The genomic sequence of the bovine T cell receptor gamma TRG loci and localization of the TRGC5 cassette

Melanie L. Conrad [b,*], Melanie A. Mawer [b], Marie-Paule Lefranc [a],
Linda McKinnell [b], Joanne Whitehead [b], Scott K. Davis [b],
Rachel Pettman [b], Ben F. Koop [b]

[a] IMGT®, Laboratoire d'ImmunoGénétique Moléculaire LIGM, Institut de Génétique Moléculaire IGH,
UPR CNRS 1142, 141 rue de la Cardonille, 34396 Montpellier Cedex 5, France
[b] University of Victoria, Centre for Biomedical Research, Victoria, British Columbia, Canada V8W 3N5

Abstract

The bovine and ovine TRG genes have previously been shown to be located in two loci, TRG1 and TRG2, in contrast to human and mouse TRG genes that are located in a single locus. The bovine TRG1 and TRG2 loci are located on chromosome 4 at 4q3.1 and 4q1.5–2.2, respectively. The complete genomic organization of the two bovine loci is described: each locus comprises three cassettes, each one includes one or several variable genes (TRGV) and one or several joining genes (TRGJ) preceding a constant (TRGC) gene. The location of the TRGC5 cassette is conclusively described in 5′ of the TRG1 locus. Analysis of 17 TRGV belonging to 10 different subgroups, 8 TRGJ and 6 TRGC genes is conducted which comprises the most comprehensive list to date.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Bovine; T cell receptor; Gamma; Genomic; DNA sequencing

1. Introduction

Interaction of T cells with antigen is mediated by the T cell receptor (TR), which is a highly specific antigen receptor that can exist in two forms: αβ or γδ (Lefranc and Lefranc, 2001). More is known about the function and structure of the αβ TR while unfortunately, less is known about the γδ TR. Both αβ and γδ T cells function similarly with respect to cytolysis and cytokine release; however, many studies have described idiosyncratic properties of γδ T cells that are not fully explained. For instance, γδ T lymphocytes specifically localize to epithelial and mucosal sites such as the skin (Hein and Dudler, 1997; Tamaki et al., 2001) and small intestine (Hein and Mackay, 1991; Komano et al., 1995), and can recognize non-peptide antigens including phosphoantigens and lipopolysaccharides (Martino et al., 2005; Yamashita et al., 2005). Recent research also demonstrates that γδ T cells can act as antigen presenting cells (Brandes et al., 2005), effectively participating in both the innate and the adaptive immune system (Holtmeier and Kabelitz, 2005).

Cells displaying the γδ TR comprise less than 5% of the peripheral blood T cell populations in humans and mice (Janeway et al., 1988). In birds (Kubota et al., 1999; Sowder et al., 1988) and ruminants (Hein and

Dudler, 1993; Mackay and Hein, 1989) γδ T cells constitute up to 40% of circulating lymphocytes. In animals where γδ T cells form a large percentage of circulating lymphocytes, the function of elevated cell populations is not well understood. Organisms that display high peripheral blood γδ T cell percentages, such as bovids, are excellent candidates for further study of γδ T cells. The genomic sequencing of the bovine TR loci can provide fundamental information to understand the relationship between γδ T cells and the immune system.

The goals of our genomic analysis were to investigate the genomic organization of the bovine T cell receptor gamma (TRG) loci and to explore gene diversity within the bovine TRG genes. Recent expression studies by Herzig et al. combined with genomic sequence (GenBank/EMBL/DDBJ and IMGT/LIGM-DB AY644517 and AY644518) from this study determined a preliminary genomic organization for the bovine TRG loci (Herzig et al., 2006). Further analysis of the aforementioned sequences in this paper reveals the complete genomic organization including the location of the TRGC5 cassette, which was not previously known. Genomic characterization will provide insight into the function and evolution of the bovine TRG loci and may help to define the role of the γδ TR.

2. Methods

2.1. BAC isolation

Scott M. Taylor from Texas A & M University kindly provided bacterial artificial chromosomes (BACs). Six TRGC[+] BAC clones were isolated from a bovine genomic library screened with primers matching the first exon of the TRGC genes. BamHI and KpnI restriction digests and BAC end sequencing demonstrated that four clones appear to cover most of the bovine TRG loci (Fig. 1). BACs 3–3, 3–1, 1–1 and 1–4 were isolated via an alkaline lysis procedure using a Clonetech Nucleobond Maxi kit as per manufacturers' instructions. BACs 3–2 and 1–2 are redundant clones that are completely overlapped by other BACs; for this reason, 3–2 and 1–2 were not sequenced.

2.2. Shotgun library construction

Shotgun libraries of the aforementioned BACs were constructed via nebulization, for random fragmentation, and sequential treatments with Mung-bean nuclease, T4 polymerase, and Klenow polymerase,
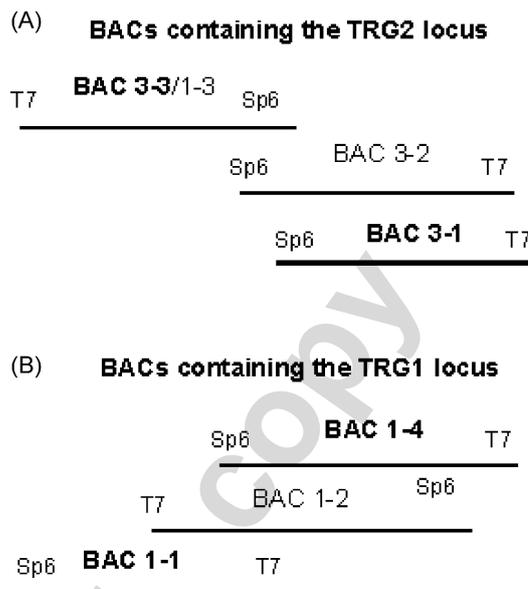
Fig. 1. TRGC[+] BAC map constructed from end sequencing and BamHI and KpnI restriction digests. (A) BACs containing TRG2 and (B) BACs containing TRG1. BACs sequenced from this map represented in bold (3–3, 3–1, 1–1 and 1–4).

using the methodology cited in Povinelli and Gibbs (1993). The subsequent blunt-ended DNA fragments were ligated with pUC19 vectors. Transformation of the pUC19 clones was conducted by electroporation. Sequencing templates were prepared using standard alkaline lysis in 96 well format and sequencing was conducted using an ABI 3800 sequencer.

Shotgun sequenced clones of each BAC were assembled into one contiguous sequence with seven-fold redundancy using the program Phred/Phrap/Consed (Ewing and Green, 1998).

