Molecular Immunology 46 (2009) 2728–2736

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

Short communication

The deduced structure of the T cell receptor gamma locus in *Canis lupus familiaris*

S. Massari[a], F. Bellahcene[b], G. Vaccarelli[c], G. Carelli[d], M. Mineccia[c], M.-P. Lefranc[b], R. Antonacci[c], S. Ciccarese[c,*]

[a] *Dipartimento di Scienze e Tecnologie Biologiche ed Ambientali, Universita' del Salento, Lecce, Italy*
[b] *IMGT, Laboratoire d'ImmunoGénétique Moléculaire, Institut de Génétique Humaine, UPR CNRS 1142, Université Montpellier 2, Montpellier Cedex 5, France*
[c] *Dipartimento di Genetica e Microbiologia, Universita' degli Studi di Bari, Bari, Italy*
[d] *Dipartimento di Sanità Pubblica e Zootecnia, Universita' degli Studi di Bari, Bari, Italy*

ARTICLE INFO

Article history:
Received 7 May 2009
Accepted 11 May 2009
Available online 17 June 2009

Keyword:
Canine genome
TRG locus
IMGT

ABSTRACT

Analyzing the recent high-quality genome sequence of the domestic dog (*Canis lupus familiaris*), we deduced for the first time in a mammalian species belonging to Carnivora order, the genomic structure and the putative origin of the TRG locus. New variable (TRGV), joining (TRGJ) and constant (TRGC) genes for a total of 40 are organized into eight cassettes aligned in tandem in the same transcriptional orientation, each containing the basic recombinational unit V-J-J-C, except for a J-J-C cassette, that lacks the V gene and occupies the 3′ end of the locus. Amphiphysin (AMPH) and related to steroidogenic acute regulatory protein D3-N-terminal like (STARD3NL) genes flank, respectively, the 5′ and 3′ ends of the canine TRG locus that spans about 460 kb. Moreover LINE1 elements, evenly distributed along the entire sequence, significantly (20.59%) contribute to the architecture of the dog TRG locus. Eight of the 16 TRGV genes are functional and belong to 4 different subgroups. Canine TRGJ genes are two for each cassette and only seven out of 16 are functional. The germline configuration and the exon–intron organization of the 8 TRGC genes was determined, six of them resulting functional. The dot plot similarity genomic comparison of human, mouse and dog TRG loci highlighted the occurrence of reiterated duplications of the cassettes during the dog TRG locus evolution. On the other hand the low ratio of functional genes to the total number of canine TRG genes (21/40), suggest that there is no correlation between the extensive duplications of the cassettes and a need for new functional genes. Furthermore the comparison revealed that the TRGC6, C7 and C8 genes are highly related across species suggesting these existed before the primate–rodent–canidae lineages diverged.

1. Introduction

In vertebrates, the T cell has a membrane bound T cell receptor (TR) that is responsible for antigen recognition in T cell-mediated immune response. There are two types of T cell populations based upon their heterodimeric receptors (αβ and γδ). The TR chains are derived from variable (V), joining (J), diversity (D), and constant (C) genes. One of each of these genes is randomly selected from the germline gene pool by a recombination mechanism to generate a wide diversity of TR for antigen recognition (Davis, 1990). The γδ TR shares many structural features common to the αβ TR; however, γδ T cells also have several unique features. For example, in mammals, γδ T cells localize to epithelial and mucosal sites (Bucy et al., 1989; Itohara et al., 1990; Hayday, 2000) and can bind antigens directly in the same way as immunoglobulins (IG) (Hayday, 2000). Furthermore, it was recently reported that γδ T cells can act as antigen presenting cells (Brandes et al., 2005), and it is suggested that γδ T cells work as a bridge between the innate and the adaptive immune system (Konigshofer and Chien, 2006). The γδ T cells compose less than 5% of peripheral blood T cell population in human and mouse (Janeway et al., 1988); however, in birds (Sowder et al., 1988; Kubota et al., 1999) and artiodactyls (Mackay and Hein, 1989; Hein and Dudler, 1993), γδ T cells compose up to 40% of circulating lymphocytes. The genomic organization of the locus encoding for the TR γ (TRG) chain is the most considerably different across species and seems to be related to the evolution of the species. The human TRG locus spans over 160 kilobases (kb) and features in its 5′ region

Abbreviations: IMGT: IMGT®, The international ImMunoGeneTics information system®; TRG, T cell receptor gamma; V-J-J-C, cassette [genomic unit in immunoglobulin (IG) or T cell receptor (TR) loci]; V, variable gene; J, joining gene; C, constant gene.

* Corresponding author at: Dipartimento di Genetica e Microbiologia, Universita' degli Studi di 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).

0161-5890/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2009.05.008

Table 1a
Description of TRGV genes of the dog genome. The position of all genes in the NW_876265 contig and their classification are also reported.

TRGV subgroup	TRGV gene name	TRGV allele name	Functionality	Positions	TRGV[a] size (bp)
TRGV1	TRGV1-1	TRGV1-1*01	P	11267189–11267633	445
TRGV2	TRGV2-1	TRGV2-1*01	F	11316511–11316957	447
	TRGV2-2	TRGV2-2*01	F	11371030–11371476	447
	TRGV2-3	TRGV2-3*01	F	11429935–11430381	447
	TRGV2-4	TRGV2-4*01	F	11485658–11486104	447
TRGV3	TRGV3-1	TRGV3-1*01	P	11325661–11326099	439
	TRGV3-2	TRGV3-2*01	P	11379568–11380007	440
	TRGV3-3	TRGV3-3*01	P	11437299–11437733	435
TRGV4	TRGV4-1	TRGV4-1*01	F	11505128–11505566	439
TRGV5	TRGV5-1	TRGV5-1*01	P	11529092–11529407	316
	TRGV5-2	TRGV5-2*01	F	11587592–11588210	619
TRGV6	TRGV6-1	TRGV6-1*01	P	11567060–11567493	434
TRGV7	TRGV7-1	TRGV7-1*01	P	11645962–11646423	462
	TRGV7-2	TRGV7-2*01	F	11649896–11650363	468
	TRGV7-3	TRGV7-3*01	F	11662469–11662940	472
TRGV8	TRGV8-1	TRGV8-1*01	P	11443810–11444257	448

[a] From L-PART1 to 3′ end of V-REGION.

