Developmental and Comparative Immunology 85 (2018) 31–43

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

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

Analysis of TCRβ and TCRγ genes in Chinese alligator provides insights into the evolution of TCR genes in jawed vertebrates

Xifeng Wang [a, 1], Peng Wang [b, 1], Renping Wang [c], Chaolin Wang [c], Jianhui Bai [b], Cuncun Ke [b], Di Yu [b], Kongpan Li [b], Yonghe Ma [b], Haitang Han [b], Yaofeng Zhao [b], Xin Zhou [a, d, **], Liming Ren [b, *]

[a] Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, 100193, People's Republic of China
[b] State Key Laboratory of Agrobiotechnology, College of Biological Sciences, National Engineering Laboratory for Animal Breeding, China Agricultural University, Beijing, People's Republic of China
[c] Administration Bureau of Chinese Alligator National Nature Reserve Protection, Anhui, People's Republic of China
[d] College of Plant Protection, China Agricultural University, Beijing, People's Republic of China

ARTICLE INFO

Article history:
Received 9 November 2017
Received in revised form
12 January 2018
Accepted 12 January 2018
Available online 21 March 2018

Keywords:
Alligator sinensis
T cell receptor
TRB genes
TRG genes
Comparative genomics
Evolution

ABSTRACT

All jawed vertebrates have four T cell receptor (TCR) chains that are expressed by thymus-derived lymphocytes and play a major role in animal immune defence. However, few studies have investigated the TCR chains of crocodilians compared with those of birds and mammals, despite their key evolutionary position linking amphibians, reptiles, birds and mammals. Here, employing an *Alligator sinensis* genomic bacterial artificial chromosome (BAC) library and available genome data, we characterized the genomic organization, evolution and expression of *TRB* and *TRG* loci in *Alligator sinensis*. According to the sequencing data, the *Alligator sinensis TRB* locus spans approximately 500 Kb of genomic DNA containing two D-J-C clusters and 43 V gene segments and is organized as $V\beta_{(39)}$-pJβ1-pCβ1-pDβ1-Dβ2- $J\beta2_{(12)}$-Cβ2-$V\beta_{(4)}$, whereas the *TRG* locus spans 115 Kb of DNA genomic sequence consisting of 18 V gene segments, nine J gene segments and one C gene segment and is organized in a classical translocon pattern as $V\gamma_{(18)}$—$J\gamma_{(9)}$-Cγ. Moreover, syntenic analysis of *TRB* and *TRG* chain loci suggested a high degree of conserved synteny in the genomic regions across mammals, birds and *Alligator sinensis*. By analysing the cloned *TRB*/*TRG* cDNA, we identified the usage pattern of V families in the expressed *TRB* and *TRG*. An analysis of the junctions of the recombined VJ revealed the presence of N and P nucleotides in both expressed *TRB* and *TRG* sequences. Phylogenetic analysis revealed that *TRB* and *TRG* loci possess distinct evolutionary patterns. Most *Alligator sinensis* V subgroups have closely related orthologues in chicken and duck, and a small number of *Alligator sinensis* V subgroups have orthologues in mammals, which supports the hypothesis that crocodiles are the closest relatives of birds and mammals. Collectively, these data provide insights into TCR gene evolution in vertebrates and improve our understanding of the *Alligator sinensis* immune system.

1. Introduction

T cells play a central role in regulator and effector functions of the vertebrate immune system, and T cell receptors (TCRs) are specifically expressed on cell surfaces to react with antigens presented by major histocompatibility complex molecules (Garcia and Adams, 2005; Reddy, 2017). Four TCR genes are located in four discrete loci on chromosomes, *TRA*, *TRB*, *TRG* and *TRD* (Davis and Bjorkman, 1988), encoding four TCR chains expressed as either an αβ or γδ heterodimer on the surface of T cells (Kronenberg et al., 1986; Rast et al., 1997). The *TRA* and *TRD* genes are located at a single chromosomal location with the *TRD* locus nested within the *TRA* locus, while the *TRB* and *TRG* loci are located at different

