

SEQUENCE REGISTER

Thierry Roger · Jean Morisset · Michel Seman

Conservation of *Tcrg-V5* and limited allelic sequence polymorphism of the other *Tcrg-V* genes used by mouse tissue-specific gd-T lymphocytes

Received: 24 May 1995 / Revised: 12 September 1995

The mouse *Tcrg* locus comprises seven *Tcrg-V*, four *Tcrg-J*, and four *Tcrg-C* segments which generate only six major types of functional g chains, Vg7-, Vg4-, Vg6-, or Vg5-Jg1-Cg1, Vg2-Jg2-Cg2, and Vg1-Jg4-Cg4 (Brenner et al. 1988). A complete analysis of restriction fragment length polymorphism (RFLP) of the *Tcrg* locus in wild and inbred mice suggested its relative conservation compared to other loci of the immunoglobulin (*Ig*) gene family (Roger et al. 1993a). Three haplotypes have been characterized in laboratory mice: *gA*, *gB*, and *gC*, represented by BALB/c, DBA/2, and AKR prototypes. *Tcr-gA* and *-gC* haplotypes are highly related. By contrast, *Tcr-gB*, likely inherited from asian mouse subspecies, appeared very different by RFLP analysis. Yet only partial sequence data have been reported on *gA* and *gB Tcrg-V* genes (Roger and Seman 1992; Roger et al. 1993c; Kalataradi et al. 1994). Here, the complete sequence of all *Tcrg-V* genes of the two haplotypes is described.

Tcrg-V1, *-V2*, and *-V3* genes from *Tcr-gB* L_i mice were co-amplified by polymerase chain reaction PCR from T lymphocytes cDNA using leader and pan *Tcrg-C* primers, cloned in M13, and subsequently sequenced (Fig. 1). When compared with BALB/c, *Tcr-gB-V1* sequence exhibits 16 nucleotide differences leading to 10 amino acid changes (n° 3, 7, 12, 23, 27, 29, 32, 42, 50, and 62). *Tcr-gB-V2* diverges from that of BALB/c by 7 nucleotides and 6 amino acids substitutions (n° 16, 24, 60, 63, 74, and 80). Finally, *Tcrg-V3* genes exhibit 6 nucleotide changes leading to 6 amino acid replacements (n° 12, 27, 31, 32, 73, and 74). As expected from RFLP analysis, the above results indicate that *Tcrg-V1*, *-V2*, and *-V3* genes present a high allelic variation superior to that found among *Tcrb-V* alleles (Smith et al. 1986; Roger et al. 1993c). *Tcrg-V1*, *-V2*, and *-V3* sequences characterized in L_i mice are identical to AKR sequences for the corresponding regions described by Kalataradi and co-workers (1994).

Tcr-gB-V4, *-V5*, *-V6*, and *-V7* from L_i display fewer variations from their *Tcr-gA* counterparts. Both *Tcrg-V4* and *-V6* exhibit 4 nucleotide substitutions leading to 3 amino acid changes (n° 12, 57, and 78 for *Tcrg-V4* and n° 26, 44, and 56 for *Tcrg-V6*). The L_i *Tcrg-V7* sequence is identical to that isolated from the BW 51-47 cell line of AKR (*Tcr-gB*) origin (Pelkonen et al. 1987) and reveals two adjacent nucleotide permutations leading to a single amino acid change (n° 62) compared with *Tcrg-A*. Of interest, *Tcrg-V5* is entirely conserved between *Tcr-gA* and *-gB* haplotypes. This gene is the first rearranged during ontogeny and is expressed on epidermal dendritic T cells in association with Vd1. Moreover, this Vg5/Vd1 Tcr exhibits no junctional diversity, suggesting a specificity for a monomorphic ligand directly depending on the V sequence. The conserved structure of *Tcrg-V5* might reflect a selective pressure necessary to preserve recognition of this monomorphic ligand likely present on keratinocytes (Havran et al. 1991).

