A Model for the Evolution of the Mammalian T-cell Receptor α/δ and μ Loci Based on Evidence from the Duckbill Platypus

Zuly E. Parra,1 Mette Lillie,2 and Robert D. Miller*1

1Department of Biology, Center for Evolutionary & Theoretical Immunology, University of New Mexico
2Faculty of Veterinary Science, University of Sydney, Sydney, Australia

*Corresponding author: E-mail: rdmiller@unm.edu.

Associate editor: Yoko Satta

Abstract

The specific recognition of antigen by T cells is critical to the generation of adaptive immune responses in vertebrates. T cells recognize antigen using a somatically diversified T-cell receptor (TCR). All jawed vertebrates use four TCR chains called α, β, γ, and δ, which are expressed as either an αβ or γδ heterodimer. Nonplacental mammals (monotremes and marsupials) are unusual in that their genomes encode a fifth TCR chain, called TCRμ, whose function is not known but is also somatically diversified like the conventional chains. The origins of TCRμ are also unclear, although it appears distantly related to TCRδ. Recent analysis of avian and amphibian genomes has provided insight into a model for understanding the evolution of the TCRδ genes in tetrapods that was not evident from humans, mice, or other commonly studied placental (eutherian) mammals. An analysis of the genes encoding the TCRδ chains in the duckbill platypus revealed the presence of a highly divergent variable (V) gene, indistinguishable from immunoglobulin heavy (IgH) chain V genes (VH) and related to V genes used in TCRμ. They are expressed as part of TCRδ repertoire (VHδ) and similar to what has been found in frogs and birds. This, however, is the first time a VHδ has been found in a mammal and provides a critical link in reconstructing the evolutionary history of TCRμ. The current structure of TCRδ and TCRμ genes in tetrapods suggests ancient and possibly recurring translocations of gene segments between the IgH and TCRδ genes, as well as translocations of TCRδ genes out of the TCRα/δ locus early in mammals, creating the TCRμ locus.

Key words: T-cell receptors, monotremes, comparative genomics.

Introduction

T lymphocytes are critical to the adaptive immune system of all jawed vertebrates and can be classified into two main lineages based on the T-cell receptor (TCR) they use (Rast et al. 1997; reviewed in Davis and Chein 2008). The majority of circulating human T cells are the αβT cell lineage which use a TCR composed of a heterodimer of α and β TCR chains. αβT cells include the familiar T cell subsets such as CD4+ helper T cells and regulatory T cells, CD8+ cytotoxic T cells, and natural killer T (NKT) cells. T cells that are found primarily in epithelial tissues and a lower percentage of circulating lymphocytes in some species express a TCR composed of γ and δ TCR chains. The function of these γδ T cells is less well defined and they have been associated with a broad range of immune responses including tumor surveillance, innate responses to pathogens and stress, and wound healing (Hayday 2009). αβ and γδ T cells also differ in the way they interact with antigen. αβ/TCR are major histocompatibility complex (MHC) “restricted” in that they bind antigenic epitopes, such as peptide fragments, bound to, or “presented” by, molecules encoded in the MHC. In contrast, γδTCR have been found to bind antigens directly in the absence of MHC, as well as self-ligands that are often MHC-related molecules (Sciammas et al. 1994; Hayday 2009).

The conventional TCR chains are composed of two extracellular domains that are both members of the immunoglobulin (Ig) domain super-family (reviewed in Davis and Chein 2008) (fig. 1). The membrane proximal domain is the constant (C) domain, which is largely invariant amongst T-cell clones expressing the same class of TCR chain, and is usually encoded by a single, intact exon. The membrane distal domain is called the variable (V) domain and is the region of the TCR that contacts antigen and MHC. Similar to antibodies, the individual clonal diversity in the TCR V domains is

![Fig. 1. Cartoon diagram of the TCR forms found in different species.](image_url)

© The Author 2012. Published by Oxford University Press on behalf of the society for Molecular Biology and Evolution. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
generated by somatic DNA recombination (Tonegawa 1983). The exons encoding TCR V domains are assembled somatically from germline gene segments, called the V, diversity (D), and joining (J) genes, in developing T cells, a process dependent upon the enzymes encoded by the recombination activating genes (RAG)-1 and RAG-2 (Yancopoulos et al. 1986; Schatz et al. 1989). The exons encoding the V domains of TCR \(\beta\) and \(\delta\) chains are assembled from all three types of gene segments, whereas the \(\alpha\) and \(\gamma\) chains use only V and J. The different combinations of V, D, and J or V and J, selected from a large repertoire of germline gene segments, along with variation at the junctions due to addition and deletion of nucleotides during recombination, contribute to a vast TCR diversity. It is this diversity that creates the individual antigen specificity of T-cell clones.