* Corresponding author.
** Corresponding author. Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, 100193, People's Republic of China.
E-mail addresses: xinzhou@cau.edu.cn (X. Zhou), renlm@cau.edu.cn (L. Ren).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.dci.2018.01.009
0145-305X/

chromosomal locations (Rast et al., 1997). Furthermore, the fifth *TCR* locus encoding an atypical TCRμ chain was found in non-placental mammals, the opossum *Monodelphis domestica* (Parra et al., 2007, 2008) and duckbill platypus (Parra et al., 2012; Wang et al., 2011), which appears to be distantly related to *TCRδ*. The structures of the *TCRδ* and *TCRμ* genes in tetrapods suggest that the *TCRμ* locus was created by *TCRδ* genes translocated out of the *TCRα/δ* locus early in mammals (Parra et al., 2007, 2012).

In general, the TCR locus exhibits a similar organization to the immunoglobulin locus. The genes encoding the TCRα chain (*TRA*) and TCRγ chain (*TRG*) consist of variable (V) segments, junction (J) segments and a constant (C) region. Moreover, the genes encoding the TCRβ chain (*TRB*) and TCRδ chain (*TRD*) also contain diversity (D) segments between the V and J segments. To obtain a comprehensive repertoire of TR, the loci encode the TCR chain through a random process of DNA rearrangements that involves somatic recombination of V (D) J gene segments and the deletion and insertion of nucleotides at V (D) J junctions during the development of T lymphocytes in the thymus (Bassing et al., 2002; Davis, 1990). The TCRs near the amino-terminus contain three complementarity-determining regions (CDRs), which are vital for antigen recognition. Furthermore, the sequence of CDR3 is more variable than the sequences of the other CDRs because of the presence of the N insertion and deletion in the junction of the V and J segments. In particular, the insertion of D segments in the CDR3 of *TRB* and *TRD* provide a broader repertoire for antigens than *TRA* and *TRG* (Arden, 1998; Borg et al., 2005).

The genomic structure of the *TRB* locus studied in all jawed vertebrates to date has a common feature; it tends to be organized in the translocon style with a pool of *TRBV* genes positioned upstream from D-J-C clusters and a single inverted *TRBV* gene at the 3' end of the last *TCRBC* gene. There are some differences in the number of *TRBV* genes and D-J-C clusters among different species. The human *TRB* locus spans 685 Kb and is largely occupied by an array of 86 TRBV genes belonging to approximately 30 subgroups, followed by two tandemly repeated D-J-C regions (Rowen et al., 1996). Like the human locus, the genomic structure of the *TCRB* loci in mice, rabbits and dogs all have two tandemly repeated D-J-C clusters downstream from various *TRBV* genes (Antonacci et al., 2014; Caccia et al., 1984; Mineccia et al., 2012). Moreover, the genome sequence and its structure around the *TRB* loci in cattle and sheep suggest that duplication events have resulted in the massive expansion of several *TRBV* subgroups and led to the generation of a third D-J-C cluster (Antonacci et al., 2008; Connelley et al., 2009; Conrad et al., 2002). The opossum, a marsupial, has four D-J-C cassettes and 36 TRBV segments that can be grouped into 28 subgroups in the *TRB* locus (Parra et al., 2008), while the chicken has two subgroups of *TRBV* genes and only one D-J-C cluster present in the *TRB* locus (Shigeta et al., 2004). By contrast, the duck *TRB* locus is organized in an unusual pattern that differs from that of the tandem-aligned clusters in mammals or the translocon organization in some teleosts but is similar to the organization in zebrafish. Overall, the organization of the *TRB* locus in duck and zebrafish contains one *TRBD* gene positioned downstream from a pool of V gene segments and has two tandemly repeated J-C clusters located downstream of the *TRBD* gene (Meeker et al., 2010; Yang et al., 2017).

