The dynamic TCRδ: TCRδ chains in the amphibian Xenopus tropicalis utilize antibody-like V genes

Zuly E. Parra¹, Yuko Ohta², Michael F. Criscitiello², Martin F. Flajnik² and Robert D. Miller¹

¹ Center for Evolutionary and Theoretical Immunology, Department of Biology, University of New Mexico, Albuquerque, NM, USA
² Department of Microbiology and Immunology, University of Maryland at Baltimore, Baltimore, MD, USA

The content and organization of the Xenopus tropicalis TCRα/δ locus was determined. This locus is highly conserved among tetrapods, with the genes encoding the TCRδ chains embedded with those encoding TCRα. However, the frog TCRα/δ is unusual in that it contains V genes that appear indistinguishable from those in the IgH locus (VH). These V genes, termed VHδ, make up 70% of the V genes at the TCRδ locus and are expressed exclusively in TCRδ chains. Finding TCRδ chains that use antibody-like V domains in frogs is similar to the situation in shark TCRδ variants and TCRλ in marsupials. These results suggest that such unconventional TCR may be more widespread across vertebrate lineages than originally thought and raise the possibility of previously unrealized subsets of T cells. We also revealed close linkage of TCRα/δ, IgH, and Igλ in Xenopus which, in combination with linkage analyses in other species, is consistent with the previous models for the emergence of these antigen receptor loci.

Key words: Amphibian • Evolution • Immunoglobulin • TCRδ

Introduction

B and T lymphocytes using somatically diversified antigen-binding receptors mediate the adaptive immune responses in all jawed vertebrates [1]. These receptors contain antigen-binding V domains encoded by exons that normally exist in a segmented form in the germ-line DNA and are assembled by DNA recombination during development. In B cells this recombination takes place at the Ig loci-encoding molecules that serve both as cell-surface BCR as well as secreted antibodies. In the case of T cells, recombination occurs at the genes encoding the TCRα, β, γ, and δ chains, which form part of the cell surface ζβ or γδ TCR complexes [2]. The genes encoding the Ig and TCR chains share similar organization and structural features consistent with their common ancestry [3]. The genes encoding the IgH, TCRβ, and TCRδ chains use V, D, and J gene segments to assemble and encode the V domain, whereas the IgL, TCRα, and TCRγ chains use just V and J. In all cases, these gene segments are flanked by conserved recombination signal sequences (RSS) that are site-specific targets of the endonuclease activity of RAG [4].

The genes encoding the TCRα and δ chains are unique among the loci undergoing V(D)J recombination in several ways. In all tetrapods examined so far, they are interspersed at a single TCRα/δ locus [5–9]. This single locus encodes two chains whose tightly regulated expression is mutually exclusive resulting in distinct T-cell lineages, the ζβ and γδ T cells [10, 11]. In most
cases, TCRβ and δ chains share a common pool of V that, depending on the chain, are recombined to either a DJδ or directly to a Jα segment.

In addition to the complex genetics of the TCRα/δ locus, TCRδ appears to have a high degree of evolutionary plasticity. Approximately one quarter of the shark TCRδ chains are expressed in an alternative isoform called New Antigen Receptor (NAR)-TCR that contains a double V structure [12]. Interestingly, each of the two V domains require V(D)J rearrangement, and the N-terminal V is more similar to the V region of an antibody discovered in the nurse shark called the IgNAR than it is to TCR V. More recently, a novel TCR locus, TCRμ, was discovered in marsupials that shares an evolutionary history with TCRδ [13]. TCRμ also encodes a double V structure and both V genes are also more similar to VH than to TCR V. Marsupials also have a conventional TCRα/δ locus with a prototypic mammalian organization and, therefore, TCRμ is not a substitute for TCRδ in these mammals [9]. However, the C regions of TCRμ do share greatest sequence similarity to Cδ and appear to have been derived from TCRδ, perhaps during the early evolution of amniotes [13]. TCRμ is also found in the duckbill platypus, consistent with its ancient origins and presence in the common ancestor of all living mammals, and thus an orthologue could still be present in some eutherian (placental) mammals, although so far none have been found [9].

