NIH Public Access
Author Manuscript
J Immunol. Author manuscript; available in PMC 2013 April 15.

Published in final edited form as:
J Immunol. 2012 April 15; 188(8): 3912–3919. doi:10.4049/jimmunol.1103521.

A second TCRδ locus in Galliformes uses antibody-like V domains: insight into the evolution of TCRδ and TCRμ genes in tetrapods[1]

Zuly E. Parra[*], **Kevin Mitchell**[*], **Rami A. Dalloul**[†], and **Robert D. Miller**[*,2]

[*]Center for Evolutionary and Theoretical Immunology, Department of Biology, University of New Mexico, Albuquerque NM 87131

[†]Avian Immunobiology Laboratory, Department of Animal & Poultry Sciences, Virginia Tech, 3170 Litton-Reaves Hall, Blacksburg, VA 24061, USA

Abstract

Analyses of the available avian genomes revealed the presence of a second *TCRδ* locus in the Galliformes. This second *TCRδ* locus is non-syntenic to the conventional *TCRα/δ* and is unusual in that the V genes are more related to IgH V genes (VH) than to TCR V genes. The second *TCRδ* is not found in another avian lineage, the passerine zebra finch. Rather the finch's conventional *TCRα/δ* locus contains VH genes that are expressed with the conventional Cδ gene, similar to what has been found in amphibians. A comparison between Galliformes and Passeriformes genomic organization suggests an origin of the second TCRδ in the former lineage involving gene duplication. Expression of these atypical TCRδ transcripts with a VH domain paired with Cδ was found in lymphoid tissues of both avian lineages. The configuration of the second TCRδ in chicken and turkey is reminiscent of the TCRδ duplication that is present in non-placental mammals and provides insight into the origin of the uniquely mammalian *TCRμ* locus.

Introduction

All jawed vertebrates, from cartilaginous fish to mammals, depend on an adaptive immune system that utilizes somatically diversified receptors (1). These receptors are the B cell receptors or Igs and the TCR, both of which consist of protein chains containing somatically diversified V and non-diversified C domains. Four TCR chains, α, β, γ and δ, are found in all jawed vertebrates (2). T cells express these chains on their surface as heterodimers of either a combination of α and β or γ and δ. These combinations are the defining hallmark of the two major T cell lineages: αβ and γδT cells (3). In both Ig and TCR the exons encoding the V domains are assembled from gene segments called the V, D and J genes for the IgH and TCRβ and δ chains or by V and J in the IgL and TCRα and γ chains (4). Somatic recombination of these gene segments is dependent on the RAG products, RAG1 and 2 (5–7). Both Ig and TCR are expressed on the surface of B and T cells, respectively, where they act as signaling receptors. Upon antigen activation, Ig can be secreted by effector B cells whereas the TCR remains a surface receptor (8). Ig and the conventional αβTCR differ in

[1]This work was made possible in part by grant No. IP20RR18754 of the National Institutes of Health, Institutional Development Award (IDeA) program of the National Center for Research Resources and by Agriculture and Food Research Initiative Competitive Grant No. 2010-65205-20412 from the USDA National Institute of Food and Agriculture.

[2]Correspondence: Robert D. Miller, MSC03 2020, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA. Fax: +1-505-277-0304. rdmiller@unm.edu.

Data deposition: The sequences presented in this article have been submitted to GenBank (http://www.ncbi.nlm.nih.gov/genbank/) under accession nos. JF936668-JF937040.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

how they recognize or bind to antigen. Ig typically bind native antigen directly, whereas αβTCR bind processed antigen presented on MHC molecules (9). Both direct and MHC restricted antigen binding have been described for γδTCR, and the role of the T cells that express this receptor remains somewhat enigmatic.

The TCRα, β, γ, and δ chains are present in all jawed vertebrates and their genes appear to be highly conserved both in sequence and organization (10). Furthermore, in the commonly studied placental mammals such as humans and mice they are the only TCR chains present (3, 11, 12). Recently, however, additional TCR forms have been described in a few distantly related non-placental species. Cartilaginous fish, for example, encode an unusual TCR chain called NAR-TCR that is expressed with three extracellular-domains, two V and one C (13). The C is the conventional Cδ, but the N-terminal V is encoded by genes that are highly similar to IgNAR V genes. IgNAR is an unusual light-chainless Ig unique to cartilaginous fish (14). The C- proximal domain is encoded by modified Vδ gene segments that lack leader peptide sequence (13).

A TCR with features analogous to NAR-TCR has also been found in marsupials and monotremes (*e.g.* opossum and platypus) (15, 16). This atypical TCR also is expressed with three extra-cellular domains, two V and a C and has been designated TCRμ, TCRμ utilizes V domains more related to VH than TCR V (15, 16). Whereas NAR-TCR is a modified conventional TCRδ chain, TCRμ is encoded by genes unlinked to the conventional TCRδ (15). TCRμ is most likely derived from a TCRδ gene duplication that occurred early in mammalian evolution, after the separation of diapsids (birds and reptiles) and synapsids (mammals) 310 million years ago (MYA).

Another variant of TCRδ has been described in the amphibian *Xenopus tropicalis* (17). The *X. tropicalis TCRα/δ*locus encodes TCRδ chains containing VH-like domains, called VHδ[3]. Like NAR-TCR in cartilaginous fish, the amphibian TCRδ variant uses a bona-fide Cδ region located in the *TCRα/δ* locus. However, like mammalian TCRμ, the *X. tropicalis* VHδ are more related to conventional Ig VH, not IgNAR V domains. Indeed the VHδ are indistinguishable from frog VH expressed in IgH chains, based on sequence, and the *TCRα/δ* and *IgH* loci are tightly linked in this species (17). There is no evidence that VHδ are used interchangeably with VH, however, and are adapted specifically for use in TCRδ chains. Unlike either NAR-TCR or TCRμ, which contain three extra-cellular Ig domains, the frog TCRδ are expressed with two extra-cellular domains, V and C, like conventional TCR.

We have recently speculated that mammalian TCRμ arose from a duplication of an ancestral *TCRδ* locus organized similar to the *X. tropicalis TCRα/δ* locus (11, 16). To test this hypothesis, we investigated the content and organization of genes encoding the TCRδ chains in those avian species for which useful genome sequence was available (18–20). Here we report the presence of a second *TCRδ* locus in the galliform species, chicken and turkey, and the organization of the *TCRα/δ* locus in a passerine, the zebra finch, that provide insight into the evolutionary history of these genes and the origins of TCRμ.

