Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 357 (2005) 103 – 114
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Genomic organization of the sheep *TRG1@* locus and comparative analyses of *Bovidae* and human variable genes

G. Vaccarelli[a,1], M.C. Miccoli[a,1], C. Lanave[b], S. Massari[c], E.P. Cribiu[d], S. Ciccarese[a,*]

[a] *Dipartimento di Anatomia Patologica e di Genetica, University of Bari, Italy*
[b] *Istituto di Tecnologie Biomediche-Sezione di Bari, CNR-Milano, Italy*
[c] *Dipartimento di Scienze e Tecnologie Biologiche ed Ambientali-Centro Ecotekne-University of Lecce, Italy*
[d] *Departement de Genetique Animale, Centre de Recherche INRA de Jouy-en-Josas, France*

Received 8 November 2004; received in revised form 23 February 2005; accepted 17 May 2005

Abstract

γδ T cells commonly account for 0.5%−5% of human (γδ low species) circulating T cells, whereas they are very common in chickens, and they may account for >70% of peripheral cells in ruminants (γδ high species). We have previously reported the ovine *TRG2@* locus structure, the first complete physical map of any ruminant animal TCR locus. Here we determined the *TRG1@* locus organization in sheep, reported all variable (*V*) gamma gene segments in their germline configuration and included human and cattle sequences in a three species comparison. The *TRG1@* locus spans about 140 kb and consists of three clusters named *TRG5*, *TRG3*, and *TRG1* according to the constant (*C*) genes. The predicted tertiary structure of cattle and sheep V proteins showed a remarkably high degree of conservation between the experimentally determined human Vγ9 and the proteins belonging to *TRG5* Vγ subgroup. However systematic comparison of primary and tertiary structure highligthed that in *Bovidae* the overall conformation of the γδ TCR, is more similar to the Fab fragment of an antibody than any TCR heterodimer. Phylogenetic analysis showed that the evolution of cattle and sheep *V* genes is related to the rearrangement process of *V* segments with the relevant *C*, and consequentely to the appartenence of the *V* genes to a given cluster. The *TRG* cluster evolution in cattle and sheep pointed out the existence of a *TRG5* ancient cluster and the occurrence of duplications of its minimal structural scheme of one *V*, two joining (*J*), and one *C*.

Keywords: T cell receptor; Locus structure and organization; *Bovidae*; Genome evolution; Variable domains peptides; Tertiary structure

1. Introduction

T-cell mediated immunity in Vertebrates is based on antigen recognition by T-cell receptors that may occur in two types: α/β or γ/δ. During lymphocyte maturation these receptors rearrange their Variable, Diversity and Joining (*V*–*D*–*J*) gene segments, thus creating a diversity both in terms of sequence and structure. Although exhibiting a number of common cell-surface proteins and similar effector capabilities, α/β and γ/δ T cells differ in some important biological properties. γ/δ T cells are most likely the first T cells emerging during thymic ontogenesis, in all Jawed Vertebrates (Hayday, 2000; Six et al., 1996; McVay et al., 1998). They do not show MHC restriction during antigenic recognition and in addition they are able to react to soluble

Abbreviations: P_B, bootstrap probability; CDR, complementarity determining regions; C, constant; D, diversity; HV, hypervariable region; J, joining; LP, leader region; MHC, major histocompatibility complex; MP, maximum parsimony method; NJ, neighbor-joining method; NMD, Non-sense-Mediated mRNA Decay; PTC, premature termination codon; PDB, protein data base; RSS, recombination signal sequences; B, residue buried in the β two strands; T, residue in the turn between the two strands; I, residue responsible for the interaction; TRA, T-cell receptor alpha; TRB, T-cell receptor beta; *TRG@*, T-cell receptor gamma locus; V, variable; *v*, vestigial form.

* Corresponding author. DAPEG-Genetics Section-Faculty of Biotechnological Sciences-University of Bari-via Amendola 165/A, 70126 Bari, Italy. Tel.: +39 080 5443384; fax: +39 080 5443386.
E-mail address: ciccarese@biologia.uniba.it (S. Ciccarese).
[1] G. Vaccarelli and M.C. Miccoli contributed equally to this work.

0378-1119/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2005.05.033

antigens including proteins and non-proteins (Carding and Egan, 2002).

TCR γδ conserves structural features common to all TCRs, e.g., the organization of each V and C domain into "Ig folds" (about 9–7 β strands, packed face to face in two antiparallel β sheets), constrained by intradomain disulphide bonding. Similarly conserved is the sequence-based subdivision of V regions into frameworks and hypervariable regions (HVs), the latter encoding surface-exposed complementary determining regions (CDRs) (Li et al., 1998). Despite the organizational conservation of TCR γδ genes, large differences exist in TCR γ/δ gene families in the different organisms, namely *V*γ genes, for which both interspecies and intraspecies divergence is strikingly high (Clark et al., 1995). On the other hand TCR γδ represents an "unique" antigen receptor, due to the peculiar structural properties of CDR1/2 and CDR3 (Wilson and Stanfield, 2001). The characteristic γ/δ pairings show a significant CDR1/2 divergence from each other. Such divergence contrasts with the constraint imposed on TCR αβ CDR1 and CDR2 by the requirements for binding to major histocompatibility complex (MHC) antigens (Garcia et al., 1996), thus CDR1/2 γ/δ sequences are more free to diverge, contributing significantly to antigen specificity.

γδ cells also represent more than 50% of the infiltrating T lymphocytes in epithelia-rich tissues such as skin, gut, and the reproductive tract (Hein and MacKay, 1991; Deusch et al., 1991). In these tissues they constitute a first defense line able to recognize not only a myriad of different microbial antigens, but above all, the unique "stress-induced antigens" which are markers of cell infection or transformation (Groh et al., 1996). However, although various immunoprotective and immunoregulating functions have been proposed for γ/δ T cells, they account only for a small percentage (1–5%) of the T lymphocytes present in the peripheral blood and in secondary lymphoid organs of most adult organisms, such as human and mouse ("γδ low" species), but more than 70% of peripheral T cells in young ruminants and chicken ("γδ high" species), this percentage decreasing to 5–25% with age (Hein and MacKay, 1991; Six et al., 1996). In cattle and sheep the *TRG* locus has two different locations on chromosome 4, one at position 4q31 (locus *TRG1@*) (Miccoli et al., 2003), and the other one at 4q22 (*TRG2@*) (in preparation). This is a peculiarity of these ruminant animals not found in the human *TRG@* locus. The ovine *TRG2@* spans over about 100 Kbp and it is structured in recombination clusters, each one comprising a *V–J–C* segment, arranged in the same transcriptional orientation. These findings support the hypothesis that a series of intra-locus gene duplications followed by a transpositional event have characterized the evolution of *TRG@* loci in *Bovidae*. On the other hand, many multigenic families are known to expand their repertoire and spread it along the entire genome by both intrachromosomal and interchromosomal gene duplications (e.g., olfactory receptors) (Glusman et al., 2001). It can be assumed that this propension to expand and contract shown by the olfactory receptors and T lymphocytes receptors gene families, as well as their propension to be involved in events of recombination, gene conversion, translocation, and positive selection for functional changes reflects a "biological demand" to adapt to a myriad of new signals, represented by the antigens in the case of the immune defense genes (Lane et al., 2002).

To allow an assessment of the contributions of the γδ high speciation mechanisms, we performed a systematic comparison of T cell receptor gamma (*TRG*) *V* gene families in cattle, sheep (*Bovidae*), and human. Our aim was: i) to determine the sheep *TRG1@* locus structural organization; ii) to identify the sequence of sheep *TRG1@* locus *V* gene segments in their germline configuration and determine them as correctly spliced transcription products; iii) to compare cattle, sheep and human V peptides tertiary structural features; iv) to perform phylogenetic analysis using cattle, sheep and human expressed *V* gene sequences only. We have established that in sheep *TRG1@* locus the *V*, *J*, and *C* genes are distributed in three recombination clusters, with a *V–J–C* structural scheme similar to that of *TRG2@* locus, but having, in addition, repeated structures within *V* gene regions. This was not surprising since the genetic complexity resulting from the organization of several repeated clusters and the higher number of genes in the ovine locus compared to the human one well correlates with the condition of "γδ high" in *Bovidae*.