fourteen variable genes, of which only six are functional, grouped into six subgroups according to the identity values (Lefranc et al., 1989; Lefranc and Rabbitts, 1990). In contrast in mice, TRG locus features four V-J-C cassettes and has a total length of about 200 kb on chromosome 13. The TRGC1 cassette, which correlates to human, has four V genes, all of which are functional, one J gene (TRGJ1) and one C gene (TRGC1) (IMGT Repertoire, http://www.imgt.org); the TRGC3, TRGC2 and TRGC4 cassettes all consist of one V, one J and one C, however, the TRGC3 cassette is not functional due to TRGC3 being a pseudogene. Genes of the TRGC2 cassette display a transcriptional orientation inverted with respect to the other three cassettes (Vernooij et al., 1993). In ruminants (cow and sheep), two TRG loci exist. The TRG1 locus maps in 4q3.1 and spans about 150 kb, the TRG2 locus maps in 4q2.2 and spans about 90 kb. Both consist of three tandemly repeated V J-J C cassettes (Miccoli et al., 2003; Conrad et al., 2007; Vaccarelli et al., 2008). The mammalian genomes that have been intensely analysed to date (human, mouse and ruminants) represent only three out of the seven orders of placental mammals. The recent high-quality draft genome sequence of the domestic dog (*Canis lupus familiaris*) (Lindblad-Toh et al., 2005) allowed us to infer for the first time in a mammalian species belonging to Carnivora order, the genomic structure and the putative origin of the TRG locus.

2. Materials and methods

2.1. Analysis of the genome

To determine TRG locus location, the dog whole genome assembly Build 2.1 was searched using the BLAST algorithm. A sequence of 482745 bp (gaps excluded) was retrieved directly from the reference sequence NW_876265 available at GenBank from positions 11256897 to 11741240. Particularly, the analysed region comprises contigs from AAEX02028365 to AAEX02028369. Locations of TRG genes are provided in Tables 1a–c.

The positions of canine sequences identical to the ones (positions 140573–141360 of the contig [GenBank: NG_001336] [GenBank: AC006033]) containing the human TRG locus enhancer (located downstream of C2 gene) were: 11301467–11302166 (3′TRGC1), 11358509–11359209 (3′TRGC2), 11410720–11411419 (3′TRGC3), 11475080–11475779 (3′TRGC4), 11495612–11496301 (5′TRGV4-1), 11553091–11553790 (3′TRGC5), 11609988–11610519 (3′TRGC6), 11688188–11688887 (3′TRGC7) and 11717021–11717720 (3′TRGC8).

The positions of STARD3NL canine gene were: 11718491–11741236.

Table 1b
Description of TRGJ genes of the dog genome. The position of all genes in the NW_876265 contig and their classification are also reported.

TRGJ gene name	TRGJ allele name	Functionality	Positions	TRG J-REGION size (bp)
TRGJ1-1	TRGJ1-1*01	P	11277597–11277662	66
TRGJ1-2	TRGJ1-2*01	F	11279970–11280028	59
TRGJ2-1	TRGJ2-1*01	P	11334867–11334932	66
TRGJ2-2	TRGJ2-2*01	F	11337664–11337722	59
TRGJ3-1	TRGJ3-1*01	P	11389757–11389822	66
TRGJ3-2	TRGJ3-2*01	ORF	11392112–11392170	59
TRGJ4-1	TRGJ4-1*01	P	11447273–11447338	66
TRGJ4-2	TRGJ4-2*01	F	11449523–11449581	59
TRGJ5-1	TRGJ5-1*01	P	11532350–11532410	61
TRGJ5-2	TRGJ5-2*01	F	11534805–11534863	59
TRGJ6-1	TRGJ6-1*01	F	11591075–11591134	60
TRGJ6-2	TRGJ6-2*01	ORF	11594105–11594165	61
TRGJ7-1	TRGJ7-1*01	P	11672308–11672370	63
TRGJ7-2	TRGJ7-2*01	F	11673958–11674017	60
TRGJ8-1	TRGJ8-1*01	F	11699329–11699388	60
TRGJ8-2	TRGJ8-2*01	P	11704089–11704146	58

Table 1c
Description of TRGC genes of the dog genome. The position of all genes in the NW_876265 contig and their classification are also reported.

TRGC gene name	TRGC allele name	Functionality	Exons	Positions	TRGC exons size (bp)
TRGC1	TRGC1*01	ORF	EX1	11287308–11287625	318
			EX2B	11297844–11297891	48
			EX3	11299806–11299948	140
TRGC2	TRGC2*01	F	EX1	11344400–11344726	327
			EX2A	11346617–11346742	126
			EX2B	11355665–11355712	48
			EX3	11356851–11356993	140
TRGC3	TRGC3*01	F	EX1	11396625–11396954	330
			EX2A	11398940–11399065	126
			EX2B	11406404–11406451	48
			EX3	11408321–11408463	140
TRGC4	TRGC4*01	F	EX1	11457580–11457906	327
			EX2A	11461197–11461322	126
			EX2B	11471197–11471244	48
			EX3	11472871–11473013	140
TRGC5	TRGC5*01	F	EX1	11539782–11540111	330
			EX2A	11542050–11542175	126
			EX3	11550601–11550740	137
TRGC6	TRGC6*01	P	EX1	11599459–11599788	330
			EX2B	11603023–11603070	48
			EX3	11605557–11605699	140
TRGC7	TRGC7*01	F	EX1	11678046–11678375	330
			EX2B	11681862–11681909	48
			EX3	11683775–11683917	140
TRGC8	TRGC8*01	F	EX1	11706758–11707087	330
			EX2B	11710599–11710634	36
			EX3	11712964–11713106	140

2.2. Sequence analysis

Computational analysis of dog TRG locus was conducted with the following programs: RepeatMasker for the identification of genome-wide repeats and low complexity regions (Smit, A.F.A., Hubley R., Green P., RepeatMasker at http://repeatmasker.org), Pipmaker (Schwartz et al., 2000; http://pipmaker.bx.psu.edu/pipmaker/) for the alignment of the dog sequence with the human (GenBank accession number NT_007819: positions 37690000–38240000) and mouse (GenBank accession number NT_039578: positions 8103509–8314280) counterparts. RepeatMasker screens DNA sequences for interspersed repeats and low complexity DNA sequences. The RepeatMasker analysis has been carried out using the Carnivora Repbase division section.

