGLV gene, can be found in the literature. This observation reveals the existence of haplotypes in the human population.

Sequence diversity amongst IGLV genes

We compiled all the available sequences of germline IGLV genes to find variants of the 39 functional or ORF IGLV genes (Table 1). Sixteen genes are monomorphic, whereas 62 alleles, corresponding to 23 polymorphic genes with two to four alleles, were found. Only four alleles are pseudogenes (IGLV1-41*02; IGLV3-9*03; IGLV3-22*02 and IGLV7-46*03). IGLV allele alignments (alignments are available from IMGT at http://imgt.cnusc.fr:8104) show 95 nucleotide changes (53 transitions, 39 transversions and 3 deletions of 1 bp), of which 69 lead to an amino acid replacement ("coding mutation") and 26 are silent. An insertion of 5 bp in allele IGLV3-22*02 was observed. Some of the Vλ sequences presented in this study were

Table 1 Human IGLV alleles. Functionality (*Fct*) of the alleles (see *box*, Fig. 1) and description of the mutations are according to IMGT (http://imgt.cnusc.fr:8104; Lefranc 1998). *R/T* indicates alleles which have been observed as rearranged sequence (*R*) or as mRNA (*T*). Accession numbers of mapped sequences are indicated in *bold* and those of the reference sequences are *underlined*. Mapped sequences are those which have been obtained from clones (such as phages, cosmids or YACs) either by subcloning or PCR, and does not apply to sequences obtained directly from genomic DNA. *Asterisks* indicate a stop codon. # indicate a frameshift due to a deletion. Mutations affecting the same codon are separated by a *vertical line*. (*F* functional, *ORF* open reading frame, *P* pseudogene)

