The EMBO Journal vol.7 no.3 pp.719-726, 1988

Variable region genes in the human T-cell rearranging gamma (TRG) locus: V-J junction and homology with the mouse genes

S.Huck, P.Dariavach and M.-P.Lefranc

Laboratoire d'Immunogénétique, UA CNRS 1191 Génétique Moléculaire, Université des Sciences et Techniques du Languedoc, Place E.Bataillon, 34060 Montpellier Cedex, France

Communicated by T.H.Rabbitts

The locus of the human T cell rearranging γ (TRG) or T cell receptor γ chain genes comprises at least 14 variable genes (TRGV) belonging to four subgroups, five joining segments (TRGJ) and two constant region genes (TRGC). Nine V γ genes belong to subgroup I, whereas subgroups II, III and IV each consists of a single gene respectively designated V9, V10 and V11. T cells expressing the γ chain (TRG$^+$) and recognized by the anti-Tiγ A monoclonal antibody have been shown to rearrange the V9 gene. In order to assess the N diversity at the V-J junction in the TRG$^+$ cells, the germline sequences of the segments involved in the V-J rearrangements must be known. In this paper, we report the sequences of the germline V9 and V10 genes. Comparison of the V-J junction and N region from transcripts or rearranged TRG genes belonging to the different subgroups shows no evidence of D segments in the human TRG locus. Sequences of the rearranged V11 gene from the JM cell line and those of the VA and VB pseudogenes, located upstream of V9 and V11 respectively, are given. Our results bring the number of human V γ genes whose sequence is known to 13 and reveal unexpected homology with the mouse V γ genes.

Key words: γ chain genes/T cell rearranging γ genes/T cell receptor/TRG/V γ subgroup

Introduction

Like the T cell receptor α and β genes, the mouse (Saito *et al.*, 1984; Hayday *et al.*, 1985) and human (Lefranc and Rabbitts, 1985) T cell Rearranging Gamma (TRG) or T cell Receptor Gamma genes rearrange specifically in T cells and consist of variable (V), joining (J) and constant (C) gene segments which must be joined during T cell differentiation (Raulet *et al.*, 1985). The human TRG gene locus which has been mapped to chromosome 7 (Rabbitts *et al.*, 1985) at band 7p15 (Murre *et al.*, 1985) comprises two human constant-region genes (TRGC*) linked to each other at 16 kb (Lefranc and Rabbitts, 1985; Lefranc *et al.*, 1986a,b) and at least 14 variable γ genes (TRGV) belonging to four subgroups and located upstream of the two Cγ genes (Figure 1A, Lefranc *et al.*, 1987). Nine Vγ genes, five of them functional and four pseudogenes, belong to subgroup I whereas subgroups II, III and IV each consists of a single gene respectively designated V9, V10 and V11 (Lefranc *et al.*, 1986a,c; Forster *et al.*, 1987). Two joining segments, J1 and J2, have been identified upstream of Cγ1 (TRGC1) and Cγ2 (TRGC2) respectively, as well as a Jγ (TRGJ) segment designated JP upstream of J1 (Lefranc *et al.*, 1986a,c). More recently two additional Jγ gene segments, JP1 and JP2, have been located upstream of the Cγ1 and Cγ2 genes respectively (Lefranc *et al.*, 1987; Quertermous *et al.*, 1987; Huck and Lefranc, 1987; Tighe *et al.*, 1988). T cells expressing the γ chain (TRG$^+$) and recognized by the anti-Tiγ A monoclonal antibody (Jitsukawa *et al.*, 1987) have recently been shown to rearrange the V9 gene to the JP segment (Triebel *et al.*, 1988). In order to assess the N diversity in the TRG$^+$ cells, the germline sequences of the segments involved in the V-J rearrangement have to be known. In this paper, we report the sequences of the germline V9 and V10 genes. We compared the N region of the transcripts and rearranged TRG genes described so far and found no evidence of D segments in the human TRG locus. Sequences of the rearranged V11 gene from the JM cell line and those of the VA and VB pseudogenes, located upstream of V9 and V11 respectively (Forster *et al.*, 1987) are given. These new V sequences bring to 13 the number of human V γ genes whose sequence is known and reveal unexpected homology with the mouse V γ genes.

[*TRGC: γ constant region genes (TRGC1 and TRGC2 refer to the two constant regions), TRGV: gamma variable region genes (TRGV1, V2, V3, V4, V5, V5P, V6, V7, V8, VA, V9, V10, VB, V11 refer respectively to the 14 variable genes in linkage order), TRGJ: γ joining region gene segments (TRGJP1, JP, J1, JP2, J2 refer respectively to the different J γ gene segments) (Lefranc *et al.*, 1987).]

Results and discussion

Germline V9 and V10 genes

The V9 and V10 genes have previously been sequenced from rearranged clones, λK20 and λA6 for V9 (Lefranc *et al.*, 1986a,c) and λR12 for V10 (Forster *et al.*, 1987). In order to be able to assess the N diversity at the V-J gene junction in the T cells expressing the γ chains (TRG$^+$), we report the germline sequences of these two genes (Figure 2). The germline V9 gene was subcloned from λSHV7, a clone isolated from a B cell lymphoblastoid cell line, and sequenced from the *Hin*dIII site located downstream of the V9 gene (Figure 1B) (Forster *et al.*, 1987). The 3′ region of the germline V10 gene was subcloned from λJM15, a clone which, as shown below, contains also the rearranged V11 gene from JM cell line and the pseudogene VB (Figure 1B). We sequenced from the *Eco*RI site at the 5′ end of the λJM15 clone, this site corresponding to the internal *Eco*RI of the V10 gene (Forster *et al.*, 1987). The subclones we isolated for sequencing correspond respectively to the VγII and VγIII-3′ probes (Figure 1B, and Materials and methods).