2.3. Nomenclature

Dog TRG genes were named following the IMGT nomenclature established for human and mouse (IMGT®, http://www.imgt.org): TRGV genes are grouped into 8 different subgroups on the basis of the percentage of nucleotide identity. TRGC are numbered (1–8) according to their location from 5′ to 3′ in the locus; TRG cassettes are named according to the constant genes. TRGJ genes are named according the cassette to which they belong.

3. Results and discussion

3.1. Genomic organization of dog TRG locus

Using data available through public databases (Lindblad-Toh et al., 2005) we were able to establish the genomic structure of the entire dog TRG locus. 482745 bp from NW_876265 contig were retrieved, and classification of annotated TRG genes and their organization on chromosome 18 was determined. Genomic sequences were identified by BLAST search using both the human TRG genes and the recombination signal (RS) sequences, and were manually annotated. 40 new (8 TRGC, 16 TRGJ and 16 TRGV) genes organized into eight cassettes aligned in tandem were identified, each containing the basic recombinational unit V-J-J-C, except for the last cassette J-J-C, that lacks the V gene and occupies the 3′ end of the locus. Amphiphysin (AMPH) and Related to steroidogenic acute regulatory protein D3-N-terminal like (STARD3NL) genes flank, respectively, the 5′ and 3′ ends of the TRG locus that spans 460 kb. In particular, the STARD3NL gene (positions 459996–482743) lies 5.2 kb downstream the last TRGC gene, in an inverted transcriptional orientation. We first analysed the compositional properties (G + C content) and identified interspersed repeated sequences. The GC content of 39.4% is almost identical to human (40.6%) and mouse (40.4%). With regard to the interspersed repeated elements, the analysis carried out through the Repeat Masker program provided the results summarized in additional Tables a (dog TRG), b (human TRG) and c (mouse TRG) (Suppl.1). The density of total interspersed repeats ranges from 29.83% in human to 39.65% in mouse with 32.94% mean value in dog. LINEs, predominantly of the LINE1 family, are the most abundant repeat elements both in dog (20.59%) and mouse (29.80%). On the other hand SINEs are the most abundant elements in human (13.95%) and only account for 2.85% of the entire locus in mouse.

Fig. 1 shows the overall organization of the dog TRG locus which encompasses eight cassettes, named according to the constant genes. All TRG cassettes lie in the same transcriptional orientation and are closely spaced. The limit dividing the eight cassettes from each other was approximately given by a space ranging from 10 to 18 kb (from the last exon of each C gene to the V gene of the downstream cassette), except one of about 35 kb between cassettes 6 and 7. Based on the percentage of nucleotide identity and on the genomic position of the genes within the locus, each TRGV, TRGJ and TRGC gene was classified and the nomenclature established according to IMGT® (Fig. 1, Tables 1a–c).

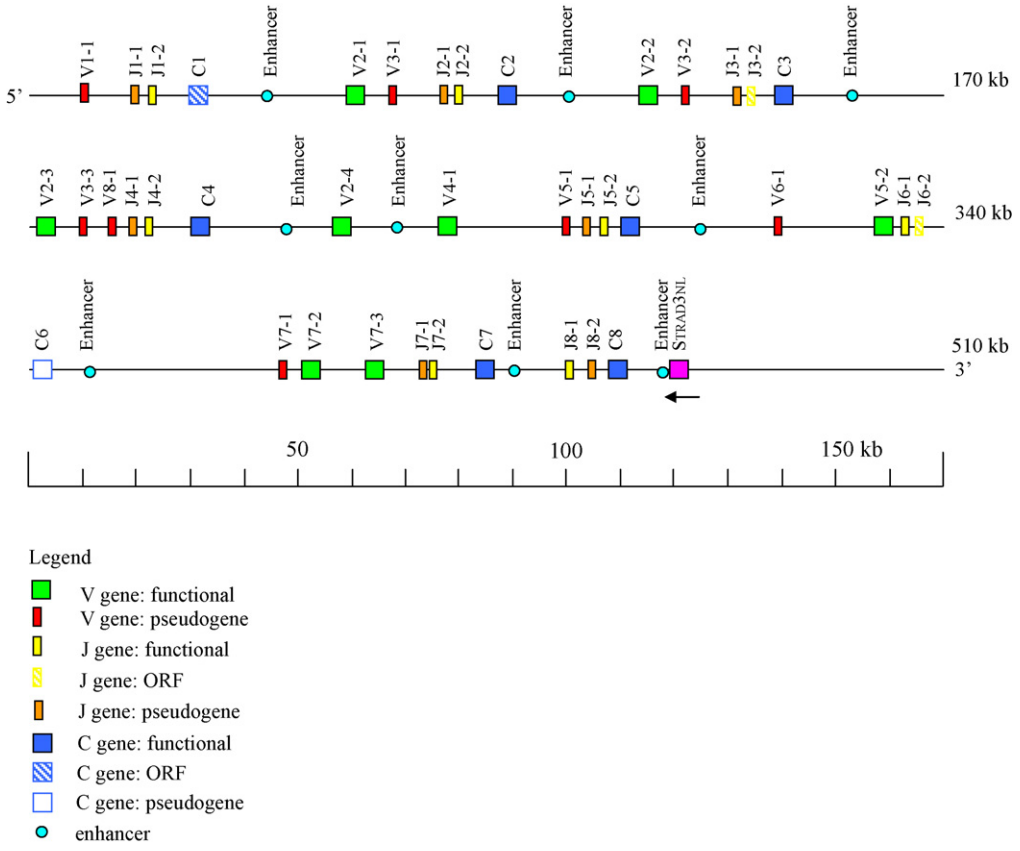

Fig. 1. Schematic representation of the genomic organization of the dog TRG locus as deduced from the genome assembly Build 2.1. The diagram shows the position of all V, J and C TRG genes indicated in according to IMGT® nomenclature. Boxes representing genes are not to scale. Exons are not shown. Enhancer-like sequences at the 3′ end of each TRGC gene derived from the dot-blot analysis (see results Section 3.4) have been inserted. As there is no "ORF" V gene, the symbol is not shown in the legend.