The TCR genes are highly conserved among species in both genomic sequence and organization (Rast et al. 1997; Parra et al. 2008, 2012; Chen et al. 2009). In all tetrapods examined, the TCR\(\beta\) and \(\gamma\) chains are each encoded at separate loci, whereas the genes encoding the \(\alpha\) and \(\delta\) chains are nested at a single locus (TCR\(\alpha/\delta\)) (Chien et al. 1987; Satyanarayana et al. 1988; reviewed in Davis and Chein 2008). The V domains of TCR\(\alpha\) and TCR\(\delta\) chains can use a common pool of V gene segments, but distinct D, J, and C genes.

Diversity in antibodies produced by B cells is also generated by RAG-mediated V(D)J recombination and the TCR and Ig genes clearly share a common origin in the jawed-vertebrates (Flajnik and Kasahara 2010; Litman et al. 2010). However, the V, D, J, and C coding regions in TCR have diverged sufficiently over the past >400 million years (MY) from Ig genes that they are readily distinguishable, at least for the conventional TCR. Recently, the boundary between TCR and Ig genes has been blurred with the discovery of non-conventional TCR isoforms that have been found that use V genes that appear indistinguishable from Ig heavy chain V (VH) (Parra et al. 2007). The second V domain is encoded by a germ-line joined, or pre-assembled, than to conventional TCR V domains. The N-terminal V domain has little or no diversity. In marsupials this V domain is encoded by a germ-line joined, or pre-assembled, V exon that is invariant (Parra et al. 2007). The second V domain in platypus is encoded by gene segments requiring somatic DNA recombination; however, only limited diversity is generated partly due to the lack of D segments (Wang et al. 2011). A TCR chain structurally similar to TCR\(\mu\) has also been described in sharks and other cartilaginous fish (fig. 1) (Criscitiello et al. 2006; Flajnik et al. 2011). This TCR, called NAR-TCR, also contains three extracellular domains, with the N-terminal V domain being related to those used by IgNAR antibodies, a type of antibody found only in sharks (Greenberg et al. 1995). The current working model for both TCR\(\mu\) and NAR-TCR is that the N-terminal V domain is unpaired and acts as a single, antigen binding domain, analogous to the V domains of light-chainless antibodies found in sharks and camels (Flajnik et al. 2011; Wang et al. 2011).

Phylogenetic analyses support the origins of TCR\(\mu\) occurring after the avian–mammalian split (Parra et al. 2007; Wang et al. 2011). Previously, we hypothesized the origin of TCR\(\mu\) being the result of a recombination between ancestral IgH and TCR\(\delta\)-like loci (Parra et al. 2008). This hypothesis, however, is problematic for a number of reasons. One challenge is the apparent genomic stability and ancient conserved synteny in the region surrounding the TCR\(\alpha/\delta\) locus; this region has appeared to remain stable over at least the past 350 MY of tetrapod evolution (Parra et al. 2008, 2010). The discovery of VH\(\delta\) genes inserted into the TCR\(\alpha/\delta\) locus of amphibians and birds has provided an alternative model for the origins of TCR\(\mu\); this model involves both the insertion of VH followed by the duplication and translocation of TCR genes. Here we present the model along with supporting evidence drawn from the structure of the platypus TCR\(\alpha/\delta\) locus, which is also the first analysis of this complex locus in a monotreme.