The presence of atypical TCRδ forms with similar features in distantly related species such as cartilaginous fish and non-eutherian mammals suggests that they may be found in other vertebrate lineages. So far, surveys of the chicken, lizard, and frog genomes failed to uncover any gene sequences bearing homology to TCRμ [9] (ZEP and RDM personal observations). However, when investigating the genome of an amphibian, Xenopus tropicalis, for evidence of TCR resembling either TCRμ or NAR-TCR, evidence for a third alternative, a variant of TCRδ, was obtained that may give insight into both the emergence and the evolution of Ig and TCR in general and the origins of TCRμ.

**Results**

**Conserved synteny at the X. tropicalis TCRα/δ locus**

As in all tetrapod species analyzed so far, the genes encoding the X. tropicalis TCRs and TCRδ chains are tightly linked, with some TCRδ genes nested among the TCRα (Fig. 1). This genomic region appears stable in tetrapods since the genes flanking the X. tropicalis TCRα/δ locus are the same as in birds and mammals, including the olfactory receptors interspersed among the Vα genes (Fig. 1) [5, 7, 9]. Individual V, D, and J gene segments in the X. tropicalis TCRα/δ locus were annotated using the convention established by the International ImMunoGeneTics database (http://www.imgt.org) and the recommendations of Koop et al. [14]. A total of 71 V gene segments were identified within the TCRα/δ locus, many of which share a high degree of sequence identity to those previously reported in Xenopus [15] (Supporting Information Table 1). Fifty-two V gene segments, in the same transcriptional orientation as the most 3’ C region (Cγ1) are Vα based on nucleotide identity (Figs. 1 and 2). All Vα appeared to be functional based on open reading frames (ORF), an upstream exon encoding a leader sequence, and a canonical RSS. They segregate into 28 subgroups designated Vα3 through Vα30 based on nucleotide identity and phylogenetic relationships (Fig. 2). Also present is a large number of Jα gene segments upstream of Cα1 similar to all mammalian species investigated (75 in the frog compared with 61 in humans, 60 in mice and 53 in the opossum) (Fig. 1 and Supporting Information Fig. 1) [9, 16].

To be consistent with the published Xenopus TCRα sequences, the 3’ most Cα was designated as Cα1 and corresponds to a sequence described previously [15] (Supporting Information Fig. 2). Upstream, within the Vα genes, is a second Cα (Cα2) that shares only 51% nucleotide identity with Cα1. Cα2 appears functional by having an ORF and canonical cysteine residues necessary to form the intra-chain disulfide bond, however it has two, three codon insertions located 5' of each cysteine and two extra cysteines in the β-strand, all of which could affect protein folding (Supporting Information Fig. 2). Attempts to isolate TCRα transcripts containing Cα2 using either 5’ RACE or targeted RT-PCR were unsuccessful using either the cDNA library or the RT-PCR on small intestine, spleen or thymus RNA (data not shown). Furthermore, neither of the two Jα gene (Jα2.1 and 2.2) upstream of Cα2 was identified in transcripts (data not shown).

Within the Vα gene segments is a Cα (G02) in the same orientation as both Cα (Figs. 1 and 3, and Supporting Information Fig. 2). Transcripts containing G02 were not detectable by Northern blot analysis of various tissues including thymus RNA, but could be amplified by 5’ RACE using thymus cDNA (data not shown). All G02 transcripts amplified contained either Vα4 or Vα5, which are the two of the four V genes located immediately upstream of G02 and in the same orientation (Fig. 1, GenBank accession numbers GQ262036–GQ262039). A comparison of the CDR3 sequences from these clones to the germ-line sequence failed to identify sequences corresponding to D segments (data not shown). In addition, all clones amplified contained the same Jα segment (Jα2.1). However, searching the genomic sequence for conserved FR4 (e.g. FGXG) and RSS motifs identified the second Jα (Jα2.2) (Fig. 1 and Supporting Information Fig. 1). Although G02 and Cα2 were found expressed at low levels or not at all they constitute a mini-locus (from Vα3 to Cα2) resembling the organization found in conventional mammalian TCRα/δ loci with V segments upstream of a Jα0-Cα0-Jα-Jα-Cα cluster. Whether this mini-locus represents an ancient or derived TCRα/δ organization remains to be determined.