Altogether, *Tcrg-V* genes located at the 5' end of the *Tcrg* locus (*Tcrg-V7*, *-V4*, *-V6*, and *-V5*) are more conserved than those genes

located at the 3' end (*Tcrg-V1* and *-V2*). Interestingly, gd T cells expressing these 5' end genes are mostly found in epithelia. Their *Tcr* usually display limited diversity (Tonegawa et al. 1989). They may also react to some ligands independently of the *Ig*-equivalent CDR3 region (Ohmen et al. 1991; Kalataradi et al. 1994). Allelic variations of *Tcrg-V* sequences might thus have consequences on the gd T cell repertoire as recently suggested (Kalataradi et al. 1994; Roger et al. 1993b; Sperling et al. 1992).

Acknowledgments The authors thank C. Thénevin and C. Joulin for expert technical assistance. This work was supported by a grant from the DRED.

References

- Brenner, M. B., Strominger, J. L., and Krangel, M. S. The gd T cell receptor. *Adv Immunol* 43: 133–192, 1988
- Garman, R. D., Doherty, P. J., and Raulet, D. H. Diversity, rearrangement and expression of murine T-cell g genes. *Cell* 45: 733–745, 1986
- Hayday, A. C., Saito, H., Gillies, S. D., Kranz, D. M., Tanigawa, G., Eusen, H. N., and Tonegawa, S. Structure, organization, and somatic rearrangement of T cell gamma genes. *Cell* 40: 259–269, 1985
- Havran, W. I., Chien, Y., and Allison, J. P. P. Recognition of self antigens by skin-derived T cells with invariant gd antigen receptor. *Science* 252: 1430–1432, 1991
- Heilig, J. S. and Tonegawa, S. Diversity of murine gamma genes and expression in fetal and adult T lymphocytes. *Nature* 322: 836–840, 1986
- Kalataradi, H., Eyster, C. L., Fry, A., Vollmer, M. K., Fu, Y. X., Born, W. K., and O'Brien, R. L. Allelic differences in TCR-g-chains alter gd T cell antigen reactivity. *J Immunol* 153: 1455–1465, 1994
- Ohmen, J. D., Barnes, P. F., Uyemura, K., Lu, S., Grisso, C. L., and Modlin, R. L. The T cell receptors of human gd T cells reactive to *Mycobacterium tuberculosis* are encoded by specific V genes but diverse V-J junctions. *J Immunol* 147: 3353–3359, 1991
- Pelkonen, J., Trauneker, A., and Karjalainen, K. A new mouse Tcr Vg gene that shows remarkable evolutionary conservation. *EMBO J* 6: 1941–1944, 1987
- Roger, T. and Seman, M. Polymorphism of the *Tcrg-V1-V2* region in mice: identification of a new *Vg1* allele in DBA/2. *Immunogenetics* 36: 67–69, 1992
- Roger, T., Pépin, L. F., Jouvin-Marche, E., Cazenave, P. A., and Seman, M. New T cell receptor gamma haplotypes in wild mice and evidence for limited *Tcrg-V* gene polymorphism. *Immunogenetics* 37: 161–169, 1993a
- Roger, T., Miconnet, I., Bruley-Rosset, M., and Seman, M. Coding sequence polymorphism of *Tcrg-V1*, *-V2*, and *-V4* genes in mice bearing *Tcr-gA* and *-gC* haplotypes. *Immunogenetics* 39: 68–70, 1993b
- Roger, T., Pépin, L. F., Couderc, J., De Franco, M., and Seman, M. Co-selection of the rare T cell receptor-gB haplotype in mouse lines selected for low antibody responsiveness to red blood cells. *Eur J Immunol* 37: 287–290, 1993c
- Roger, T., Miconnet, I., Bruley-Rosset, M., and Seman, M. Rearrangement by chromosomal inversion in the T cell receptor gamma locus in a murine ab T cell clone. *Mol Immunol* 30: 1617–1620, 1993d
- Smith, L. R., Plaza, A., Singer, P. A., and Theophilopoulos, A. N. Coding sequence polymorphisms among Vb T cell receptor genes. *J Immunol* 144: 3234–3237, 1990
- Sperling, A., Cron, R., Decker, D., Stern, D., and Bluestone, J. Peripheral T cell receptor γδ variable gene repertoire maps to the T cell receptor loci and is influenced by positive selection. *J Immunol* 149: 3200–3207, 1992
- Tonegawa, S., Berns, A., Bonneville, M., Farr, A., Ishida, I., Ito, K., Itohara, S., Janeway, C. A. Jr., Kanagawa, O., Katshuki, M., Kubo, R., Lafaille, J., Mombaerts, P., Murphy, D., Nakanishi, N., Takagaki, Y., Van Kaer, L., and Verbeek, S. Diversity, development, ligands, and probable functions of gd T Cells. *Symp Quant Biol* 54: 31–44, 1989