2. Materials and methods

2.1. BAC clones isolation, variable and joining mapping and subcloning

A sheep BAC library, constructed by Vaiman et al. (1999), was screened by PCR with primers designed according to the published sequences: two primer pairs (C5A1/C5B; C5A/C5B) for sheep *C5* (AY147901) were used: C5A1: 5′-GAGGAAGTAAAACGCCATAG-3′, C5B: 5′-CCCATTGCCTTTTTAGTCAG-3′, C5A: 5′-GGCTTGATGGAGACTTGTTT-3′. Following the screening, BAC clones, address 655D3, 634F4, and 403D4 were isolated and DNAs were then subjected to PCR to confirm the presence of *C5* gene using the previously designed primer pair. A second primer pair (VG3A/VG3B) designed according to the published sequence for sheep *V3.1* (Z13003) was used; VG3A: 5′-TTCACATTTTTCTCGTTCTGG-3′ and VG3B: 5′-CTGTGGAGGGGGAGATTTG-3′. The presence of *V3.1* on the isolated BAC 202C5 DNA was confirmed by PCR. A third primer pair (CB1/CC1) designed according to the published sequence for sheep *C1* (Z12964) was used; CB1: 5′-ATTACACATACCTCCTCCTC-3′ and CC1: 5′-ATGACATAAAAGGGGAAAAA-3′. The presence of *C1* on the isolated BAC 637E8

DNA was confirmed by PCR. Isolated BAC clones DNA was digested with various rare-cutting restriction enzymes. DNA samples separation was performed on 1.2% agarose gel. Gels were run on a CHEF for 15 h, at 14 °C, 6 V/cm, 120° and 2 s switching time in 0.5 × TBE. The gels were blotted onto a nylon membrane (Amersham), which was then hybridized by using the DIG High Prime DNA Labeling and Detection Kit (Roche Diagnostic GmbH) to various *V* specific cDNA probes. The cDNA clones were a gift from Dr. WR Hein (Basel Institute for Immunology, Basel, Switzerland). Amplified DNA fragments and subclone inserts were purified using the High Pure PCR Product Purification Kit (Roche Diagnostic GmbH). Nucleotide sequences were determined using a commercial service and the resulting sequences were compared to existing ones using the Basic BLASTN research program.

The distances between *V*, *J* and *C* elements were determined by PCR on BAC DNAs with primers based on published cDNA sequences. Approximately 0.7 μg of each BAC DNA template were used in the PCR reactions. The standard buffer reaction consisted of 10 × PCR buffer (Roche Diagnostic GmbH); 0.2 mM of each dNTP; primers at 0.05 μM; 2.5U Expand Long Template Polymerase (Roche Diagnostic GmbH) was added in a 50 μL final volume. The thermal cycling protocol was: 4 min 94 °C, 30 s 94 °C, 40 s 58 °C, 5 min 68 °C for 30 cycles. The final cycle was extended for 15 min at 68 °C.

2.2. RT-PCRs (*V* transcription assay)

To test for the presence of correctly spliced transcription products, 6 *V*-specific 18- to 21-mer primers were designed. Five *C*-specific primers were also synthesized. Total RNAs were extracted from adult and fetus ovine spleen (Trizol method-Invitrogen) and polyA(+)RNAs were purified (Dynabeads oligo(dT)$_{25}$-Dynal). First strand cDNA synthesis was performed using *C* third exon specific primers (cDNA synthesis kit-Invitrogen). Ten microliters of the purified cDNA (50 μL) were used together with two primers, one for *V* and one for the relevant *C* first exon, in a standard 50 μL-PCR reaction. The reactions were cycled 35 times at temperatures of 94 °C for 30 s, 56–58 °C for 45 s, and 72 °C for 1 min, followed by a final extension at 72 °C for 7 min. The first PCR products were reamplified using the same *V* primer and an internal *J*-specific primer in opposite orientation. Secondary PCR reactions were stopped during the exponential phase (20 cycles) and at the end point (35 cycles). The results were run on a 2.0% agarose gel, stained, and photographed. The following primers were used, for *V3.1*: 5′-GCATTCACATTTTTCTCGTT-3′ and 5′-CAAGTTTTGTCCCATCGCC-3′ (*J5.2*), for *V3.2*: 5′-ATGCTCAAATGTTCGCTTC-3′ and 5′- CAAGTTTTGTCCCATCGCC-3′ (*J5.2*), for *V7*: 5′-TCCTCTTCTCCTCCCTCT-3′ and 5′-CAAGTTTTGTCCCATCGCC-3′ (*J5.2*), for *VP*: 5′-GATCTTCCAATAACACAAC-3′ and 5′-ACGAGCTTAGTTCCTTCACC-3′ (*J5.1*).

Total RNAs were extracted from young adult bovine spleen (Trizol method-Invitrogen) and polyA(+)RNAs were purified (Dynabeads oligo(dT)$_{25}$-Dynal). The treated RNA was used to generate full-length cDNA. First strand cDNA synthesis (cDNA synthesis kit-Invitrogen) was performed using bovine *C4* first exon specific primer (C4BR1: 5′-GGTCTTCATGGTATTTCCCT-3′). The 5′ RACE PCR were performed with a second *C4* specific primer (C4BR3: 5′-GTTGCAGCCTGTTGATGGTT-3′) and an ovine *V6* specific primer (V6: 5′-CAACTTGAGCCCCACCACC-3′). PCR products obtained were purified with High Pure PCR Product Purification Kit (Roche Diagnostic GmbH) and cloned into pCR-XL-TOPO vector (TOPO XL PCR cloning kit; Invitrogen).

2.3. Sequence alignment and computational analysis

The nucleotide variable sequences were multialigned on the basis of amino acids alignment using CLUSTALW (Thompson et al., 1994; website: http://www.ebi.ac.uk/clustalw/), with manually optimization by using GeneDoc program (Nicholas et al., 1997; website: http://www.psc.edu/biomed/genedoc/). For comparative analysis of genomic sequences mVISTA tool was used (website: http://gsd.lbl.gov/vista/index.shtml). The prediction of the structural features of cattle and sheep variable peptides was obtained using the Swiss-Model Automated Protein Modeling Server (Peitsch et al., 2000) (website: http://www.expasy.org/spdbv/).

2.4. Sequence analysis and drawing of dendrograms

TRGV sequences used in the analysis were: human (*Homo sapiens*): V1S2 (M13429), V1S3 (X04038), V1S4 (M36285), V1S5 (M36286), V1S8 (M13434), V2 (X07205), Vδ1 (M22198); cattle (*Bos taurus*): V2S1 (D13648), V5S8 (D13649), V1S1 (D16119), V1S2 (D16123), V5S11 (D16126), V5S6 (D16129), V5S7 (D16130), V1S3 (D16131), V5S16 (D16133), V3S1 (U73186), V3S2 (U73187), V4S1 (U73188), V6S1 (AY560834); sheep (*Ovis aries*): V1S1 (Z12998), V2S1 (Z12999), V2S2 (Z13000), V2S3 (Z13001), V2S4 (Z13002), V3S1 (Z13003), V3S2 (AY348328), V4S1 (Z13004), V5S1 (Z13005), V5S2 (Z13006), V6 (Z13007), V7 (AY362775); chicken (*Gallus gallus*): V1S3 (U78210), V1S4 (U78212), V1S5 (U78213), V1S8 (U78216), V2S7 (U78225), V2S8 (U78226), V2S9 (U78227), V3S3 (U78230), V3S4 (U78231), V3S8 (U78235).

The sequences were multialigned on the basis of amino acids alignment using the PILEUP program of GCG package (Wisconsin Package Version 9.1, Genetics Computer Group, Madison, Wisc.), with optimization of alignment by hand. Phylogenetic analysis was done on aminoacidic aligned sequences using PROTML from Molphy (Adachi and Hasegawa, 1992). ML reconstructions

utilized PROTML program with the Jones–Taylor–Thornton (JTT-f). The analysis on nucleotide aligned sequences was performed using maximum parsimony method (MP) on non-synonymous positions (1+2 codon position). We used the neighbor-joining (NJ; Saitou and Nei, 1987) method to reconstruct the phylogenetic trees. MP as implemented in PAUP*4.04b (Swofford, 2000) was also used to examine the reliability of topologies generated by the NJ method. The reliability of trees was examined by the bootstrap test (Felsestein, 1985) and the interior-branch test (Rzhetsky and Nei, 1992; Sitnikova, 1996), which produced the bootstrap probability (P_B) values, for each interior branch in the tree.