2.3. Verification of correct assembly

Correct computer sequence assemblies were verified through comparison of in silico and BamHI and PvuII restriction digests. To confirm that the four TRGC[+] BACs correctly represent bovine genomic DNA, Southern blot analysis was performed using 300–500 bp probes that were constructed using PCR from each BAC clone.

2.4. Sequence analysis

The programs Align (Blattner and Schroeder, 1999a) and PipMaker (Schwartz et al., 2000) were used to create dot plot sequence comparisons. These programs were used both to verify sequence assembly, and to annotate genes. Annotation of V, J and C genes was accomplished by using nrBLAST (BCM Search Launcher) (Smith

et al., 1996) to retrieve TRG cDNA sequences from various species, creating a dot plot with Align and manually locating splice sites. Further annotation of TRGJ genes was accomplished by manually searching the sequence for the conserved motif FGXG using EditSeq (Blattner and Schroeder, 1999b), and then determining the recombination signal sequences (RSS) by searching for conserved DNA motif CACAGTG in proximity to the J gene.

The complete sequence of the bovine TRG loci was determined from the four BACs sequenced for this study plus a 35 kb section of DNA sequenced by the Baylor College of Medicine Human Genome Sequencing Centre (BCM HGSC) www.hgsc.bcm.tmc.edu/projects/bovine. The sequences can be located in GEDI (for GenBank/EMBL/DDBJ and IMGT/LIGM-DB): AY644517 (TRG1) and AY644518 (TRG2) and in NCBI Entrez Gene NW_937068.

3. Results and discussion

3.1. Bovine TRG genomic organization

The human and mouse TRG loci are located on a single locus in the chromosome (Lefranc, 1990; Lefranc et al., 1989; Lefranc and Rabbitts, 1989; Vernooij et al., 1993) (IMGT Repertoire, http://imgt.cines.fr). Research into ruminant genes using fluorescent in situ hybridization (FISH), however, has demonstrated that the TRG genes are localized at two different loci on the chromosome in both sheep and cattle (Massari et al., 1998). Recently, much analysis has been performed on the ovine TRG locus (Miccoli et al., 2005; Vaccarelli et al., 2005); the aim of this research is to expand TR knowledge in ruminants with analysis of TRG in bovids.

The genomic sequencing performed in this study confirms the results by Massari et al. that the bovine TRG genes are localized at two different loci on chromosome 4, named TRG1 and TRG2, respectively (Massari et al., 1998). Four BACs were sequenced for this study, shown in Fig. 1, plus an additional 35 kb of BCM HGSC data analyzed; together, these sequences comprise the bovine TRG genes. The 3–3 and 3–1 BACs form a 188 kb contig (AY644518); 103 kb of this sequence covers the bovine TRG2 locus located at 4q1.5–2.2. BAC 1–1 and 1–4 overlap to form 178 kb of the TRG1 locus (AY644517) located on chromosome 4q3.1. Situated 5′ to this locus is a gapped 35 kb contig (NW_937068) that contains TRG1 genes but does not overlap. Together, the two loci total approximately 316 kb.

The size and genomic organization of the bovine TRG loci is notable when compared with other species. The entire human TRG locus spans 160 kb (Lefranc et al., 1989) and the mouse TRG locus covers 205 kb (Vernooij et al., 1993). In bovine, there are two separate TRG loci and together they are substantially larger than those of either human or mouse, spanning over 316 kb.

The gene organization of the two bovine TRG loci consists of three tandemly repeated V, J, C cassettes at each locus. The sheep TRG loci display a similar organization, and in order to facilitate comparative studies between ruminants, identical gene names have been assigned to the TRGC genes by the IMGT Nomenclature Committee (Lefranc et al., 2003a, 2005a). A preliminary investigation of the bovine TRG organization was conducted, however, the TRGC5 cassette could not be definitively placed (Herzig et al., 2006). This research completes the description of the genomic organization of the bovine TRG1 locus: it identifies the previously unknown location and full sequence of the TRGV3-1, TRGV3-2, TRGV7-1, TRGV10-1 and TRGV4-1 genes in the TRGC5 cassette, and of the TRGV8-3, TRGV8-4, TRGV9-1, TRGV9-2 in the TRGC3 cassette, and for the first time conclusively describes the location of the TRGC5 cassette. The location of these genes within the bovine TRG loci may aid in finding the genomic location of TRG homologues in sheep.

It was initially postulated that the TRGC5 cluster might be located on the 3′ end of the TRG1 locus (Herzig et al., 2006). Analysis of the genomic sequence AY644517 indicates that the only gene located in the 72 kb of sequence 3′ of the TRGC4 cassette is unrelated to TRG (U6 snRNA-associated Sm-like protein LSM8); whereas analysis of the 5′ region of the TRG1 locus reveals three exons of the TRGC5 gene. As Herzig et al. grouped the TRGV3, TRGV7 and TRGV4 genes with the TRGC5 cassette, analysis indicates that the 35 kb BCM HGSC sequence (NW_937068) is located at the 5′ end of the TRG1 locus, illustrated in Fig. 2. The BCM HGSC sequence (NW_937068) contains two unresolved gaps, which may contain additional V and J genes, however these gaps do not interfere with the genes annotated in this study.

Analysis of the BAC clones that cover the TRG2 locus showed that they contained the TRGC1, TRGC2 and TRGC6 cassettes, interestingly with two rearranged V–J genes: TRGV5-2 rearranged to TRGJ2-1 (in the TRGC2 cassette), and TRGV6-2 rearranged to TRGJ6-1 (in the TRGC6 cassette). PCR analysis of this region will further clarify if any V or J genes were eliminated due to these TRG V–J rearrangements and if the
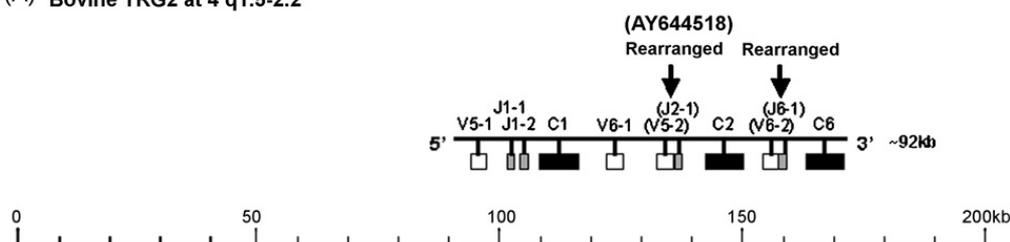
Fig. 2. Organization of the bovine TRG1 and TRG2 loci: (A) bovine TRG2 at 4q1.5–2.2 and (B) bovine TRG1 at 4q3.1. The most recent annotations are from NW_937068 (sequenced by www.hgsc.bcm.tmc.edu), AY644517 v.2 and AY644518 (sequenced for this study). Note that a preliminary TRG1 contig was constructed from a clone containing a TRGV8-2–TRGJ3-1 rearrangement (AY644517 v.1, 125,381 bp; gi:49424744, submitted on 2 June 2004). In the final contig described in this manuscript, the rearranged part of the sequence was replaced with a germline sequence originating from another clone and containing the TRGV8-2 and TRGJ3-1 in germline configuration and the TRGV2-1, TRGV9-1 and TRGV9-2 previously deleted by the rearrangement. The sequence resubmitted to GenBank in April 19, 2006 is AY644517 v.2, 243,699 bp; gi:92918655. AY644517 v.1 is still available in the GEDI archives. Parentheses indicate rearranged V–J genes in AY644518. All gene names approved by the IMGT Nomenclature Committee (http://imgt.cines.fr) (Lefranc et al., 2003a, 2005a).