3.2. Classification of canine TRGV, TRGJ and TRGC genes

The functionality of V, J and C genes was predicted through the manual alignment of sequences adopting the following parameters: (a) identification of the leader sequence at the 5′ of TRGV genes; (b) determination of proper recombination signal (RS) sequences located at 3′ and 5′ ends of TRGV and TRGJ genes, respectively; (c) determination of correct acceptor and donor splicing sites; (d) estimation of the expected length of the coding regions; (e) determination of the functionality of the coding regions (absence of frameshifts and stop signals in functional genes). Results are shown in Tables 1a–c. Forty genes were identified, 16V, 16J and 8C. Eight of the 16 TRGV genes are functional and belong to 4 different subgroups (Fig. 2a). The potential TRGV canine repertoire would comprise 4 genes belonging to the TRGV2 subgroup, 1 gene to the TRGV4 and TRGV5 subgroups, and 2 genes to the TRGV7 subgroup (Table 1a, Fig. 2a). Eight other genes are pseudogenes.

Canine TRGJ genes (Table 1b), identified also for J core motifs, are two for each cassette, in the same transcriptional orientation as the relevant C. Seven of the 16 TRGJ genes are functional, 7 are pseudogenes and 2 are ORF. The functional TRGJ are typically 59 or 60 bases in length and are flanked by RS at the 5′ end and by an RNA splice site at the 3′ end. On the physical map, five of them occupy the genomic position closest to the C gene in the J-J-C block (Fig. 1). Pseudogenes and ORF TRGJ vary in length from a minimum of 58 to a maximum of 66 bp.

The germline configuration and the exon–intron organization of the 8 TRGC genes (TRGC1–TRGC8) was determined. Six of them (TRGC2 to TRGC5, TRGC7 and TRGC8) are functional, whereas TRGC1 is an open reading frame (ORF), and TRGC6 is a pseudogene. The first exon (EX1) encodes the C domain that comprises 110 amino acids or is slightly shorter (105 aa for TRGC1, 109 aa for TRGC2 and TRGC4) (Table 1c, Fig. 2b). The first part of the connecting region is encoded by one or two exons (EX2A and/or EX2B), a situation that reminds the EX2 polymorphism observed in the human TRGC2 gene (Buresi et al., 1989), and in bovine (Takeuchi et al., 1992) and sheep (Miccoli et al., 2003) TRGC genes. Thus, the canine TRGC2, C3 and C4 genes have both EX2A and EX2B exons (encoding 42 and 16 aa, respectively), whereas the TRGC1, C6, C7 and C8 genes have only a single EX2B exon (16 aa or, for TRGC8, 12 aa) and TRGC5, a single EX2A exon (42 aa). The remaining part of the connecting region (16 aa), the transmembrane region (24 aa) and the cytoplasmic region (7 aa or, for TRGC5, 6 aa) are encoded by EX3 (47 aa or, for TRGC5, 46 aa). In contrast to the bovine and ovine TRG loci, the extensive duplication of the TRG cassettes does not seem to match a real need of the adaptive immune response. Indeed the reiterated cassette duplication in canine TRG locus resulted in a total of 40 genes, with 21 of them functional and 19 pseudogenes or ORF. On the contrary, in ovine TRG locus, out of 32 genes, 24 genes are functional with only 7 pseudogenes (Vaccarelli et al., 2008).

3.3. Genomic architecture of dog TRG locus is invaded by LINE1

Fig. 3 shows the alignment of the dog, human and mouse masked sequences as percentage identity plot (pip) (Schwartz et al., 2000). The position of all canine identified V, J and C genes, together with the location and orientation of the interspersed repeats, is represented in the pip. It is noteworthy to note how LINE1 elements, evenly distributed along the entire sequence, significantly contribute to the architecture of the dog TRG locus. The L1 family of non-LTR retrotransposons has literally infested mammalian genomes account for 19% of the genome in mouse and for 17% of