Materials and Methods

Genome analyzes

The chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*) and zebra finch (*Taeniopygia guttata*) genomes were carefully scanned for TCRδ and VHδ-like genes. Cδ and VHδ genes were compared to all the three avian genomes using the BLAST/BLAT tool from Ensembl (www.ensembl.org). The assembly's versions used were Chicken 2.1 (WASHUC2), Turkey

[3]**Abbreviations:** VHδ, V genes that are similar to VH but found expressed with Cδ.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

2.01 (UMD2) and zebra finch 1.1 (taeGut3.2.4). The chicken (Build2.1) and zebra finch (Build1.1) genomes were also examined using the whole genome BLAST available at NCBI (www.ncbi.nlm.nih.gov/). Chromosomes found to contain genes of interest were retrieved from Ensembl and further examined using the BLAST algorithm. To physically confirm the location of the VHδ near Vδ3 gene in the zebra finch assembly, PCR on genomic DNA was performed. Genomic DNA was extracted from spleen of zebra finch using the QIAGEN DNeasy blood and tissue kit (QIAGEN Sciences, Germantown, MD) and PCR was performed using primers located in the Vδ3 and VHδ genes. The Vδ3 forward primer 5′-TCCGGCTTCACCTTCGAGAATCA-3′ and the VHδ reverse primer 5′-GGTGGCTGTGTCTGCAGCTACTGG-3′.

RT-PCR

All procedures involving live animals were approved under institutional protocol number 10-100515-MCC. Thymuses were collected from day 4 and day 21 old chickens. Thymus and spleen were collected from a zebra finch male approximately six months old. Tissues were preserved in RNA*later* (Ambion, Austin, TX). Total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA). RT-PCR was performed using GenAmp RNA PCR core kit (Applied Biosystems, Foster City, CA). PCR reactions were performed using the AdvantageTM-HF-2 kit (BD Biosciences, CLONTECH Laboratories, Palo Alto, CA). The PCR products were cloned using TopoTA cloning® kit (Invitrogen, Carlsbad, CA). Sequencing was performed using the BigDye terminator cycle sequencing kit v3 (Applied Biosystems, Foster City, CA) 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). Chromatograms were analyzed using the Sequencher TM4.9 software (Gene Codes Corporation, Ann Arbor, MI). Data have been entered on GenBank under accession numbers JF936668 –JF937040.

5′ RACE was performed on zebra finch thymus and spleen using the GeneRacer kit from Invitrogen (Invitrogen, Carlsbad, CA). Nested 5′ RACE amplification was performed using the following primers: Cδ1 Outside primer: 5′-GGCCATGCAGGTCACCTCTGTGT-3′; Cδ1 Nested primer: 5′-GCTTTCCCTGTGCTTCCCCCTTC-3′. Cδ2 Outside primer: 5′-GCCGTGCAGGTCACCTCTGTGTC-3′. Cδ2 Nested primer: 5′-TGCTTTTCCAGAGCTTCCCCCTTC-3′. Cα Outside primer 5′-TCCTCGCTGTTCTCCATGGTTGC-3′. Cα Nested primer: 5′-TTCCAGACCCTGGTGGGGACAAT-3′.

Additionally, primers specific for VHδ gene were paired with primers specific for both Cδ and Cα genes. Forward primer used for VHδ: 5′-GGTTCACCTGTCACATCTCTGGTG-3′.

For chicken, primers complementary to VHδ were paired with primers complementary to Cδ2: Forward on VHδ 5′-CCCAGGGAAGGGACAGTTTCTGG-3′; Reverse on Cδ2 5′-GTCACCCTTGGGCCCATCAAGAC-3′.

Phylogenetic analyzes

For the V genes analyzes, the nucleotide sequence from FR1 to FR3 regions, including CDR1 and CDR2, were aligned using BioEdit (21) and the accessory application ClustalX (22). Alignments were based on amino acid sequence, using the toggle translation on BioEdit (21). Alignments were corrected by visual inspection when necessary and were then analyzed using the MEGA Software (23). 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. Exon 1 of the C genes, encoding IgC domain, were aligned and phylogenetic

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

trees were generated as described for the V genes. GenBank accession numbers for sequences used in the construction of the phylogenetic trees presented in this paper are shown in Supplemental data Table 1.

Results

TCRδ transcripts using VH genes are present in birds

Little is known regarding avian TCRδ genetics or genomics. The finch *TCRα/δ* locus was located on chromosome 27 and analyzed for its V, D, J, and C content. Two Cδ genes sharing 95% nucleotide identity and a single Cα were identified (Fig. 1). Upstream of the most 5′ Cδ (Cδ1) were a single D and two J segments. The second Cδ (Cδ2) had a single associated upstream D and J. At the 5′ end of the locus were 10 Vα and three Vδ, with a fourth Vδ located between the two Cδ genes. Surprisingly, also present immediately upstream of the Cδ1 cluster was a V gene segment that shared greater identity to Ig VH genes. Indeed, phylogenetic analysis of this VH-like gene revealed that it clustered within Ig VH clan I when compared to Ig and TCR V sequences from mammals, birds and amphibians (Fig. 2). Following the nomenclature established for *X. tropicalis*, this finch VH-like gene was designated as VHδ (17). Given the atypical nature of finding VH genes in TCR loci, further confirmation that a VHδ is indeed located within the zebra finch TCRα/δ locus was sought. In the current finch genome assembly, the VHδ is only 633 bp from one of the Vδ genes (Vδ3) (Fig. 1). PCR was performed on zebra finch genomic DNA using primers located within these two gene segments to see if this region could be amplified. A predicted 1130 bp fragment that included part of Vδ3 and VHδ and the intervening sequence was amplified. This product was sequenced and found to match the genome assembly, physically confirming that there is a VHδ located 633 bp downstream of Vδ3 in the zebra finch TCRα/δ locus. Therefore its presence in the TCRα/δ locus is not an assembly artifact.