presence of two V–J rearrangements in AY644518 results from joining independent clones during the contig construction.

Multi-species comparison demonstrates that ruminants maintain a larger pool of C genes than other organisms. The bovine TRB locus contains three functional TRBC genes (Conrad et al., 2002) whereas the ovine (Miccoli et al., 2005) and bovine TRG loci contain 5 and 6 functional TRGC genes, respectively. An illustration of human and mouse TRG organization compared with bovine and ovine is located in Fig. 3. The utilization of different TRGC genes is not fully understood, however, one hypothesis takes into account preferential gene usage and the developmentally regulated appearance of certain γδ T cell populations.

Preferential gene usage has been demonstrated extensively in the tissues of mice and humans; for example, 70–90% of peripheral blood γδ lymphocytes in human uses the TRGV9 and TRDV2 genes (Casorati et al., 1989; Haas et al., 1993). In addition to tissue type determining gene usage, TRG V, J and C preference has also been observed at different developmental stages in ruminants (Hein and Dudler, 1993). The existence of six functional TRGC genes in bovids indicates an enhanced

repertoire over humans and mice as well as increased possibility for specific developmental functions or tissue localization.

Preferential gene usage in tissues, coupled with developmentally related expression and functionally different γδ T cell populations in ruminants indicate that differential TRG expression is significant for the proper functioning of the bovine immune system.

3.2. Gene analysis

3.2.1. TRGV genes

The entire bovine TRG loci contain 17 TRGV genes and Fig. 4 represents the most comprehensive comparison of deduced amino acid sequence to date. The germline TRG V-REGION is characterized, as that of the other antigen receptor V genes, by three hypervariable regions or complementarity determining regions (CDR-IMGT) separated by less variable regions or framework regions (FR-IMGT) (Lefranc and Lefranc, 2001; Lefranc et al., 2003b). Standardized graphical two-dimensional representation of the bovine TRG V-REGION based on the IMGT unique numbering or IMGT Collier de Perles are available in IMGT Repertoire (http://imgt.cines.fr). Table 1 shows for the first time the complete list of the

Species Comparison – TRG Loci
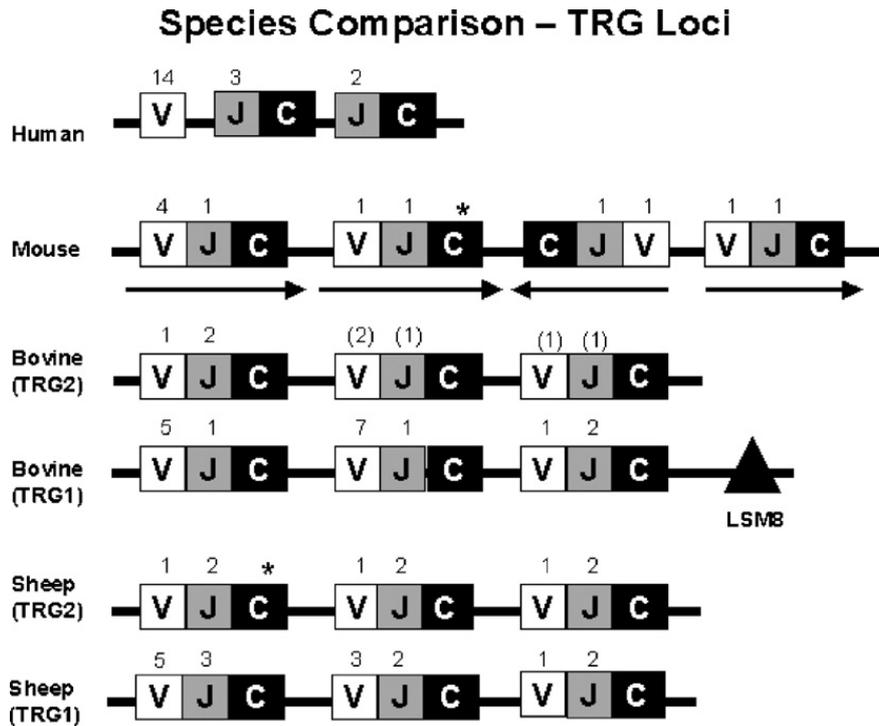
Fig. 3. Schematic representation of the genomic organization of the TRG genes in different species; the numbers above the squares represent the number of V or J genes in the cluster. The TRG genes are organized in one locus in human (Lefranc et al., 1989; Lefranc and Rabbitts, 1989) and mouse (Vernooij et al., 1993), whereas they are organized in two loci in ruminants: bovine TRG1 (NW_937068, AY644517) and TRG2 (AY644518), and ovine TRG1 and TRG2 (Miccoli et al., 2003, 2005; Vaccarelli et al., 2005). Asterisks (*) indicate pseudogenes and parenthesis () indicate genomically rearranged areas that may have more V and/or J genes. It is not excluded that the presence of two V–J rearrangements in AY644518 may result from joining independent clones during the contig construction. The triangle represents the U6 snRNA-associated Sm-like protein (LSM8) gene located 3′ of the bovine TRG1 locus.

CDR-IMGT of the germline bovine TRGV genes, with the particularities of each TRGV gene in term of gaps, and for TRGV2-1, of additional amino acids, according to the IMGT unique numbering for V-REGION (Lefranc et al., 2003b).