As shown in Figure 2, both V9 and V10 genes have conserved heptamer-nonamer sequences separated by a 23 bp

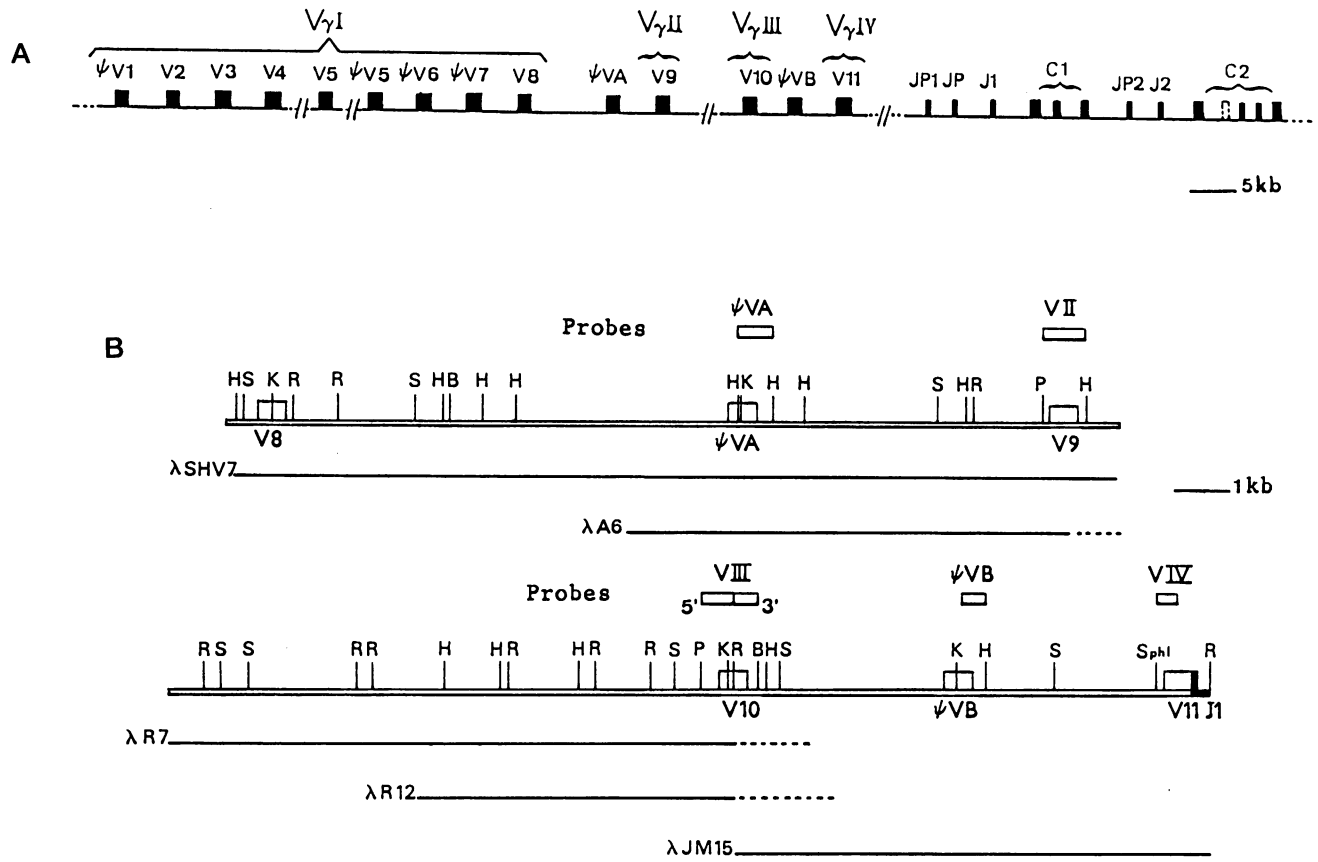

Fig. 1. (**A**) Schematic representation of the human T-cell rearranging γ (TRG) locus (Lefranc *et al.*, 1987). For detailed maps, see Lefranc *et al.* (1986b,c), Forster *et al.* (1987) and Huck and Lefranc, 1987. (**B**) Restriction enzyme maps of the region encompassing the ψVA, V9, V10, ψVB and V11 genes. λJM15 was isolated from a library of JM DNA (in lambda ongC-RI) using the J1 probe pH60. λA6, λSHV7, λR7 and λR12 have previously been described (Lefranc *et al.*, 1986c; Forster *et al.*, 1987). The presence of the *Eco*RI site downstream of V11 results from the V11-J1 rearrangement in λJM15. B, *Bam*HI; H, *Hin*dIII; K, *Kpn*I; R, *Eco*RI; S, *Sac*I. Only the *Pst*I (P) or *Sph*I sites relevant to the probes are shown.

A

Germline V9

V E K Q D I A T Y Y C A L W E V

GTA GAG AAA CAG GAC ATA GCT ACC TAC TAC TGT GCC TTG TGG GAG GTG CACAGCAGCAGACAGTTTGAGCCATCCCATTCAATAAATGTTTATTGA

B

Germline V10

V E K E D M A V Y Y C A A W W V

GTA GAG AAA GAA GAC ATG GCC GTT TAC TAC TGT GCT GCG TGG TGG GTG GCCACATACTAGAACTGTTGAAACAACATGCACAAAATCCCCTCCCAG

Fig. 2. Sequence of the 3′ region of the germline V9 (**A**) and V10 (**B**) genes, respectively single members of the VγIII and VγIV subgroups. The heptamer−nonamer sequences are underlined. The cysteine involved in the intrachain disulfide bond is circled.

spacer. Partial sequences of the germline V9 (from λSHV7) and V10 (from λJM15) genes are identical (except for their 3′ end) to the corresponding region of the rearranged V9 (from λA6) and V10 (from λR12) genes respectively (Forster *et al.*, 1987). These results are in agreement with our previous data which showed that the TRG diversity originated only from V−J joining and N-region diversity and that there was no evidence of somatic mutation by comparison of the complete sequences (leader, intervening sequence and variable region) of three rearranged TRGV genes with their germline counterpart (Lefranc *et al.*, 1986c). The nucleotide sequences of the 3′ end of the germline V9 and V10 genes (Figure 2) will allow us to assess the N-region diversity in the TRG^+ cells, and more particularly that of the Tiγ A^+ cells, since these cells have been shown to rearrange functionally and express the V9 gene (Triebel *et al.*, 1988).