All V gene segments, including the VHδ are flanked by canonical 23 bp spacer recombination signal sequences (RSS), typical of V genes. The D genes have 12 bp spacers at the 5′ end and 23 bp spacers at the 3′ end, also typical of D segments used in TCRδ. This asymmetrical organization of RSS allows D to D recombination as is common in TCRδ chains (24). The J genes have 12bp spacer RSS at the 5′ end and a conserved splice site at the 3′ end (15, 17). The finch VHδ also encodes canonical cysteine residues necessary for intra-chain disulfide bonds and appears completely functional (Fig. 3). To investigate the use of the VHδ in zebra finch TCRδ transcripts, 5′ RACE PCR were performed on thymus and spleen RNA using Cδ specific primers. One hundred twenty individual, unique TCRδ transcripts were analyzed. The majority of these transcripts contained V regions encoded by Vδ (n =117) (GenBank accession numbers JF936922 – JF937040). Three clones, all from thymus RNA, used the VHδ, consistent with this V gene being used in V(D)J recombination and expressed as part of the TCRδ repertoire. Of the three VHδ clones, one was a non-productive rearrangement (clone JF936921 in Fig. 4A). Two types of VHδ transcripts were amplified. The first type contained VHδ recombined with Dδ1 and Jδ1.1 and transcribed with Cδ1 from the most 5′ D-J-C cluster (Fig. 4A). The second type contained VHδ recombined with both Dδ1 and Dδ2 gene segments and used the Jδ2 and Cδ2 genes (Fig 4A). One of the clones using a Dδ1 alone has an unpaired cysteine present in the CDR3. This is due to N-additions in the junction and, therefore, may be random rather than performing a conserved function. Similar to *X. tropicalis* TCRδ using VHδ, the finch clones would encode a TCRδ chain with two extra-cellular Ig domains, VHδ and Cδ, not three as found in shark NAR-TCR and mammalian TCRμ. The conventional TCRδ clones, using Vα and Vδ genes, also use the same two clusters of D and J segments for V(D)J recombination (not shown). The recombination and expression of VHδ in TCRδ chain transcripts also supports its presence in the TCRα/δ locus not being an assembly artifact. In other words the

VHδ is clearly in the same locus with the D and J segments used with conventional Vα and Vδ genes.

A second TCRδ locus in chickens uses only VHδ

To investigate VHδ genes in other avian species, the chicken *TCRα/δ* locus was examined. The conventional chicken *TCRα/δ* locus was identified on chromosome 27 in the latest genome assembly (SD Fig. 1). Chicken chromosome 27 shares a great deal of conserved synteny with that of the zebra finch, however, some rearrangements, mostly inversions have occurred. No evidence of a VHδ gene in the chicken *TCRα/δ* locus was found. Rather, when comparing the finch VHδ gene segments to the whole chicken genome, a homologue was found on the tip of chromosome 10 unlinked to *TCRα/δ* locus (Fig. 5). Closer examination of the genes on chicken chromosome 10 revealed the presence of a second *TCRδ* locus comprised of a single cluster of VHδ, Dδ, Jδ, and Cδ genes (Fig. 5). The chicken and finch VHδ shared 80 percent nucleotide identity and is also related to clan I VH genes in other vertebrates (Figs. 2 and 3). In addition to the canonical cysteine residues necessary for intra-chain disulfide bond, the chicken VHδ gene encodes an unpaired cysteine residue in the region between CDR1 and FR2 (Fig. 3). Similar extra cysteine residues have been observed in frog VHδ (17). The organization of the RSS were as in the finch *TCRα/δ* locus and also typical of conventional TCRδ. The Cδ region located on chromosome 10, hereafter named Cδ2, shares only 59% nucleotide identity with the Cδ (Cδ1) located in the conventional *TCRα/δ* locus on chromosome 27 (Fig. 3 and 6). The chicken Cδ2 gene is also unusual in that it encodes an extra, unpaired cysteine residue that may facilitate additional inter-chain bonds (Fig. 3). The region homologous to chicken chromosome 10 was also examined in the finch genome. This region corresponds to finch chromosome 10 as well (SD Fig. 2). Although there is substantial conserved synteny between these avian chromosomes 10, there was not evidence of TCR related genes on the finch chromosome (SD Fig. 2).

Primers complementary to VHδ and Cδ2 were used in RT-PCR on chicken thymus RNA obtained from day 4 and day 21 post-hatch chicks to investigate expression from the second *TCRδ* locus. Products cloned and sequenced confirm the presence of a single V-D-J gene segment combination (Fig. 4B and 5). Thymic transcripts were isolated and compared from two day 4 and four day 21 post-hatch chicks and almost no nucleotide variation in the V, D, J and C genes was found (Fig. 4B). Moreover, even though there was variation in the CDR3 length, the D gene segment was relatively easy to identify (Fig. 4B). These results are consistent with the presence of single gene copies, low allelic variation and a strong selective pressure to maintain these invariant alleles in chickens.

Other avian species with have a second TCRδ

Given that a passerine (zebra finch) and a galliform (chicken), which last shared a common ancestor approximately 90 to 120 MYA (25) both encode TCRδ using VHδ, but are organized in radically different ways, it was of interest to investigate other avian species. The remaining avian genome available was that of another galliform, the turkey (*Meleagris gallopavo*). Turkey chromosome 29 contains the conventional *TCRα/δ* locus and is clearly homologous to chicken and finch chromosomes 27 (SD Fig. 1). However, in the current assembly the turkey *TCRα/δ* locus contains many gaps or regions that were not completely sequenced. Although Cδ, Cα and conserved flanking genes were found, we were not able to identify any V genes in this locus. Turkey chromosome 10, however, was found to contain a *TCRδ* locus homologous and similarly organized to that found on chicken chromosome 10 (Fig. 5). Turkey and chicken VHδ and Cδ2 each share greater than 94% nucleotide identity and are likely orthologous (Fig. 2, 3 and 6). Turkey VHδ and Cδ2 genes also encode unpaired cysteine residue in the CDR1-FR2 boundary of VHδ and in exon 1 of the Cδ2 (Fig. 3). Turkey D and J gene segments also share a high percent identity when compared to the

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

chicken (97 and 96% respectively, Fig. 4C and D). The high percent similarity and the similar organization observed between chicken and turkey is not surprising. These two species shared a last common ancestor approximately 40 MYA (26). However, some reorganization of the second TCRδ locus has occurred. Conserved nucleotides in turkey and finch germ-line Dδ genes, when compared with chicken and duck Dδ, suggest that in the past two D genes segments became germ-line joined to produce a longer D segment (Fig. 4D). Whereas in the chicken and duck this fused Dδ segment appears to have become germ-lined joined to the Jδ (Fig. 4D). Although both chickens and turkeys have a second TCRδ locus on their respective chromosomes 10, these chromosomes are not completely homologous. Rather the majority of chicken chromosome 10 shares conserved synteny with turkey chromosome 12 (SD Fig. 2).