**An unconventional TCRδ**

At the 5’ end of the Xenopus TCRα/δ locus is a cluster of gene segments in reverse transcriptional orientation, encoding a TCRδ
chain (Fig. 1). A single Cδ (Gδ1), the most 5′ TCR gene in the locus, shares 61% nucleotide identity with Cδ2 (Figs. 1 and 3 and Supporting Information Fig. 2). Present are 19 V genes of which three are Vα (Vα1, Vα2.1, and 2.2) and two are Vδ (Vδ1.1 and 1.2) (Figs. 1 and 2). The three Vα genes belong to two subgroups (Vα1 and Vα2) sharing 63% nucleotide identity and fall into the same phylogenetic clade as X. tropicalis Vα subgroups 10 and 11 (group I, Fig. 2). Phylogenetic analyses revealed that the Vδ fall within a well-supported clade with opossum Vδ2 and Vδ3, human Vδ2 and mouse Vδ4 (group F, Fig. 2). X. tropicalis Vδ gene segments share the highest similarity with opossum Vδ3 (45.9%) when compared with mammalian and avian Vα/δ sequences. X. tropicalis Vδ genes also have an unpaired cysteine residue located in the e-strand of FR3 (Fig. 4). Some of the Vδ gene segments encoding supporting V domains in shark NAR-TCR also contain a cysteine in the e-strand (GenBank accession no. DQ022688 and DQ022691) [12].

The 14 remaining V genes were compared with Ig and TCR V genes from other vertebrates and found to be most similar to VH based on nucleotide identity and phylogenetic analyses. These genes have been designated VHδ due to their similarity to VH but their location in the TCRα/δ locus and exclusive expression with Gδ1 (Figs. 1, 5, 6). VHδ genes segregated into five sub-groups based on 80% or greater nucleotide identity (Fig. 1 and Supporting Information Table 2). Two VHδ (VHδ2 and VHδ3) are pseudo-genes due to in-frame stop codons; however, the rest appear functional. In phylogenetic analyses, the Xenopus VHδ all clustered with clan II VH gene segments (Fig. 5). The VHδ genes form two distinct clades, one containing VHδ1 through 4 and the other VHδ5 only (Fig. 5). VHδ1 through 4 share less than 60% nucleotide identity with the Xenopus clan II VH subgroups, VH5, 9 and 11 [17, 18]. In contrast, the clade that includes VHδ5 also contains VH genes from X. laevis and X. tropicalis (i.e. V that have been identified in IgH transcripts). Indeed, VHδ5 genes in the TCRα/δ locus share a surprisingly high degree of identity (91–96%) with VH5 genes in the IgH locus, consistent with the VHδ5 genes being paralogous to VH5 genes.

Upstream of Gδ1 are three Jδ genes identified by first searching the TCRα/δ genomic region for conserved RSS and investigating flanking regions for sequences encoding conserved TCR Jδ motifs (Fig. 1). Both the Jδ were confirmed in transcripts containing Gδ1 (Fig. 6).

Expression of Gδ1 with Vα, Vδ, and VHδ

To determine which V genes are used in Gδ1 transcripts, 5′ RACE was performed. As expected, the Vα and Vδ genes oriented upstream of Gδ1 were recombined and transcribed with Gδ1 (Fig. 6, Table 1). Transcripts containing Vδ1 genes could be amplified from the cDNA library, as well as small intestine, spleen, and thymus RNA individually (Table 1, Fig. 6). Gδ1 transcripts containing Vα2 were amplified from the cDNA library, small
intestine, spleen, and thymus, whereas Vα1 transcripts were found only in the cDNA library and small intestine (Table 1). Significantly, Cδ1 transcripts containing recombined VHδ genes were also amplified by 5′RACE from all tissues (Fig. 6, Table 1). Only VHδ1, 4 and 5 were amplified, consistent with these three being functional (Fig. 1, Table 1). These results confirm that Xenopus TCRδ chains can be expressed with VH-like V genes, as well as Vα and Vδ.

Using nucleotide sequences corresponding to CDR3 from Cδ1 transcripts, two Dδ (Dδ1.1 and Dδ1.2) genes were identified (Fig. 1). Transcripts containing Vα, Vδ, and VHδ all used these same two Dδ, which are similar in length (12 and 11 bp, respectively) and share a region of micro-homology, three nucleotides in length (Fig. 6 and 7). This micro-homology appears to facilitate recombination in clones lacking P or N nucleotides in the junction between D segments (Fig. 6). Both Dδ have ORF in all three reading frames and all three are used (Fig. 7 and data not shown). However, there was a slight bias toward the first reading frame of Dδ1.1 (12 out of 26 clones) and third reading frame Dδ1.2 (10 out of 25 clones) both of which encode a tyrosine and two hydrophobic amino acids (Fig. 7). The Dδ are flanked by RSS with 12bp spacers on the 5′ side and 23bp on the 3′ side, as is typical of TCR D (Fig. 7) [19]. However, the RSS are not well conserved, having four different heptamer sequences (Fig. 7).