The nucleotide sequence data reported in this paper have been submitted to the EMBL nucleotide sequence database and have been assigned the accession numbers Z48588–Z48594 and Z49051

T. Roger · J. Morisset · M. Seman (✉)
Laboratoire d'Immunodifférenciation, Tour 54, BP 7124, Université Denis Diderot, 2 place Jussieu, 75251 Paris cedex 05, France

<i>Tcrg-V1</i>	1	Gln Leu Lys Gln Thr Glu Val Ser Val Thr Arg Ala Thr Asp Glu Ser Ala Glu Ile Ser
<i>Tcr-gB</i>	CAG CTA AAG CAA ACT GAA <u>GAA</u> TCG GTC ACC AGA GCG ACA GAT GAG AGT GGG CAA ATA TCC	
<i>Tcr-gA</i>	CAG CTC GAG CAA ACT GAA TTA TCG GTC ACC AGA GCG ACA GAT GAG AGT GGG CAA ATA TCC	
	Glut Leu Glu	
21	Cys Ile Ala Ser Leu Pro Asp Phe Gly Asn Thr Glu Ile His Trp Tyr Arg Glu Lys Ala	
	TGT ATA GCA TCT CTT CCA <u>GAC</u> TTC GGC AAC ACA GAA ATA <u>CAC</u> TGG TAC CGG CAA AAA GCA	
	TGT ATA GTT TCT CTT CCA TAT TCC AAC ACA GCT ATA CAT TGG TAC CGG CAA AAA GCA	
	Val Tyr Ser Ala	
41	Lys Gin Phe Glu Tyr Leu Ile Tyr Val Gin Thr Asn Tyr Asn Glu Arg Pro Leu Glu	
	AAA CAG TTT GAG TAT CTA ATA TAT GTC <u>CAC</u> ACA AAC TAC AAT CAA CGA CCC TTA GGA GGG	
	AAA CAG TTT GAG TAT CTA ATA TAT GTC TCA ACA AAC TAC AAT CAA CGA CCC TTA GGA GGG	
	Lys Ser	
61	Lys His Lys Lys Ile Glu Ala Ser Lys Asp Phe Glu Thr Ser Thr Ser Thr Leu Lys Ile	
	Arg CAC AAA AAA ATT GAA GCA AGT AAA GAT TTT CAA ACT TCT ACC TCA ACC TTG AAA ATA	
	Arg AAC AAA AAA ATT GAA GCA AGT AAA GAT TTT CAA ACT TCT ACC TCA ACC TTG AAA ATA	
	Asn	
81	Asn Tyr Leu Lys Lys Glu Asp Glu Ala Thr Tyr Tyr Cys Ala Val Trp	
	ATT ACT TTG AAG AAA GAA GAT GAA GCC ACC TAC TCA TGT GCA GTC TTG AT	
	ATT ACT TTG AAG AAA GAA GAT GAA GCC ACC TAC TCA TGT GCA GTC TTG AT	
<i>Tcrg-V3</i>	1	Gln Leu Glu Gln Thr Glu Leu Ser Val Thr Arg Ala Thr Asp Glu Ser Ala Glu Ile Ser
<i>Tcr-gB</i>	CAG CTG GAG CAA ACT GAA TTA TCG GTC ACC AGA GCG ACA GAT GAG AGT GGG CAA ATA TCC	
<i>Tcr-gA</i>	CAG CTG GAG CAA ACT GAA TTA TCG GTC ACC AGA GCG ACA GAT GAG AGT GGG CAA ATA TCC	
	Glut	
21	Cys Ile Val Phe Leu Pro Tyr Phe Ser Asn Thr Ala Ile His Trp Tyr Arg Glu Lys Thr	
	TGT ATA GTT <u>TCT</u> CTT CCA TAT TTC TCC AAC ACA GCT ATA CAT TGG TAC CGG CAA AAA GCA	
	TGT ATA GTT TAT CCA TAT TTC TCC AAC ACA GCT ATA CAT TGG TAC CGG CAA AAA GCA	
	Tyr	
41	Asn Gin Phe Glu Tyr Leu Ile Tyr Val Ala Thr Asn Tyr Asn Glu Arg Pro Leu Glu	