In addition to the whole genome sequences, we searched for cDNA sequences that contain atypical TCRδ genes in the EST and GenBank databases available at NCBI. Remarkably, we found a complete cDNA sequence obtained from the spleen of a white pekin duck (*Anas platyrhynchos*). This sequence contains 5′ UTR, leader sequence, VHδ, D and J gene segments and Cδ region with connecting peptide, transmembrane and cytoplasmic regions (GenBank AF415216). Alignment of the cDNA duck sequence with the TCRδ genes from chicken and turkey, demonstrate that this duck TCRδ transcript is homologous to that encoded by the second TCRδ in Galliformes (Fig. 2, 3, 4C/D and 6). Duck VHδ shares high percent nucleotide identity to the turkey and chicken VHδ genes (84% and 85%, respectively), and fall in the same phylogenetic group (Fig. 2). The duck Cδ shares 81 and 85 percent nucleotide identity with the turkey and chicken Cδ2 sequences, respectively (Fig. 3 and 6). Curiously, the duck VHδ and Cδ genes lack the extra cysteine residues found in the Galliformes (Fig. 3). The duck CDR3 sequence was aligned with germ-line gene segments from chicken (Fig. 4C). Sequences corresponding to V, D and J gene segments can be easily identified given the high percent identity with the chicken germ-line genes (Fig. 4C). These results are consistent with ducks clearly having TCRδ that use VHδ, but the location of these genes relative to the conventional TCRδ is not known. However, the duck VHδ and Cδ share greater similarity to that of the second TCRδ locus in chicken than the TCRα/δ locus in finch and, although speculative at this point, the duck is likely to have similar organization as that of the chicken.

Since the duck TCRδ transcript containing a VHδ was full-length, it enabled analysis of the feature found in the connecting peptide and transmembrane regions. Conserved in the duck connecting peptide region is a cysteine that forms inter-chain disulfide bonds with TCRγ (not shown) (3). Also conserved are the lysine and arginine residues in the transmembrane region that play a role in the association with the CD3 complex (not shown) (3). The exons encoding the connecting peptide and transmembrane regions were also identified in the chicken genome and also encode these conserved sites (not shown). These features are consistent with avian TCRδ chains using VHδ forming heterodimers with a second chain, most likely TCRγ.

Evolution of TCRδ genes

Comparison of all the Cδ genes from zebra finch, chicken, turkey and duck to the respective ones from placental mammals, marsupials, monotremes, amphibians and fish, resulted in a well-supported cluster that contains all avian Cδ sequences (100% bootstrap support, Fig. 6 and Fig. 7). However, within the avian cluster, two distinct groups were formed. One group contained the conventional Cδ regions from chicken and turkey (Cδ1) and the zebra finch Cδ1 and Cδ2 (Fig. 6). The other group contained the Cδ2 from chicken and turkey (Fig. 6). The only duck Cδ sequence available fell in the same cluster with the Galliformes Cδ2, further supporting the organization of the duck TCRδ genes being similar to that of the Galliformes Cδ2 (Fig. 6).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The avian Cδ cluster was sister to a clade containing mammalian Cδ and Cμ (Fig. 6). Placental and marsupial Cδ formed a well-supported group (98% support) and the platypus and opossum Cμ are clearly closely related. However the phylogenetic position of the Cδ from platypus is not well defined and varies between being basal to all the mammalian Cδ, or being basal to all mammalian Cδ and Cμ regions (Fig. 6).

Discussion

The organization of the *TCRα/δ* locus is highly conserved across amphibians, birds and mammals (11, 17). This includes a high degree of conserved synteny with genes flanking the *TCRα/δ* locus making it appear to be an evolutionarily stable genomic region. When the frog *X. tropicalis TCRα/δ* locus was characterized, the first for an amphibian, V genes that were indistinguishable from antibody VH were found within the locus in a transcriptional orientation upstream of Cδ1, one of the two Cδ genes present in this species. Further characterization revealed these VHδ genes underwent VDJ recombination and used in TCRδ chains and, although Vα and Vδ are available and used, the majority of V genes used in *X. tropicalis* TCRδ are VHδ (17). In frogs, *IgH* and *TCRα/δ* are closely linked and an ancient synteny between these loci may have facilitated translocation of VH into the *TCRα/δ* locus in tetrapod ancestors (17). Indeed, the VHδ used in TCRδ chains are highly similar to VH genes used in IgH chains, although the VHδ appear dedicated for use in TCRδ and there is no evidence of trans-locus VDJ recombination (17). The discovery that the zebra finch *TCRα/δ* locus, like that of amphibians, also contains VHδ on its own is not surprising, and confirms this as the likely ancestral state for the locus in tetrapods. In fact, the platypus *TCRα/δ* locus contains a single VHδ gene segment located downstream of Vα/δ genes and upstream of the Cδ, within the otherwise conserved *TCRα/δ* locus (Fig. 7). However, the *TCRα/δ* loci of humans, mice, and opossums have been characterized in detail and they do not contain VHδ and therefore, these genes have been lost in marsupials and placental mammals (11, 27, 28).

The genes encoding TCRδ chains using VHδ appear to be genomically mobile in the avian lineage (Fig. 7). Instead of being part of the *TCRα/δ* locus as in passerines, the genes have translocated to a separate chromosome in the Galliformes, and also likely in the Anseriformes, creating a second *TCRδ* locus containing a single VHδ, D, J, and Cδ cassette. In the case of the chicken they were translocated to the tip of chromosome 10. Chromosome 10 in the turkey underwent subsequent rearrangement such that now it is not homologous to that of chicken. All VHδ were simultaneously or subsequently lost from the Galliformes *TCRα/δ* locus. Consequently, the chicken *TCRα/δ* locus appears structured more like that of marsupials and placental mammals, whereas the finch locus was more similar to that of *Xenopus.* This would suggest that in birds there has been some evolutionary pressure to retain TCRδ chains using antibody-like V genes, whether in or out of the *TCRα/δ* locus. Given the incomplete nature of the turkey genome sequence, caution must be applied to the results presented here. It is acknowledged that the turkey assembly was helped by the availble chicken genome, however three independent turkey linkage maps were available and used (19). A more recent comparative analysis of the chicken and turkey genomes revealed a high degree of chromosomal stability among the subfamily of birds that includes these species further supporting the analyses presented here (29).