This asymmetrical RSS organization facilitates the use of multiple D in individual recombinants, as has been described previously for TCRδ chains [13, 19]. The majority of Cδ1 transcripts (24 out of 32) encoded CDR3 using both Dδ genes irrespective of the V or J genes being used (Fig. 6). The remaining eight clones contained only a single Dδ. Four were Dδ1.1, three were Dδ1.2, and one was

**Figure 2.** Phylogenetic tree of X. tropicalis Vα and Vδ compared with mammalian and avian genes using the neighbor joining method. Similar results were found using the minimum evolution method. X. tropicalis sequences are shown in bold. Vα and Vδ genes fall into nine groups that are indicated by brackets. Bootstrap values shown are based on 1000 replicates. A distance bar is shown below the tree.
indeterminate due to excessive trimming. Therefore, when single Dδ are used there does not seem to be a preference.

It is noteworthy that all of the X. tropicalis TCRδ clones characterized encoded single V and C domains, irrespective of the type of V being used. There is no evidence for X. tropicalis TCRδ chains having three extracellular Ig domains, two V and one C, as has been found for shark NAR-TCR and mammalian TCRγδ.

The TCRα/β and IgH loci are linked in Xenopus

The evidence for Xenopus TCRγδ using VH-like genes raises questions regarding linkage between TCRα/β and IgH in frogs. This was investigated by RFLP analysis for Cδ2, VH, and Cε genes using pedigreed X. laevis. All loci analyzed in X. tropicalis share the same linkage patterns as X. laevis, making the latter a suitable proxy for frogs. This was confirmed using a housekeeping gene, SLC7A7, adjacent to IgF on scaffold 972 from the Xenopus genome, which also co-segregated with TCRδ and IgH on chromosome 1.

No evidence for trans-locus somatic recombination between IgH and TCRδ in X. tropicalis

Linkage between Xenopus IgH and TCRδ/β raised the possibility of the use of a common pool of V genes. To investigate this, VHδ were used to search Xenopus expressed sequence tag (EST) databases for evidence of their expression with different C genes. A single Cδ1 X. laevis EST (EB477571) containing a VHδ was identified (VHδ751 in Fig. 5). All other EST from both X. tropicalis and X. laevis containing V sequence with greater than 90% nucleotide identity to VHδ contained FR4 with the WGXG motif conserved in JH genes, and several contained sufficient C region sequence to confirm that they were frog IgM or IgX transcripts, not TCRδ or TCRγδ.

Discussion

For much of the past two decades, the story of TCR structure and genetics has been one of conservation. This was recently
illustrated when the genomic organization of the sandbar shark TCRα locus was found to have the translocon organization similar to that of mammals [23]. Furthermore, the genomic region containing TCRα/δ has been relatively stable over the 340 million years of evolution separating anamniotes from amphibians [7, 9, 24]. This made all the more surprising the discovery of VH-related genes within the Xenopus TCRα/δ locus, a characteristic not found previously in any tetrapod. It should be stressed that the VHδ genes are not orthologues of VH genes in the IgH locus [17, 18]. Rather, the results support that they are paralogues of VH genes expressed exclusively in TCRδ chains.

The discovery of NAR-TCR in sharks and TCRμ in marsupials provides precedence for TCR chains using antibody-like V genes [12, 13]. One question is whether the characteristics of mammalian TCRμ, frog TCRδ and shark NAR-TCR are the result of convergent evolution or homology by descent. This can be addressed by considering the likely origins of VH-like genes in the TCRδ locus. It is noteworthy that Xenopus is one of the genera with a diverse pool of VH genes at their IgH locus. Xenopus has retained all three ancient VH clans (I, II, and III) and these VH genes segregate into eleven subgroups [17, 18]. This level of diversity is in contrast to many species such as chicken, platypus, opossum, and rabbit that have only single VH subgroup, most often clan III [25–27]. The X. tropicalis VHδ are most related to clan II VH. Platypus TCRμ V genes, on the other hand, are related to VH clan III. In contrast, the V genes used in marsupial TCRμ fall outside the clan I, II, or III designations, and are from a sister clade to the VH [13]. Alternatively, shark NAR-TCR create the N-terminal V domain using V genes related to IgNAR, a type of antibody unique to cartilaginous fish [12, 13]. Therefore, the V genes being used in each case appear independently derived, consistent with the convergent mode of evolution.