IGLV gene name	IGLV allele names	Fct	R/T	Accession number of IGLV genes having the same V-region sequence	Positions of the mutations (regarding the sequence used as reference)
1-36	IGLV1-36*01	F	+	**Z73653**/**D87009**/D87010/ Z22187/U03900/U03901	
1-40	IGLV1-40*01	F	+	**M94116**/**Z73656**/**D87010**/ Z22194	
	IGLV1-40*02	F	+	X53936/Z22193	g9>c I c10>g, L4>VI
	IGLV1-40*03	F	+	Z22192	g9>c I c10>g, L4>V I a253>g, T85>A I
1-41	IGLV1-41*01	ORF		**M94118**/**X14615**/**Z73655**/ Z22212	
	IGLV1-41*02	P		**D87010**	g295>t, E99>* I c332>t, P111>L I
1-44	IGLV1-44*01	F	+	**Z73654**/**D86999**/Z22188/ X59707/U03902	
1-47	IGLV1-47*01	F	+	**Z73663**/Z22189/M94114	
	IGLV1-47*02	F	+	**D87016**	g168>t, R56>S I
1-50	IGLV1-50*01	ORF		**M94112**/**Z73662**/**D87018**/ Z22195	
1-51	IGLV1-51*01	F	+	**Z73661**/**D87018**/Z22191/ U03870	
	IGLV1-51*02	F	+	**M30446**	t162>c I c168>a, D56>E I
2-8	IGLV2-8*01	F	+	**X97462**/D87021/Y12417	
	IGLV2-8*02	F	+	L27695	g37>a, G13>R I
	IGLV2-8*03	F		Y12418	c230>t, S77>F I
2-11	IGLV2-11*01	F	+	**Z73657**/**D86998**/Y12414	
	IGLV2-11*02	F	+	Z22198	t96>g I
	IGLV2-11*03	F		Y12415	g132>a I
2-14	IGLV2-14*01	F	+	**Z73664**/**D87015**/Z22197/ L27693	
	IGLV2-14*02	F	+	L27822	c87>t I t93>g I g94>a, G32>S I t103>c, a104>t, Y35>L I t170>g, V57>G I t198>g, N66>K I
	IGLV2-14*03	F		Y12412	g132>a I g168>t, E56>D I
	IGLV2-14*04	F		Y12413	g168>t, E56>D I
2-18	IGLV2-18*01	F	+	**Z73642**/**D87007**/D87015/ Z22199/L27689	
	IGLV2-18*02	F	+	L27697/Y12416	t317>c, L106>S I
	IGLV2-18*03	F	+	L27694	t272>c, I91>T I t317>c, L106>S I
	IGLV2-18*04	F	+	L27692	t227>c, F76>S I t249>c I t317>c, L106>S I
2-23	IGLV2-23*01	F	+	**X14616**	
	IGLV2-23*02	F	+	**Z73665**/Y12411	g170>t, G57>V I a339>c, L113>F I
	IGLV2-23*03	F	+	**D86994**/Z22196/L27688	a339>c, L113>F I
2-33	IGLV2-33*01	ORF		**Z73643**/**D87014**/Z22200/ L27687	
	IGLV2-33*02	ORF		L27823	a3>g I
	IGLV2-33*03	ORF		L27691	t96>c I a256>g, M86>V I
3-1	IGLV3-1*01	F	+	**X57826**/**Z73647**/**D87023**/ Z22208/L26403/ L26402	
3-9	IGLV3-9*01	F	+	**X97473**/**D87021**	
	IGLV3-9*02	F	+	X74288	a52>g, T18>A I a82>c, I28>L I a88>t, g89>a, S30>Y I a95>g, N32>S I g173>a, S58>N I
	IGLV3-9*03	P		X51754	a52>g, T18>A I a82>c, I28>L I a88>t, g89>a, S30>Y I a95>del#, N32>del# I g173>a, S58>N I
3-10	IGLV3-10*01	F	+	**X97464**/**D87021**	
	IGLV3-10*02	F	+	L29166	g166>a, E56>K I c207>a I t270>c I c299>a, A100>D I a319>g, T107>A I a325>t, g326>a, S109>Y I
3-12	IGLV3-12*01	F		**Z73658**	
	IGLV3-12*02	F		**D86998**	a259>g,T87>A I
3-16	IGLV3-16*01	F	+	**X97471**/**D87015**	
3-19	IGLV3-19*01	F	+	**X56178**/**D87007**/Z22202/ M94113/L35919	
3-21	IGLV3-21*01	F	+	**X71966**	
	IGLV3-21*02	F	+	**D87007**	a27>g I a49>c, K17>Q I a160>g, I54>V I t166>g, Y56>D I c324>t I
	IGLV3-21*03	F	+	M94115	a27>g I a160>g, I54>V I t166>g, Y56>D I c324>t I
3-22	IGLV3-22*01	F	+	**Z73666**/**D87007**	

Table 1 (continued)

IGLV gene name	IGLV allele names	Fct	R/T	Accession number of IGLV genes having the same V-region sequence	Positions of the mutations (regarding the sequence used as reference)
	IGLV3-22*02	P		**X71967**	c53>a, T18>K I g88>a, E30>K I g140>del#, G47>del# I c148>t, c149>g, t150>a, P50>* I 150^151>ins^tatac#, 50^51>ins^Y# I g322>a, D108>N I c330>t I
3-25	IGLV3-25*01	F	+	**X97474**	
	IGLV3-25*02	F	+	**D86994**	t14>c, g15>a, M5>T I
	IGLV3-25*03	F		L29165	t14>c, g15>a, M5>T I t294>c I
3-27	IGLV3-27*01	F	+	**D86994**	
3-32	IGLV3-32*01	ORF		**Z73645**/**D87014**	
4-3	IGLV4-3*01	F	+	**X57828**/**Z73652**/**D87024**/ Z22211	
4-60	IGLV4-60*01	F	+	**Z73667**	
	IGLV4-60*02	F	+	**D87000**	a288>t, L96>F I
4-69	IGLV4-69*01	F	+	**Z73648**/**D86993**/L29806	
	IGLV4-69*02	F	+	U03868	c198>t I
5-37	IGLV5-37*01	F	+	**Z73672**/**D87009**/D87010	
5-39	IGLV5-39*01	ORF		**Z73668**	
5-45	IGLV5-45*01	F	+	**Z73670**	
	IGLV5-45*02	F	+	**Z73671**/U93494	g22>t, A8>S I g75>a I
	IGLV5-45*03	F	+	**D86999**	g22>t, A8>S I
5-48	IGLV5-48*01	ORF		**Z73649**/**D87016**	
5-52	IGLV5-52*01	F	+	**Z73669**/**D87018**	
6-57	IGLV6-57*01	F	+	**Z73673**/**D86996**/M87320/ X92337/X92338	
7-43	IGLV7-43*01	F	+	**X14614**/**D86999**/**Z73659**/ Z22204/X01015	
7-46	IGLV7-46*01	F	+	**Z73674**/Z22205	
	IGLV7-46*02	F	+	**D86999**	c275>t, S92>L I
	IGLV7-46*03	P		Z22210	c271>del#, L91>del# I
8-61	IGLV8-61*01	F	+	**Z73650**/**D87022**/S39395/ U03639/U03635/ Z22206	
	IGLV8-61*02	F	+	U03637	c223>t, R75>C I
9-49	IGLV9-49*01	F	+	**Z73675**/Z22207	
	IGLV9-49*02	F	+	**D87016**	g291>a I
	IGLV9-49*03	F	+	U03869	g153>a I
10-54	IGLV10-54*01	F	+	**Z73676**	
	IGLV10-54*02	F	+	**D86996**	a86>t, N29>I I a228>c, L76>F I g320>t, W107>L I
	IGLV10-54*03	F		S70116	g33>c I t125>c, g126>t, L42>P I c127>g, Q43>E I
11-55	IGLV11-55*01	F		**D86996**	