Rearranged V11 gene from the JM cell line

JM cells (equivalent rearrangement to Jurkatt) have two rear-

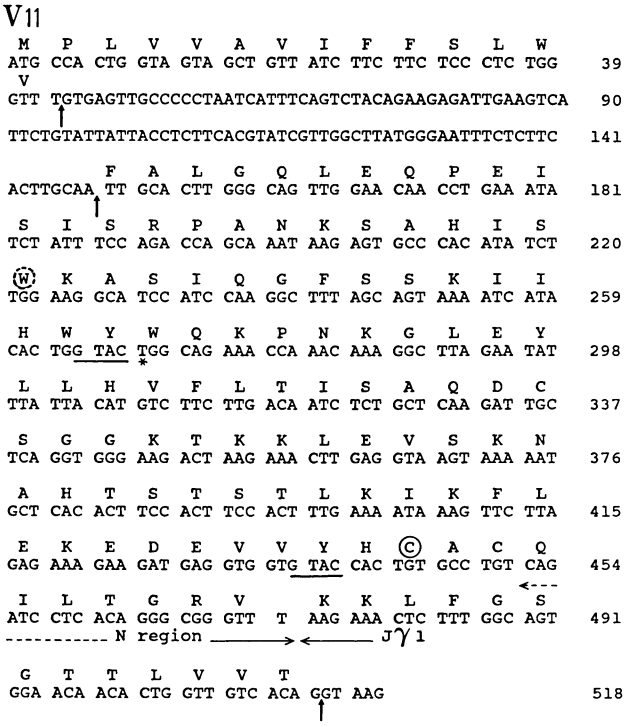

Fig. 3. Nucleotide sequence and deduced amino acid sequence of V11 non-productively joined to J1 (see text). The tryptophan residue which replaces the cysteine normally involved in the intra-chain disulfide bond is in a dotted circle. An asterisk indicates the nucleotide substitution which alters the *Kpn*I site. The *Rsa*I sites are underlined. The N region is indicated by a dotted line. The splice sites are shown by arrows.

ranged TRG alleles. The DNA cut with *Bam*HI, *Eco*RI or *Hin*dIII and hybridized with the Jγ1 probe pH60 (Lefranc and Rabbitts, 1985, and Materials and methods) shows two rearrangements, one of which has been identified as a rearranged V8 gene whereas the other has been assigned to a new gene V11, belonging to a new subgroup VγIV (Forster *et al.*, 1987). Therefore we isolated the rearranged Vγ gene from a λ phage library made from JM DNA by screening with the pH60 probe. One 9.5 kb *Eco*RI clone (λJM15) containing the rearranged V11 gene was analyzed and the restriction map appears in Figure 1B. This clone also contains the pseudogene VB and the 3′ region of the germline V10 gene (Forster *et al.*, 1987).

A 1.0 kb *Sph*I−*Eco*RI fragment containing the rearranged V11-J gene was subcloned. The sequence of this gene is shown in Figure 3. The conserved tryptophan−tyrosine residues (codons 39 and 40) are present but there is no associated *Kpn*I site due to a single base substitution at the first nucleotide of codon 41 (the *Kpn*I site GGTACC found in all the Vγ I genes and in V10 overlaps the codons 39−41). Although there is no deletion and no stop codon in the coding sequence of V11, it is unknown if V11 encodes a functional γ polypeptide: first, the splice site in 5′ of the variable region is poorly conserved; second, one of the two cysteines responsible for the intrachain disulfide bond in the V region is replaced with a tryptophan residue due to a single base pair substitution. Nevertheless, it has been shown that a functional antibody exists in which the cysteine in the VH domain is replaced with a tyrosine residue (Rudikoff and Pumphrey, 1986) and therefore we have to await the identification of a γ chain with a V11 variable domain to assess its functionality.

The V11−J junction is out of frame in the JM cell line and the N region resembles more those of the non-functional cDNA isolated from thymocyte and peripheral blood lymphocyte cDNA libraries (Yoshikai *et al.*, 1987) by its length and the rather important deletion (11 bp) of the J region (see Table I for compilation of the V−J junctions). A deletion of at least 9 bp seems to have occurred in the variable region since the tryptophan residue conserved in the other mouse or human TRGV genes is absent; the next residue in 5′ is a cysteine which probably belongs to the variable region of V11 since this residue is found in the same position in the mouse V5 and V6 sequences (Table II and Figure 7). About 40 nucleotides encompassing the 3′ end of the variable and N region of the rearranged gene from Jurkatt cell line have previously been published (Quertermous *et al.*, 1986). The JM sequence is identical to Jurkatt except for a permutation of two nucleotides. The codon CAC which encodes a histidine in JM differs by one nucleotide of the conserved TAC codon found in all the other TRGV genes whereas the codon CCA in Jurkatt differs by three substitutions. Despite that difference, our data confirm that JM and Jurkatt are clones from the same cell line.

The TRGJ segment involved in the JM V11 rearrangement was identified as J1. Indeed, the leucine residue, most in 3′ of J, is encoded by CTG like in one allelic form of J1 (Lefranc *et al.*, 1986a). To confirm this assignment to J1, we sequenced 250 bp upstream of the *Eco*RI site located at the 3′ of λJM15 (Figure 1B) (data not shown). Eight nucleotide differences have previously been noted between the J1- and J2-containing fragments in that region (Lefranc *et al.*, 1986a) (the nucleotide C at position 612 in that reference is not taken into account since it can be found in either J1 or J2 region; M.-P.Lefranc, unpublished data). The JM sequence displays the same nucleotides as M13H60 (a J1-containing fragment) in these eight positions, confirming that V11 rearranged to J1.

Southern hybridization using a V11 probe (a 270 bp *Sph*I−*Rsa*I fragment containing the 5′ region of the V11 gene) (Figure 3) detects a single hybridizing band with *Bam*HI-, *Eco*RI-, *Hin*dIII-, *Kpn*I- and *Sac*I-digested DNAs (data not shown). However, these fragments have a large size and we therefore hybridize the V11 probe to *Sph*I-digested DNAs in order to confirm that V11 is the single member of the subgroup IV (Figure 4A).

ψVA and ψVB

We previously located two pseudogenes VA and VB 6 kb upstream of V9 and 4.3 kb upstream of V11 respectively (Forster *et al.*, 1987). We now report the complete sequence of VA and VB which have been isolated from λA6 (Lefranc *et al.*, 1986c; Forster *et al.*, 1987) and from λJM15 respectively (Figure 1B). Sequences of VA and VB are shown in Figure 5. A translation stop codon (TAG) occurs nine codons upstream of the conserved cysteine codon in the VA sequence as previously shown in a partial sequence of that gene (Forster *et al.*, 1987). Three stop codons (TGA: codon 2 of the leader, TAG: codons 23 and 105 of the variable region) are found in the VB segment. Moreover there is a 2 bp insertion (codon 64) which results into a frameshift.