**Materials and Methods**

**Identification and Annotation of the Platypus TCR\(\alpha/\delta\) Locus**

The analyses were performed using the platypus (*Ornithorhynchus anatinus*) genome assembly version 5.0.1 (http://www.ncbi.nlm.nih.gov/genome/guide/platypus/).
The platypus genome was analyzed using the whole-genome BLAST available at NCBI (www.ncbi.nlm.nih.gov/) and the BLAST/BLAT tool from Ensembl (www.ensembl.org). The V and J segments were located by similarity to corresponding segments from other species and by identifying the flanking conserved recombination signal sequences (RSS). V gene segments were annotated 5′ to 3′ as Vα or Vβ followed by the family number and the gene segment number if there were greater than one in the family. For example, Vα15.7 is the seventh Vα gene in family 15. The D segments were identified using complementarity-determining region-3 (CDR3) sequences that represent the V-D-J junctions, from cDNA clones using VHδ. Platypus TCR gene segments were labeled according to the IMGT nomenclature (http://www.imgt.org/). The location for the TCRα/β genes in the platypus genome version 5.0.1 is provided in supplementary table S1, Supplementary Material online.

Confirmation of Expression of Platypus VHδ
Reverse transcription PCR (RT–PCR) was performed on total splenic RNA extracted from a male platypus from the Upper Barnard River, New South Wales, Australia. This platypus was collected under the same permits as in Warren et al. (2008). The cDNA synthesis step was carried out using the Invitrogen Superscript III-first strand synthesis kit according to the manufacturer’s recommended protocol (Invitrogen, Carlsbad, CA, USA). TCRδ transcripts containing VHδ were targeted using primers specific for the Cδ and VHδ genes identified in the platypus genome assembly (Warren et al. 2008). PCR amplification was performed using the QIAGEN HotStar HiFidelity Polymerase Kit (BD Biosciences, CLONTECH Laboratories, Palo Alto, CA, USA) in total volume of 20 μl containing 1× Hotstar HiF PCR Buffer (containing 0.3 mM dNTPs), 1 μM of primers, and 1.25U Hotstar HiFidelity DNA polymerase. The PCR primers used were 5′-GTACCGCCAACCACCAGGAAA G-3′ and 5′-CAGTTCACTGCTCCATCGCTTTCA-3′ for the VHδ and Cδ, respectively. A previously described platypus spleen cDNA library constructed from RNA extracted from tissue from a Tasmanian animal was also used (Vernersson et al. 2002).

PCR products were cloned using TopoTA cloning® kit (Invitrogen). Sequencing was performed using the BigDye terminator cycle sequencing kit version 3 (Applied Biosystems, Foster City, CA, USA) and according to the manufacturer recommendations. Sequencing reactions were analyzed using the ABI Prism 3100 DNA automated sequences (PerkinElmer Life and Analytical Sciences, Wellesley, MA, USA). Chromatograms were analyzed using the Sequencher 4.9 software (Gene Codes Corporation, Ann Arbor, MI, USA). Sequences have been archived on GenBank under accession numbers JQ664690–JQ664710.

Phylogenetic Analyses
Nucleotide sequences from FR1 to FR3 of the V genes regions, including CDR1 and CDR2, were aligned using BioEdit (Hall 1999) and the accessory application ClustalX (Thompson et al. 1997). Nucleotide alignments analyzed were based on amino acid sequence to establish codon position (Hall 1999). Alignments were corrected by visual inspection when necessary and were then analyzed using the MEGA Software (Kumar et al. 2004). Neighbor joining (NJ) with uncorrected nucleotide differences (p-distance) and minimum evolution distances methods were used. Support for the generated trees was evaluated based on bootstrap values generated by 1000 replicates. GenBank accession numbers for sequences used in the tree construction are in supplementary table S2, Supplementary Material online.

Results and Discussion
The TCRα/δ locus was identified in the current platypus genome assembly and the V, D, J, and C gene segments and exons were annotated and characterized (fig. 2). The majority of the locus was present on a single scaffold, with the remainder on a shorter contig (fig. 2). Flanking the locus were SALL2, DAD1 and several olfactory receptor (OR) genes, all of which share conserved synteny with the TCRα/δ locus in amphibians, birds, and mammals (Parra et al. 2008, 2010, 2012). The platypus locus has many typical features common to TCRα/δ loci in other tetrapods (Satyanarayana et al. 1988; Wang et al. 1994; Parra et al. 2008, 2010, 2012). Two C region genes were present: a Vc that is the most 3′ coding segment in the locus, and a Cδ oriented 5′ of the Jα genes. There is a large number of Jα gene segments (n = 32) located between the Cδ and Cα genes. Such a large array of Jα gene segments are believed to facilitate secondary Vα to Jα rearrangements in developing αβT cells if the primary rearrangements are nonproductive or need replacement (Hawwari and Krangel 2007). Primary TCRα V–J rearrangements generally use Jα segments towards the 5′-end of the array and can progressively use downstream Jα in subsequent rearrangements. There is also a single Vδ gene in reverse transcriptional orientation between the platypus Cδ gene and the Jα array that is conserved in mammalian TCRα/δ both in location and orientation (Parra et al. 2008).