obtained by PCR amplification, but *Taq* polymerase errors cannot explain all the differences observed. Moreover, most of the sequences are genomic and from independent studies.

On average, there are three nucleotide changes between the reference sequence and one of its alleles. The diversity from one allele to another is therefore rather limited, but the number of IGLV alleles in the human population is high. A total of 62 alleles for 23 functional or ORF human IGLV genes are reported in this study, and it is likely that more alleles will be described when IGLV genes from individuals belonging to different populations have been sequenced. Up to now, 11 alleles have been described for the four functional IGLC genes (data not shown).

IGLV CDRs contain 51% of the coding mutations and 27% of the silent mutations. As the length of the CDRs represents 20% of the length of the IGL V-regions, we can conclude that most of the coding mutations are clustered in the CDRs. Therefore, coding mutations increase the allelic diversity of the CDRs and, consequently, the global diversity of the Vλ repertoire in the human population. The substitution rate at the codons implicated in the CDRs of IGLV genes is thus heavily biased towards nonsynonymous substitutions, which is indicative of natural positive selection for polymorphism. Similarly, in the HLA genetic system, mutations are also concentrated in defined areas of the genes, and the antigen recognition site (ARS) of HLA molecules is subject to natural selection favoring amino acid replacements, whereas the non-ARS region is subject to functional constraint (Imanishi and Gojobori 1992). Note that the coding mutations observed in the CDRs do not change the known Vλ loop canonical structures.

Table 2 RFLP in the human IGL locus at 22q11.2. This table summarizes all the data gathered from the literature and from our own studies. DNAs were extracted from the PBL of 18 French, 24 Tunisian and 98 Senegalese individuals. The Senegalese people belong to the Niokholo Mandenka population, a geographically, culturally, and genetically well-defined population from eastern Senagal known to be highly polymorphic for the HLA class II and IGHG genes (Langaney and Gomila 1973; Lalouel and Langaney 1976; Tiercy et al. 1992; Dard et al. 1996). DNA extraction and Southern blot analysis were performed as described by Ghanem et al. (1988), except that DNA transfer was performed onto Nylon N+ filters (Amersham) in alkaline conditions. Vλ subgroup specific probes used for Southern blot analysis were described by Frippiat et al. (1995). Probes were labeled by incorporation of α32P-dCTP during a PCR reaction using the Vλ specific primers defined by Frippiat et al. (1995). Membranes were washed under high stringency conditions. Restriction endonucleases were purchased from Boehringer Mannheim (Germany). Alleles were designated according to IMGT (Lefranc et al. 1998). IGLC alleles correspond to duplications of IGLC2 and/or IGLC3. Note that several independent studies give similar frequencies for IGLV2-8, IGLV7-43 and IGLC alleles. Sizes of the RFLP alleles may vary slightly from one author to another. *Asterisks* indicate an association between a disease and an IGLV or an IGLC RFLP allele frequency. Assignment of IGLV genes to RFLP polymorphisms was based on restriction maps and IGLV sequence analysis. (*DCM* dilated cardiomyopathy, *SLE* systemic lupus erythematosus, *RA* rheumatoid arthritis)