The *TRG* locus is the smallest and least complex of the three conventional TCR loci and is the most considerably different across species. *TRG* loci in some species, such as humans (Glusman et al., 2001; Lefranc et al., 1989), rabbits (Massari et al., 2012), dolphins (Linguiti et al., 2016), opossums (Deakin et al., 2006) and sharks (Chen et al., 2009), are organized in a "translocon" configuration, in which many tandemly duplicated V genes are localized in a chromatin domain followed by arranged J-C clusters. Duplication of the region of V genes has led to a substantial increase in the number of V genes. By contrast, *TRG* loci are organized into several cassettes in tandem, each containing the basic recombinational unit V-J-C, in some species such as cattle, sheep, mice and dogs. It is notable that duplication events of the V, J and C regions occurred in these species. The human *TRG* locus spans 160 Kb and contains 14 V gene segments in its 5′ region, followed by two J-C clusters (Glusman et al., 2001; Lefranc et al., 1989). In contrast, in mice, the *TRG* locus contains four V-J-C cassettes and has a total length of 205 Kb on chromosome 13 (Vernooij et al., 1993). In cattle and sheep, two paralogous *TRG* loci have been found, and both loci consist of three tandemly repeated V-J-J-C cassettes (Vaccarelli et al., 2008). The duck *TRG* locus, which was first studied clearly in birds, spans approximately 100 Kb and exhibits a simple structure with 15 Vγ genes, 5 Jγ genes and a single Cγ gene (Yang et al., 2017). However, two different TCR loci were identified in Atlantic salmon. The first locus, *TRG1*, spans 260 Kb and contains four tandemly repeated clusters, and the second locus, *TRG2*, contains a single V-J-C cluster that was not expressed (Yazawa et al., 2008). The *TRG* of the sandbar shark is present as a single locus arranged in a classic translocon pattern that contains at least five V region genes, three J segment genes, and one C segment (Chen et al., 2009).

As the sister group to mammals, reptiles occupy a key phylogenetic position for understanding the evolution of the TCR. Crocodilians, as reptiles, are only distantly related to lizards and are thought to be the closest living relatives of birds and mammals and thus occupy an important position in evolution (Kumar and Hedges, 1998). In addition, crocodilians live in marshes, lakes, and rivers and often suffer from serious injuries when males fight for mates and females battles for nests. However, they appear to recover quickly from open wounds in water and are thought to have robust immune systems that resist microbial infection (Merchant et al., 2006). Recently, IgH genes of crocodilians were identified; the results indicated that there are multiple μ genes and that IgM subclasses can be expressed through class-switch recombination. The crocodilian α genes were the first IgA-encoding genes identified in reptiles, and their presence suggests that reptiles and birds share a common ancestral organization (Cheng et al., 2013; Magadan-Mompo et al., 2013). Analysis of the IgL gene locus and expression suggests that the IgL gene repertoire is highly diverse and complex in *Alligator sinensis* (Wang et al., 2016). However, knowledge of the *TRB* and *TRG* locus in reptiles is limited because of a lack of available genomic sequence information.

Most recently, Wan et al. reported the 2.3 Gb genome of the Chinese alligator (*Alligator sinensis*) (Wan et al., 2013), but the *TRB* and *TRG* loci were not described in that study. Therefore, in this study, we report the detailed genomic organization and repertoire diversity of *TRB* and *TRG* loci in *Alligator sinensis*, an endangered reptile species, to provide insight into our understanding of the crocodilian immune system and the evolution of *TRB* and *TRG* in vertebrates.

2. Materials and methods

2.1. Sample collection, DNA and RNA extraction and reverse transcription

Blood samples of Chinese alligators (*Alligator sinensis*) were collected from the Beijing Zoo. Genomic DNA was extracted from blood cells according to a routine protocol. Various tissues of *Alligator sinensis* were provided by the Administration Bureau of Chinese Alligator National Nature Reserve Protection of Anhui province. Total RNA was extracted from the spleen and thymus using a TRIzol kit (TIANGEN BIOTECH, Beijing) according to the manufacturer's instructions. Reverse transcription was conducted