Southern hybridization using a ψVA probe (a 600 bp *Hin*dIII fragment) and a ψVB probe (a 500 bp *Kpn*I−*Hin*dIII

Table I. Comparison of the V−J junction (N region) of the rearranged TRG genes and transcripts with their germline counterparts

				Variable region	N region	Joining region	
V.I	(Ref.)						
				C A T W D G		G Q E L G	
GL V2	(1)		g	TGT GCC ACC TGG GAC GGG		GGG CAA GAG TTG GGC	JP (1)
λF1	(1)	V2-JP	- r	TGT GCC ACC TGG GAC GGG	CCA AG*	--- --- GAG TTG GGC	
						N Y Y K K	
						AAT TAT TAT AAG AAA	J1 or 2 (2)
HGT26	(3)	V2-J1	(+)c	TGT GCC ACC TGG GAC ---	GCT CGA AG	--T TAT TAT AAG AAA	
						T T G W F	
					S G Y	ACC ACT GGT TGG ATC	JP1 (10)
λM4r1	(11)	V2-JP1	+ r	TGT GCC AC- --- --- ---	T TCG GGA TAT	--- --- 21bp--- ---	
						S S D W I	
						AGT AGT GAT TGG ATC	JP2 (8)
λM4r2	(11)	V2-JP2	- r	TGT GCC ACC TGG GA- ---	G GGG GGT GTC T*	--- --T GAT TGG ATC	
				C A T W D R		N Y Y K K	
GL V3	(1)		g	TGT GCC ACC TGG GAC AGG		AAT TAT TAT AAG AAA	J1 or 2 (2)
λF6	(1)	V3-J1	+ r	TGT GCC ACC TGG --- ---		AAT TAT TAT AAG AAA	
λS12(λS1)	(1)	V3-J2	+ r	TGT GCC ACC TGG --- ---	R T		
					CGG ACG	AAT TAT TAT AAG AAA	
pTγ-1	(4)	V3-J2	- c	TGT GCC ACC TGG GAC AGG	CAA A**	AAT TAT TAT AAG AAA	
HGP03	(3)	V3-J2	- c	TGT GCC ACC TGG GAC ---	CCA GGC T**	--- TAT TAT AAG AAA	
				C A T W D G			
GL V4	(1)		g	TGT GCC ACC TGG GAT GGG			
λS13(λS6)	(1)	V4-J2	- r	TGT GCC ACC TGG GAT GGG	CAA TA*	AAT TAT TAT AAG AAA	
				C A T W D R			
GL V8	(1)		g	TGT GCC ACC TGG GAC AGG			
					K D		
Pγ1	(5)	V8-J2	+ c	TGT GCC ACC TGG --- ---	AAG G	-AT TAT TAT AAG AAA	
					T D R I		
pM17c64	(11)	V8-J2	+ c	TGT GCC ACC TGG --- ---	ACA GAT AGG ATC	--- TAT TAT AAG AAA	
						S S D W I	
						AGT AGT GAT TGG ATC	JP2 (8)
pM17c54	(11)	V8-JP2	- c	TGT GCC --- --- --- ---	CCC CTG GAG G**	--- --- --- TGG ATC	
V.II							
				C A L W E V		N Y Y K K	
GL V9	(this paper)		g	TGT GCC TTG TGG GAG GTG		AAT TAT TAT AAG AAA	J1 or 2 (2)
					R G		
λK20	(2)	V9-J1	+ r	TGT GCC TTG --- --- ---	CGA GG	--T TAT TAT AAG AAA	
λA6	(1)	V9-J1	- r	TGT GCC TTG TGG GAG GTG	C**	--T TAT TAT AAG AAA	
				(L)	G		
PBLC1.15	(6)	V9-J1	+ c	TGT GCC TTG TTG GAG ---	GGA	AAT --- TAT AAG AAA	
pTγ-10	(7)	V9-J	- c	TGT GCC TTG TGG GAG GTG		-AT TAT TAT AAG AAA	
					G		
IDP2.11	(6)	V9-J2	+ c	TGT GCC TTG TGG GAG ---	GGT	--- TAT TAT AAG AAA	
Pγ12	(5)	V9-J2	- c	TGT GCC TT- --- --- ---	C CGG CCC G**	--- --- --- AAG AAA	
						G Q E L G	
					A Q	GGG CAA GAG TTG GGC	JP (1)
HGP02	(3)	V9-JP	+ c	TGT GCC TTG TGG GAG G--	CG CAG	GGG --- GAG TTG GGC	
						T T G W F	
						ACC ACT GGT TGG ATC	JP1 (10)
					R	T T G W F	
λRBN	(8)	V9-JP1	+ r	TGT GCC TTG TGG GAG GTG	CGC	ACC ACT GGT TGG ATC	
V.III							
				C A A W W V		N Y Y K K	
GL V10	(this paper)		g	TGT GCT GCG TGG TGG GTG		AAT TAT TAT AAG AAA	J1 or 2 (2)
λP12	(9)	V10-J2	- r	TGT GCT GCG TGG --- ---	GGT	-AT TAT TAT AAG AAA	
pTγ-15	(7)	V10-J2	- c	TGT GCT GCG TGG --- ---	GAT TAT CAG	--T TAT TAT AAG AAA	
HGT25	(3)	V10-J1	- c	TGT --- --- --- --- ---	GTC CCC CCG GT	--- --- 28bp--- ---	
HGP08	(3)	V10-J2	- c	TGT GCT GCG TGG --- ---	GAT TAT AAT TGG GG*	--- --- --- --- AAA	
HGP10	(3)	V10-J2	(+)c	---.--- 27bp--- --- ---	GAC	--- --- --- AAG AAA	
HGP32	(3)	V10-J2	(+)c	TGT GCT GCG TGG --- ---	GAT CGT TG	--- --T TAT AAG AAA	
						S S D W I	
						AGT AGT GAT TGG ATC	JP2 (8)
HGP06	(3)	V10-JP2	(+)c	TGT GCT GCG TGG --- ---	GCC CCA TAT	AGT AGT GAT TGG ATC	
V.IV							
				C A C [Q I L]			
pTγR4	(7)	V11-J	- r	TGT GCC TGT [CAG ATC CTC]	ACA GGG CGG GTT	--- --- --T AAG AAA	
λJM15	(this paper)	V11-J1	- r	TGT GCC TGT [CAG ATC CTC]	ACA GGG CGG GTT	--- --- --T AAG AAA	

(1)(Lefranc et al.,1986c) (4)(Dianylas et al.,1986) (7)(Quertemous et al.,1986) (10)(Huck and Lefranc,1987)
(2)(Lefranc et al.,1986a) (5)(Littman et al.,1987) (8)(Quertemous et al.,1987) (11)(Tighe et al.,1987)
(3)(Yoshikai et al.,1987) (6)(Krangel et al.,1987) (9)(Forster et al.,1987)

g, germline; r, rearranged; c, cDNA; −, out of frame V−J junction; +, in-frame V−J junction; (+), in-frame V−J junction but non-productive transcript.