There are 99 conventional TCR V gene segments in the platypus TCRα/δ locus, 89 of which share nucleotide identity with Vα in other species and 10 that share identity with Vδ genes. The Vδ genes are clustered towards the 3′-end of the locus. Based on nucleotide identity shared among the platypus V genes they can be classified into 17 different Vα families and two different Vδ families, based on the criteria of a V family sharing >80% nucleotide identity (not shown, but annotated in fig. 2). This is also a typical level of complexity for mammalian Vα and Vδ genes (Giudicelli et al. 2005; Parra et al. 2008). Also present were two Dδ and seven Jδ gene segments oriented upstream of the Cδ. All gene segments were flanked by canonical RSS, which are the recognition substrate of the RAG recombinase. The D segments were symmetrically flanked by an RSS containing at 12 bp spacer on the 5′-side and 23 bp spacer on the 3′-side, as has been shown previously for TCR D gene segments in other species (Carroll et al. 1993; Parra et al. 2007, 2010). In summary, the overall content and organization of the platypus TCRα/δ locus appeared fairly generic.
What is atypical in the platypus TCRα/β locus was the presence of an additional V gene that shared greater identity to antibody VH genes than to TCR V genes (figs. 2 and 3). This V gene segment was the most proximal of the V genes to the D and J genes and was tentatively designated as VHα. VHα are, by definition, V genes indistinguishable from Ig VH genes but used in encoding TCRδ chains and have previously been found only in the genomes of birds and frogs (Parra et al. 2008, 2010, 2012).

VH genes from mammals and other tetrapods have been shown to cluster into three ancient clans and individual species differ in the presence of one or more of these clans in their germ-line IgH locus (Tutter and Riblet 1989; Ota and Nei 1994). For example, humans, mice, echidnas, and frogs have...
VH genes from all three clans (Schwager et al. 1989; Ota and Nei 1994; Belov and Hellman 2003), whereas rabbits, opossums, and chickens have only a single clan (McCormack et al. 1991; Butler 1997; Johansson et al. 2002; Baker et al. 2005). In phylogenetic analyses, the platypus VHβ was most related to the platypus Vβ genes found in the TCRβ locus in this species (fig. 3). Platypus VHβ, however, share only 51–61% nucleotide identity (average 56.6%) with the platypus Vβ genes. Both the platypus Vβ and VHβ clustered within clan III (fig. 3) (Wang et al. 2011). This is noteworthy given that VH genes in the platypus IgH locus are also clan III and, in general, clan III VH are the most ubiquitous and conserved lineage of VH (Johansson et al. 2002; Tutter and Riblet 1989). Although clearly related to platypus VH, the VHβ gene share only 34–65% nucleotide identity (average 56.9%) with the bona fide VH used in antibody heavy chains in this species.