The D segments found in TCRδ of all vertebrates have an asymmetrical RSS that allows for D to D recombination (24). D to D recombination in avian TCRδ using VHδ is only possible in the finch where there are two Dδ to recombine with. Curiously, the one of the finch VHδ clones had an unpaired cysteine in the CDR3, apparently due to random N-additions. Whether there is a functional significance to this cysteine is not known. However, in chicken and turkey there are unpaired cysteines at the CDR1-FR2 boundary. Unpaired cysteines

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

have also been found in V domains of shark NAR-TCR where they are thought to provide inter-domain stability (13). Whether they perform a similar role in avian TCRδ remains to be determined.

As in frogs, the extra-cellular domains of avian TCRδ using VHδ are structured like conventional TCR chains with a single V and single C domain. Furthermore, the C region used is a conventional TCRδ C and most likely pairs with TCRγ chains on γδT cells. One prediction would be that γδTCR containing VHδ might bind antigen directly, like antibody, rather than through MHC restriction, similar to what has been found in some γδTCR in mammals (30). Whether this is the case or not remains to be seen but might be the basis of selective pressure to retain these atypical TCRδ.

The VHδ genes used in TCRδ are clearly derived from Ig VH genes. Phylogenetic analyses of amphibian, avian, and mammalian VH have long revealed three ancient lineages, or clans, of VH genes designated I, II, and III (31). In *X. tropicalis*, the VHδ are related to clan II VH (Fig. 2), which are found in the frog *IgH* locus, and we previously hypothesized that close linkage between *TCRα/δ* and *IgH* maintained the similarity between VHδ and clan II VH either through gene conversion or gene replacement (17). All avian VHδ group with clan I VH (Fig. 2). Curiously, all known avian VH are clan III and birds are thought to only have clan III VH genes. It is likely that avian ancestors contained all three VH clans, or at least I and III given all three are present in amphibians and mammals. The avian VHδ must be evolving very slowly or are under purifying selection in that they are still clearly clan I-related although there are no clan I VH genes remaining in the avian *IgH* locus (Fig. 2). Therefore, close linkage between *IgH* and *TCRα/δ* is not necessary to maintain the presence of VHδ in TCR genes nor is the *IgH* locus necessary to constrain their evolution; at least not in birds.

The second Galliform *TCRδ* locus, unlinked to *TCRα/δ* and using VHδ exclusively, provides a potential "missing link" to understanding the origins of TCRμ in mammals. TCRμ, which is uniquely mammalian, was first discovered in marsupials but later shown to be present in monotremes as well (15, 16). TCRμ C regions are most related to TCRδ and it is likely derived from the duplication and divergence of TCRδ genes. TCRμ's phylogenetic distribution suggests it was present in the ancestors of all living mammals although lost in the placental mammals (11, 16). The V domains of opossum TCRμ are more closely related to VH genes than TCR V genes, but are clearly distinguishable from Ig VH. In the platypus, the TCRμ V domains and the VHδ that remains in the *TCRα/δ* locus are clearly related to clan III VH (Fig. 2 and 6, 16).

Previously, we proposed a model for the origin of TCRμ that predicted its derivation from a primordial gene cluster similar to the second Galliform *TCRδ* locus (11). Given the phylogenetic position of the Galliformes within the avian tree, it is likely that the current arrangement of TCRδ genes in chickens and turkeys is a derived form unique to this order or possibly the Galloanserae (including ducks, chickens and turkeys), and is not the ancestor of TCRμ. Rather, the avian arrangement demonstrates that the TCRδ genes using VHδ are mobile and can be translocated and retained (Fig. 7). The phylogenetic analyses of Cμ relative to Cδ are consistent with Cμ having diverged from Cδ after the split between mammals and birds/reptiles 310 MYA (15, 16). This independent origin of TCRμ and the second Galliformes TCRδ locus is also supported by the apparent separate origins of their V genes as detailed above. TCRμ likely, therefore, evolved from its own translocation event that occurred early in the lineage leading to mammals.

TCRμ has also undergone substantial re-organization and expansion by gene duplication in mammals compared to the second *TCRδ* locus in chickens and turkeys. This has resulted in a

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

more complex genomic structure with multiple tandem TCRμ clusters, each containing genes organized as Vμ-D_n-J-Vμj-Cμ in the opossum and Vμ-D_n-J-Vμ-J-Cμ in the platypus (15, 16). Opossum Vμj is a complete germ-line joined V gene most likely generated by retrotransposition. The repeating V(D)J organization in both mammalian lineages allows for encoding a mature form of TCRμ with three extra-cellular domains: double V domains and a C (15, 16). In this regard, TCRμ shares structural features analogous to those found in shark NAR-TCR, which also contains three extra-cellular domains (13). The N-terminal V in NAR-TCR is related to IgNAR antibody V domains and it is likely that NAR-TCR represents an evolutionary gene lineage that evolved uniquely in sharks. However, the N-terminal V of TCRμ and NAR-TCR, which have somatic diversity, are unpaired and likely function similarly as a single antigen-binding domain analogous to what has been shown for IgNAR antibodies in sharks and light-chainless IgG in camels (32). The similarity between NAR-TCR and TCRμ however is likely due to convergent evolution.