IGLV genes affected by the RFLP	Population analysed	Number of individuals	Restriction enzyme	RFLP allele names	Allele size (kb)	Frequence	Reference
1-41	Caucasoid	34	*Kpn*I	IGLV1-41*A1	10.5	0.47	Chuchana et al.1993
				IGLV1-41*A2	4.6	0.31	
				IGLV1-41*A3	17.0	0.19	
				IGLV1-41*A4	24.0	0.03	
2-8	Caucasoid	34	*Hin*dIII	IGLV2-8*A1	16.0	0.50	Blancher et al. 1990
				IGLV2-8*A2	9.0	0.50	
	American (racially mixed)	59	*Hin*dIII	IGLV2-8*A1	20.0	0.39	Paul et al. 1991
				IGLV2-8*A2	10.0	0.61	
	American with SLE (racially mixed)	57	*Hin*dIII	IGLV2-8*A1	20.0	0.48	Paul et al. 1991
				IGLV2-8*A2	10.0	0.52	
3-19	Caucasoid	34	*Hin*dIII	IGLV3-19*A1	7.8	0.69	Chuchana et al. 1990b
				IGLV3-19*A2	3.2	0.31	
5-39	Mandenka (east Senegal)	48	*Hin*dIII	Deletion of IGL V5-39		0.35	This study
6-57	Mandenka (east Senegal)	48	*Hin*dIII	IGLV6-57*A1	5.2	0.98	This study
				IGLV6-57*A2	7.4	0.01	
				IGLV6-57*A3	7.6	0.01	
7-43	Caucasoid	34	*Bam*HI	IGLV7-43*A1	4.7	0.52	Chuchana et al. 1993
				IGLV7-43*A2	3.0	0.48	
	Japanese	127	*Bam*HI	IGLV7-43*A1	5.3	0.55	Nishi et al. 1992
				IGLV7-43*A2	3.3	0.45	Honda et al. 1993
	Japanese with silicosis*	46	*Bam*HI	IGLV7-43*A1	5.3	0.38	Honda et al. 1993
				IGLV7-43*A2	3.3	0.62	
	Japanese with DCM*	61	*Bam*HI	IGLV7-43*A1	5.3	0.39	Nishi et al. 1992
				IGLV7-43*A2	3.3	0.61	
7-43	Caucasoid	34	*Kpn*I	IGLV7-43*B1	16.0	0.47	Chuchana et al. 1993
				IGLV7-43*B2	13.0	0.31	
				IGLV7-43*B3	17.0	0.19	
				IGLV7-43*B4	24.0	0.03	
7-43	Caucasoid	96	*Taq*I	IGLV7-43*C1	12.9	0.46	Sidebottom et al. 1993
				IGLV7-43*C2	9.2	0.34	
				IGLV7-43*C3	4.2	0.17	
	Caucasoid with RA	104	*Taq*I	IGLV7-43*C1	12.9	0.49	Sidebottom et al. 1993
				IGLV7-43*C2	9.2	0.36	
				IGLV7-43*C3	4.2	0.15	
8-61	Mandenka (east Senegal)	98	*Hin*dIII	No polym detected			This study
	French	18	*Hin*dIII	No polym detected			This study
	Tunisian	24	*Hin*dIII	Deletion of IGLV8-61		0.04	This study
9-49	Mandenka (east Senegal)	48	*Hin*dIII	No polym detected			This study

Table 2 (continued)