It was necessary to rule out that the VHβ gene present in the platypus TCRβ locus was not an artifact of the genome assembly process. One piece of supporting evidence would be the demonstration that the VHβ is recombined to downstream Dβ and Jβ segments and expressed with Cβ in complete TCRβ transcripts. PCR using primers specific for VHβ and Cβ was performed on cDNA synthesized from splenic RNA from two different platypuses, one from New South Wales and the other from Tasmania. PCR products were successfully amplified from the NSW animal and these were cloned and sequenced. Twenty clones, each containing unique nucleotide sequence, were characterized and found to contain the VHβ recombined to the Dβ and Jβ gene segments (fig. 4A). Of these 20, 11 had unique V, D, and J combinations that would encode 11 different complementarity-determining regions-3 (CDR3) (fig. 4B). More than half of the CDR3 (8 out of 11) contained evidence of using both D genes (VDDJ) (fig. 4B). This is a common feature of TCRβ V domains where multiple D genes can be incorporated into the recombination due to the presence of asymmetrical RSS (Carroll et al. 1993). The region corresponding to the junctions between the V, D, and J segments, contained additional sequence that could not be accounted for by the germ-line gene segments (fig. 4B). There are two possible sources of such sequence. One are palindromic (P) nucleotides that are created during V(D)J recombination when the RAG generates hairpin structures that are resolved asymmetrically during the re-ligation process (Lewis 1994). The second are non-templated (N) nucleotides that can be added by the enzyme terminal deoxynucleotidyl transferase (TdT) during the V(D)J recombination process. An unusual feature of the platypus VHβ is the presence of a second cysteine encoded near the 3′-end of the gene, directly next to the cysteine

Fig. 3. Phylogenetic tree of mammalian VH genes including the platypus VHβ and monotreme Vγ. The three major VH clans are bracketed. The platypus VHβ and echidna Vγ μ is in bold and indicated by a smaller bracket in VH clan III. The three-digit numbers following the VH gene labels are the last three digits of the GenBank accession number referenced in supplementary table S2, Supplementary Material online. The numbers following the platypus and echidna Vγ μ labels are clone numbers. The tree presented was generated using the Minimum Evolution method. Similar topology was generation using the Neighbor Joining method.
predicted to form the intra-domain disulfide bond in Ig domains (fig. 4A). Additional cysteines in the CDR3 region of VH domains have been thought to provide stability to unusually long CDR3 loops, as has been described for cattle and the platypus previously (Johansson et al. 2002). The CDR3 of TCR chains is controversial. Nonetheless, it has been invoked to explain the variation in expressed TCR chains that exceeds the apparent gene copy number in sharks, and has also been postulated to occur in salmonids (Yazawa et al. 2008; Chen et al. 2009). Therefore, it does not seem to be out of the realm of possibility that somatic mutation is occurring in platypus VHδ. Indeed, the mutations appear to be localized to the V region with no variation in the C region (fig. 4A). This may be due to its relatedness of VHδ to Ig VH genes where somatic hyper-mutation is well documented. Such somatic mutation contributes to overall affinity maturation in secondary antibody responses (Wysocki et al. 1986). The pattern of mutation seen in platypus VHδ however, is not localized to the CDR3, which would be indicative of selection for affinity maturation, but was also found in the framework regions. Furthermore, in the avian genomes where there is also only one VHδ locus (fig. 4). A VHδ gene isolated from platypus spleen RNA. The germ-line sequence of the 3' end of VHδ, the two Dδ, are shown at the top. The germ-line Jδ sequences are shown on the right-hand side of the alignment interspersed amongst the cDNA sequences using each. Nucleotides in the junctions between the V, D, and J segments, shown italicized, are most likely N-nucleotides added by TdT.
more than a single VHδ does not alter the principal conclusions of this study.

Previously, we hypothesized the origin of TCRμ in mammals involving the recombination between and ancestral TCRα/δ locus and an IgH locus (Parra et al. 2008). The IgH locus would have contributed the V gene segments at the S′-end of the locus, with the TCRδ contributing the D, J, and C genes at the S-end of the locus. The difficulty with this hypothesis was the clear stability of the genome region surrounding the TCRα/δ locus. In other words, the chromosomal region containing the TCRα/δ locus appears to have remained relatively undisrupted for at least the past 360 million years (Parra et al. 2008, 2010, 2012). The discovery of VHδ genes within the TCRα/δ loci of frog and zebra finch is consistent with insertions occurring without apparently disrupting the local syntenic region. In frogs, the IgH and TCRα/δ loci are tightly linked, which may have facilitated the translocation of VH genes into the TCRα/δ locus (Parra et al. 2010). However, close linkage is not a requirement since the translocation of VH genes appears to have occurred independently in birds and monotremes, due to the lack of similarity between the VHδ in frogs, birds, and monotremes (Parra et al. 2012). Indeed, it would appear if the acquisition of VH genes into the TCRα/δ locus occurred independently in each lineage.