At the nucleotide level, the bovine TRGV8 subgroup genes have a percentage of identity of 94% with the *O. aries* V gamma2.3 gene (AY348329), whereas the bovine TRGV9 subgroup genes have a similar percentage of identity (93%) with the *O. aries* Vgamma2.4 gene
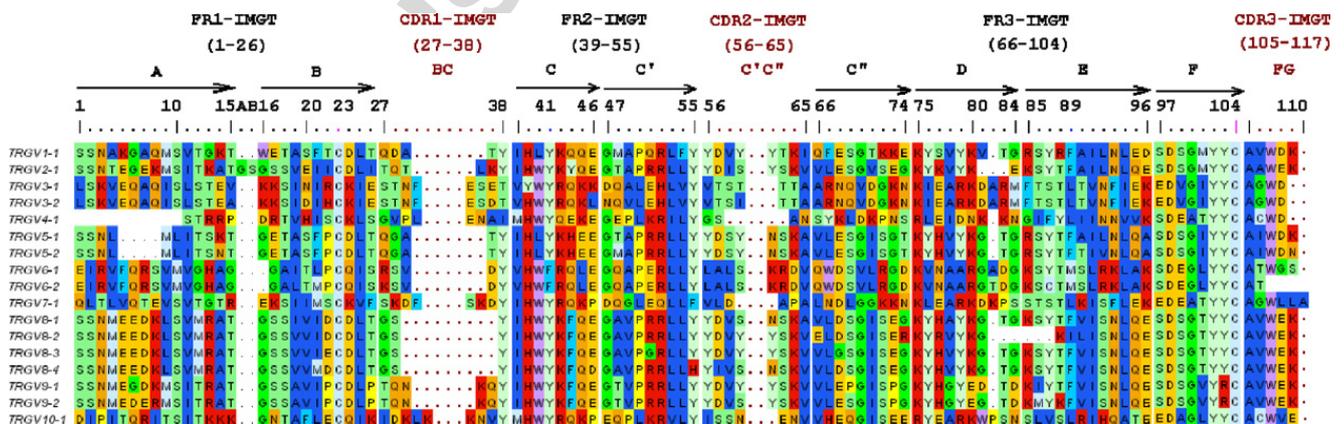
Fig. 4. Alignment of the deduced amino acid sequence of the 17 TRGV genes from the bovine TRG loci according to amino acid physicochemical characteristics. Except for TRGV4-1 that is from a cDNA sequence (DG179592) (see text), the other sequences are translation of the germline sequences in NW_937068, AY644517 and AY644518. Blank indicates "partial" (N-terminal end of TRGV4-1, and C-terminal end of TRGV6-2 owing to its rearrangement). C-terminal end of TRGV4-1 is from NW_937068. The initial alignment with ClustalW (Thompson et al., 1994) was refined with IMGT/ V-QUEST (Giudicelli et al., 2004) and using the rules of the IMGT unique numbering for V-REGION (Lefranc et al., 2003b). Viewed with JalView (Clamp et al., 2004). The color coding is according to the IMGT amino acid physicochemical characteristics classes (Pommié et al., 2004) (Appendix 1).

Table 1
CDR-IMGT lengths of the germline bovine TRGV genes

Bovine TRGV subgroups	Bovine TRGV genes	CDR-IMGT lengths	FR-IMGT gaps	FR-IMGT additional amino acids
TRGV1	TRGV1-1	[5.8.5]	1 (82)	
TRGV2	TRGV2-1	[6.8.5]	3 (81, 82, 83)	2 (15A, 15B)
TRGV3	TRGV3-1	[8.7.4]	0	
	TRGV3-2	[8.7.4]	0	
TRGV4	TRGV4-1	[8.4.4]	1 (82)	
TRGV5	TRGV5-1	[5.8.5]	5 (5–8, 82)	
	(TRGV5-2)	[5.8.5]	5 (5–8, 82)	
TRGV6	TRGV6-1	[5.8.5]	1 (16)	
	(TRGV6-2)	[5.8.x]	1 (16)	
TRGV7	TRGV7-1	[8.6.6]	0	
TRGV8	TRGV8-1	[3.8.5]	1 (82)	
	TRGV8-2	[3.8.5]	6 (82–87)	
	TRGV8-3	[3.8.5]	1 (82)	
	TRGV8-4	[3.8.5]	1 (82)	
TRGV9	TRGV9-1	[6.8.5]	1 (82)	
	TRGV9-2	[6.8.5]	1 (82)	
TRGV10	TRGV10-1	[9.7.5]	0	

TRGV5-2 and TRGV6-2 are in parentheses which indicates that these genes are involved in a genomic rearrangement. The CDR-IMGT lengths are according to the IMGT unique numbering for V-REGION (Lefranc et al., 2003b). The lengths of the CDR1-IMGT, CDR2-IMGT and CDR3-IMGT are shown between brackets with lengths separated with dots. For example, the TRGV1-1 gene [5.8.5] has a CDR1-IMGT of five amino acids (at positions 27–29, 37, 38 in the IMGT Collier de Perles), a CDR2-IMGT of eight amino acids (at positions 56–59 and 62–65) and a CDR3-IMGT of five amino acids (at positions 105–109). FR-IMGT gaps are shown with a first number indicating the number of gaps and between parentheses the positions of the gaps. Similarly the two FR-IMGT additional amino acids of TRGV2-1 are shown with their positions 15A and 15B.

(AY348331), thus defining the TRGV8 and TRGV9 subgroups in sheep. TRGV4-1 in NW_937068 shows three deletions of one nucleotide at codons 48 (g), 77 (g) and 91 (a) and three substitutions at codons 39 (a > g), 52 (g > a) and 84 (t > c) (positions according to the IMGT unique numbering for V-REGION (Lefranc et al., 2003b)), compared to a recently identified cDNA containing TRGV4-1 (DQ179592) (Herzig et al., 2006). These discrepancies may be due to sequencing errors in the genomic sequences and/or to polymorphisms, although the possibility of recently duplicated genes cannot be excluded. The cDNA, although partial, was included in the alignment (Fig. 4), as TRGV4-1 displays interesting homology with the mouse TRGV5 gene and shares the same germline CDR-IMGT lengths [8.4.4], as shown by IMGT/V-QUEST analysis (Giudicelli et al., 2004).
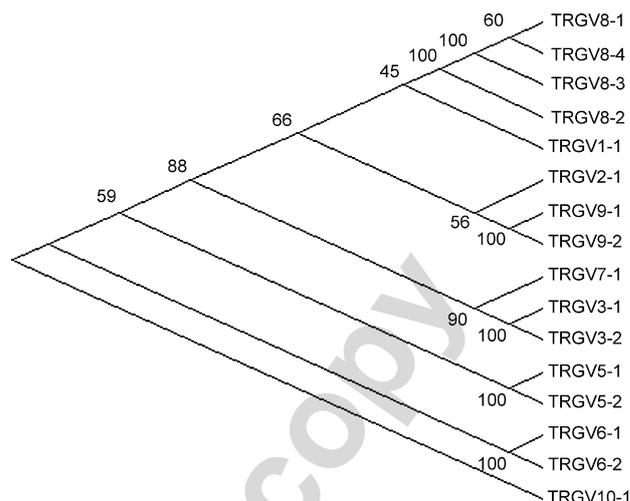
Fig. 5. Phylogenetic analysis of the 16 bovine TRGV genes. Maximum parsimony tree constructed in Mega 3, bootstrap 500. TRGV4-1 is not included in the tree as there is only partial sequence available for this gene.