The three residues in brackets in the V11 sequence probably belong to the N region (see text and Figure 7). (L) in PBLC1.15 clone is assumed to be due to a reverse transcriptase error (Krangel *et al.*, 1987).

Table II. Correspondence between the different mouse Vγ gene nomenclatures

	Hayday *et al.* (1985)	Traunecker *et al.* (1986)	Garman *et al.* (1986)	Heilig and Tonegawa (1986)	Iwamoto *et al.* (1986)	Pelkonen *et al.* (1987)
V1	V10.8A	V1	V1.2	V2	V2	V1
V2	V10.8B	V2	V1.1	V1	V1	V2
V3	V5.7	V3	V1.3	V3	V3	V3
V4		V4	V2	V4	V6	V4.3
V5			V4		V5	V4.2
V6			V3	V5	V4	V4.1
V7						V4.4

Sequences of the genes which are underlined have been published in the references mentioned above.

fragment) (Figure 1B, and Materials and methods) detects a unique band when digested with *Hin*dIII (0.6 and 4 kb respectively) (Figure 4B and C). These results indicate that both genes exist as a single copy in the genome and that there is no cross-hybridization between them or with the other TRGV genes; however, no subgroup was assigned to these genes since they are not functional.

V−J junction

In Table I are shown the rearranged TRG genes and transcripts sequenced so far with their germline counterparts. The transcripts which encode a γ polypeptide at the surface of the cells are underlined (Krangel *et al.*, 1987; Littman *et al.*, 1987; Tighe *et al.*, 1988). From our data, it seems likely that the HGP02 transcript (Yoshikai *et al.*, 1987) might be functional and encode a V9-JP-C1 polypeptide since the TRG^+ cells recognized by the anti-TiγA monoclonal antibody (Jitsukawa *et al.*, 1987) have been shown to express a V9-JP rearranged gene (Triebel *et al.*, 1988). Note that a cluster of four or five nucleotides present either in the variable region (GGAC, GAGG) or in the J segment (CAAG, AAT, ATTAT) may be found as part of the N region as shown in λS12, λK20, λF1, λS13 or HGP08 respectively (Table I). This might be explained either by the insertion or deletion of nucleotides which 'isolate' these clusters from their V or J gene segments or by the duplication of these clusters as shown in clone pTγ15 (Quertermous *et al.*, 1986) (Table I) followed, or not, by a deletion of the corresponding nucleotides in the V or J segments. The N regions in Table I show no evidence for the existence of D segments in the human TRG locus, indeed they seem to be the result of a random insertion or deletion of nucleotides at the V−J junction, occasionally with a duplication of a few nucleotides belonging to either the V or J segments as discussed above. It should be noted that, in the N region of mouse V−J rearrangements, an AT dinucleotide is frequently found which is thought to be the result of a duplication of the AT dinucleotide present at the 5′ end of the J segment (Traunecker *et al.*, 1986).

Homology between the human and the mouse Vγ genes

Amino acid sequences of the human and mouse Vγ genes are aligned in Figure 7 and percentage of homology between the different Vγ genes is shown in Table III. Between the human VγI-subgroup genes, the homology ranges from 62 to 93%, the lowest values being associated with the pseudogene V6 (62−69%), whereas the highest homology is

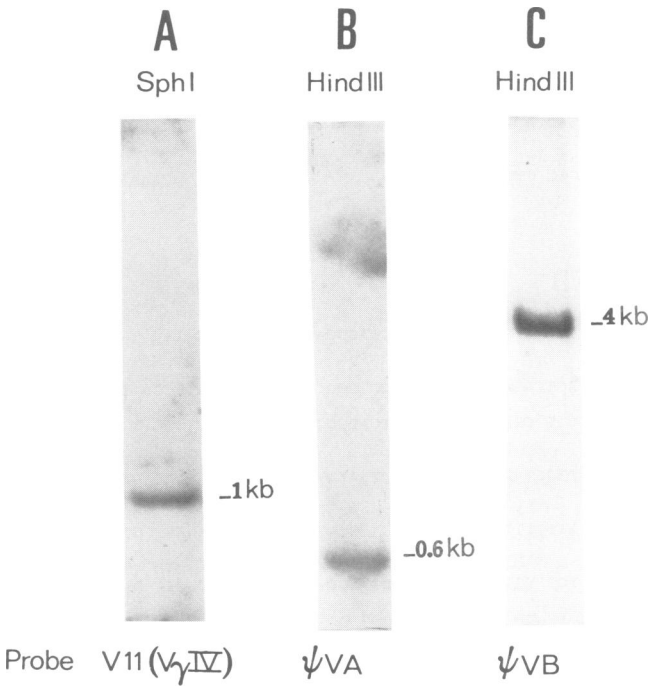

Fig. 4. Southern filter hybridization of genomic DNA digested (**A**) with *Sph*I and hybridized to the VγIV (V11) probe, (**B**) with *Hin*dIII and hybridized to the ψVA probe, (**C**) with *Hin*dIII and hybridized to the ψVB probe.

observed between V2 and V4 (93%) (Table IIIA). These genes result from a recent gene duplication as also shown by the similarity of the restriction maps in the region encompassing them (Lefranc *et al.*, 1986c) and by the identical size of the rearranged V2 and V4 bands obtained with *Hin*dIII and *Eco*RI digests (Forster *et al.*, 1987). The homology between the human Vγ subgroups ranges from 20 to 44% (Table IIIB), this last value being found between V10 and V11, respectively single members of the VγIII and VγIV subgroups, located most in 3′ in the human locus (Figure 1A).

Recently Pelkonen *et al.* (1987) noticed the relatively high homology between the mouse V7 gene and the human VγI subgroup genes, more particularly the V8 gene (49%) (Table IIID), homology which is much higher than that observed between the mouse V7 gene and the other mouse Vγ genes (17−23%) (Table IIIC). Interestingly, our data show that the human V11 gene is 52% homologous to the mouse V4 gene and 51% homologous to the mouse V2 and V3 genes