3. Results and discussion

3.1. Isolation of ovine TRG@ germline V and J genes

We have previously described the germline structure of the constant (*C*) genes in sheep (Miccoli et al., 2001) and the genomic organization of the ovine *TRG2@* locus (Miccoli et al., 2003). In this paper we present an analysis of the organization and structure of the ovine *TRG1@* locus and an overview of the isolated variable (*V*) and joining (*J*) gene segments of the two ovine *TRG@* loci in terms of their germline structure. As shown in Table 1a, *TRG1@ V* segments can be divided into three distinct subgroups and *TRG2@ V* segments into two subgroups. This division depends on the rearrangment of *V* segments with their relevant *J*(*C*) (Hein and Dudler, 1993). The *V1* gene was isolated from the partially characterized λ6A3 phage clone (Massari et al., 1998) using the *V1* cDNA as a probe. Using *C5*-, *V3.1*- and *C1*- specific primer sets, five clones, 634F4, 403D4, 655D3, 202C5 and 637E8 were isolated from an ovine genomic BAC library (Vaiman et al., 1999). Hybridization experiments (Fig. 1a, b) between a *V2.2* cDNA probe and digested BAC 655D3 genomic clone DNA detected three (*V2.1/V2.2, V2.3, V2.4*) of the four variable genes found rearranged with *J3(C3)* in mature transcripts by Hein and Dudler (1993). Isolation and sequencing of these genes supported the hypothesis that *V2.1/V2.2* transcripts result from two alleles of a single gene.

Two signals were observed after hybridization of a *V3* cDNA probe with *EcoR*V digested BAC 202C5 clone DNA (not shown). Relevant bands were cloned and subclones sequenced. The resulting sequences represent two variable genes: one is genomic *V3.1*; the other is a newly identified gene that, in accord with the current nomenclature for bovine ortholog, we named *V3.2*. It exhibits 98% nucleotide identity with *V3.1*.

Table 1
Gene name and rearrangement is derived from cDNA sequences (Hein and Dudler, 1993)

Table 1a

Locus	Gene name	Rearrangement	Genomic clone	Acc. number	
TRG1@	*V1*	*J1*(*C1*)	λ6A3*	AY362770	This paper
	V2.1/2.2	*J3*(*C3*)	BAC 655D3	AY348330	“
	V2.3		“	AY348329	“
	V2.4		“	AY348331	“
	V3.1	*J5.1*(*C5*)	BAC 202C5	AY348332	“
	V3.2		“	AY348328	“
	V4	*J5.2*(*C5*)	BAC 655D3	AY362775	“
	V7		“	“	“
	#VP		“	“	“
TRG2@	*V5.1*	*J2.1*(*C2*)	BAC 201E2	AY135487	Miccoli et al., 2003
	V5.2		“	AF540880	“
	V6	*J4.1*(*C4*)	“	AF540881	“

Table 1b

TRG1@	*J1.1*		λ6A3*	AY362774	This paper
	#J1.2	(*C1*)	“	AY362778	“
	J3.1	(*C3*)	λ3U3*	AY362776	“
	v J3.2		“	AY362777	“
	J5.1	(*C5*)	pBSC57.3*	AY147900	“
	J5.2		“	“	“
	v J5P		“	“	“
TRG2@	*J2.1*	(*C2*)	BAC 201E2	AF540883	Miccoli et al., 2003
	J2.2		BAC 201E2*	AY362771	This paper
	J6.1	(*C2*)	λ17A5	AF540882	Miccoli et al., 2003
	J6.2		λ17A5*	AY362773	This paper
	J4.1	(*C4*)	λ4P1	AF540884	Miccoli et al., 2003
	J4.2		BAC 201E2*	AY362772	This paper

(*) indicates genomic clones partially described in Miccoli et al., 2001, 2003; *TRG2@ V5.1, V5.2, V6, J2.1, J6.1* and *J4.1* were not isolated in this study, but were placed based on previous report. Newly identified *V* and *J* genes are underlined; (#) coding region has stop codon; (*v*) vestigial form.

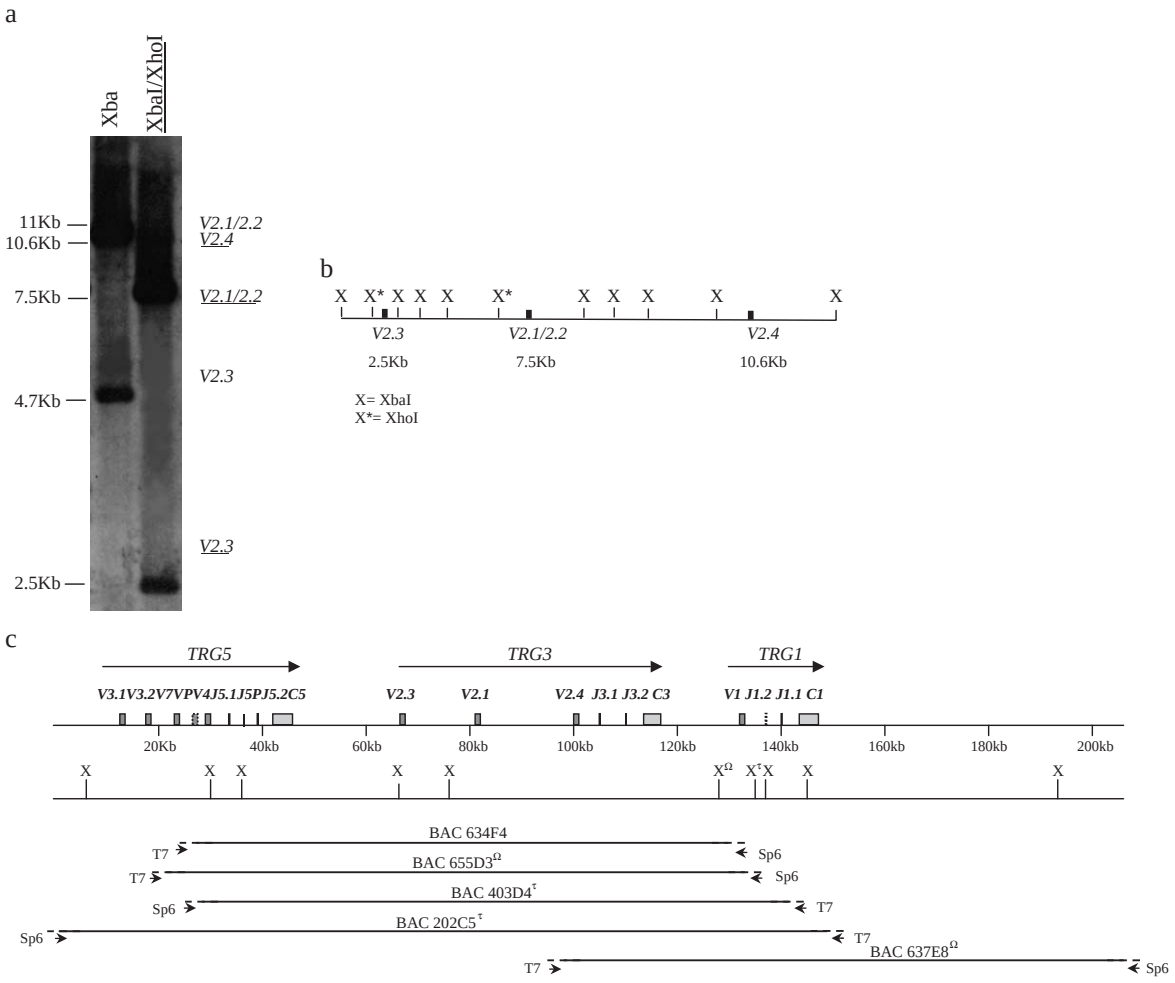

Fig. 1. (a) Southern hybridization experiments to investigate the genomic organization of variable genes of ovine *TRG3* cluster. Genomic BAC 655D3 was digested with *Xba*I and *Xba*I/*Xho*I enzymes, separated by pulse field gel electrophoresis and hybridized with a specific cDNA (*V2.2*) probe. (b) Schematic map of *TRG3* cluster variable region. Genes are indicated as boxed filled with gray shades (*V* segments). Only the dimension of *Xba*I/*Xho*I fragments (underlined in Fig. 1a) is indicated: *V2.4* (10.6 Kb), *V2.1/2.2* (7.5 Kb), *V2.3* (2.5 Kb). (c) Schematic map of sheep *TRG1@* locus. Three separated screenings were necessary for the analysis of the whole locus: (i) two primer pairs for sheep *C5* allowed the isolation of BAC 655D3, 634F4 and 403D4; (ii) a primer pair for *V3.1* made it possible to isolate BAC 202C5 and (iii) a primer pair for *C1* allowed the isolation of BAC 637E8. On the top line the genes (*V*, *J*, and *C*) are indicated as boxes filled with different grey shade. Vestigial gene is indicated as a vertical bar, while genes having stop codons but found transcribed are indicated with dotted lines. *Xho*I restriction map is represented on the line below; *Xho*I restriction site polymorphisms, (Ω) and (τ), are indicated. Five bars represent the BAC clones from which the locus physical map was determined; (➤) indicates the BAC insert orientation in pBeloBAC11 vector.