All bovine TRGV gene subgroups are expressed in either the bovine system or have homologues that are expressed in the ovine immune system with the exception of the TRGV10 subgroup, represented by a single gene, TRGV10-1 (Herzig et al., 2006; Miccoli et al., 2005). In contrast to the ovine TRGV10-1 (VgammaP, AY362775), which is a pseudo-gene owing to a nucleotide deletion (g at codon 31 (nt 93), according to the IMGT unique numbering for V-REGION (Lefranc et al., 2003b) and to IMGT-V-QUEST (Giudicelli et al., 2004) and a defective splicing site of the leader exon L-PART1, the bovine TRGV10-1 has an open reading frame and a L-PART1 DONOR-SPLICE in the expected splicing frame 1 (IMGT Aide-mémoire, http://imgt.ci-nes.fr). It would be of interest to determine if the bovine TRGV10-1 gene contributes to the expressed TRG repertoire as it has a highly divergent predicted amino acid sequence, shown in Fig. 4 and the nucleotide sequence of TRGV10-1 being represented as an outgroup in the maximum parsimony tree illustrated in Fig. 5. For extensive analysis of the bovine TRGV regions please see Herzig et al. (2006).

TRGV–TRGJ rearrangements depend, as those of the other antigen receptor V, D and J genes, on specific recombination signal sequences (RS) (Lefranc and Lefranc, 2001). RSS consist of conserved palindromic heptamer (CACAGTG) and nonamer sequences (AATTTTTGT or ACAAAAATT) separated by either 12 or 23 base pair (bp) non-conserved spacers (Lefranc and Lefranc, 2001; Sakano et al., 1979) (IMGT Repertoire, http://imgt.cines.fr). TRG V-RS are comprised of a conserved V-HEPTAMER, a 23 bp V-SPACER and a conserved V-NONAMER. Alignment
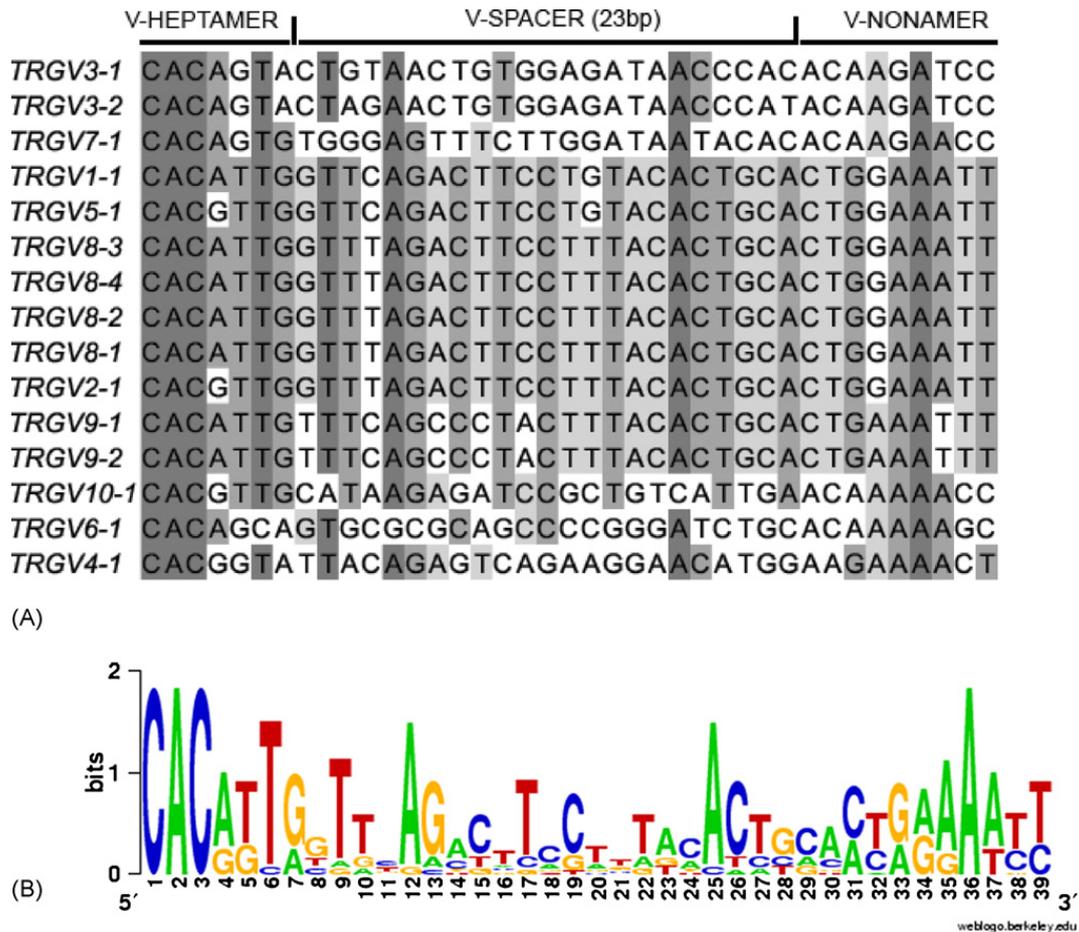
Fig. 6. (A) A comparison of the recombination signal sequences for 15 bovine TRGV genes. TRGV5-2 and TRGV6-2 were not included due to their rearrangements to TRGJ2-1 and to TRG6-1, respectively, in AY644518 (TRG2 locus). Constructed with ClustalW (Thompson et al., 1994), viewed with JalView (Clamp et al., 2004). (B) Graphical representation of the nucleotide frequency at each position of the TRG V-RS. Overall height of the nucleotide stack is equal to the frequency of that nucleotide at a certain position. Produced in WebLogo (Crooks et al., 2004), using IMGT color menu (IMGT Scientific chart, http://imgt.cines.fr).

of the bovine TRG V-RS demonstrates that the first three nucleotides (CAC) in the V-HEPTAMER are highly conserved throughout all V-RS as well as the sixth nucleotide (A) in the V-NONAMER, Fig. 6.