An obvious question is whether mammalian TCRμ chains and shark NAR-TCR are functionally analogous to TCRδ chains in frogs and birds that use VHδ. Based on the predicted structure, with their unpaired N-terminal V domains, TCRμ and NAR-TCR are likely to have a very different antigen binding interaction than do avian or amphibian γδTCR containing VHδ. However, all three TCR types may endow T cells with the common feature of direct antigen binding. Why neither form currently exists in placental mammals is not known. Perhaps the evolution of direct antigen binding in conventional γδTCR cells reduced the evolutionary pressure to retain them.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Flajnik MF, Kasahara M. Origin and evolution of adaptive immune system: genetic events and selective pressures. Nat Rev Genet. 2010; 11:47–59. [PubMed: 19997068]
2. Rast JP, Anderson MK, Strong SJ, Luer C, Litman RT, Litman GW. Alpha, beta, gamma, and delta T cell antigen receptor genes arose early in vertebrate phylogeny. Immunity. 1997; 6:1–11. [PubMed: 9052832]
3. Davis, MM.; Chein, YH. T cell antigen receptors. In: Paul, WE., editor. Fundamental Immunology. 6. Lippincott; Philadelphia: 2008. p. 313-345.
4. Litman GW, Rast JP, Fugmann SD. The origins of vertebrate adaptive immunity. Nat Rev Immunol. 2010; 10:543–53. [PubMed: 20651744]
5. Tonegawa S. Somatic generation of antibody diversity. Nature. 1983; 302:575–581. [PubMed: 6300689]
6. Yancopoulos GD, Blackwell TK, Suh H, Hood L, Alt FW. Introduced T cell receptor variable region gene segments recombine in pre-B cells. Evidence that B and T cells use a common recombinase. Cell. 1986; 44:251–259. [PubMed: 3484682]
7. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell. 1989; 59:1035–1048. [PubMed: 2598259]
8. Pierce SK, Liu W. The tipping points in the initiation of B cell signaling: how small changes make big differences. Nat Rev Immunol. 2010; 10:767–777. [PubMed: 20935671]
9. Huseby ES, White J, Crawford F, Vass T, Becker D, Pinilla C, Marrack P, Kappler JW. How the T cell repertoire becomes peptide and MHC specific. Cell. 2005; 122:247–260. [PubMed: 16051149]
10. Chen H, Kshirsagar S, Jensen I, Lau K, Covarrubias R, Schluter SF, Marchalonis JJ. Characterization of arrangement and expression of the T cell receptor gamma locus in the sandbar shark. Proc Natl Acad Sci USA. 2009; 106:8591–8596. [PubMed: 19439654]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

11. Parra ZE, Baker ML, Hathaway J, Lopez AM, Trujillo J, Sharp A, Miller RD. Comparative genomic analysis and evolution of the T cell receptor loci in the opossum *Monodelphis domestica*. BMC Genomics. 2008; 9:111. [PubMed: 18312668]
12. Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. Annu Rev Immunol. 1999; 17:109–147. [PubMed: 10358755]
13. Criscitiello MF, Saltis M, Flajnik MF. An evolutionarily mobile antigen receptor variable region gene: doubly rearranging NAR-TcR genes in sharks. Proc Natl Acad Sci USA. 2006; 103:5036–5041. [PubMed: 16549799]
14. Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA. Maturation of shark single-domain (IgNAR) antibodies: evidence for induced-fit binding. J Mol Biol. 2007; 367:358–372. [PubMed: 17258766]
15. Parra ZE, Baker ML, Schwarz RS, Deakin JE, Lindblad-Toh K, Miller RD. A unique T cell receptor discovered in marsupials. Proc Natl Acad Sci USA. 2007; 104:9776–9781. [PubMed: 17535902]
16. Wang X, Parra ZE, Miller RD. Platypus TCRμ Provides Insight into the Origins and Evolution of a Uniquely Mammalian TCR Locus. J Immunol. 2011; 187:5246–5254. [PubMed: 21976776]
17. Parra ZE, Ohta Y, Criscitiello MF, Flajnik MF, Miller RD. The dynamic TCRδ: TCRδ chains in the amphibian *Xenopus tropicalis* utilize antibody-like V genes. Eur J Immunol. 2010; 40:2319–2329. [PubMed: 20486124]
18. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004; 432:695–716. [PubMed: 15592404]
19. Dalloul RA, et al. Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. PLoS Biol. 2010; 8:e1000475. [PubMed: 20838655]
20. Warren WC, et al. The genome of a songbird. Nature. 2010; 464:757–762. [PubMed: 20360741]
21. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999; 41:95–98.
22. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 1997; 24:4876–4882. [PubMed: 9396791]
23. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform. 2004; 5:150–163. [PubMed: 15260895]
24. Carroll AM, Slack JK, Mu X. V(D)J recombination generates a high frequency of nonstandard TCR delta-associated rearrangements in thymocytes. J Immunol. 1993; 150:2222–2230. [PubMed: 8383718]
25. Van Tuinen M, Hedges SB. Calibration of avian molecular clocks. Mol Biol Evol. 2001; 18:206–213. [PubMed: 11158379]
26. Van Tuinen M, Dyke GJ. Calibration of galliform molecular clocks using multiple fossils and genetic partitions. Mol Phylogenet Evol. 2004; 30:74–86. [PubMed: 15022759]
27. Satyanarayana K, Hata S, Devlin P, Roncarolo MG, De Vries JE, Spits H, Strominger JL, Krangel MS. Genomic organization of the human T-cell antigen-receptor a/d locus. Proc Natl Acad Sci USA. 1988; 85:8166–8170. [PubMed: 3186718]
28. Wang K, Klotz JL, Kiser G, Bristol G, Hays E, Lai E, Gese E, Kronenberg M, Hood L. Organization of the V gene segments in mouse T-cell antigen receptor alpha/delta locus. Genomics. 1994; 20:419–428. [PubMed: 8034314]
29. Zhang Y, Zhang X, O'Hare TH, Payne WS, Dong JJ, Scheuring CF, Zhang M, et al. A comparative physical map reveals the pattern of chromosomal evolution between the turkey (*Meleagris gallopavo*) and chicken (*Gallus gallus*) genomes. BMC Genomics. 2011; 12:447. [PubMed: 21906286]
30. Hayday AC. γδ T cells and the lymphoid stress-surveillance response. Immunity. 2009; 31:184–196. [PubMed: 19699170]
31. Ota T, Nei M. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin VH gene family. Mol Biol Evol. 1994; 11:469–482. [PubMed: 8015440]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

32. Flajnik MF, Deschacht N, Muyldermans S. A case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? PLoS Biol. 2011; 9:e1001120. [PubMed: 21829328]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