A

ψVA

```
 M   V   G   C   P   A   F   L   W   V   I   L   F
ATG GTG GGC TGC CCT GCC TTC CTT TGG GTC ATC CTG TTT    39

 P
CCA GGTCAGTCTCCTACCTAAGGACCTGTGTAGAGGTGGCGAGGCTTTGGG  90
    ↑
                                    A   G   W   W   L
ACATCCCTACCTCACTGCTCTCTCCCTGACAG CT  GGC TGG TGG CTC  136
                                    ↑
 I   R   P   E   Q   L   A   H   V   L   G   H   *
ATC AGG CCG GAG CAG CTG GCC CAT GTC CTG GGG CAC TAG    175

 G   S   L   V   I   L   Q   C   V   V   R   T   R
GGA AGC TTG GTC ATC CTG CAG TGC GTG GTC CGC ACC AGG    214

 I   S   Y   T   H   W   Y   Q   Q   K   G   Q   V
ATC AGC TAC ACC CAC TGG TAC CAG CAG AAG GGC CAG GTC    253

 P   E   A   L   H   Q   L   A   M   S   K   L   D
CCT GAG GCA CTC CAC CAG CTG GCC ATG TCC AAG TTG GAT    292

 V   Q   W   D   S   I   L   K   A   D   K   I   I
GTG CAG TGG GAT TCC ATC CTG AAA GCA GAT AAA ATC ATA    331

 A   K   D   G   S   S   S   I   L   A   V   L   K
GCC AAG GAT GGC AGC AGC TCT ATC TTG GCA GTA CTG AAG    370

 L   E   T   G   I   E   G   M   N   Y   C   T   T
TTG GAG ACA GGC ATC GAG GGC ATG AAC TAC TGC ACA ACC    409

 W   A   L
TGG GCC CTG CGCAGCCTTGCATGCTGCCCCAGCCCTACACAAAAGGAC   457
```

B

ψVB

```
 M   *   S   A   L   L   G   L   P   K   Y   C   D
ATG TGA TCT GCC CTC CTT GGC CTC CCA AAG TAC TGT GAT    39

 Y
TAC AGTCATGAGCCACCACACCTGGCCAGATGAAAGTGTTTCATTATCATT  90
    ↑
TCCTCTTTCTATGTCCCAGGTTCATGGGGATTTCTCTTTTTTGTTTGCAGTT 142

                I   K   F   K   A   I   K   N   V   N
GAACTTGGTAAG TT AAA TTT AAA GCA ATA AAA AAA GTC AAC    183
                ↑
 Y   I   F   V   N   R   A   T   D   K   S   V   *
TAC ATT TTT GTC AAC AGA GCA ACA GAT AAA AGT GTC TAG    222

 V   S   C   V   V   S   T   E   D   F   V   N   I
GTA TCT TGT GTG GTG TCC ACT GAA GAC TTT GTA AAT ATA    261

 V   I   H   W   Y   Q   Q   K   L   N   L   N   Y
GTT ATA CAC TGG TAC CAG CAG AAA CTG AAT CTA AAT TAT    300

 E   P   S   D   N   Y   H   F   N   N   K   P   Q
GAG CCA TCT GAC AAC TAT CAC TTC AAT AAC AAG CCC CAG    339

 F       V   C   L   G   E   K   N   K   K   L   E
TTC -GA GTT TGC TTA GGT GAG AAA AAC AAG AAA CTT GAG    378

 A   R   T   N   F   Q   M   S   T   S   V   F   T
GCA AGA ACA AAT TTT CAA ATG TCT ACT TCA GTC TTT ACC    417

 I   N   F   I   G   K   E   D   E   A   I   Y   Y
ATA AAC TTC ATA GGA AAG GAA GAT GAG GCC ATT TAC TAC    456

 C   T   A   *   D
TGC ACT GCT TAG GAC CCACAGCATCAGTGCCACACTGTCCCACACA   502

ACAACCTCT                                             511
```

Fig. 5. Nucleotide sequence and deduced amino acid sequence of the pseudogenes VA (**A**) and VB (**B**). Cysteines probably involved in the intra-chain disulfide bonds are circled. Heptamer and nonamer sequences are underlined. The dotted line in the VA sequence indicates a *Hin*dIII site whereas the underlined GGTACC nucleotides correspond to the *Kpn*I site found in the VA and VB genes. Splicing sites are shown by arrows. Asterisks indicate stop codons. Insertion of two nucleotides at codon 64 results in a frameshift indicated by a dash.

(and slightly less to V1 which belongs to the same subgroup). In both cases, the homology between the human and mouse genes is higher than that found between the most closely related human subgroup genes, VγIII and VγIV (44%), the conserved regions being mainly located in the framework

Table III. Percentage of homology between the amino acid sequences of the human and mouse Vγ genes

	Human						
A. Human	V2	V3	V4	V5P	V6	V7	V8
V1	75	76	79	72	62	74	72
V2		76	93	69	68	72	78
V3			77	87	67	76	75
V4				67	63	72	79
V5P					65	74	72
V6						66	69
V7							78

	Human				
	VA	VII	VIII	VB	VIV
B. Human	VA	V9	V10	VB	V11
V8(VI)	22	26	23	21	25
VA(VA)		20	20	21	21
V9(VII)			33	21	32
V10(VIII)				31	44
VB(VB)					32

	Mouse					
C. Mouse	V2	V3	V4	V5	V6	V7
V1	88	90	48	22	31	21
V2		94	47	22	31	22
V3			50	23	30	23
V4				26	24	22
V5					30	19
V6						17

	Mouse						
D. Human	V1	V2	V3	V4	V5	V6	V7
V8	23	26	24	21	25	21	49
VA	29	26	29	23	22	15	20
V9	30	33	34	31	27	29	24
V10	38	40	43	43	27	27	27
VB	38	40	40	34	22	23	17
V11	46	51	51	52	19	27	23

Homology between (**A**) the human VγI subgroup genes; (**B**) the human TRGV genes of the different subgroups including VA and VB; (**C**) the mouse Vγ genes; (**D**) the mouse Vγ genes and the human TRGV genes of the different subgroups. In (**B**) and (**D**) the human V8 gene was chosen as a representative of the VγI subgroup as being the most 3′ gene in the Vγ I subgroup and therefore the nearest of the other subgroup genes (Forster *et al.*, 1987). Moreover this gene is, among the VγI subgroup genes, the most homologous to the mouse V7 (Pelkonen *et al.*, 1987). The mouse V1, V2 and V3 belong to the same subgroup. The other mouse V genes are single members of different subgroups.

Fig. 6. Schematic representation of the mouse TRG locus. For detailed maps, see Hayday *et al.* (1985); Garman *et al.* (1986); Traunecker *et al.* (1986); Pelkonen *et al.* (1987). Due to the complexity of the various nomenclatures for the mouse locus, we had to use the numbers 1–7 for the mouse V genes according to Table II.

regions of the genes (Figure 7). Note that the other genes located in the 3′ region of the human TRGV locus (VA, V9, V10, VB) share more conserved amino acids with the mouse genes than with the human subgroup VγI genes.