3.2.2. TRGJ genes

In expressed antigen receptors, the variable domains correspond to V–J-REGION and V–D–J-REGION, resulting from V–J or V–D–J gene rearrrangement (Lefranc and Lefranc, 2001). The CDR3-IMGT of rearranged sequences corresponds to the V–J or V–D–J JUNCTION. In αβ TR, which bind to peptide/MHC (pMHC), CDR1-IMGT and CDR2-IMGT are mainly involved in TR/pMHC complex stabilization, but also play a minor role in peptide recognition (Kaas and Lefranc, 2005). CDR3-IMGT is the main site of peptide identification due to its high variability (Chlewicki et al., 2005; Garcia et al., 1999). The CDR-IMGT of the domains of the γδ TR show similar structural features although the recognized antigens are non-peptide

antigens presented by non-classical MHC (Adams et al., 2005; Kaas et al., 2004).

The eight bovine TRGJ genes, which range in size from 15 to 19 amino acids, have varying sequence conservation, shown in Fig. 7. Interestingly, TRGJ2-1 and TRGJ6-1 are rearranged in AY644518. By comparison to TRGJ1-1, the TRGJ2-1 amino acid sequence in Fig. 7 is complete and corresponds to the germline sequence; in contrast the two amino acids at the beginning of TRGJ6-1 may result from the N-diversity (Lefranc and Lefranc, 2001) and the N-terminal end of TRGJ6-1 needs to be confirmed with a germline sequence. Highlighting the alignment in JalView with the IMGT amino acid classes (Pommié et al., 2004) indicates that although the sequence is not highly conserved, the physicochemical characteristics between amino acids are conserved. The phenylalanine residue (position 118, according to the IMGT unique numbering for V-DOMAIN (Lefranc et al., 2003b) is completely conserved as well as three hydrophobic
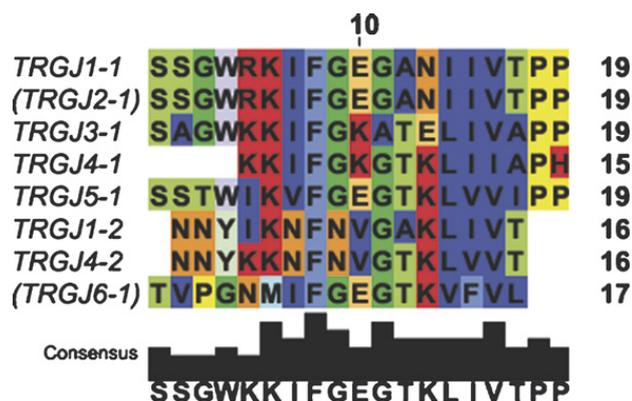
Fig. 7. Amino acid physicochemical comparison of eight bovine TRG J-REGIONs. The corresponding nucleotide sequences end with a canonical DONOR-SPLICE (splicing frame 1) (IMGT Aide-mémoire, http://imgt.cines.fr). Parentheses () indicate that TRGJ2-1 and TRGJ6-1 are rearranged. Constructed with ClustalW (Thompson et al., 1994), viewed with JalView (Clamp et al., 2004). The color coding is according to the 11 IMGT amino acid physicochemical characteristics classes (Pommié et al., 2004) (Appendix 1).

Table 2
Expression of the bovine TRGJ genes

Bovine TRGJ genes	mRNA Genbank accession	Species
TRGJ1-1	D90417[a]	Bovine
TRGJ1-2	D90413	Bovine
TRGJ2-1	D90417[a]	Bovine
TRGJ3-1	D90414	Bovine
TRGJ4-1	AY362778	Sheep
TRGJ4-2	X	X
TRGJ5-1	AY735449[b]	Bovine
TRGJ6-1	X	X

X indicates TRGJ genes that were not located in mRNA or EST sequence.

[a] As the TRGJ1-2 and TRGJ2-1 sequences are identical, the cDNA can be assigned to one or the other TRGJ.

[b] This clone has a TRGV7-1-TRGJ5-1 rearrangement spliced to TRGC5, in agreement with these three genes being located in the same cassette.

residues. Three of the eight TRGJ genes have changes in the "FGXG" J-REGION motif. Expression analysis indicates that TRGJ3-1 ("FGKA") and TRGJ1-2 ("FNVG") are both transcribed, therefore, both glycine (G) residues are not necessarily required for functionality. Expression data for the TRGJ sequences is located in Table 2.

A comparison of RS demonstrates the relative conservation of TRGJ nonamer and heptamer sequences. Alignment of six TRG J-RS indicates that these sequences have a classical J-RS structure with the J-HEPTAMER being more conserved than the

J-NONAMER, as demonstrated in Fig. 8. The J-NONAMER (gXTTXttg) demonstrates little conservation at nucleotides 2 and 5, whereas the J-HEPTAMER remains the most conserved with six nucleotides, CAXtGTG.