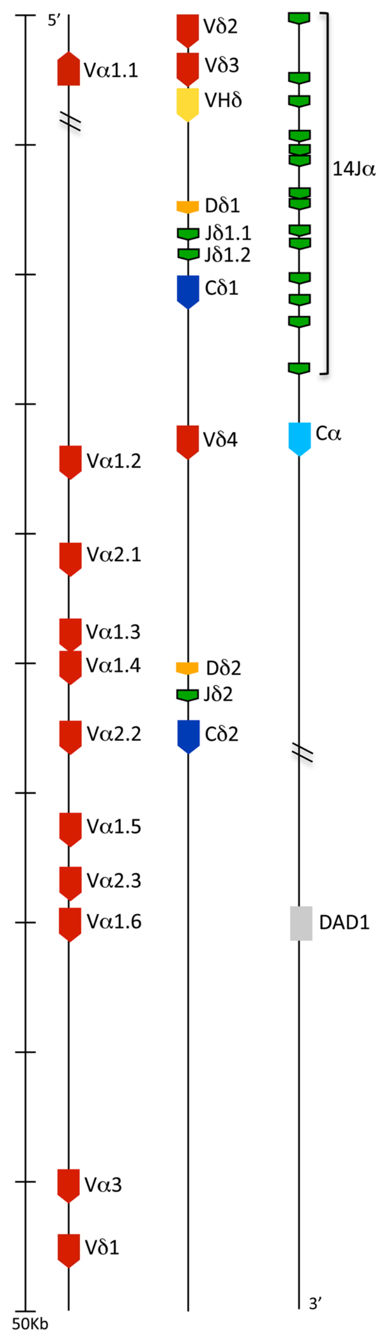

Figure 1.
Zebra finch TCRα/δ locus. V, D and J gene segments and C regions were numbered based on their position on the locus (5′ to 3′) and were color coded as follow: V, red; D, orange; J, green; Cδ, dark blue; Cα, light blue. The VHδ gene segment is indicated in yellow. Genes with conserved synteny in tetrapods are shaded light grey. Transcriptional orientation is shown with the direction of the arrow in each segment.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

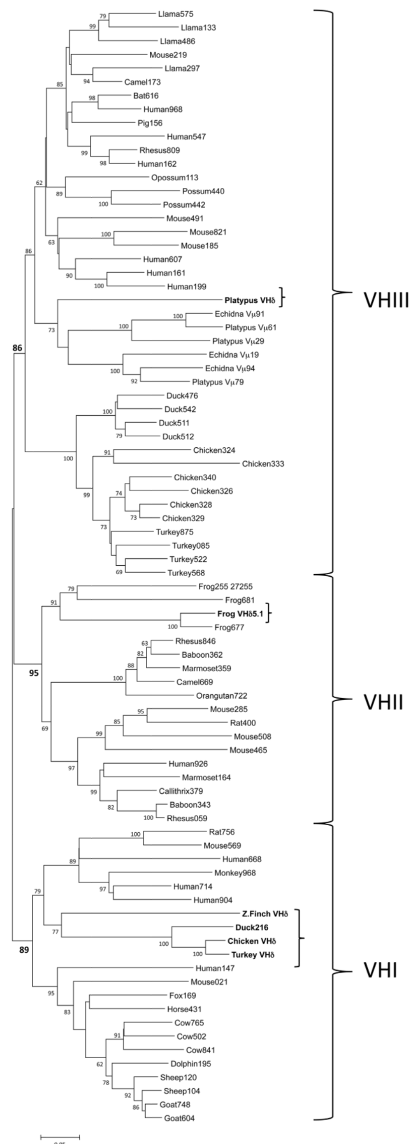

Figure 2.

Phylogenetic tree of the avian VHδ compared with VH genes from birds, mammals and amphibians. Avian VHδ genes are in bold and bracketed. The tree shown was generated using the Minimum evolution distance method. Similar results were achieved using Neighbor Joining. Bootstrap values are based on 1000 replicate samples. The last three digits of the accession number are indicated for those sequences taken from GenBank. The three VH clans are indicated with brackets on the right and their bootstrap support is shown in bold. A distance bar is shown below the tree.

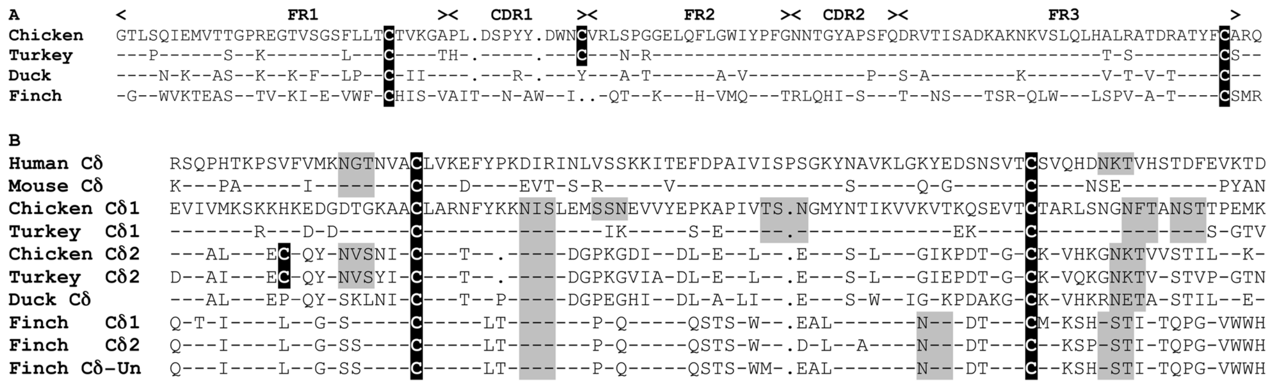

Figure 3.

Predicted amino acid alignment of atypical avian TCRδ genes contain residues conserved in conventional Ig domains. Translations were based on nucleotide sequence from the genomic assemblies, except for duck that is cDNA. **A**. Predicted amino acid alignment of frameworks (FR) FR1 to FR3 regions of the VHδ from four different avian species. FR and CDR are indicated at the top of the alignment. Dashes indicate identity with the first sequence. Dots indicate gaps. Cysteine residues involved in intrachain disulfide bond and extra cysteine residues are highlighted in black with white letters. **B**. Alignment of Ig-C domain of conventional and atypical TCRδ. Human and mouse sequences are shown as references on top of the alignment. Dashes in the mouse sequence indicate identity to the human sequence. Dashes in all avian Cδ sequences indicate identity with the chicken Cδ1 sequence. Gaps are indicated with dots. Cysteine residues are indicated as in **A**. Potential glycosylation sites are highlighted in light gray.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