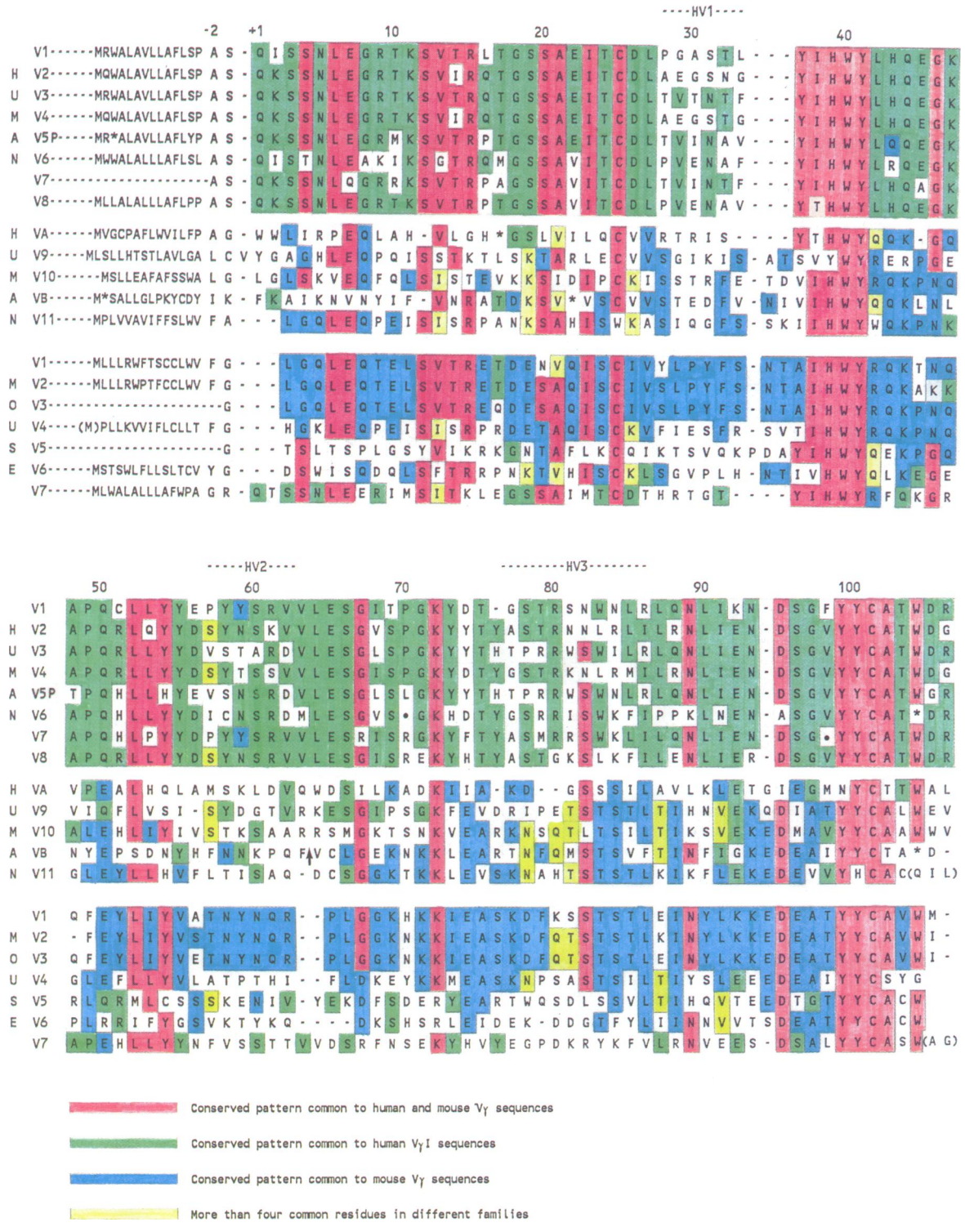

Fig. 7. Amino acid sequence comparison of the human and mouse variable genes in the T-cell rearranging γ (TRG) locus. Numbers above the sequences and hypervariable regions (HV1, HV2, HV3) refer to the human VγI subgroups (Lefranc *et al.*, 1986c). In position −2, for all the Vγ genes, are aligned the amino acid residues which result from the splicing between the leader and the variable region. (M) in the mouse V4 leader sequence indicates that the insertion of a nucleotide G would create a methionine in that position; this nucleotide being deleted, a ATG codon located upstream is used, which results in a longer leader (for its sequence, see Garman *et al.*, 1986). Gaps (indicated by dashes) have been introduced into the genes to optimize sequence homology. Dots in human V6, V7 indicate a single base pair deletion and an arrow in VB corresponds to a 2 bp insertion resulting in a frameshift; asterisks indicate stop codons (human V5P, V6, VA, VB). Amino acid residues which might belong to the N region are in brackets. At position 31 in human V9, can be found, either Thr (T) as in λA6 (Forster *et al.*, 1987) or Lys (K) as shown in that table (λK20, Lefranc *et al.*, 1986c). These substitutions correspond respectively to allele A1 and allele A2 of TRGV9. References for the human Vγ gene sequences; TRGV1, V2, V3, V4, V5P, V6, V7, V8, V9*A2 (Lefranc *et al.*, 1986a,c); V9*A1, V10 (Forster *et al.*, 1987); germline V9 and V10, VA, VB, V11 (this paper). References for the mouse Vγ genes; see Table II and Rupp *et al.* (1986).

The homology of the human V8 gene (Vγ I subgroup) to the mouse V7 on the one hand, and that of the human V11 gene (VγIV subgroup) to the mouse V4 on the other hand, suggest that these human and mouse variable γ genes derived from an ancestor gene which duplicated before the separation of the species.

Materials and methods

λ phage isolation and mapping

A partial genomic library was prepared from *Eco*RI- digested JM cell line DNA, in λ ongC-RI (Genofit). Rearranged TRG genes were isolated using a J1 probe [this probe, pH60, contains the 700 bp *Hin*dIII−*Eco*RI from M13H60 (Lefranc and Rabbitts, 1985; Lefranc *et al.*, 1986a) subcloned in pUC]. Phages were mapped by standard procedures and relevant subclones prepared in pUC or M13 vectors (Vieira and Messing, 1982).