	ATT CAA CAG TTT GAG TAT CTA ATA TAT GTC GCA ACA AAC TAC AAT CAA CGA CCC TTA GGA GGG	
	ATT CAA CAG TTT GAG TAT CTA ATA TAT GTC GCA ACA AAC TAC AAT CAA CGA CCC TTA GGA GGG	
	Asn AGC AAC AAA AAA ATT GAA GCA AGT AAA GAT TTT CAA ACT TCT ACC TCA ACC TTG AAA ATA	
	Asn AGC AAC AAA AAA ATT GAA GCA AGT AAA GAT TTT CAA ACT TCT ACC TCA ACC TTG AAA ATA	
	His	
61	Gly Lys Asn Lys Lys Ile Glu Ala Ser Lys Asp Phe Lys Thr Ser Thr Ser Thr Leu Lys	
	GGG AGG AAC AAA ATT GAA GCA AGT AAA GAT TTT <u>AAA</u> ACT TCT ACC TCA ACC TTG AAA ATA	
	GGG AGG AAC AAC AAA ATT GAA GCA AGT AAA GAT TTT CAA ACT TCT ACC TCA ACC TTG AAA ATA	
	Gly	
81	Ile Asn Tyr Leu Lys Lys Glu Asp Glu Ala Thr Tyr Tyr Cys Ala Val Trp Met	
	ATA ATT TAC TTG AAG AAA GAA GAT GAA GCC ACC TAC TAC TGT GCA GTC TTG ATG AG	
	ATA ATT TAC TTG AAG AAA GAA GAT GAA GCC ACC TAC TAC TGT GCA GTC TTG ATG AG	
<i>Tcrg-V4</i>	1	Lys Leu Glu Gln Pro Glu Ile Ser Ile Ser Arg Ala Arg Asp Glu Thr Ala Glu Ile Ser
<i>Tcr-gB</i>	CAA CCT GAA ATA TCA ATT TCC AGA <u>GCA</u> AGA GAT GAG ACT GCA CAA ATA TCC	
<i>Tcr-gA</i>	AGA TTG GAG CAA CCT GAA ATA TCA ATT TCC AGA CCA AGA GAT GAG ACT GCA CAA ATA TCC	
	Pro	
21	Cys Ile Val Phe Ile Glu Ser Phe Arg Ser Val Thr Ile His Trp Tyr Arg Glu Lys Pro	
	TGT AAA GTT TTC ATC GAA AGC TTT AGG AGT GTA ACC ATA CAT TGG TAC CGG CAG AAA CCA	
	TGT AAA GTT TTC ATC GAA AGC TTT AGG AGT GTA ACC ATA CAT TGG TAC CGG CAG AAA CCA	
41	Asn Glu Gly Leu Glu Phe Leu Leu Tyr Val Leu Ala Thr Pro Thr His Ile Phe Leu Asp	
	AAC CAA GGT TTA GAG TTT CTA TTA TAT GTC CTT GCA ACC CCT ACC CAT <u>AAT</u> TTC TTA GAT	
	ACG CAA GGT TTA GAG TTT CTA TTA TAT GTC CTT GCA ACC CCT ACC CAT GAT TTC TTA GAT	
	Val	
61	Lys Glu Tyr Lys Met Glu Ala Ser Lys Asn Pro Ser Ala Ser Thr Ser Thr Leu Thr	
	AGG AGG TAC AAG AAA ATT GAG GCA ACC TGT AAA CCT AGT GCT TCT ACA TCG <u>ACA</u> TIG ACA	
	AGG TAC AAG AAA ATT GAG GCA AGT AAA CCT AGT GCT TCT ACA TCG ATA TIG ACA	
	His	
81	Ile Tyr Ser Leu Glu Glu Asp Glu Ala Ile Tyr Tyr Cys Ser Tyr Gly TER	
	ATA ATT TCC TTG GAG GAA GAA GAC GAA GCT ATC TAC TAC TGT TCC TAC GGC TAA AG	
	ATA ATT TCC TTG GAG GAA GAC GAA GCT ATC TAC TAC TGT TCC TAC GGC TAA AG	
<i>Tcrg-V6</i>	1	Lys Leu Thr Ser Pro Leu Glu Ser Tyr Val Ile Lys Arg Lys Gly Asn Thr Ala Phe Leu Lys
<i>Tcr-gB</i>	TCA CCT CTG GGG TCA TAT GTC ATC AAG AGG AAA GGA ATT AGC GCT TTT CTC AAA	