The biological significance of Xenopus TCRδ using VH-like genes related to clan II VH (VHS, 9 and 11) is unknown. The VH clans appear to evolve at different rates with clan III being more conserved and widespread [28]. However, it is noteworthy that, in Xenopus, B cells expressing clan II VH appear later in ontogeny [29]. A question to be addressed is whether there is a similar late expression of VHδ genes in developing Xenopus γδT cells. This would seem unlikely since VH and VHδ are distinct sets of V genes in separate loci and expressed in distinctly different lymphocyte lineages. Furthermore, the majority of V genes available for rearrangement and expression of TCRδ chains are VHδ. Whether they ultimately comprise the majority of γδT cells and the order of their appearance in frog ontogeny both remain to be determined.

If the V genes used in shark NAR-TCR, frog TCRδ, and marsupial TCRμ represent convergent evolution then the origin of VHδ genes in the X. tropicalis TCRα/δ locus needs to be considered. The relationship between X. tropicalis VHδ and clan II VH is consistent with a block duplication inserting multiple VH into the TCRα/δ locus. VHδ genes in particular share a high degree of similarity with VHδ genes in the IgH locus, consistent with either a recent introgression into the TCRα/δ locus or a more ancient duplication that has been subject to

---

**Figure 5.** Phylogenetic tree showing the X. tropicalis VHδ gene segments compared with Ig VH gene segments from different species. X. tropicalis VHδ gene segments are indicated in bold. The VHδ pseudogenes, VHδ2 and VHδ3, are not included in this tree but cluster with VHδ1 and VHδ4 subgroups when included. Frog VH sequences indicated with an asterisk are those from X. tropicalis, with the remaining being from X. laevis. Three distinct VH clans are indicated by brackets. The tree was generated using the minimum evolution method. Similar results were found using the neighbor joining method. Numbers at the nodes are bootstrap values based on 1000 replicates. A distance bar is shown at the bottom of the tree.
Figure 6. (A) Diagram of expressed isoforms of \(X. \) tropicalis C\( \delta 1\). (B) CDR3 nucleotide alignment of representative sequences. Germline D sequences are labeled on top of the alignment. V\( \alpha \), V\( \delta \), and VH\( \delta \) gene segments are indicated on the left–hand side and J gene segments on the right. Highlighted in gray is a 5 bp sequence present in both D\( \delta \) segments. Microhomologies in the germline and CDR3 regions are underlined; P and N nucleotides (shown in white letters) are highlighted in black and gray, respectively. GenBank accession numbers are shown in parenthesis on the right.

Table 1. C\( \delta 1\) transcripts amplified from different \(X. \) tropicalis tissues

<table>
<thead>
<tr>
<th>Tissue/subgroup</th>
<th>V( \alpha )</th>
<th>V( \delta )</th>
<th>VH( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA library</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>NF(^{b})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small intestine</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Spleen</td>
<td>NF</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Thymus</td>
<td>NF</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

\(^{b}\)Not found.

Figure 7. The germline \(X. \) tropicalis D\( \delta \) gene segments. Highlighted in gray is a 5 bp sequence present in both D\( \delta \) segments. Reading frames and RSS of the two D\( \delta \) gene segments found recombined with V\( \alpha \), V\( \delta \), and VH\( \delta \).