Probes

VγII probe, a 700 bp *Pst*I−*Hin*dIII fragment isolated from λSHV7 (Forster *et al.*, 1987 and this paper) containing the germline V9 gene (Lefranc *et al.*, 1986a); VγIII-5′ probe, a 600 bp *Pst*I−*Eco*RI fragment isolated from λR12 (Forster *et al.*, 1987); VγIII-3′ probe, a 400 bp *Eco*RI−*Bam*HI fragment isolated from λJM15; VγIV probe, a 270 bp *Sph*I−*Rsa*I fragment containing the 5′ region of the V11 gene; ψVA probe, a 600 bp *Hin*dIII fragment containing the 3′ region of VA and its 3′ flanking region (Forster *et al.*, 1987); ψVB probe, a 500 bp *Kpn*I−*Hin*dIII fragment containing the 3′ region of VB and its 3′ flanking region.

Hybridization and sequencing procedure

Southern filter hybridization was carried out with 10 μg genomic DNA using nick-translated probes (Rigby *et al.*, 1977). Conditions for hybridization, washing and monitoring were previously described (Lefranc *et al.*, 1986c). Nucleotide sequence analysis was carried out by dideoxy-chain termination procedure in M13 vectors (Sanger *et al.*, 1977; Bankier and Barrell, 1983) deploying exonuclease III−nuclease S1 methods (Guo *et al.*, 1983) or directed sequencing using known restrictions enzyme sites.

Acknowledgements

We are very grateful to Dr T.H.Rabbitts for providing us with the JM cell line, to A.Forster for his help in the sequencing of the ψVA gene and to Dr G.Lefranc for helpful discussion. We would like to thank Véronique Ramond for typing the manuscript. This research was supported by Centre National de la Recherche Scientifique, Association pour la Recherche sur le Cancer, Fondation pour la Recherche Médicale, Ligue Française contre le Cancer and Ministère de la Recherche et de l'Enseignement Supérieur.

References

Bankier,A.T. and Barrell,B.G. (1983) *Nucleic Acids Biochem.*, **B508**, 1−34.

Dialynas,D.P., Murre,C., Quertermous,T., Boss,J.M., Leiden,J.M., Seidman,J.G. and Strominger,J.L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2619−2623.

Forster,A., Huck,S., Ghanem,N., Lefranc,M.-P. and Rabbitts,T.H. (1987) *EMBO J.*, **6**, 1945−1950.

Garman,R.D., Doherty,P.J. and Raulet,D.H. (1986) *Cell*, **45**, 733−742.

Guo,L.H., Yang,R.C.A. and Wu,R. (1983) *Nucleic Acids Res.*, **11**, 5521−5540.

Hayday,A.C., Saito,H., Gillies,S.D., Kranz,D.M., Tanigawa,G., Eisen,H.N. and Tonegawa,S. (1985) *Cell*, **40**, 259−269.

Heilig,J.S. and Tonegawa,S. (1986) *Nature*, **322**, 836−840.

Huck,S. and Lefranc,M.-P. (1987) *FEBS Lett.*, **224**, 291−296.

Iwamoto,A., Rupp,F., Ohashi,P.S., Walker,C.L., Pircher,G., Joho,R., Hengartner,H. and Mak,T.W. (1986) *J. Exp. Med.*, **163**, 1203−1212.

Jitsukawa,S., Faure,F., Lipinski,M., Triebel,F. and Hercend,T. (1987) *J. Exp. Med.*, **166**, 1192−1197.

Krangel,M.S., Band,H., Hata,S., McLean,J. and Brenner,M.B. (1987) *Science*, **237**, 64−67.

Lefranc,M.-P. and Rabbitts,T.H. (1985) *Nature*, **316**, 464−466.

Lefranc,M.-P., Forster,A. and Rabbitts,T.H. (1986a) *Nature*, **319**, 420−422.

Lefranc,M.-P., Forster,A. and Rabbitts,T.H. (1986b) *Proc. Natl. Acad. Sci. USA*, **83**, 9596−9600.

Lefranc,M.-P., Forster,A., Baer,R., Stinson,M.A. and Rabbitts,T.H. (1986c) *Cell*, **45**, 237−246.

Lefranc,M.-P., Forster,A. and Rabbitts,T.H. (1987) In Kappler,J. and Davis,M. (ed.), *The T Cell Receptor*. Alan Liss, New York, Vol. 73, pp. 25−29.

Littman,D.R., Newton,M., Crommie,D., Ang,S.-L., Seidman,J.G., Gettner,S.N. and Weiss,A. (1987) *Nature*, **326**, 85−88.

Murre,C., Waldmann,R.A., Morton,C.C., Bongiovanni,K.F., Waldmann,T.A., Shows,T.B. and Seidman,J.G. (1985) *Nature*, **316**, 549−552.

Pelkonen,J., Traunecker,A. and Karjalainen,K. (1987) *EMBO J.*, **6**, 1941−1944.

Quertermous,T., Strauss,W., Murre,C., Dialynas,D.P., Strominger,J.L. and Seidman,J.G. (1986) *Nature*, **322**, 184−187.

Quertermous,T., Strauss,W., van Dongen,J.J.M. and Seidman,J.G. (1987) *J. Immunol.*, **138**, 2687−2690.

Rabbitts,T.H., Lefranc,M.-P., Stinson,M.A., Sims,J.E., Schroder,J., Steinmetz,M., Spurr,N.L., Solomon,E. and Goodfellow,P.N. (1985) *EMBO J.*, **4**, 1461−1465.

Raulet,D.H., Garman,R.D., Saito,H. and Tonegawa,S. (1985) *Nature*, **314**, 103−107.

Rigby,P.W.J., Dieckmann,M., Rhodes,C. and Berg,P. (1977) *J. Mol. Biol.*, **113**, 237−251.

Rudikoff,S. and Pumphrey,J.G. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7875−7878.

Rupp,F., Frech,G., Hengartner,H., Zinkernagel,R.M. and Joho,R. (1986) *Nature*, **321**, 876−878.

Saito,H., Kranz,D.M., Takagaki,Y., Hayday,A.C., Eisen,H.N. and Tonegawa,S. (1984) *Nature*, **309**, 757−762.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463−5467.

Tighe,L., Forster,A., Clark,D., Boylston,A., Lavenir,I. and Rabbitts,T.H. (1988) *Eur. J. Immunol.*, in press.

Traunecker,A., Oliveri,F., Allen,N. and Karjalainen,K. (1986) *EMBO J.*, **5**, 1589−1593.

Triebel,F., Faure,F., Graziani,M., Jitsukawa,S., Lefranc,M.-P. and Hercend,T. (1988) *J. Exp. Med.*, **167**, in press.

Vieira,J. and Messing,J. (1982) *Gene*, **19**, 259−268.

Yoshikai,Y., Toyonaga,B., Koga,Y., Kimura,N., Griesser,H. and Mak, T.W. (1987) *Eur. J. Immunol.*, **17**, 119−126.

Received on November 24, 1987