IGLC genes affected by the RFLP	Population analysed	Number of individuals	Restriction enzyme	RFLP allele names	Allele size (kb)	Frequence	Reference
IGLC2 and/or IGLC3	Caucasoid	110	*Eco*RI	IGLC*A1	8	0.75	Taub et al. 1983
				IGLC*A2	13	0.05	
				IGLC*A3	18	0.20	
				IGLC*A4	23	0.01	
	English	104	*Eco*RI	IGLC*A1	8	0.83	Sidebottom et al. 1987
				IGLC*A2	13	0.02	
				IGLC*A3	18	0.13	
				IGLC*A4	23	0.01	
	English with RA*	108	*Eco*RI	IGLC*A1	8	0.92	Sidebottom et al. 1987
				IGLC*A2	13	0.02	
				IGLC*A3	18	0.06	
				IGLC*A4	23	0.00	
	French	27	*Eco*RI	IGLC*A1	8	0.80	Ghanem et al. 1988
				IGLC*A2	13	0.04	
				IGLC*A3	18	0.15	
				IGLC*A4	23	0.02	
	Lebanese	23	*Eco*RI	IGLC*A1	8	0.72	Ghanem et al. 1988
				IGLC*A2	13	0.07	
				IGLC*A3	18	0.20	
				IGLC*A4	23	0.02	
	Tunisian	33	*Eco*RI	IGLC*A1	8	0.65	Ghanem et al. 1988
				IGLC*A2	13	0.11	
				IGLC*A3	18	0.23	
				IGLC*A4	23	0.02	
	Black African	26	*Eco*RI	IGLC*A1	8	0.63	Ghanem et al. 1988
				IGLC*A2	13	0.11	
				IGLC*A3	18	0.21	
				IGLC*A4	23	0.02	
	Caucasoid	112	*Eco*RI	IGLC*A1	8	0.7	Williams et al. 1988
				IGLC*A2	13	0.00	
				IGLC*A3	18	0.22	
				IGLC*A4	23	0.00	
	Caucasoid with Graves' disease*	119	*Eco*RI	IGLC*A1	8	0.86	Williams et al. 1988
				IGLC*A2	13	0.00	
				IGLC*A3	18	0.14	
				IGLC*A4	23	0.00	
	American (racially mixed)	54	*Eco*RI	IGLC*A1	8	0.79	Paul et al. 1991
				IGLC*A2	13	0.00	
				IGLC*A3	18	0.18	
				IGLC*A4	23	0.03	
	Amerian with SLE (racially mixed)	66	*Eco*RI	IGLC*A1	8	0.79	Paul et al. 1991
				IGLC*A2	13	0.01	
				IGLC*A3	18	0.18	
				IGLC*A4	23	0.02	
	Japanese	46	*Eco*RI	IGLC*A1	8	0.21	Kay et al. 1992
				IGLC*A2	13	0.22	
				IGLC*A3	18	0.50	
				IGLC*A4	23	0.05	
				IGLC*A5	28	0.02	
	Venezuelan	70	*Eco*RI	IGLC*A1	8	0.65	Blasisni et al. 1996
				IGLC*A2	13	0.04	
				IGLC*A3	18	0.30	
				IGLC*A4	23	0.10	

Interestingly, of 42 nucleotide changes in the CDRs, 8 destroy 5 AGC or AGT Ser codons defined as a target for somatic hypermutation (Wagner et al. 1995) and 3 mutations create 3 new AGT codons. The loss of targets for somatic hypermutation will probably result in a less efficient affinity maturation of the antibodies made from these alleles. An increase in the number of AGT or AGC codons in the CDRs would have the opposite effect.

There is also no link between expression level and polymorphism, since genes expressed at low levels are not more or less polymorphic than highly expressed genes.