The similarity between the platypus VHδ and V genes in the TCRμ locus is, so far, the clearest evolutionary association between the TCRμ and TCRδ loci in one species. From the comparison of the TCRα/δ loci in frogs, birds, and monotremes, a model for the evolution of TCRμ and other TCRδ forms emerges (fig. 5), which can be summarized as follows:

1) Early in the evolution of tetrapods, or earlier, a duplication of the D-J-Cδ cluster occurred resulting in the presence of two Cδ each with its own set of Dδ and Jδ segments (fig. 5A).

2) Subsequently, a VH gene or genes was translocated from the IgH locus and inserted into the TCRα/δ locus, most likely to a location between the existing Vα/Vδ and the S′-proximal D-J-Cδ cluster (fig. 5B). This resulted in the configuration like that which currently exists in the zebra finch genome (Parra et al. 2012).

3) In the amphibian lineage there was an inversion of the region containing VHδ-Dδ-Jδ-Cδ cluster and an expansion in the number of VHδ genes (fig. 5C). Currently, X. tropicalis has the greatest number of VHδ genes, where they make up the majority of V genes available in the germ-line for use in TCRδ chains (Parra et al. 2010).

4) In the galliform lineage (chicken and turkey), the VHδ-Dδ-Jδ-Cδ cluster was trans-located out of the TCRα/δ locus where it currently resides on another chromosome (fig. 5D). There are no Vα or Vδ genes at the site of the second chicken TCRα/δ locus and only a single Cδ gene remains in the conventional TCRα/δ locus (Parra et al. 2012).

5) Similar to galliform birds, the VHδ-Dδ-Jδ-Cδ cluster was trans-located out of the TCRα/δ locus in presumably the last common ancestor of mammals, giving rise to TCRμ (fig. 5E). Internal duplications of the VHδ-Dδ-Jδ genes gave rise to the current (V-D-J) – (V-D-J) – Cδ organization necessary to encode TCR chains with double V domains (Parra et al. 2007, Wang et al. 2011). In the platypus, the second V-D-J cluster, encoding the supporting V, has lost its D segments and generates V domains with short CDR3 encoded by direct V to J recombination (Wang et al. 2011). The whole cluster appears to have undergone additional tandem duplication as it exists in multiple tandem copies in the opossum and also likely in the platypus (Parra et al. 2007, 2008; Wang et al. 2011).

6) In the therian lineage (marsupials and placentals), the VHδ was lost from the TCRα/δ locus (Parra et al. 2008). This is consistent with marsupials retaining TCRμ, however the second set of V and J segments, encoding the supporting V domain in the protein chain, were replaced with a germ-line joined V gene, in a process most likely involving germ-line V(D)J recombination and retro-transposition (fig. 5F) (Parra et al. 2007, 2008).