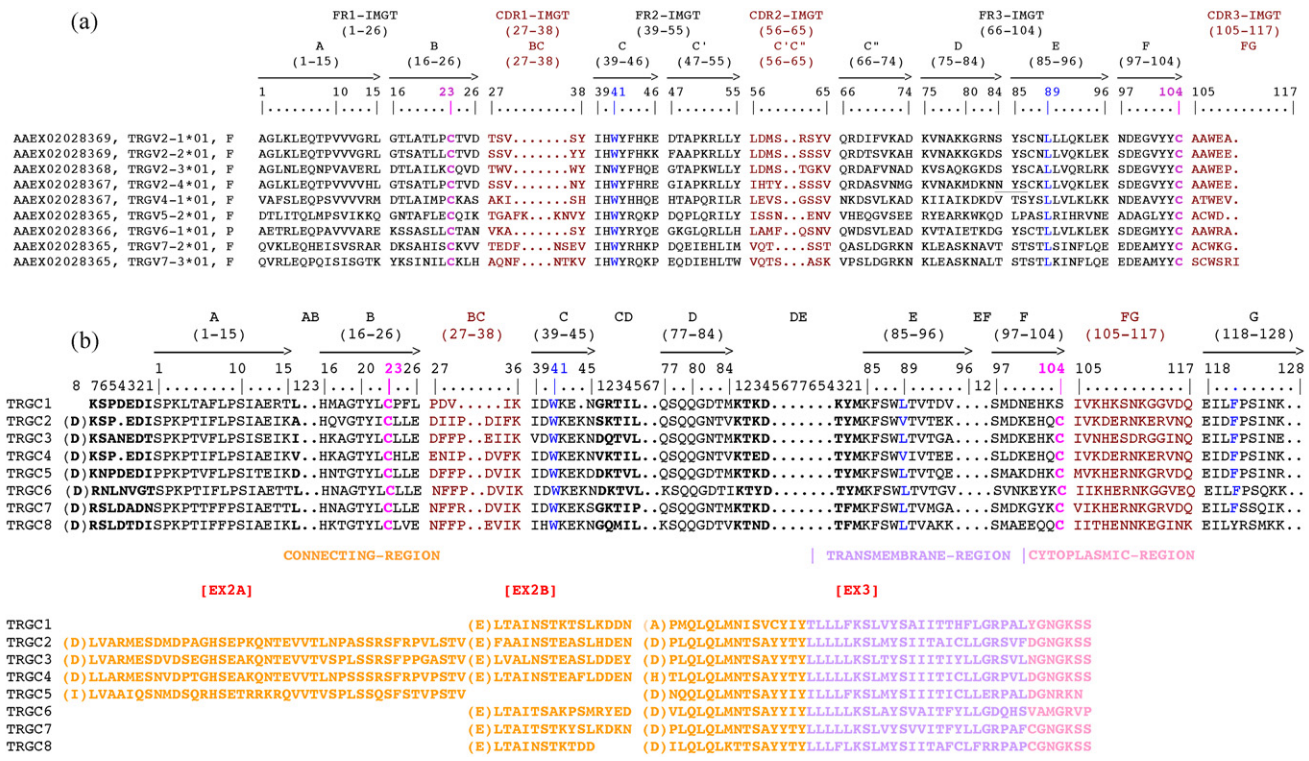

Fig. 2. IMGT Protein display of the dog TRGV and TRGC genes. (a) TRGV genes. Description of the strands and loops is according to the IMGT unique numbering for V-DOMAIN (Lefranc et al., 2003). (b) TRGC genes. Description of the strands and loops is according to the IMGT unique numbering for C-DOMAIN (Lefranc et al., 2005).

the genome in human (Böhne et al., 2008). So far we are not aware of the corresponding percentage of the genome in dog. In mouse and man TRG loci there seems to be no correlation concerning the percentage of LINE1, since a considerable presence of it in mouse (29.80%) is matched by a moderate presence in man (7.56%) (Suppl. 1). Furthermore in the dog, human and mouse TRG locus comparison, the absence of similarity (not necessarily correlated to the presence of LINE1), is evident in the region between the first and the third exon of TRGC1, C2, C3, C4 and C5 genes (boxes). Furthermore the comparison revealed that the TRGC6, C7 and C8 genes are highly related across species suggesting these existed before the primate–rodent–canidae lineages diverged (Fig. 3).

3.4. Comparison of human and canine TRG loci

Fig. 4 shows the dotplot matrix comparing the sequence of dog locus against human. As expected for loci consisting of a series of clusters of correlated genes, the matrixes show numerous homologous regions of various sizes. When comparing the dog genomic organization described above to the human TRG locus (IMGT®, the international ImMunoGeneTics information system®, http://www.imgt.org) (Lefranc et al., 1989; Lefranc and Rabbitts, 1990; Lefranc and Lefranc, 2001), we find that in dog a significant number of duplicative events has led to a substantial increase in the number of V, J and C genes, while duplications in human TRG locus mainly concern the V genes, and more particularly, the TRGV1 subgroup.

PipMaker dotplot matrix of the dog TRG locus, against the human TRG locus, displays two identity diagonals in line with the J-C regions of each canine cassette (blue rectangles). This result reflects the organization of the two J-C regions in man, which are superimposed “end to end”: the first consists of three J genes (JP1, JP and J1) and one C gene (C1) while the second consists of two J genes (JP2 and J2) and one C gene (C2), all functional, the longest diagonals referring to block JP2-J2-C2. The apparent greater length is due to the small diagonals located downstream of each canine C and corresponding to the single human enhancer sequence (yellow circles in Fig. 4) (see accession number and positions in the sequence for the human TRG enhancer sequence in Section 2). From now on this sequence will be referred to as the enhancer-like sequence.

Moreover canine J-C5, J-C1 and J-C6 share longer similarity regions with human J-C2 at their 3′, due to the presence, though partial, of the STARD3NL gene in the relevant cassettes (violet squares). On the other hand, the canine STARD3NL gene at the 3′ end of the locus is recognized in all of its length by the human ortholog, while it seems to have been lost in the J-C7 block and is only marginally detected as similarity spots in TRGC2, TRGC3 and TRGC4 cassettes (violet squares). The analysis of the matrix suggests a duplicative mechanism that involved either the entire V-J-J-C unit, or single V genes (TRGV2-4 and TRGV4-1) as in the TRGC5 cassette, where spots of similarity with the enhancer-like element and the STARD3NL gene are evident. It seems that TRGV2-4 originated through an extended duplication that mostly concerned its 3′, containing TRGV4-1 preceded by an enhancer-like element.

These observations and preliminary data of comparative genomics involving other mammalian species (in preparation) suggest how early duplicative events gave rise to the same locus thus imprinting an expansion rhythm starting from its 3′ end.

The nine variable genes of the human TRGV1 subgroup, all located in the 5′ part of the locus, seem to display no similarity traits with the canine TRGV genes (Fig. 4). The single TRGVA (pseudogene) features one similarity diagonal, including both its coding and non-coding parts, against five functional canine genes (TRGV2-1, V2-2, V2-3, V2-4 and V4-1) and one pseudogene (TRGV6-1) (green squares in Fig. 4). The human TRGVB (pseudogene) and TRGV10 and TRGV11 (ORF) recognize the areas corresponding to the canine TRGV1-1, V3-1, V3-2, V3-3, V7-1, V7-2 and V7-3, of which only the last two are functional (red rectangles in Fig. 4). The total lack of similarity was evident for canine TRGV8-1, V5-1 and V5-2 genes. The concomitant presence of a line of similarity with the absence