V4, *V7* and *VP* germline structure was established by sequencing a plasmid clone (PV410) identified by the hybridization of a *V4* cDNA probe with *Xho*I/*Not*I digested BAC 655D3 clone DNA. In particular, *V7* was identified thanks to the high (93,7%) nucleotide identity with bovine *V4S1* (U73188).

Table 1b provides an overview of all the *J* segments; they are divided according to locus and in terms of the *C* gene to which they are spliced in mature transcripts (Hein and Dudler, 1993) and which defines the recombination cluster they are part of. The sequence analyses of the genomic clones obtained enabled us to identify two *J* gene segments for each recombination cluster in the sheep genome, except for the *TRG5* cluster which has three of them. Individual *J* gene segments are numbered in the chronological order of their discovery since *J* segments we refer to as *J*×*.1* were previously identified as cDNAs by Hein and Dudler, while those referred to as *J*×*.2* correspond to genes never previously isolated. The only exception is the *J5P*, whose name reflects the homology to the human *JP*. As to the isolation of *J* segments belonging to *TRG1@* locus, the sequence of plasmid subclones from the λ6A3 phage, enabled us to detect *J1.1* and *J1.2* gene segments. The *J3.1* segment, previously isolated as cDNA, was detected in a λ3U3 subclone from which a vestigial form (*J3.2* gene), has also been isolated. Finally, the complete sequencing of the pBSC57.3 clone showed the presence of three *J5* gene segments, *J5.1*, *J5P* and *J5.2*. With regard to the *TRG2@* locus, reported here is the isolation of the *J2.2*, *J6.2*, and *J4.2* underlined in Table 1b; see companion paper for genomic description and mapping of *J* segments on *TRG2@*.

3.2. Genomic structure of the ovine TRG1@ locus and sequence analysis of V segments

Two BAC clones (637E8 and 202C5) were indicated to contain the entire *TRG1@* locus as a whole. Southern blotting, subcloning, sequencing and PCR data of these BAC DNA inserts, allowed to locate the *V*, *J* and *C* genes on the physical map of the *TRG1@* locus (4q3.1), as reported in Fig. 1c. The three clusters (*TRG5*, *TRG3*, and *TRG1*), having a total length of about 140 Kbp, are distributed over 220 Kb and lie in the same transcriptional orientation. The *TRG5* cluster (*V3.1*, *V3.2*, *V7*, *VP*, *V4-J5.1*, *J5P*, *J5.2-C5*) spans about 41 Kbp and contains a larger number of *V* segments than either *TRG3* (*V2.3*, *V2.1*, *V2.4-J3.1*, *J3.2-C3*) or *TRG1* (*V1-J1.2*, *J1.1-C1*). Now if the chromosomal mapping and the organization of the ovine *TRG1@* locus are compared with the counterparts of the human locus (IMGT Database), we can see first that *Bovidae* (cattle and sheep) *TRG1@* is included within a region of homology (GDB, 2001) with human *TRG@* locus on HSA7; second, the difference concerning genomic structure lies in that *TRG1@* locus is substantially more complex in gene number and organization, since in man it is only in the region of *V* genes that the duplications are considerable. In particular, the human *TRGV1* subfamily consists of nine closely related members, only five of which encode functional peptides. The *TRG2* subfamily contains a single member (*V9h*), increasing the total number of six functional peptide-encoding *V* segments (Arden et al., 1995). The single-member subfamilies *V3*, *V4*, *V5*, and *V6* have been classified as pseudogenes, because the leader introns are unspliced in all transcripts. The absence of splicing is due to the fact that splice donor and acceptor sites are different from the splice consensus sequences (Zhang et al., 1994).

Fig. 2 presents the alignment of *TRG1@* variable genes in their germline configuration. All of them except *VP* encode functional peptides. The *V* elements are made up of two exons: the first one, which encodes the initial part of the leader region (LP1), is followed by an intron whose length ranges from 79 bp for *V4.1* to 136 bp for *V2.4*. The second exon encodes the remaining portion of the leader (LP2) and the entire V region. The splice consensus sequences are readily detectable in the intron, while characteristic heptamer/nonamer recombination signal sequences (RSSs), showing only slight differences with the CACAGTG/ACAAAAACC consensus sequence (Ramsden et al., 1994), are visible immediately downstream of the 3′ end of each *V* element coding region. The deduced amino acid sequence of *V3.2* shows 95,8% identity with *V3.1*. However, *V3.2* differs significantly from *V3.1* at the 3′ end, particularly for the deletion of three nucleotides in the heptamer. *VP* consists of two exons of 46 and 321 bp, which are separated by a 222 bp-intron. It contains an ATG start codon, splice and RSS sites, but has a stop codon in position 196. According to the IMGT definition of functionality, a *V* gene is considered a pseudogene if the coding region has a stop codon in L-PART1 and/or the V-EXON. Therefore *VP* might be considered a pseudogene (Fig. 2).

3.3. Transcription assay of newly isolated V genes

To examine whether the newly identified *V* genes (all belonging to *TRG5* cluster and underlined in Table 1a), are transcribed and spliced, we developed a rapid RT-PCR-based assay. *V*- and *J*-specific primers and a single primer located in the first exon of each C were used in RT-PCR reactions performed on first-strand cDNA prepared from adult and fetus spleen polyA(+)RNA. PCR reactions were stopped during the exponential phase (20 cycles) (Fig. 3a) and at the end point (35 cycles) (Fig. 3b). Reactions included *V3.1* as control since it was previously reported as a cDNA isolated from a sheep library by Hein and Dudler, 1993. Results of these experiments show that all the newly identified *V* genes were transcribed at a similar extent in adult and foetus. A strong band is present in the *VP* lane. It is well known that mRNAs with a premature translation termination codon are unstable, since they are subjected to a selective degradation by the Nonsense-Mediated mRNA Decay (NMD) mechanism (Maquat, 2004). *VP* features a premature termination codon (PTC) at a position that is more than 50–55 nucleotides upstream of the splicing-generated exon-exon junction between *J* and *C* exons; it therefore seems to escape from this mechanism, since its RT-PCR product indicates an abundant and healthy mRNA. The presence of a second *VP* exon without the stop codon is excluded: a) because the PCR primer was designed in the exon bearing the PTC and b) the nucleotide sequence of PCR products shown in Fig. 3 (*VP* lane) has confirmed the presence of the STOP codon. A mechanism other than the alternatively spliced transcript formation (Wang et al., 2002) has probably been adopted in the case of the *VP* gene to escape NMD and to generate a functional mRNA.

3.4. Alignment of cattle, sheep and human Vγ peptides and comparison of their tertiary structures

We focused on the structural and functional inferences of the multialignment of the cattle, sheep (included the sequences reported in this paper), and human *V*γ polypeptide sequences as deduced from their germline configuration. The protein sequences were trimmed of leader peptides as IMGT suggestion, assuming that different selection mechanisms may be involved in the evolution of *V* gene segments and leader peptide sequences. *V* segments of all the six *TRG* recombination clusters comprised in the sheep *TRG1@* and *TRG2@* loci (with exception of VP) were included in the analysis. In cattle, all members of subgroups 2, 3 and 4 were included and assigned to *TRG3* (subgroup 2) and to *TRG5* (subgroups 3 and 4). Two out of three sequences of subgroup 1 and nine out of twenty of subgroup 5 were assigned to *TRG1* and *TRG2/6*, respec-