Figure 8. Close linkage of TCR\( \alpha/\delta \), IgH, and Ig\( \lambda \) loci in the Xenopus genome. Probes and restriction enzyme giving the RFLP are indicated on the left. Arrows on the right indicate polymorphic bands with VH that segregate with the bottom band of TCR\( \delta \). Arrows below indicate recombinants (siblings 4 and 9) between TCR\( \delta \) and Ig\( \lambda \). Size markers are indicated on the right. The bands for SLC7A7 and Ig\( \lambda \) are found between 2.3 and 4.3 kb marker bands. Sibling 5 (asterisk) is a natural triploid. P, paternal; M, maternal.
ongoing nonhomologous recombination or gene-conversion that preserved the high-sequence similarity. Given the close linkage between IgH and TCRα/δ in Xenopus, such nonhomologous recombination is highly likely. The similarity between VH65 and VH5 genes and their designation as the respective “fifth” subgroup in each locus is purely coincidental. VH5 were identified as the fifth new subgroup among a collection of Xenopus IgH transcripts [18]. VH65, on the other hand, was annotated as the fifth subgroup in linear order in the TCRα/δ genomic sequence using conventional TCR nomenclature (Materials and Methods). In contrast to VH65, the VH61 through four subgroups are less similar to any of the known Xenopus VH subgroups, sharing less than 60% nucleotide identity with the latter. This suggests that either these VH6 are older introgressions of VH genes into the TCRα/δ locus and were later lost in the IgH locus, or are genes that have diverged at a higher rate than VH65.

These results also bear on the link between γdTCR and IgH noted previously, including recent evidence that TCRγ genes in sharks may undergo somatic mutation similar to that of Ig [23]. In mice and humans, Vδ are more similar to VH in CDR3 length [30]. Ig VH and TCR Vδ both have CDR3 heterogeneous in size and associated with partners (IgL and TCRγ, respectively) with generally shorter CDR3. In contrast, the CDR3 of both Vβ and Vδ are relatively similar, presumably due to the constraints of MHC restriction. These observations led to proposals that γdT cells recognize antigens differentially than γdT cells, such as recognizing free, unprocessed antigens [30]. Indeed, while conventional γdT cells are MHC restricted, some γdT cells have been shown to respond in an MHC un-restricted manner to unprocessed viral antigens [31]. Furthermore, so-called “innate” γdT cells with limited TCR diversity in mice and humans can recognize self-molecules such as the MHC-class I chain-related MICA (in humans) and RAE-1 (in mice) or metabolic by-products such as isoprenylphosphate [32]. If TCR with antibody-like V domains are products of convergence, it is likely antigen recognition is driving this evolution. We predict that such T cells, including Xenopus γdT cells using VH6, may bind microbial or other pathogen-associated epitopes in a manner more like B cells. In the case of γdTCR, the δ chain would be performing the role similar to H chain in antibodies, where the H chain often takes a predominant role in antigen recognition [33]. It seems unlikely that shark NAR-TCR β T cells or mammalian TCRδ T cells would be the equivalent of the “innate” γdT cells in eutherians, given the diversity of their receptors. Rather, this model provides some explanation for the consistent finding that, over evolutionary time, the TCRδ locus has demonstrated extraordinary plasticity, and we predict that the unusual TCRδ chains in sharks and frogs, and TCRδ in marsupials and monotremes, are involved in adaptive immunity.

The capacity to bind free antigen in solution may have provided a strong pressure for rapid evolution of the VH repertoire, resulting in the clans and families that emerged early in the Ig. Perhaps what is being observed in sharks, frogs, and non-eutherian mammals has been the TCRδ advantage of this unique VH diversity and incorporating gene segments into its own repertoire, which in mammals resulted in the evolution of a separate locus, TCRδ. Why this has not been found in any eutherian mammal, such as humans and mice, is not clear [9]. It is possible that γdT cells entirely satisfy the role of direct antigen recognition in eutherians.

An atypical feature shared by NAR-TCR and TCRδ is the expression of an extracellular form containing three Ig superfamily domains [12, 13]. There was no evidence for Xenopus TCRδ chains expressed in a double V form, nor does it appear that the Xenopus TCRα/δ locus could encode such a form. Double Vs require tandem array of rearranging V, D, and J segments as has been found in shark NAR-TCR and the platypus TCRδ homologue or, alternatively, the presence of a second germ-line joined V, as is the case for marsupial TCRδ [12, 13]. Although transcripts encoding double V were not found for X. tropicalis TCRδ, many of the V segments expressed with C51 encode un-paired cysteine residues in positions similar to that of VNAR and the supporting Vδ that are used in the NAR-TCR (Fig. 4). These cysteines are presumed to form inter-domain disulfide bonds that help to stabilize the double V structure. What role they may play in Xenopus TCRδ chains remains to be determined. None of the other V genes in the locus contained unpaired cysteines, although the majority of the Vδ expressed with C51 and are not used in the TCRδ chain. In addition, such unusual V genes also raise the obvious question of whether all of the Xenopus TCRδ chains are associated with TCRγ or are found in other complexes. This is a question that remains to be answered for NAR-TCR and TCRδ as well.