3.2.3. TRGC genes

A comparison of the six functional bovine TRGC genes, shown in Fig. 9, reveals conserved blocks of similarity throughout all regions with the exception of the connecting region, located structurally between the C-DOMAIN and the transmembrane region (Lefranc and Lefranc, 2001; Lefranc et al., 2005b). The length of
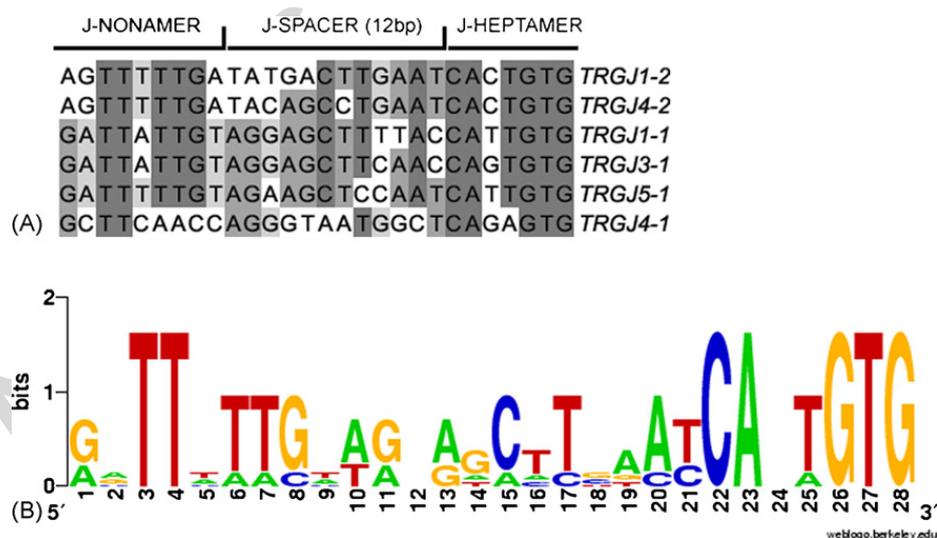
Fig. 8. (A) A comparison of the recombination signal sequences for six TRGJ genes. Constructed with ClustalW (Thompson et al., 1994), viewed with JalView (Clamp et al., 2004). TRGJ2-1 and TRGJ6-1 were not included as they are rearranged. (B) Graphical representation of the nucleotide frequency at each position of the TRG J-RS. Overall height of the nucleotide stack is equal to the frequency of that nucleotide at a certain position. Produced in WebLogo (Crooks et al., 2004), using IMGT color menu (IMGT Scientific chart, http://imgt.cines.fr).

Fig. 9. Alignment of the six functional bovine TRG C-REGIONs highlighted by blocks of physicochemical similarity. From positions 1 to 39 the TRGC5 sequence is from the cDNA AY735449. Alignment with ClustalW (Thompson et al., 1994), viewed with JalView (Clamp et al., 2004). The color coding is according to the 11 IMGT amino acid physicochemical characteristics classes (Pommié et al., 2004) (Appendix 1). IMGT protein display according to the IMGT unique numbering for C-DOMAIN (Lefranc et al., 2005b) is available in IMGT Repertoire (http://imgt.cines.fr).

the connecting region is dependent on a ruminant specific repeated motif TTEPP/TTKPP (with variations) that may confer greater flexibility in binding to antigen (Adams et al., 2005; Hein and Dudler, 1993; Kaas et al., 2004). This motif is observed once in TRGC3, TRGC4, and TRGC6, twice in TRGC2, and four times in TRGC1. Similar to the ovine TRGC5, bovine TRGC5 does not contain this motif (Miccoli et al., 2001). A correlation may exist between the function of this particular repeat

and the high percentages of peripheral blood γδ T cells seen in this species.

4. Conclusion

This research has expanded on the preliminary work by Herzig et al. (2006) to produce a comprehensive genomic map of the bovine TRG1 and TRG2 loci. The previously unknown order of bovine genes TRGV3-1,

Fig. A.1. Color menu for the 11 IMGT physicochemical characteristics classes.

TRGV3-2, TRGV7-1 TRGV10-1 and TRGV4-1 was identified, as well as localization of TRGV8-3, TRGV8-4, TRGV9-1, TRGV9-2 to the TRGC3 cassette. In addition to this information, for the first time the location of the TRGC5 cassette is conclusively described. The results of this work will aid in further annotation of the ovine TRG loci and provide researchers with detailed comparisons of the bovine V, J and C genes for utilization in further studies.

Acknowledgements

Appendix A

See Fig. A.1.

References

Adams, E.J., Chien, Y.H., Garcia, K.C., 2005. Structure of a gamma-delta T cell receptor in complex with the nonclassical MHC T22. Science 308, 227–231.

Blattner, F., Schroeder, J. 1999a. Align (https://www.dnastar.com/web/r13.php).

Blattner, F., Schroeder, J. 1999b. EditSeq (https://www.dnastar.com/web/r13.php).

Brandes, M., Willimann, K., Moser, B., 2005. Professional antigen-presentation function by human gammadelta T cells. Science 309, 264–268.

Casorati, G., De Libero, G., Lanzavecchia, A., Migone, N., 1989. Molecular analysis of human gamma/delta+ clones from thymus and peripheral blood. J. Exp. Med. 170, 1521–1535.

Chlewicki, L.K., Holler, P.D., Monti, B.C., Clutter, M.R., Kranz, D.M., 2005. High-affinity, peptide-specific T cell receptors can be generated by mutations in CDR1, CDR2 or CDR3. J. Mol. Biol. 346, 223–239.

Clamp, M., Cuff, J., Searle, S.M., Barton, G.J., 2004. The Jalview Java alignment editor. Bioinformatics 20, 426–427.

Conrad, M.L., Pettman, R., Whitehead, J., McKinnel, L., Davis, S.K., Koop, B.F., 2002. Genomic sequencing of the bovine T cell receptor beta locus. Vet. Immunol. Immunopathol. 87, 439–441.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–194.

Garcia, K.C., Teyton, L., Wilson, I.A., 1999. Structural basis of T cell recognition. Annu. Rev. Immunol. 17, 369–397.

Giudicelli, V., Chaume, D., Lefranc, M.P., 2004. IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V–J and V–D–J rearrangement analysis. Nucleic Acids Res. 32, W435–W440.

Haas, W., Pereira, P., Tonegawa, S., 1993. Gamma/delta cells. Annu. Rev. Immunol. 11, 637–685.

Hein, W.R., Dudler, L., 1993. Divergent evolution of T cell repertoires: extensive diversity and developmentally regulated expression of the sheep gamma delta T cell receptor. EMBO J. 12, 715–724.

Hein, W.R., Dudler, L., 1997. TCR gamma delta+ cells are prominent in normal bovine skin and express a diverse repertoire of antigen receptors. Immunology 91, 58–64.

Hein, W.R., Mackay, C.R., 1991. Prominence of gamma delta T cells in the ruminant immune system. Immunol. Today 12, 30–34.

Herzig, C., Blumerman, S., Lefranc, M.-P., Baldwin, C., 2006. Bovine T cell receptor gamma variable and constant genes: combinatorial usage by circulating gammadelta T cells. Immunogenetics 58, 138–151.

Holtmeier, W., Kabelitz, D., 2005. Gammadelta T cells link innate and adaptive immune responses. Chem. Immunol. Allergy 86, 151–183.

Janeway Jr., C.A., Jones, B., Hayday, A., 1988. Specificity and function of T cells bearing gamma delta receptors. Immunol. Today 9, 73–76.