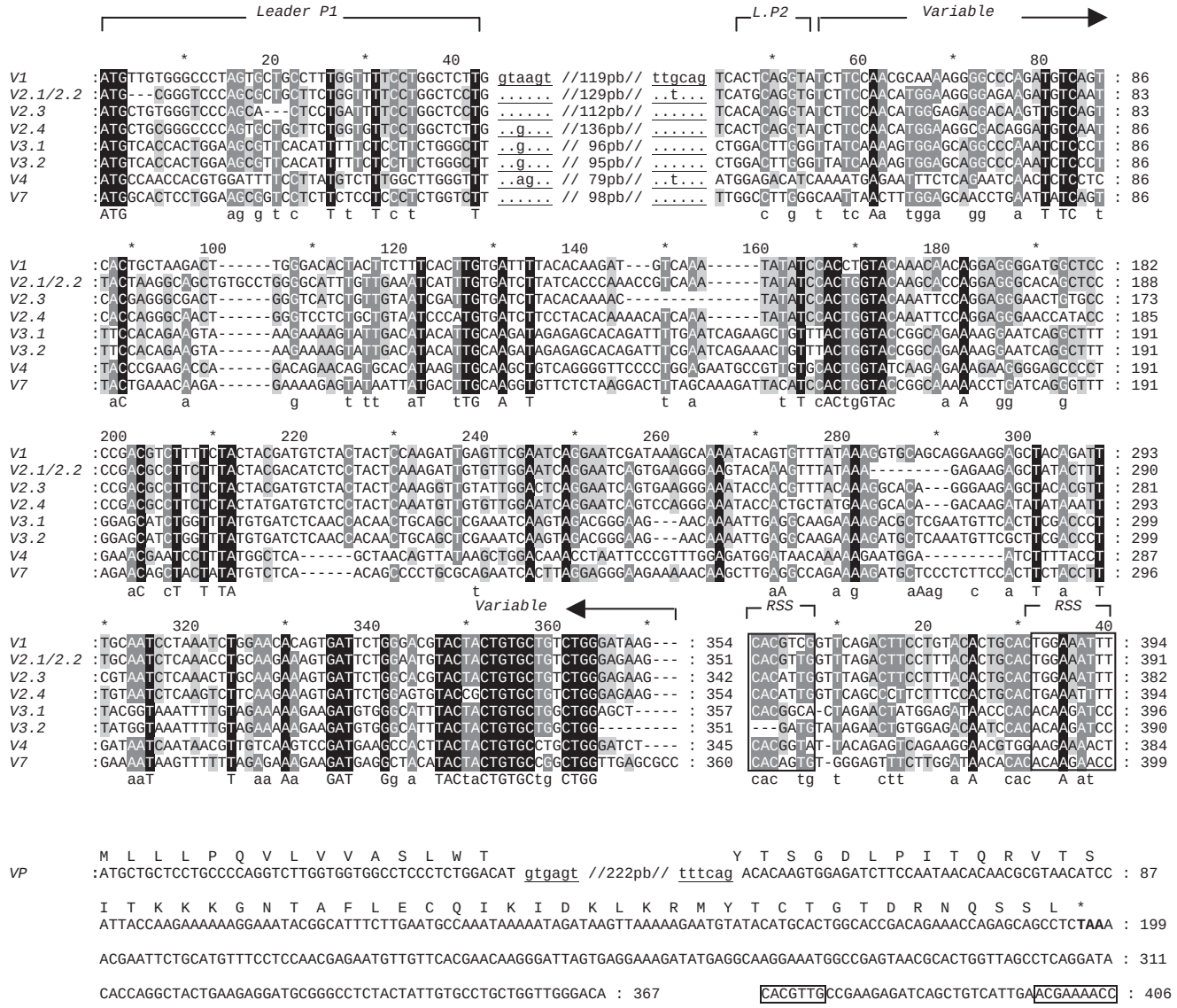

Fig. 2. The sheep *TRG1@ V* germline genes nucleotide alignment. The most abundant character at each position is shown as *white* on a *black* background; the second most abundant nucleotide has a *grey* background and the third most common character is shown as *black* on a different *grey* shade background. Gaps are presented by dashes. The donor and acceptor splice sites are in lowercase and underlined and heptamers and nonamers in the 3′ RSS are boxed. Nucleotides coding for Leader P1, Leader P2 and Variable regions are indicated. *VP* is shown separately as a pseudogene: the translation stop codon in the amino acid sequence is indicated (*).

tively. *V6*c was newly isolated by us as cDNA from the bovine spleen (AY560834) and assigned to *TRG4*. The appartenance to a given cluster resulted from the *C* with which the rearranged *V–J* was found spliced in mature transcripts (Ishiguro et al., 1993; Hein and Dudler, 1997; Takeuchi et al., 1992; this paper). From the alignment reported in Fig. 4, it firstly appears a clear separation between the *TRG4, TRG3, TRG2/6, TRG1 V* peptides and those belonging to *TRG5*, differences being particularly straightforward for the latter.

The complementarity determining regions (CDR1 to CDR3) in cattle and sheep were inferred on the basis of human IMGT variable sequences. The analysis of the changes in the residues that, according to Chothia et al. (1988) determine the structural features of the Ig domain, (residues buried in the β strands (B), residues in the turn between the two strands (T) and residues responsible for the interaction (I) between the variable domains of H and L chains), shows that although in most of the cases these residues are conserved, in some cases there are changes which alter the type of the aminoacid. In particular the buried residue G^{3C} is Q in *TRG5* and *TRG4* and it is absent in *TRG2/6*; for the residues in the turn G^{15} is E, D, K or S in *TRG5* and W in *TRG1*; Y^{65} is I or L in *TRG5* and V in *TRG4*. The rest of the deviations occurs for the I-type residues: H^{37} is Q or E in *TRG5* and Q in *TRG1* and *TRG4* and P^{43} residue is L in *TRG5*. In order to examine more closely the structural features of cattle and sheep variable peptides, we took advantage of the knowledge of the crystal structure of two human αβ TCRs (Garcia et al., 1999) and

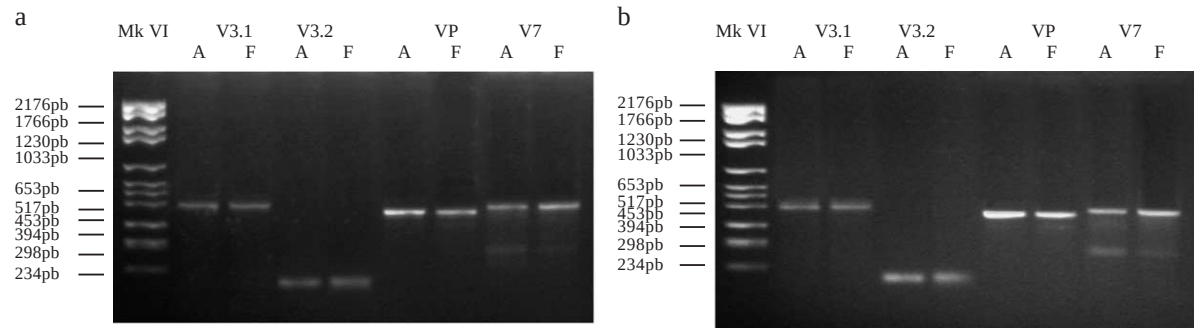

Fig. 3. RT-PCR assay products of newly isolated *V* genes rearranging with *C5*; *V3.1* is taken as a control. The results of RT-PCR performed on polyA+RNA extracted from adult (A) and foetus (F) spleen, meant do determine whether they are represented in TCR γ chain repertoire, are shown. The gene-specific oligonucleotide primers used are described in Materials and Methods; their efficiency was determined on genomic clones and negative or low results were repeated twice. The products were analyzed by electrophoresis in 2% agarose gel with a molecular weight marker. (a) PCR products at 20 cycles; (b) PCR products at 35 cycles.

of one γδ TCR (Allison et al., 2001), and used the Swiss-Model Automated Protein Modeling Server (Peitsch et al., 2000) to predict the three-dimensional structure of the protein aligned in Fig. 4. The lengths and the position of the A, B, E and D β strands of the lower β sheet and the C, C′, C″ and F β strands of the upper β sheet are shown in Table 2. It is interesting to observe that the human V1S2h, V1S4h, and V1S3h (colored orange in Table 2) structures were predicted not on the basis of the homology with the published structure of V9h (Allison et al., 2001) as one