IGLV restriction fragment length polymorphisms

To further investigate the polymorphisms existing within the human IGL locus, we compiled previously published RFLP data, and we searched for RFLP in the Mandenka population (eastern Senegal, West Africa) which is known to be highly polymorphic for the HLA class II and IGHG genes (Dard et al. 1996). *Hin*dIII-digested genomic DNA of 48 healthy Mandenkalu individuals was successively hybridized in Southern blot analysis at high stringency to probes specific for the IGLV5, 6, 8 and 9 subgroups. Our results for these IGLV subgroups, as well as data from the literature for the IGLV1, 2, 3 and 7 subgroups and for the IGLC genes are reported in Table 2. Of the 48 Mandenkalu individuals, 17 were homozygous for a polymorphic deletion of the IGLV5-39 gene. Two individuals were heterozygous for rare *Hin*dIII RFLP alleles of the 6–57 gene (Table 2). One had the *A1/*A2 (5.2 kb/7.4 kb) genotype; the second had the *A1/*A3 (5.2 kb/7.6 kb) genotype. No RFLP polymorphism was detected for the IGLV8 and IGLV9 genes.

No RFLP associated with the IGLV8 subgroup was detected when, in an additional analysis, *Hin*dIII-digested genomic DNA from 24 Tunisians, 18 French and 50 other Mandenkalu individuals were hybridized, at high stringency, to the IGLV8 subgroup-specific probe. However, an homozygous deletion of IGLV8-61 was observed in one Tunisian individual. Except for that deletion and the polymorphism by insertion/deletion of the IGLV5-39 gene, the overall number of human IGLV genes is rather constant in all individuals analyzed. Our data and those in the literature (Table 2) show that IGLV genes, which belong to small subgroups (IGLV6, 8 and 9 subgroups comprise 1 or 2 genes), display less RFLP polymorphism than IGLV genes which belong to larger subgroups (IGLV1, 2, 3, 5 and 7 subgroups comprise 8, 9, 23, 5 and 3 genes, respectively). Allele frequencies are almost identical in independent studies. The only remarkable difference is found in the Japanese population where five RFLP alleles were found for the IGLC2 and/or IGLC3 genes at unique frequencies. Thus, the number of RFLP and insertions/deletions in the IGL locus is low and the allele frequencies are almost the same in different human populations.

In a previous paper (Frippiat et al. 1997), we suggested that the human IGL locus may have undergone less evolutionary shuffling than the human IGH and IGK loci because there are only 2 Vλ genes outside the major IGL locus, while there are 24 VH, a cluster of DH and at least 23 Vκ outside the major IGH and IGK loci. This hypothesis seems confirmed by the low number of RFLPs and insertions/deletions observed in the IGL locus. It could also explain the clustered organization of the Vλ genes.

In conclusion, the diversity of the lambda light chain repertoire in the human population is increased by 23 functional or ORF IGLV genes which are polymorphic (they possess from two to four alleles). In contrast, RFLP and polymorphism by insertion/deletion is limited in the IGL locus. Analysis of large populations, from different parts of the world, for allelic polymorphism and RFLP will allow an evaluation of the implications of these polymorphisms for the immune response in normal and pathological situations.

Acknowledgements The authors would like to thank Pr. G. Lefranc (Laboratoire d'ImmunoGénétique Moléculaire, Université Montpellier II, France), Dr. P. Dard (Département d'Anthropologie et d'Ecologie, Université de Genève, Switzerland) and Dr. Z. Soua (Laboratoire de Biochimie, Faculté de Médecine de Sousse, Tunisia) for the DNA from French, Senegalese and Tunisian individuals, respectively. This work was supported by the European Union's Biotech programme (BIO4CT96-0037), the CNRS (Centre National de la Recherche Scientifique), the MENRT (Ministère de l'Education Nationale, de la Recherche et de la Technologie), the ARC (Association pour la Recherche sur le Cancer), the FRM (Fondation pour la Recherche Medicale), Ligue Nationale contre le Cancer, the University Montpellier II and the University Nancy 1.