<i>Tcr-gA</i>	CTC ACG TCA CCT CTG GGG TCA TAT GTC ATC AAG AGG AAA GGA ATT AGC GCT TTT CTC AAA	
21	Cys Glu Ile Lys Thr Gly Val Glu Lys Pro Asp Ala Tyr Ile His Trp Tyr Glu Glu Lys	
	TGT CAA ATA AAA ACA <u>GCG</u> GTT CAG AAC CCC GAT GCA TAC ATA CAC TGG TAC CAA GAG AGG	
	TGT CAA ATA AAA ACA AGT GTT CAG AAC CCC GAT GCA TAC ATA CAC TGG TAC CAA GAG AGG	
	Ser	
41	Pro Gly Glu His Leu Glu Arg Met Leu Cys Ser Ser Lys Glu Asn Ile Val Tyr Glu	
	CCA GGC CAG CAT CTC CAA AGA ATG CTG TGT AGT TCT TCA AAA GAA AAC ATT GTC TAT GAG	
	CCA GGC CAG CAT CTC CAA AGA ATG CTG TGT AGT TCT TCA AAA GAA AAC ATT GTC TAT GAG	
	Arg	
61	Lys Asp Phe Ser Asp Glu Arg Tyr Glu Ala Arg Thr Trp Glu Ser Asp Leu Ser Ser Val	
	AAA GAT TTT AGT GAC GAA AGA TAT GAG GCA AGG ACA TGG CAG AGT GAT TTG TCT TCA GTC	
	AAA GAT TTT AGT GAC GAA AGA TAT GAG GCA AGG ACA TGG CAG AGT GAT TTG TCT TCA GTC	
81	Leu Thr Ile His Glu Val Thr Glu Glu Asp Thr Glu Thr Tyr Tyr Cys Ala Cys Trp Asp	
	CTC ACC ATA CAC CAA GTG ACA GAA GAG GAC GGA ACT TAT TAC TGT GCA TGC TGG GAT	
	CTC ACC ATA CAC CAA GTG ACA GAA GAG GAC GGA ACT TAT TAC TGT GCA TGC TGG GAT	
<i>Tcrg-V7</i>	1	Asn Leu Glu Glu Arg Ile Met Ser Ile Thr Lys Leu Glu Glu Ser Ser Ala Ile Met Thr
<i>Tcr-gB</i>	AAC TTG GAA GAA AGA ATA ATG TCA ATC ACC AAG CTA GAG GGG TCC TCT TGT GCT ATA ATG ACT	
<i>Tcr-gA</i>	AAC TTG GAA GAA AGA ATA ATG TCA ATC ACC AAG CTA GAG GGG TCC TCT TGT GCT ATA ATG ACT	
21	Cys Asp Thr His Arg Thr Gly Thr Tyr Ile His Trp Tyr Arg Phe Glu Lys Gly Arg Ala	
	TGT GAT ACT CAC AGA ACA GCA GGC ACT TAC ATC CAC TGG TAC CGA TTC CAG AAA GGG AGG GGC	
	TGT GAT ACT CAC AGA ACA GCA GGC ACT TAC ATC CAC TGG TAC CGA TTC CAG AAA GGG AGG GGC	
41	Pro Glu His Leu Leu Tyr Tyr Asn Phe Val Ser Ser Thr Thr Val Val Asp Ser Arg Phe	
	CCA GAG CAC CCT CTC TAC TAT AAC TTC GTC AGT TCC ACA ACT GTG GTG GAT TCC AGA TTC	
	CCA GAG CAC CCT CTC TAC TAT AAC TTC GTC AGT TCC ACA ACT GTG GTG GAT TCC AGA TTC	
61	Asn Ser Glu Lys Tyr His Val Tyr Glu Gly Pro Asp Lys Arg Tyr Lys Phe Val Leu Arg	
	ATT <u>TCA</u> GAG AAA TAT CAT GTT TAT GAA GGC CGG GAC AGG AGG TAT AAA TTT GTG CTT CGG	
	ATT TTG GAG AAA TAT CAT GTT TAT GAA GGC CGG GAC AGG AGG TAT AAA TTT GTG CTT CGG	
	Leu	
81	Asn Val Glu Glu Ser Asp Ser Ala Leu Tyr Tyr Cys Ala Ser Trp Na	
	ATT GTG GAG GAG TCC GAT TCT GCT CTC TAC TGT GGC TCC TGG GCT	
	ATT GTG GAG GAG TCC GAT TCT GCT CTC TAC TGT GGC TCC TGG GCT	