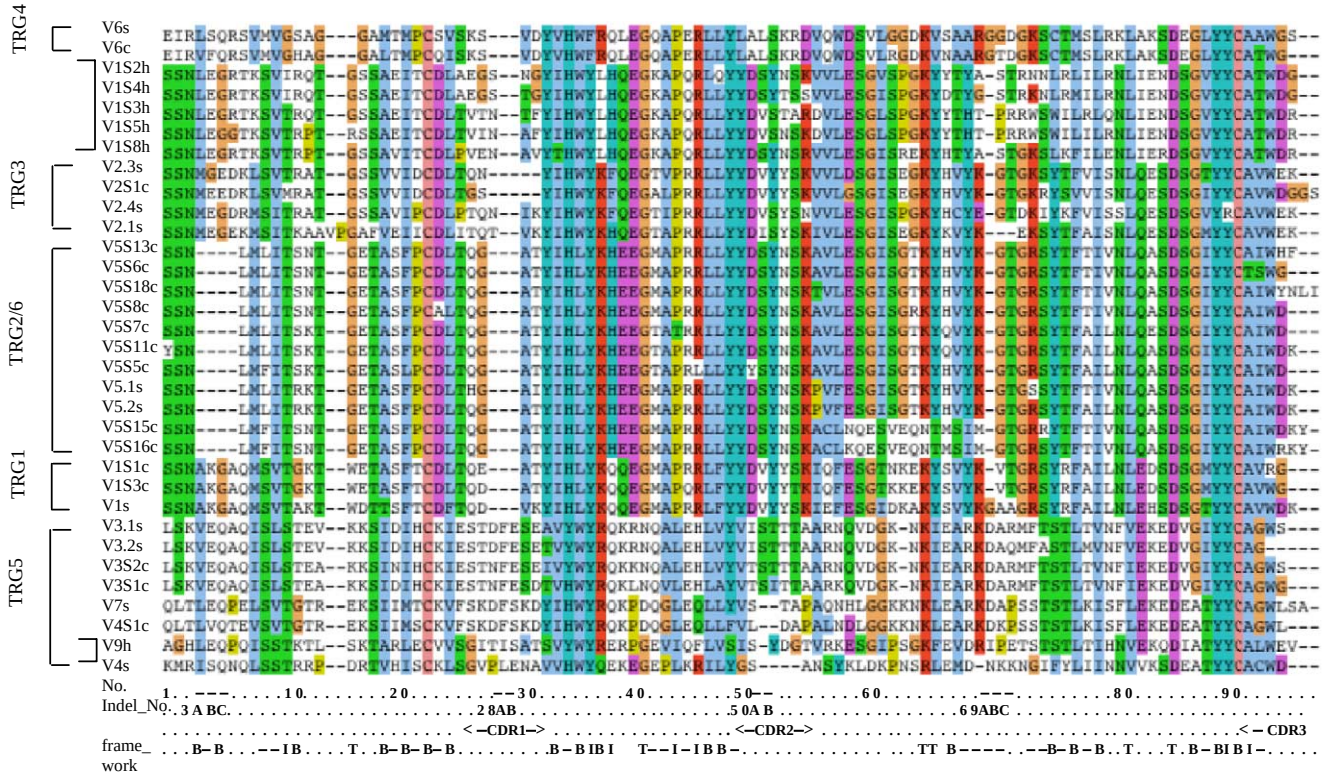

Fig. 4. The Vγ peptide sequence alignment of cattle (c), sheep (s) and human (h). The IMGT nomenclature to refer to the gene segments is used. The numbering system for amino acid residues refers to protein sequences trimmed of leader peptides (IMGT Database). Amino acid residues are color-coded (in online version only) to highlight functionally conserved features. Hydrofobic residues are *light blue*; tyrosine (Y) and histidine (H) are *blue* in a hydrophobic preference. Asparagine, glutamine, serine, and threonine (N, Q, S, and T) are *green*; aspartate and glutamate (D and E) are *purple*; arginine and lysine (R and K) are *red*; cysteine (C) is *pink*. More than 50% occurrence of a property results in coloring. (S and T are also colored in a hydrophobic preference). All glycines (G, *orange*) and prolines (P, *yellow*) are colored. Lengths of CDR1 and CDR2 are given. CDR3 length is shown in part since the V segment is in its germline structure. Insertion or deletion are accommodated by appending characters to the corresponding numbers. Residues at positions homologous to those of Ig residues that determine conserved structural features of the Ig domain are indicated by characters. B, residues buried between the β-sheets. T, residues in turns. I, inter-domain sites.—, residues that are not conserved in Ig but part of the β-sheet framework.

Table 2
The length and position of β-strands of the tertiary structure of all proteins aligned in Fig. 5 are predicted by the Swiss-Model Automated Protein Modeling Server (Peitsch et al., 2000; website: http://www.expasy.org/spdbv/)

			CDR1			CDR2					CDR3		
	A (L)	B (L)		C (U)	C' (U)		C" (U)	D (L)	E (L)	F (U)		protein	pdb-code
V6s	np	12-24		33-37	44-50B		54-58	66-69C	71-80	87-91		Ig	1otu
V6c	np	12-24		33-37	44-50B		54-58	66-69C	71-80	87-91		Ig	1otu
V9h	3-3B	18-25		32-37	44-50		53-57	64-69B	71-78	87-91		V9	1hxm
V1S2h	3-4	18-25		30-34	42-49		58-61	_	63-71	_		Ig	1fgv
V1S4h	3-4	15-25		31-37	44-48		55-57	64-69B	72-80	87-93		Ig	1fgv
V1S3h	3-4	21-25		32-36	44-46		50B-52	_	_	_		Ig	1a7q
V1S5h	_	_		_	_		_	_	_	_		no protein	
V1S8h	_	_		_	_		_	_	_	_		no protein	
V2.3s	np	18-20		35-46	48-55		55-57	64-66	75-78	87-89		Ig	1fgv
V2S1c	np	15-24		21-37	np		np	64-69C	72-80	87-90		Ig	1fdl
V2.4s	3B-4	18-24		33-37	44-46		np	64-68	73-79	87-91		Ig	1ad0
V2.1s	_	_		_	_		_	_	_	_		no protein	
V5S13c	1-3	18-24		33-37	44-46		np	64-68	71-78	87-91		Ig	1ad0
V5S6c	1-3	18-24		33-37	np		np	64-68	71-78	87-91		Ig	1ad0
V5S18c	1-3	18-24		33-37	44-46		np	64-68	71-78	87-91		Ig	1ad0
V5S8c	1-3	18-24		33-37	44-46		np	64-69	71-77	87-91		Ig	1ad0
V5S7c	1-3	18-24		33-37	np		np	64-68	71-78	87-91		Ig	1ad0
V5S11c	1-3	18-24		33-37	np		np	64-68	71-78	87-91		Ig	1ad0
V5S5c	1-3	18-24		33-37	44-46		np	64-68	71-78	87-91		Ig	1ad0
V5.1s	1-3	18-24		33-37	44-46		np	64-68	71-78	87-91		Ig	1ad0
V5.2s	1-3	18-24		33-37	44-46		np	64-68	71-78	87-91		Ig	1ad0
V5S15c	1-3	18-24		33-37	44-50B		53-59	65-69C	71-77	87-91		Ig	1ad0
V5S16c	1-3	18-24		32-37	44-50B		53-59	65-69	71-78	87-91		Ig	1ad0
V1S1c	_	_		_	_		_	_	_	_		no protein	
V1S3c	_	_		_	_		_	_	_	_		no protein	
V1s	np	15-24		33-37	np		np	64-66	71-80	87-91		Ig	1ad0
V3.1s	3-3B	18-25		32-37	44-50		52-56	64-69B	71-78	87-91		V9	1hxm
V3.2s	3-3B	18-25		32-37	44-50		52-56	64-69B	71-78	87-91		V9	1hxm
V3S2c	3-3B	18-25		32-37	44-50		52-56	64-69B	71-78	87-91		V9	1hxm
V3S1c	3-3B	18-25		32-37	44-50		52-56	64-69B	71-78	87-91		V9	1hxm
V7s	3-3B	18-25		32-37	44-49		55-57	64-69B	71-78	87-91		V9	1hxm
V4S1c	3-3B	18-25		32-37	44-50		53-57	64-69B	71-78	87-91		V9	1hxm
V4s	3-3B	18-25		32-37	np		np	64-69B	71-78	87-91		V9	1hxm

On the left cattle and sheep *TRGV* subgroups are indicated with different colours (in online version only). The V9h structure experimentally determined for human VGAMMA9/VDELTA2 (Allison et al., 2001) has 1 hxm.pdb entry. As an example, the PDB entries retrieved by the program in searching sequences of known 3D structures for the following peptides are: V1S2h (1fgv.pdb), V1S3h (1a7q.pdb) (orange); V5S16c, V5S18c (blue) and V2.4s (ligth blue) (1ad0.pdb), V6s (1otu.pdb) (pink). The lower (L) β sheet contains four strands labeled A, B, E, and D and the upper (U) β sheet contains four strands C, C′, C″ and F. On the right the PDB entries retrieved by the program are reported. Not present (np) indicates lack of β-strand,- indicates no protein. The majority of cattle and sheep *V* peptides belonging to the *TRG3* and *TRG2/6* clusters do not have a C″ strand.