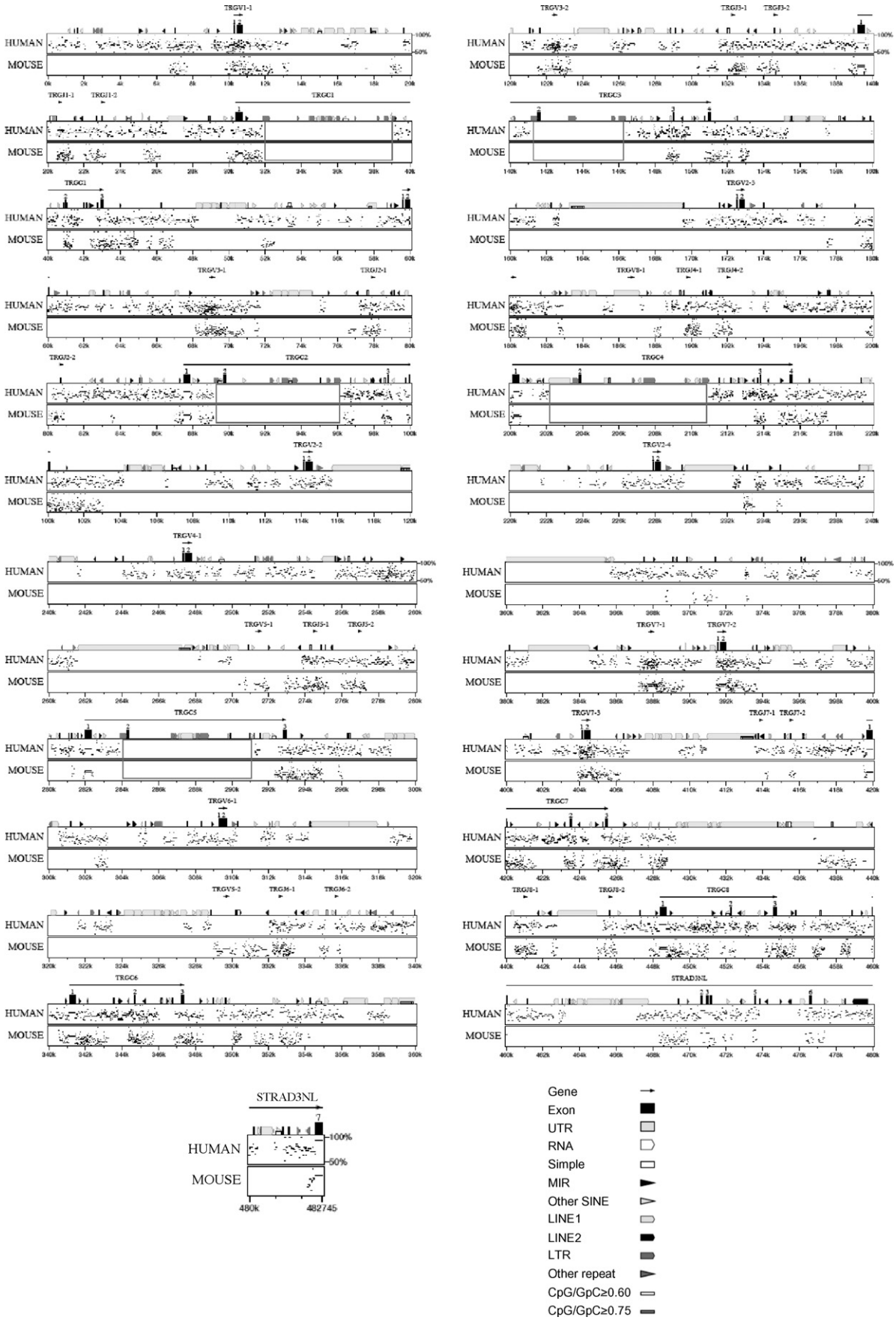

Fig. 3. Pip of the dog, human and mouse TRG locus comparison. The dog sequence is shown on the horizontal axis together with the position and orientation of all genes and repetitive sequences. Horizontal lines represent ungapped alignments at the percentage similarity corresponding to the scale on the right to the human (top plot) and mouse (lower plot) sequences. Boxes indicate the absence of similarity within intronic regions of TRGC1, C2, C3, C4 and C5 genes (see results Section 3.3).

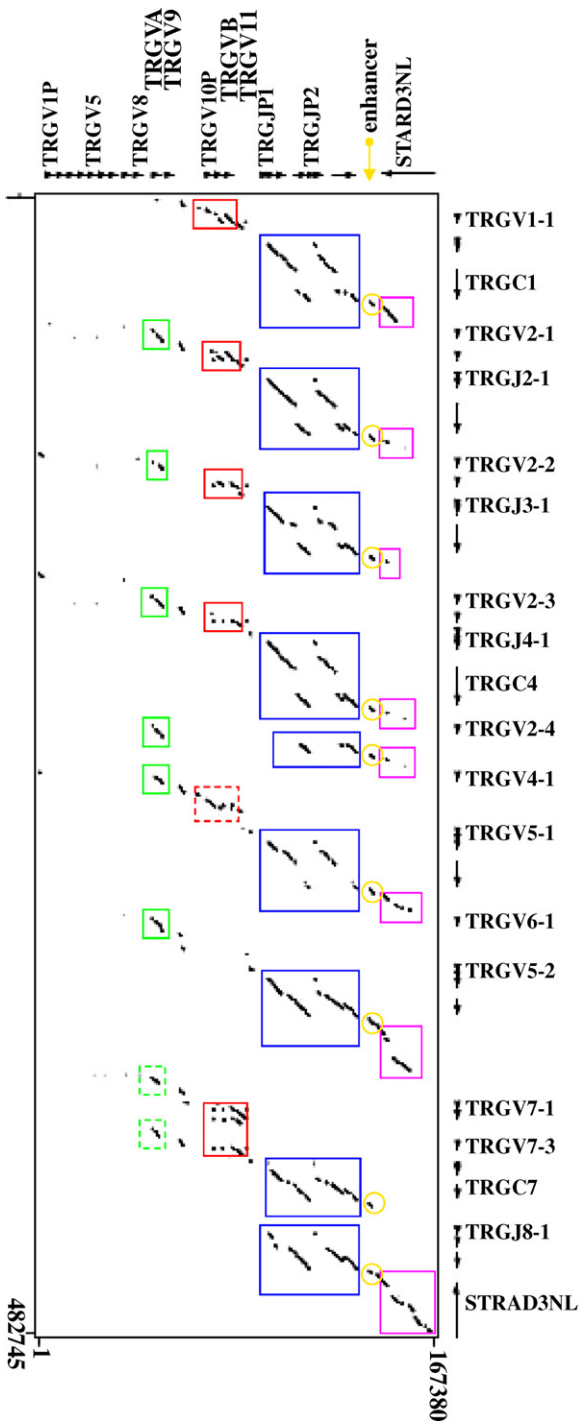

Fig. 4. Dotplot matrix of canine/human TRG genomic comparison. Using the PipMaker program canine TRG has been plotted against human. The transcriptional orientation of each gene is indicated by arrows and arrow-heads. Coloured rectangles enclose J-C blocks (blue), TRGV genes (red and green) and STARD3NL gene (violet). Intermittent line squares (green and red) indicate line of similarity of human TRGV gene in the absence of gene content in the canine sequence. Yellow circles indicate the position of the enhancer-like sequences in dog TRG locus (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.).