TCR forms such as TCRμ, which contain three extracellular domains, have evolved at least twice in vertebrates. The first was in the ancestors of the cartilaginous fish, the NAR-TCR (Crisci et al. 2006) and the second in the mammals as TCRμ (Parra et al. 2007). NAR-TCR uses an N-terminal V domain related to the V domains found in IgNAR antibodies, which are unique to cartilaginous fish (Greenberg et al. 1995; Crisci et al. 2006), and not closely related to antibody VH domains. Therefore, it appears that NAR-TCR and TCRμ are more likely the result of convergent evolution rather than being related by direct descent (Parra et al. 2007; Wang et al. 2011). Similarly, the model proposed in fig. 5 posits the direct transfer of VH genes from an IgH locus to the TCRα/δ locus. But it should be pointed out the VHδ found in frogs, birds, and monotremes are not closely related (fig. 3); indeed, they appear derived each from different, ancient VH clans (birds, VH clan I; frogs VH clan II; platypus VH clan III). This observation would suggest that the transfer of VHδ into the TCRα/δ loci occurred independently in the different lineages. Alternatively, the transfer of VH genes into the TCRα/δ locus may have occurred frequently and repeatedly in the past and gene replacement is the best explanation for the current content of these genes in the different tetrapod lineages. The absence of VHδ in marsupials, the highly divergent nature of Vμ genes in this lineage, and the absence of conserved syteny with genes linked to TCRμ in the opossum, provide little insight into the origins of TCRμ and its relationship to TCRδ or the other conventional TCR (Parra et al. 2008). The similarity between VH, VHδ, and Vμ genes in the platypus genome, which are all clan III, however is striking. In particular, the close relationship between the platypus VHδ and Vμ genes lends greater support for the model presented in fig. 5E, with TCRμ having been derived from TCRδ genes.
Fig. 5. A model of the stages of evolution of the TCRα/β loci in tetrapods and the origins of TCRμ in mammals. A color key of the gene segments is presented at the bottom. (A) Depiction of the Dβ-Jβ-Cβ duplication in an ancestral TCRα/β locus that provides a second Cβ gene found in frogs and zebra finch. (B) Depiction of the insertion of a VH gene into the TCRα/β locus producing a current organization as it is found in zebra finch. (C) Depiction of the inversion/translocation and VHβ gene duplication that yielded the current organization found in frogs. (D) Depiction of the translocation of a VHδ-Dδ-Jδ-Cδ cluster to a location outside the TCRα/β locus generating a second TCRδ locus as it is currently found in chicken and turkey. (E) Depiction the translocation that took place in mammals giving rise to the TCRμ locus. (F) Loss of TCRμ in placental mammals, loss of D gene segments in cluster encoding the support V domain, retro-transposition to form a germ-line joined V in marsupials, and duplication of TCRμ clusters in both monotremes and marsupials.
The presence of TCR chains that use antibody-like V domains, such as TCRα using VH5, NAR-TCR or TCRμ are widely distributed in vertebrates with only the bony fish and placental mammals missing. In addition to NAR-TCR, some shark species also appear to generate TCR chains using antibody V genes. This occurs via trans-locus V(D)J recombination between IgM and IgW heavy chain V genes and TCRα and TCRα D and J genes (Criscitelli et al. 2010). This may be possible, in part, due to the multiple clusters of Ig genes found in the cartilaginous fish. It also illustrates that there has been independent solutions to generating TCR chains with antibody V domains in different vertebrate lineages. In the tetrapods, the VH genes were trans-located into the TCR loci where they became part of the germ-line repertoire. Whereas in cartilaginous fish something equivalent may occur somatically during V(D)J recombination in developing T cells. Either mechanism suggests there has been selection for having TCR using antibody V genes over much of vertebrate evolutionary history.

The current working hypothesis for such chains is that they are able to bind native antigen directly. This is consistent with a selective pressure for TCR chains that may bind or recognize antigen in ways similar to antibodies in many different lineages of vertebrates. In the case of NAR-TCR and TCRμ, the N-terminal V domain is likely to be unpaired and bind antigen as a single domain (fig. 1), as has been described for IgNAR and some IgG antibodies in camels (recently reviewed in Flajnik et al. 2011). This model of antigen binding is consistent with the evidence that the N-terminal V domains in TCRμ are somatically diverse, while the second, supporting V domains have limited diversity with the latter presumably performing a structural role rather than one of antigen recognition (Parra et al. 2007; Wang et al. 2011). There is no evidence of double V domains in TCRα chains using VH5 in frogs, birds, or platypus (fig. 1) (Parra et al. 2010, 2012). Rather, the TCR complex containing VH5 would likely be structured similar to a conventional γδTCR with a single V domain on each chain. It is possible that such receptors also bind antigen directly, however this remains to be determined.

A compelling model for the evolution of the Ig and TCR loci has been one of internal duplication, divergence and deletion; the so-called birth-and-death model of evolution of immune genes promoted by Nei and colleagues (Ota and Nei 1994; Nei et al. 1997). Our results in no way contradict that the birth-and-death mode of gene evolution has played a significant role in shaping these complex loci. However, our results do support the role of horizontal transfer of gene segments between the loci that has not been previously appreciated. With this mechanism T cells may have been able to acquire the ability to recognize native, rather than processed antigen, much like B cells.

**Supplementary Material**

Supplementary tables S1 and S2 are available at Molecular Biology and Evolution (http://www.mbe.oxfordjournals.org/).

**Acknowledgments**

The authors would like to thank Kathy Belov, University of Sydney, for providing access to spleen RNA from the NSW platypus. This work was made possible in part by grant No. P20 GM103452-09 from the National Institutes of Health, Institutional Development Award (IDeA) program.

**References**