would expect, but on the basis of the homology with the human Fab three-dimensional structure. Of all the examined sequences in *Bovidae*, only the *TRG5 V*s (colored green in Table 2) structure was predicted based on the homology with V9h; this is in agreement with the phylogenetic trees obtained by the NJ method analyzing *V* genes from cattle, sheep, human and chicken, where human *V9* gene was shown to be the ancestral form in human, as is the case for *TRG5 V*s in cattle and sheep (Fig. 5). It is remarkable to note that the structures inferred from the sequences of all (except for *TRG5*) clusters (pink, light blue, blue and gray) in *Bovidae* have been predicted on the basis of the homology with the human Fab three-dimensional structure. The program was not able to predict the tertiary structures of the most diversified variable γ chain encoded by genes derived from more recent duplications in the two loci both in human (V1S5h and V1S8h) and in *Bovidae* (V2.1s, V1S1c and V1S3c). The above reported data show, in *Bovidae* like in human, that the overall conformation of the γδ TCR is certainly more similar to the Fab fragment of an antibody than any TCR heterodimer (Wilson and Stanfield, 2001).

With regard to the loops that contain the complementarity determining regions CDR1 and CDR2 (CDR3 length is shown in part since the *V* segment is in its germline structure), on the basis of the conservation of the loop

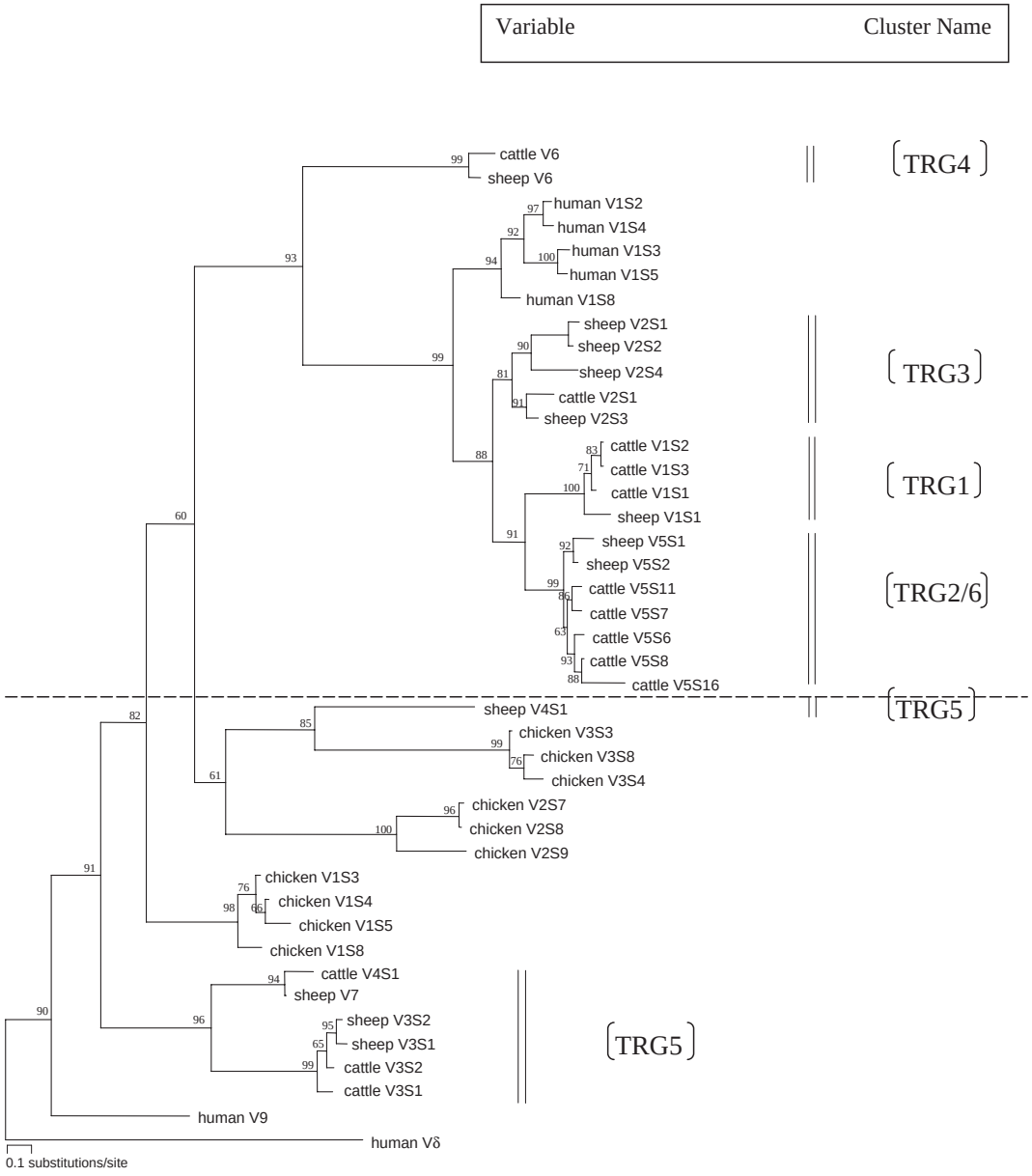

Fig. 5. The phylogenetic tree reported in the figure is the most significant tree that ML selected among those calculated using NJ method on the base of distances (PROTML in MOLPHY package), analyzing amino acid sequences by JTTF-f of variable (*V*) genes from cattle and sheep (Ishiguro et al., 1993; Hein and Dudler, 1993; Hein and Dudler, 1997; Miccoli et al., 2003; this paper), human (IMGT Database, 2002), and chicken (Six et al., 1996). On nodes, the most significant bootstrap values out of 100 replicates are shown. The extent of the compared sequences is 75 amino acid positions. The recommended IMGT nomenclature for the cattle, sheep and chicken genes is used. The nomenclature for the human genes is from the literature. All the cattle and sheep variable sequences examined (double vertical lines) are shown with the relevant *TRG* cluster (*TRG1* to *TRG5*).

length and of some residues located at certain key positions (Chothia et al., 1992), the alignment (Fig. 4) shows a large range of length variation in the C^{22}–(W/L)^{34} region, which contains the CDR1 domain. The sequences included between V6s and V1s have a loop length from 10 to 13 aa, whereas *TRG5* family members have a loop 15 aa long. In this case it is likely that extraresidues form a longer hairpin loop, lacking an ordered conformation and extending away from the domain (Arden, 1998). The CDR1 loop contains one hydrophobic residue in position 29, which penetrates in the protein framework and interacts with the highly conserved hydrofobic residue located at both ends of the loop, in position 24 and 32; this interaction could be important to stabilize the conformation of the loop (Arden, 1998). A similar conformation occurs in IgHV, whereas in TRAV a Pro in position 30 is involved in the stabilization and in TRBV histidine which is exposed to the solvent in pos. 29. In the alignment (Fig. 4) the CDR2 loop appears to be more structurally constrained than the CDR1; in fact the CDR2 loop of *TRG* is almost always 8 residues long compared to 7 residues for TRA and TRB.

3.5. Evolution of TRG clusters in cattle and sheep

A phylogenetic analysis of V sequences from human, cattle, sheep and chicken is reported (Fig. 5). A variable segment belonging to the human family *Vδ1* (Accession number M22198) was used as an outgroup. The dash line (– in Fig. 5) that divides the tree in two parts is useful for understanding the evolutionary relationship among *TRGV* genes. The upper part features a cluster only with human (subfamily *TRGV1*s members) and *Bovidae* sequences. The cluster in the lower part consists in the human *V9*, cattle and sheep *TRG5 V* subgroup and the chicken sequences. Cattle and sheep *V* genes were classified into six main groups, each one of which refers to a recombination cluster, according to the *C* name. This classification was based on bootstrap probability (P_B) values ranging from 60% to 93%.