References

Blancher A, Chuchana P, Lefranc G, Lefranc M-P (1990) *Hin*dIII restriction fragment alleles of the human immunoglobulin Vλ II subgroup genes. Nucleic Acids Res 18:6178

Blasini AM, Delgado MB, Valdivieso C, Guevara P, Ramirez JL, Stekman IL, Rodriguez MA, Williams RC Jr (1996) Restriction fragment length polymorphisms of constant region genes of immunoglobulin lambda chains in Venezuelan patients with systemic lupus erythematosus. Lupus 5:300–302

Chapelle A de la, Lenoir A, Boue G, Gallano P, Huerre C, Szajnert MF, Jeanpierre M, Lalouel M, Kaplan JC (1983) Lambda Ig constant region genes are translocated to chromosome 8 in a Burkitt's lymphoma (8,22). Nucleic Acids Res 11:1133–1142

Chothia C, Lesk AM (1987) Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol 196:901–917

Chuchana P, Blancher A, Brockly F, Alexandre D, Lefranc G, Lefranc M-P (1990a) Definition of the human immunoglobulin variable lambda (IGLV) gene subgroups. Eur J Immunol 20:1317–1325

Chuchana P, Blancher A, Lefranc G, Lefranc M-P (1990b) *Hin*dIII RFLP of the human immunoglobulin VλIII subgroup genes. Nucleic Acids Res 18:6179

Chuchana P, Frippiat J-P, Blancher A, Lefranc G, Lefranc M-P (1993) Haplotypes of the human immunoglobulin lambda IGLV7S1 and IGLV1S1 genes defined by restriction-site polymorphism and insertion/deletion of a 6 kb DNA fragment. Am J Hum Genet 53:518–525

Cook GP, Tomlinson IM (1995) The human immunoglobulin VH repertoire. Immunol Today 16:237–242

Dard P, Sanchez-Mazas A, Dugoujon JM, De Lange G, Langaney A, Lefranc M-P, Lefranc G (1996) DNA analysis of the immunoglobulin IGHG loci in a Mandenka population from eastern Senegal: correlation with Gm haplotypes and hypotheses for the evolution of the Ig CH region. Hum Genet 98:36–47

Erikson J, Martinis J, Croce CM (1981) Assignment of the genes for human λ immunoglobulin chains to chromosome 22. Nature 294:173–175

Felsenstein J (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). Cladistics 5:164–166

Frippiat J-P, Williams SC, Tomlinson IM, Cook GP, Cherif D, Le Paslier D, Collins JE, Dunham I, Winter G, Lefranc M-P (1995) Organization of the human immunoglobulin lambda light chain locus on chromosome 22q11.2. Hum Mol Genet 4:983–991

Frippiat J-P, Dard P, Marsh S, Winter G, Lefranc M-P (1997) Immunoglobulin lambda light chain orphons on human chromosome 8q11.2. Eur J Immunol 27:1260–1265

Ghanem N, Dariavach P, Bensmana M, Chibani J, Lefranc G, Lefranc M-P (1988) Polymorphism of immunoglobulin lambda constant region genes in populations from France, Lebanon and Tunisia. Exp Clin Immunogenet 5:186–195

Honda K, Kimura A, Dong RP, Tamai H, Nagato H, Nishimura Y, Sasazuki T (1993) Immunogenetic analysis of silicosis in Japan. Am J Respir Cell Mol Biol 8:106–111

Ignatovich O, Tomlinson IM, Jones PT, Winter G (1997) The creation of diversity in the human immunoglobulin Vλ repertoire. J Mol Biol 268:69–77

Imanishi T, Gojobori T (1992) Patterns of nucleotide substitutions inferred from the phylogenies of the class I major histocompatibility complex genes. J Mol Evol 35:196–204

Kawasaki K, Minoshima S, Mine E, Shibuya K, Shintani A, Schmeits JL, Wang J, Shimizu N (1997) One-megabase sequence analysis of the human immunoglobulin lambda gene locus. Genome Res 7:250–261

Kay PH, Moriuchi J, Ma PJ, Saueracker E (1992) An unusual allelic form of the immunoglobulin lambda constant region genes in the Japanese. Immunogenetics 35:341–343

Lalouel JM, Langaney A (1976) Bedik and Niokholonko of Senegal: intervillage relationship inferred from migration data. Am J Phys Anthropol 45:453–466

Langaney A, Gomila J (1973) Bedik and Niokholonko: intra- and inter-ethnic migration. Hum Biol 45:137–150

Lefranc M-P (1998) IMGT locus on focus. A new section of experimental and clinical immunogenetics. Exp Clin Immunogenet 15:1–7