Fig. 1 Nucleotide and deduced amino acid sequences of *Tcrg-V* genes. Codons presenting divergent nucleotide between *Tcr-gA* and *Tcr-gB* sequences are underlined. Only divergent amino acids are noticed under *Tcr-gA* genes. *Tcr-gB-V1*, -V2, -V3, -V4, -V5, -V6 and -V7 and *Tcr-gA-V7* sequences were determined in this work, whereas *Tcr-gA-V1*, -V2, -V3, -V4, -V5 and -V6 genes (M12832, M13337, M12831, X04397, M13336, X04315, M13338, M30171) were previously described (Hayday et al. 1985; Garman et al. 1986; Heilig and Tonegawa 1986; Kalafatadi et al. 1994). Numbering corresponds to the coding region in BALB/c (Heilig and Tonegawa 1986). To determine sequences, cDNA preparations from thymus, spleen, or skin were amplified by PCR as previously described (Roger et al. 1993d) using leader or 5' *Tcrg-V* gene primers (*Tcr-gV1*-V2-V3: 5'-TGCTGCTCTGGGTTGGGCT-3'; *Tcr-gV4*: 5'-CAACCT-GAAATATCAATTTC-3'; *Tcr-gV5* (5'*Tcrg-V5* intron): 5'-GATGTCTCTTCCTTGCTTT-3'; *Tcr-gV6*: 5'-TCACCTCTGGGGT-CATATGT-3'; *Tcr-gV7*: 5'-GCAGACAAACATCCTCCAAC-3') and a pan-*Tcrg-C* primer (5'-GGAAAGAACCTTTCAAGGA-3'). Products of the appropriate size were purified by electroelution, then cloned in M13mp19. DNA from at least three different positive plates was sequenced using Sequenase® (US Biochemicals, Cleveland, OH)