The results of the organization of the ovine *TRG1@* locus and of the phylogenetic analyses obtained on *V* genes indicate that in *Bovidae* the *V* evolutionary behaviour is such that all the *V* members distribute in the phylogenetic tree according to their belonging to a given *TRG* recombination cluster. This opens the way to deducting the evolutionary history of the *Bovidae* loci which, starting from an ancestral cluster (the *TRG5*), after a series of duplications, has led to the origin of the most recent clusters (presumably the *TRG2/6* and the *TRG1*).

4. Conclusions

The main objective of this study was to focus on the inferred evolution and structural features of variable (*V*) gene segments of both *TRG* loci in cattle and sheep compared to the human locus. Particularly when considering the objectives set in previous sections, conclusions can be drawn in the following points: i) the structural genomic analysis of the sheep *TRG1@* locus reiterates the peculiarity of the organization of *Bovidae* γ loci in clusters, each containing the basic structure *V–J–J–C* segment; on the other hand, the human γ locus has duplicated *J–C* ensembles separated from the set of *V* genes; ii) the expression profiles of correctly spliced and transcripted *TRGV* segments indicate the adoption of a mechanism that allows to escape the Nonsense Mediated mRNA Decay mechanism, in presence of a premature termination codon as is the case of *VP*; iii) comparison of both analyses of V proteins primary and tertiary structure highligthed, in *Bovidae* like in human, that the overall conformation of the γδ TCR is more similar to the Fab fragment of an antibody than any TCR heterodimer. Moreover analysis of cattle and sheep *V* tertiary structure, inferred from the comparison with the published crystallographic structures present in database, showed an unquestionable and remarkably high degree of conservation between the proteins belonging to *TRG5V*γ subgroup from cattle and sheep and the human *V*γ9; iv) phylogenetic data obtained from the analysis of *TRGV* genes from cattle, sheep, human and chicken has shown *Vγ9* to be the ancestral form in human, as is the case for *TRG5 V*s in cattle and sheep. When considering that the γδ T population in human is confined to a small number of specific *Vγ9* and *Vδ2* chain pairings that are usually tissue specific, and that in peripheral blood, up to 90% of the 5% of γδ TCRs use a specific combination of *Vγ9* and *Vδ2* gene segments, it seems to us that investigations on the tissue specific expression of *TRG5V*γ subgroup *V* in *Bovidae*, should be carried out. Finally further studies should be performed on comparing large genomic portions of cattle and sheep *TRG* and *TRD* chromosomal regions with their human counterparts in order to provide a genomic perspective for understanding γδ loci evolution in *Bovidae* that fully represent a "γδ -high" species condition.

Acknowledgments

The financial support of Progetti di Interesse Nazionale (PRIN) and Ministero per l'Università e la Ricerca Scientifica e Tecnologica (MIUR) is gratefully acknowledged. We thank P. Barsanti for critical review in the preparation of this manuscript.

References

Adachi, J., Hasegawa, M., 1992. Amino acid substitution of proteins coded for mitochondrial DNA during mammalian evolution. Jpn. J. Genet. 67, 187–197.

Allison, T.J., Winter, C.C., Fournie, J.J., Bonneville, M., Garboczi, D.N., 2001. Structure of a human γδ T-cell antigen receptor. Nature 411, 820–824.

Arden, B., 1998. Conserved motifs in T-cell receptor CDR1 and CDR2: implications for ligand and CD8 co-receptor binding. Curr. Opin. Immunol. 10, 74–81.

Arden, B., Clark, S.P., Kabelitz, D., Mak, T.W., 1995. Human T-cell receptor variable gene segment families. Immunogenetics 42, 455–500.

Carding, S.R., Egan, P.J., 2002. γδ T cells: functional plasticity and heterogeneity. Nat. Immunol. 2, 336–345.

Chothia, C., Boswell, D.R., Lesk, A.M., 1988. The outline structure of the T-cell αβ receptor. EMBO J. 7, 3745–3755.

Chothia, C., et al., 1992. Structural repertoire of the human V_H segments. J. Mol. Biol. 227, 799–817.

Clark, S.P., Arden, B., Kabelitz, D., Mak, T.W., 1995. Comparison of human and mouse T-cell receptor variable gene segment subfamilies. Immunogenetics 42, 531–540.

Deusch, K., Luling, F., Reich, K., Classen, M., Wagner, H., Pfeffer, K., 1991. A major fraction of human intra-epithelial lymphocytes simultaneously expresses the gamma delta T cell receptor, the CD8 accessory molecule, and preferentially uses the Vδ1 gene segment. Eur. J. Immunol. 21, 1053–1059.

Felsestein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Garcia, K.C., et al., 1996. An αβ T cell receptor structure at 2.5 angstrom and its orientation in the TcR-MHC complex. Science 274, 209–219.

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

Glusman, G., Yanai, I., Rubin, I., Lancet, D., 2001. The complete human olfactory subgenome. Genome Res. 11, 685–702.

Groh, V., Bahram, S., Herman, A., Beauchamp, M., Spies, T., 1996. Cell stress-regulated human major hystocompatibility complex class I gene expressed in gastrointestinal epithelium. Proc. Natl. Acad. Sci. 93, 12445–12450.

Hayday, A.C., 2000. γδ cells: a right time and a right place for a conserved third way of protection. Annu. Rev. Immunol. 18, 975–1026.

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

Hein, W.R., Dudler, L., 1997. TCR γδ 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 γδ T cell in the ruminant immune system. Immunol. Today 12, 30–34.

IMGT Database: http://imgt.cines.fr:8104/.

Ishiguro, N., Aida, Y., Shinagawa, M., 1993. Molecular structures of cattle T-cell receptor γ and δ chains predominantly expressed on peripheral blood lymphocytes. Immunogenetics 38, 437–443.

Lane, R.P., et al., 2002. Genomic analysis of the olfactory receptor region of the mouse and human T-cell receptors α/δ loci. Genome Res. 12, 81–87.

Li, H., Lebedeva, M.I., Llera, A.S., Fields, B.A., Brenner, M.B., Mariuzza, R.A., 1998. Structure of the V δ domain of human γδ T-cell antigen receptor. Nature 391, 502–506.

Maquat, L.E., 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat. Rev., Mol. Cell Biol. 5, 89–99.

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. Chomosom. Res. 6, 419–420.

McVay, L.D., Jaswal, S.S., Kennedy, C., Hayday, A., Carding, S.R., 1998. The generation of human gamma delta T cells repertoires during fetal development. J. Immunol. 160, 5851–5860.

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., et al., 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.

Nicholas, K.B., Nicholas Jr., H.B., Deerfield, D.W. II., 1997. GeneDoc: analysis and visualization of genetic variation, EMBNEW. News 4, 14.

Peitsch, M.C., Schwede, T., Guex, N., 2000. Automated protein modelling—the proteome in 3D. Pharmacogenomics 1 (3), 257–266.

Ramsden, D.A., Baetz, K., Wu, G.E., 1994. Conservation of sequence in recombination signal sequence spacers. Nucleic Acids Res. 22, 1785–1796.

Rzhetsky, A., Nei, M., 1992. A simple method for estimating and testing minimum-evolution trees. Mol. Biol. Evol. 9, 945–967.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sitnikova, T., 1996. Bootstrap method of interior-branch test for phylogenetic trees. Mol. Biol. Evol. 13, 605–611.

Six, A., et al., 1996. Characterization of avian T-cell receptor γ genes. Proc. Natl. Acad. Sci. 93, 15329–15334.

Swofford, D.L., 2000. PAUP. Phylogenetic Analysis Using Parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, MA.

Takeuchi, N., Ishiguro, N., Shinagawa, M., 1992. Molecular cloning and sequence analysis of bovine T-cell receptor γ and δ chain genes. Immunogenetics 35, 89–96.

The Genome Database (GDB) (2001) http://www.gdb.org/.

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 (22), 4673–4680.

Vaiman, D., et al., 1999. Construction and characterization of a sheep BAC library of three genome equivalents. Mamm. Genome 10, 585–587.

Wang, J., Hamilton, J.I., Carter, M.S., Li, S., Wilkinson, M.F., 2002. Alternatively spliced TCR mRNA induced by disruption of reading frame. Science 297, 108–110.

Wilson, I.A., Stanfield, R.L., 2001. Unraveling the mysteries of γδ T cell recognition. Nat. Immunol. 2, 579–581.

Zhang, X.M., Tonnelle, C., Lefranc, M.P., Huck, S., 1994. T cell receptor γ cDNA in a human fetal liver and thymus: variable regions of γ chains are restricted to Vγ1 or V9, due to the absence of splicing of the V10 and V11 leader intron. Eur. J. Immunol. 24, 571–578.