Lefranc M-P, Giudicelli V, Busin C, Bodmer J, Müller W, Bontrop R, Lemaitre M, Malik A, Chaume D (1998) IMGT, the internaional ImMunoGeneTics database. Nucleic Acids Res 26: 299–305

Nagaoka H, Ozawa K, Matsuda F, Hayashida H, Matsumara R, Haino M, Shin EK, Fukita Y, Imai T, Anand R, Yokoyama K, Eki T, Doeda E, Honjo T (1994) Recent translocation of variable and diversity segments of the human immunoglobulin heavy chain from chromosome 14 to chromosome 15 and 16. Genomics 22:189–197

Nishi H, Kimura A, Fukuta S, Kusukawa R, Kawamura K, Nimura Y, Nagano M, Yasuda H, Kawai C, Sugimoto T, Okada R, Yazaki Y, Tanaka H, Harumi K, Koga Y, Sasazuki T, Toshima H (1992) Genetic analysis of dilated cardiomyopathy: HLA and immunoglobulin genes may confer susceptibility. Jpn Circ J 56:1054–1061

Paul E, Livneh A, Manheimer-Lory AJ, Diamond B (1991) Characterization of the human Ig VλII gene family and analysis of VλII and Cλ polymorphism in systemic lupus erythematosus. J Immunol 147:2771–2776

Reynaud C-A, Mackay CR, Müller RG, Weill J-C (1991) Somatic generation of diversity in a mammalian primary lymphoid organ: the sheep ileal Peyer's patches. Cell 64:995–1005

Sheehan KM, Mainville CA, Willert S, Brodeur PH (1993) The utilization of individual VH exons in the primary response of adult BALB/c mice. J Immunol 151:5364–5375

Sidebottom D, Grennan DM, Sanders PA, Read A (1987) Immunoglobulin lambda light chain genes in rheumatoid arthritis. Ann Rheum Dis 46:587–589

Sidebottom D, Bate AS, Grennan DM (1993) Lack of association between immunoglobulin light chain constant and variable region polymorphisms (Cκ, VκII, VλVII and VHIII) and rheumatoid arthritis. Exp Clin Immunogenet 10:129–136

Solomon A, Frangione B, Franklin EC (1982) Bence Jones proteins and light chains of immunoglobulins. Preferential association of the V lambda VI subgroup of human light chains with amyloidosis AL (lambda). J Clin Invest 70:453–460

Taub RA, Hollis GF, Hieter PA, Korsmeyer SJ, Waldmann TA, Leder P (1983) Variable amplification of immunoglobulin λ light chain genes in human populations. Nature 304:172–174

Tiercy JM, Sanchez-Mazas A, Excoffier L, Shi-Isaac X, Jeannet M, Mach B, Langaney A (1992) HLA-DR polymorphism in a Senegalese Mandenka population: DNA oligotyping and population genetics of DRB1 specificities. Am J Hum Genet 51:592–608

Tomlinson IM, Cook GP, Carter NP, Elaswarapu R, Smith S, Walter G, Buluwela L, Rabbitts TH, Winter G (1994) Human immunoglobulin VH and D segments on chromosomes 15q11.2 and 16p11.2. Hum Mol Genet 3:853–860

Wagner SD, Milstein C, Neuberger MS (1995) Codon bias targets mutation. Nature 376:732

Williams RC, Marshall NJ, Kilpatrick K, Montano J, Brickell PM, Goodall M, Ealey PA, Shine B, Weetman AP, Craig RK (1988) Kappa/lambda immunoglobulin distribution in Graves' thyroid-stimulating antibodies. J Clin Invest 82:1306–1312

Williams SC, Frippiat J-P, Tomlinson IM, Ignatovich O, Lefranc M-P, Winter G (1996) Sequence and evolution of the human germline V lambda repertoire. J Mol Biol 264:220–232

Wu S, Cygler M (1993) Conformation of complementarity determining region L1 loop in murine IgG λ light chain extends the repertoire of canonical froms. J Mol Biol 229:597–601

Zachau HG (1993) The immunoglobulin κ locus – or what has been learned from looking closely at one-tenth of a percent of the human genome. Gene 135:167–173