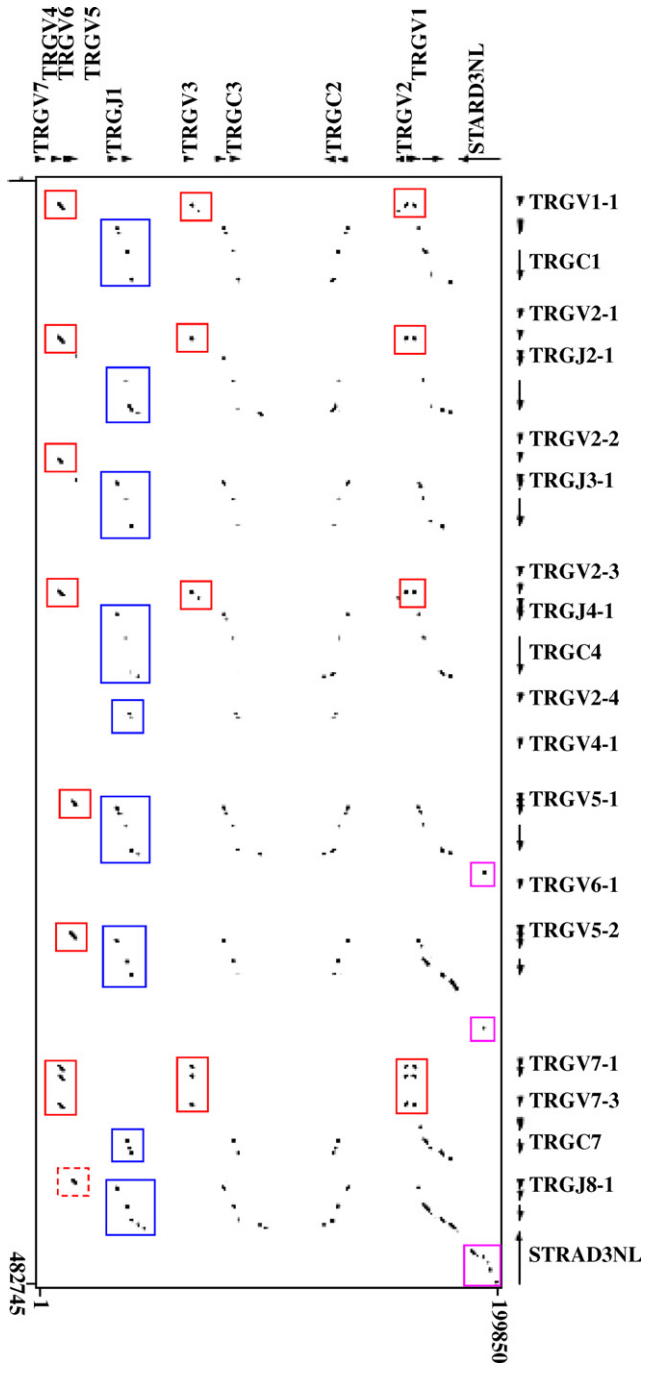

Fig. 5. Dotplot matrix of canine/mouse TRG genomic comparison. Using the PipMaker program canine TRG has been plotted against human. The transcriptional orientation of each gene is indicated by arrows and arrow-heads. Coloured rectangles enclose J-C blocks (blue), TRGV genes (red) and STARD3NL gene (violet). For sake of simplicity only J-C blocks in dog/mouse TRGC1 cassette comparison are indicated in blue rectangles. Intermittent line squares (red) indicate line of similarity of mouse TRGV gene in the absence of gene content in the canine sequence (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.).

of gene content is indicated with intermittent line squares. When looking at the human vs. canine matrix, the importance that the J-C blocks must have had during the evolution of these loci emerges clearly. Indeed, these regions, by preserving their intergenic portions, behave quite differently from other regions including the V genes, for which a strong intra-species and inter-species sequence divergence has been reported (Vaccarelli et al., 2005).

3.5. Comparison of mouse and canine TRG loci

The genomic organization of the murine TRG locus [GenBank: accession number NT_039578: positions 8103509–8314280]

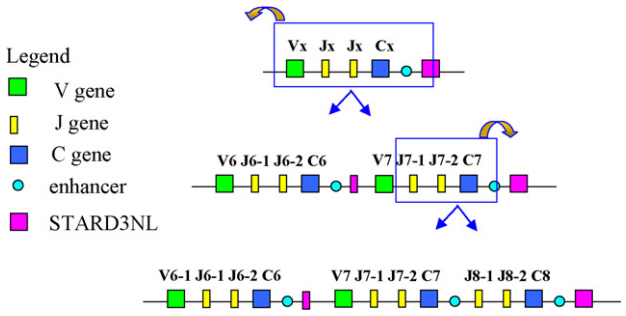

Fig. 6. Hypothetic evolutionary model explaining the origin of dog TRG locus. The genes are represented (not in scale) as small colored boxes; the enhancer-like sequences and STARD3NL gene are indicated. Rectangles are used to indicate the regions involved in duplicative events. Duplications of a cassette are indicated by arrows. The ancestral TRG locus consisted of the TRGC6 and TRGC7 forerunners cassettes, lying at the 3′ end of the locus. TRGC6 and TRGC7 cassettes were originated by a duplication containing the enhancer-like element and part of STARD3NL gene.

(Vernooij et al., 1993) is apparently similar to that of canidae, as it features four V-J-C cassettes and a sequence putatively serving as an enhancer is found downstream of each C gene. However, the resemblance to the dog genomic organization is only apparent because in PipMaker dotplot matrix obtained by comparison of dog TRG locus with each murine TRG cassette, only short similarity traits corresponding to exonic regions are evident (Fig. 5). In the blue rectangles for each canine cassette three spots are indicated, which correspond to the J genes and to the first and last exon of the constant genes, respectively.

The duplications of the 3′ region of the STARD3NL gene are indicated. The similarity lines sharing dog/mouse variable genes corresponding to the mouse TRGC1 cassette are boxed in the figure. In particular the functional murine TRGV4 gene recognizes the areas corresponding to the canine TRGV1-1, V3-1, V3-2, V3-3, V7-1, V7-2 and V7-3, of which only the last two are functional (red squares in Fig. 5). Moreover, the functional murine TRGV6 is recognized by the canine TRGV5-1 (pseudogene) and TRGV5-2 (functional). The similarity relationship between V genes in a three species comparison (Figs. 4 and 5; Suppl. 2) highlights the absence in the murine locus of the TRGVA and TRGV9 human genes and the absence in the canine locus of the nine variable genes of the human TRGV1 subgroup. Instead, the latter are recognized by the murine TRGV7 functional gene; likewise, the human TRGVB (pseudogene) and TRGV10 and TRGV11 (ORF) are recognized by the murine TRGV4 functional gene (Suppl. 2).