Kaas, Q., Lefranc, M.-P., 2005. T cell receptor/peptide/MHC molecular characterization and standardized pMHC contact sites in IMGT/3Dstructure-DB. In Silico Biol. 5, 505–528.

Kaas, Q., Ruiz, M., Lefranc, M.-P., 2004. IMGT/3Dstructure-DB and IMGT/StructuralQuery, a database and a tool for immunoglobulin,

T cell receptor and MHC structural data. Nucleic Acids Res. 32, D208–D210.

Komano, H., Fujiura, Y., Kawaguchi, M., Matsumoto, S., Hashimoto, Y., Obana, S., Mombaerts, P., Tonegawa, S., Yamamoto, H., Itohara, S., et al., 1995. Homeostatic regulation of intestinal epithelia by intraepithelial gamma delta T cells. Proc. Natl. Acad. Sci. U.S.A. 92, 6147–6151.

Kubota, T., Wang, J., Gobel, T.W., Hockett, R.D., Cooper, M.D., Chen, C.H., 1999. Characterization of an avian (*Gallus gallus domesticus*) TCR alpha delta gene locus. J. Immunol. 163, 3858–3866.

Lefranc, M.-P., 1990. The human T-cell receptor delta genes. The mouse T-cell receptor gamma genes. Res. Immunol. 141, 692–695.

Lefranc, M.-P., Chuchana, P., Dariavach, P., Nguyen, C., Huck, S., Brockly, F., Jordan, B., Lefranc, G., 1989. Molecular mapping of the human T cell receptor gamma (TRG) genes and linkage of the variable and constant regions. Eur. J. Immunol. 19, 989–994.

Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Chaume, D., 2003. IMGT, the international ImMunoGeneTics information system, http://imgt.cines.fr: the reference in immunoinformatics. Stud. Health Technol. Inform. 95, 74–79.

Lefranc, M.-P., Giudicelli, V., Kaas, Q., Duprat, E., Jabado-Michaloud, J., Scaviner, D., Ginestoux, C., Clement, O., Chaume, D., Lefranc, G., 2005a. IMGT, the international ImMunoGeneTics information system. Nucleic Acids Res. 33, D593–D597.

Lefranc, M.-P., Lefranc, G., 2001. The T Cell Receptor FactsBook. Academic Press, London, UK, pp. 1–398.

Lefranc, M.-P., Pommié, C., Kaas, Q., Duprat, E., Bosc, N., Guiraudou, D., Jean, C., Ruiz, M., Da Piedade, I., Rouard, M., Foulquier, E., Thouvenin, V., Lefranc, G., 2005b. IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. Dev. Comp. Immunol. 29, 185–203.

Lefranc, M.-P., Pommié, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V., Lefranc, G., 2003b. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev. Comp. Immunol. 27, 55–77.

Lefranc, M.-P., Rabbitts, T.H., 1989. The human T-cell receptor gamma (TRG) genes. Trends Biochem. Sci. 14, 214–218.

Mackay, C.R., Hein, W.R., 1989. A large proportion of bovine T cells express the gamma delta T cell receptor and show a distinct tissue distribution and surface phenotype. Int. Immunol. 1, 540–545.

Martino, A., Casetti, R., D'Alessandri, A., Sacchi, A., Poccia, F., 2005. Complementary function of gamma delta T-lymphocytes and dendritic cells in the response to isopentenyl-pyrophosphate and lipopolysaccharide antigens. J. Clin. Immunol. 25, 230–237.

Massari, S., Lipsi, M.R., Vonghia, G., Antonacci, R., Ciccarese, S., 1998. T-cell receptor TCRG1 and TCRG2 clusters map separately in two different regions of sheep chromosome 4. Chromosome Res. 6, 419–420.

Miccoli, M.C., Antonacci, R., Vaccarelli, G., Lanave, C., Massari, S., Cribiu, E.P., Ciccarese, S., 2003. Evolution of TRG clusters in cattle and sheep genomes as drawn from the structural analysis of the ovine TRG2@ locus. J. Mol. Evol. 57, 52–62.

Miccoli, M.C., Lipsi, M.R., Massari, S., Lanave, C., Ciccarese, S., 2001. Exon-intron organization of TRGC genes in sheep. Immunogenetics 53, 416–422.

Miccoli, M.C., Vaccarelli, G., Lanave, C., Cribiu, E.P., Ciccarese, S., 2005. Comparative analyses of sheep and human TRG joining regions: evolution of J genes in Bovidae is driven by sequence conservation in their promoters for germline transcription. Gene 355, 67–78.

Pommié, C., Levadoux, S., Sabatier, R., Lefranc, G., Lefranc, M.-P., 2004. IMGT standardized criteria for statistical analysis of immunoglobulin V-REGION amino acid properties. J. Mol. Recogn. 17, 17–32.

Povinelli, C.M., Gibbs, R.A., 1993. Large-scale sequencing library production: an adaptor-based strategy. Anal. Biochem. 210, 16–26.

Sakano, H., Huppi, K., Heinrich, G., Tonegawa, S., 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 280, 288–294.

Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., Miller, W., 2000. PipMaker—a web server for aligning two genomic DNA sequences. Genome Res. 10, 577–586.

Smith, R.F., Weise, B.A., Wojzyynski, M.K., Davison, D.B., Worley, K.C., 1996. BCM search launcher—an integrated interface to molecular biology data base search and analysis services available on the world wide web. Genome Res. 6, 454–462.

Sowder, J.T., Chen, C.L., Ager, L.L., Chan, M.M., Cooper, M.D., 1988. A large subpopulation of avian T cells express a homologue of the mammalian T gamma/delta receptor. J. Exp. Med. 167, 315–322.

Tamaki, K., Sugaya, M., Tada, Y., Yasaka, N., Uehira, M., Nishimoto, H., Nakamura, K., 2001. Epidermal and dermal gamma-delta T cells. Chem. Immunol. 79, 43–51.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Vaccarelli, G., Miccoli, M.C., Lanave, C., Massari, S., Cribiu, E.P., Ciccarese, S., 2005. Genomic organization of the sheep TRG1@ locus and comparative analyses of Bovidae and human variable genes. Gene 357, 103–114.

Vernooij, B.T., Lenstra, J.A., Wang, K., Hood, L., 1993. Organization of the murine T-cell receptor gamma locus. Genomics 17, 566–574.

Yamashita, S., Tanaka, Y., Tsutsumi, S., Aburatani, H., Minato, N., Ihara, S., 2005. Analysis of mechanism for human gammadelta T cell recognition of nonpeptide antigens. Biochem. Biophys. Res. Commun. 334, 349–360.