The common origin of Ig and TCR is beyond doubt when considering gene organization, conserved RSS, and common enzymatic machinery mediating somatic recombination [34]. Although many of the characteristics in Xenopus TCRδ, NAR-TCR, and TCRβ appear to be the result of convergent evolution, the structure of Xenopus TCRα/δ locus and the linkage analyses are consistent with prevailing models of the origin and evolution of Ig and TCR genes [35, 36]. For example, the inverted genomic region containing VH6 and C61 is reminiscent of the duplication and inversion event predicted to give rise to the current TCRα/δ locus configuration in mammals [9, 35]. The tight linkage of IgH, TCRα/δ, and Igλ in Xenopus is consistent with what is found in many species, arguing for a common ancestral arrangement [35]. This synteny is consistent with a model where the genes encoding both chains of an ancestral antigen receptor heterodimer were linked and then duplicated en bloc to create two gene complexes [35]. If the original antigen receptor was TCR-like, then the two complexes, βγ and αδ, may have been derived from the second genome-wide duplication believed to have occurred early in vertebrate history [37]. The αδ complex has remained linked in the tetrapods and some fish species, whereas βγ has been separated [35, 38, 39]. The presence of an ancestral Igλ chain in the original complex is parsimonious, since it shares the RSS with a 23 bp spacer type with all other antigen receptor V genes. Different lineages of receptor genes would have been free to independently evolve their own RSS configurations, most likely through inversions, resulting in the symmetrical 12 bp spacer RSS flanking DH genes.
and the 12 bp spacer RSS flanking V genes in the Igx and Igσ loci [36, 40].

In conclusion, NAR-TCR in sharks, TCRμ in non-eutherian mammals, and now TCRδ in Xenopus comprise what appears to be a growing list of TCRδ or δ-like chains utilizing antibody-like V domains. How long this list of species will become is unknown; however, these observations may reveal a new mechanism for antigen recognition by T cells that has remained undiscovered because of its absence in commonly studied eutherian mammals.

Materials and methods

Genome screening and identification of V–D–J gene segments

Version 4.1 of the X. tropicalis genome assembly was analyzed using the BLAST algorithm at www.ncbi.nlm.nih.gov/BLAST along with the Xenopus EST BLAST server at http://www.sanger.ac.uk/ and ENSEMBL at www.ensembl.org. Known Vα and Jα gene segments and the Cα1 genes were identified using available sequences [15, 41]. Novel Vα, Jα, and Cα genes and all Vδ, Dδ, Jδ, and Cδ genes were identified by similarity to homologues from other species. Conserved RSS were also used to identify V, D, and J gene segments. The location for all TCRα/δ genes in version 4.1 of the X. tropicalis assembly is provided in Supporting Information Tables 1 and 2.

Annotation of the X. tropicalis TCRα/δ and VH regions

The TCRα/δ locus is located on scaffold number 539 (length 791,252 bp). V and D gene subgroups and individual gene segments were annotated in 5′-3′ order in the genomic sequence in accordance with the International ImMunoGeneTics nomenclature (http://www.imgt.cines.fr/). For example, Vα2.1 refers to the most 5′ Vα in subgroup 2. J genes were numbered 3′-5′ also according to the established nomenclature [14]. Cα1 was identified as being identical to Cα2 genes identified previously in X. laevis [15]. Cα2 and both Cαδ were identified in the genomic sequence as being a C region gene with similarity to homologues in other species. The two Cαδ were numbered based on Cαδ1 being more frequently isolated among TCRδ transcripts (data not shown). Scaffold 1168 containing VH genes most closely resembling the VHδ found on scaffold 539 was also characterized. The V genes found on this scaffold are utilized in IgH transcripts and this genomic sequence is likely part of the IgH locus in X. tropicalis (see Results section).