Matrices in Figs. 4 and 5 led us to assume the origin of dog TRG locus (Fig. 6). We suggest that after the duplication of a minimum ancestral cassette consisting of one V, two J and one C genes, the ancestral TRG locus consisted of two cassettes, which probably were the forerunners of the TRGC6 and TRGC7 cassettes. Subsequently a duplication of the J-J-C region containing the 3′C7 sequence inclusive of the enhancer-like element but excluding the STARD3NL gene gave rise to the TRGC8 cassette. This hypothesis is supported also by data presented in Fig. 3.

4. Conclusion

The recent high-quality draft genome sequence of the domestic dog has allowed us to define the genomic structure and the putative origin of the TRG locus taken into consideration for the first time in an organism (*C. lupus familiaris*) belonging to Carnivora order. The locus extension is more significant for the number of kilobases (approximately 500 on chromosome 18) compared to the orthologs (human, mouse, sheep and cow) analysed so far. The structural organization includes 40 (16 TRGV, 16 TRGJ and 8 TRGC) genes organized into eight cassettes aligned in tandem, each containing the basic recombinational unit V-J-J-C, except for the most downstream cassette that lacks the V gene. Moreover LINE1 elements, evenly distributed along the entire sequence, significantly contribute (20.59% of the entire sequence) to the architecture of the dog TRG locus. The dot plot similarity genomic comparison of human, mouse and dog TRG loci highlighted the occurrence of reiterated duplications of the cassettes during the dog TRG locus evolution. Detailed analysis of the contig and gene structures and the low ratio of functional genes to the total number of canine TRG genes (21/40), suggest that there is no correlation between the genomic duplications of the cassettes and a need for new functional genes. Results presented here do not allow dog to be classified as a γδ high species rather than a γδ low species. To this end, it is clear that functional and molecular assays are necessary, which are aimed at identifying the percentage of γδ cells circulating in peripheral blood. Comparative genomic and phylogenetic analyses with other species of mammals in which the complete sequence of the TRG locus is available are also necessary to further clarify the yet not well defined role of γδ cells in the cell-mediated immune response.

Acknowledgement

The financial support of Ministero per l'Università e la Ricerca Scientifica e Tecnologica (MIUR) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molimm.2009.05.008.

References

Böhne, A., Brunet, F., Galiana-Arnoux, D., Schultheis, C., Volff, J.N., 2008. Transposable elements as drivers of genomic and biological diversity in vertebrates. Chromosome Res. 16, 203–215.
Brandes, M., Willimann, K., Moser, B., 2005. Professional antigen-presentation function by human gammadelta T cells. Science 309, 264–268.
Bucy, R.P., Chen, C.L., Cooper, M.D., 1989. Tissue localization and CD8 accessory molecule expression of T gamma delta cells in humans. J. Immunol. 142, 3045–3049.
Buresi, C., Ghanem, N., Huck, S., Lefranc, G., Lefranc, M.-P., 1989. Exon duplication and triplication in the human T-cell receptor gamma constant region genes and RFLP in French, Lebanese, Tunisian, and Black African populations. Immunogenetics 29, 161–172.
Conrad, M.L., Mawer, M.A., Lefranc, M.P., McKinnell, L., Whitehead, J., Davis, S.K., Pettman, R., Koop, B.F., 2007. The genomic sequence of the bovine T cell receptor gamma TRG loci and localization of the TRGC5 cassette. Vet. Immunol. Immunopathol. 115, 346–356.
Davis, M.M., 1990. T cell receptor gene diversity and selection. Annu. Rev. Biochem. 59, 475–496.
Hayday, A.C., 2000. Gammadelta 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 repertoires: extensive diversity and developmentally regulated expression of the sheep gammadelta T cell receptor. EMBO J. 12, 715–724.
Itohara, S., Farr, A.G., Lafaille, J.J., Bonneville, M., Takagaki, Y., Haas, W., Tonegawa, S., 1990. Homing of a gamma delta thymocyte subset with homogeneous T-cell receptors to mucosal epithelia. Nature 343, 747–754.
Janeway Jr., C.A., Jones, B., Hayday, A., 1988. Specificity and function of T cells bearing gamma delta receptors. Immunol. Today 9, 73–77.
Konigshofer, Y., Chien, Y.H., 2006. Gammadelta T cells—innate immune lymphocytes? Curr. Opin. Immunol. 18, 527–533.
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., 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., Rabbitts, T.H., 1990. Genetic organization of the human T-cell receptor gamma and delta loci. Res. Immunol. 141, 565–577.
Lefranc, M.-P., Lefranc, G., 2001. The T cell receptor. In: FactsBook. Academic Press, London, pp. 1–398.
Lefranc, M.-P., Pommié, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V., Lefranc, G., 2003. 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., 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., 2005. IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. Dev. Comp. Immunol. 29, 185–203.
Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kama, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., Evan Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E.A., Ponting, C.P., Galibert, F., Smith, D.R., deJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.-W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.-P., Parker, H.G., Pollinger, J.P., Stephen, M.J., Searle, S.M.J., Sutter, N.B., Thomas, R., Webber, C., Lander, E.S., Broad Institute Genome Sequencing Platform, 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803–819.
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.
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.
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.
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.
Takeuchi, N., Ishiguro, N., Shinagawa, M., 1992. Molecular cloning and sequence analysis of bovine T-cell receptor gamma and delta chain genes. Immunogenetics 35, 89–96.
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.
Vaccarelli, G., Miccoli, M.C., Antonacci, R., Pesole, G., Ciccarese, S., 2008. Genomic organization and recombinational unit duplication-driven evolution of ovine and bovine T cell receptor gamma loci. BMC Genomics 9, 81.
Vernooij, B.T., Lenstra, J.A., Wang, K., Hood, L., 1993. Organization of the murine T-cell receptor gamma locus. Genomics 17, 566–574.