RNA extraction, cDNA synthesis, and RT-PCR

The combination intestine, liver, spleen, and thymus cDNA library used was constructed using the Orient Express Oligo(dT) cDNA library kit (Novagen, Madison, WI) [42]. Total RNA was extracted from small intestine, spleen, and thymus from X. tropicalis as described previously and under the approval of animal protocol number 1009003 to MFF [42]. Total RNA was treated to remove DNA contamination using the TURBO DNA-free kit (Ambion, Austin, TX) and then used for RT-PCR. RT-PCR were performed using GeneAmp RNA PCR core kit (Applied Biosystems, Foster City, CA). PCR were performed using the AdvantageTM-HF-2 kit using manufacturer-recommended conditions (BD Biosciences, CLONTECH Laboratories, Palo Alto, CA). PCR products were cloned using the TopoTA cloning kit (Invitrogen, Carlsbad, CA) and sequenced using the BigDye Terminator Cycle Sequencing kit v3 (Applied Biosystems, Foster City, CA). Sequencing reactions were analyzed on the ABI Prism 3100 DNA automated sequenced (PerkinElmer Life and Analytical Sciences, Wellesley, MA). Chromatograms were analyzed using the SequencherTM 4.3 program (Gene Codes, Ann Arbor, MI).

Primers for Cα1 and Cα2 semi-nested PCR and 5′ RACE were Cα1 outside primer, 5′-GGGCCGTAGTTCATTGAGGAGGACAGC-3′; Cα1-nested primer: 5′-CCAGACAGGCAGTGGGAGAGTCACC-3′; Cα2 outside primer: 5′-GAAATCTGTTGGTAAACATCCTCC-3′; Cα2-nested primer: 5′-CAGCGGCACAGAGACACCTCC-3′.

To investigate whether VHδ genes are utilized in IgH chains, VHδ1 and VHδ5 primers were paired in PCR with a primer specific for Xenopus IgM C region using the spleen cDNA as target. IgM C region primer: 5′-AATGGCCGATAAGTGGTATGGTAG-3′; VHδ1.2 and VHδ1.3 primer: 5′-GTTCCTGGAAAAGATTCT-3′; VHδ1.5 and VHδ1.6 primer: 5′-CTTGGGCTGAAACTTCTCCTG-3′.

The phylogenetic analyses

The GenBank accession numbers of all sequences used are in Supporting Information Table 3. Nucleotides corresponding to FR1 through FR3 of all V genes and exon 1 of all C genes were aligned using BioEdit [43, 44]. Phylogenetic trees were constructed using the neighbor joining and minimum evolution methods with 1000 bootstrap replicates using MEGA version 3.0 [45].

Linkage analyses

Erythrocyte genomic DNA from 20 X. laevis siblings was digested with various restriction enzymes and blotted onto nitrocellulose membranes (Whatman, ME). Blots were hybridized under high stringency conditions with 32P-dCTP radiolabeled probes and exposed to the X-ray film [46]. Probes used for the experiments included the exons encoding the Cα2 domain and 3′ UTR, the C domain of Igα, and a V domain from the VH1 subgroup in the IgH locus (GenBank NM_001086404) [17, 18]. VH1 was chosen to type the IgH locus since it shares less than 60% nucleotide identity to any VHδ and, therefore would not
cross-hybridize. To further confirm Ig-TCR linkage, the SLC7A7 gene near IgF in scaffold-972 was used as probe, specifically nucleotide position 140–1537 of X. laevis SLC7A7 (BC072040) [21, 22]. Tetraploid X. laevis was used for this analysis due to the availability of pedigreed families for this species [47].

Acknowledgements: Support was provided by National Institutes of Health (IP20RR18754) and the National Science Foundation (IOS-0641382) to R. D. M., and the National Institutes of Health (RR00603) to M. F. F.

Conflict of interest: The authors declare no financial or commercial conflict of interest.

References


Abbreviations: EST: expressed sequence tag · NAR: New Antigen Receptor · ORF: open reading frame · RSS: recombination signal sequences · VH: V genes that are indistinguishable from VH but found in the TCR locus

Full correspondence: Prof. Robert D. Miller, MSC03 2020, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA Fax: +1-505-277-0304 e-mail: rdmiller@unm.edu

Current address: Professor Michael F. Criscitiello, Department of Veterinary Pathobiology, Texas A&M University, College Station, TX 77843, USA

Received: 22/3/2010
Revised: 3/5/2010
Accepted: 12/5/2010
Accepted article online: 18/5/2010

© 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim www.eji-journal.eu