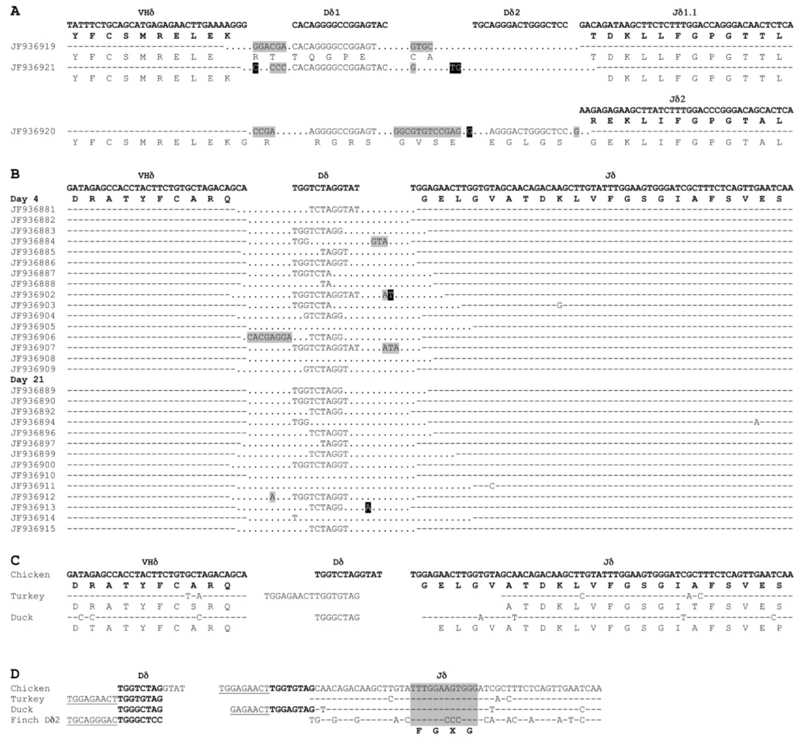

Figure 4.

Junctional diversity in avian atypical TCRδ CDR3 regions from different avian species. GenBank accession numbers are shown on the left of each sequence. P and N nucleotides are highlighted in black and grey, respectively. V-D-J germline sequences are at the top of the figure and indicated in bold. Gaps are shown with dots. Dashes indicate identity with the sequence at the top of each alignment for V and J gene segments. **A.** Adult zebra finch unique thymus cDNA sequences. D and J gene segments are indicated above the sequences. **B**. Unique V-D-J recombination sequences obtained from day 4 and day 21 chicken thymuses. Germline sequences are shown as described above. **C.** The 3′ end of the germ-line VHδ, the complete germ-line D, and complete germ-line J genes (un-recombined) from the chicken turkey genome assembles compared with the corresponding regions from the duck cDNA sequence (AF415216). **D.** Comparison of Dδ and Jδ germline gene segments from chicken, turkey, zebra finch (Dδ2 and Jδ2) and duck cDNA sequence. Conserved motif in the Jδ is highlighted in grey. Similar sequences between Dδ and Jδ gene segments are indicated in bold or underlined.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

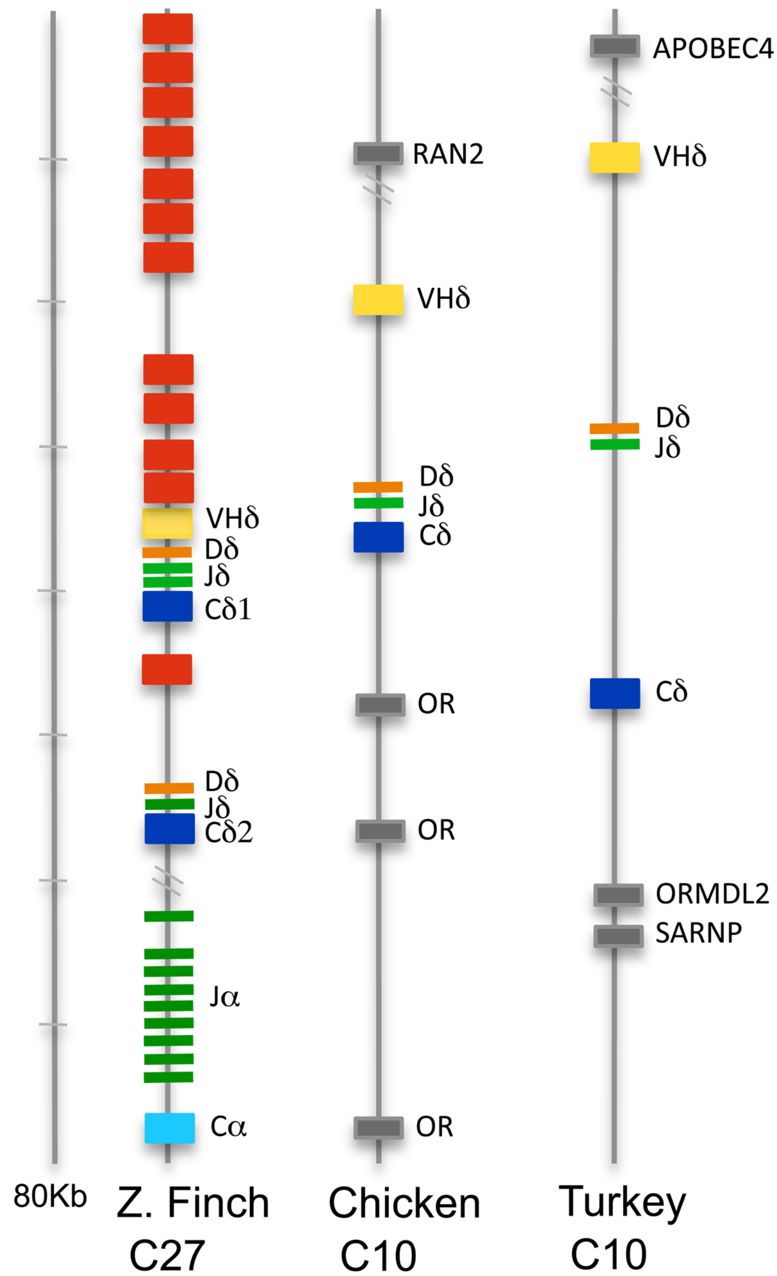

Figure 5.
Representation of the zebra finch *TCRα/δ* locus and of a second *TCRδ* locus in chicken and turkey. V, D and J gene segments and C regions were color coded as in Fig. 1.

NIH-PA Author Manuscript

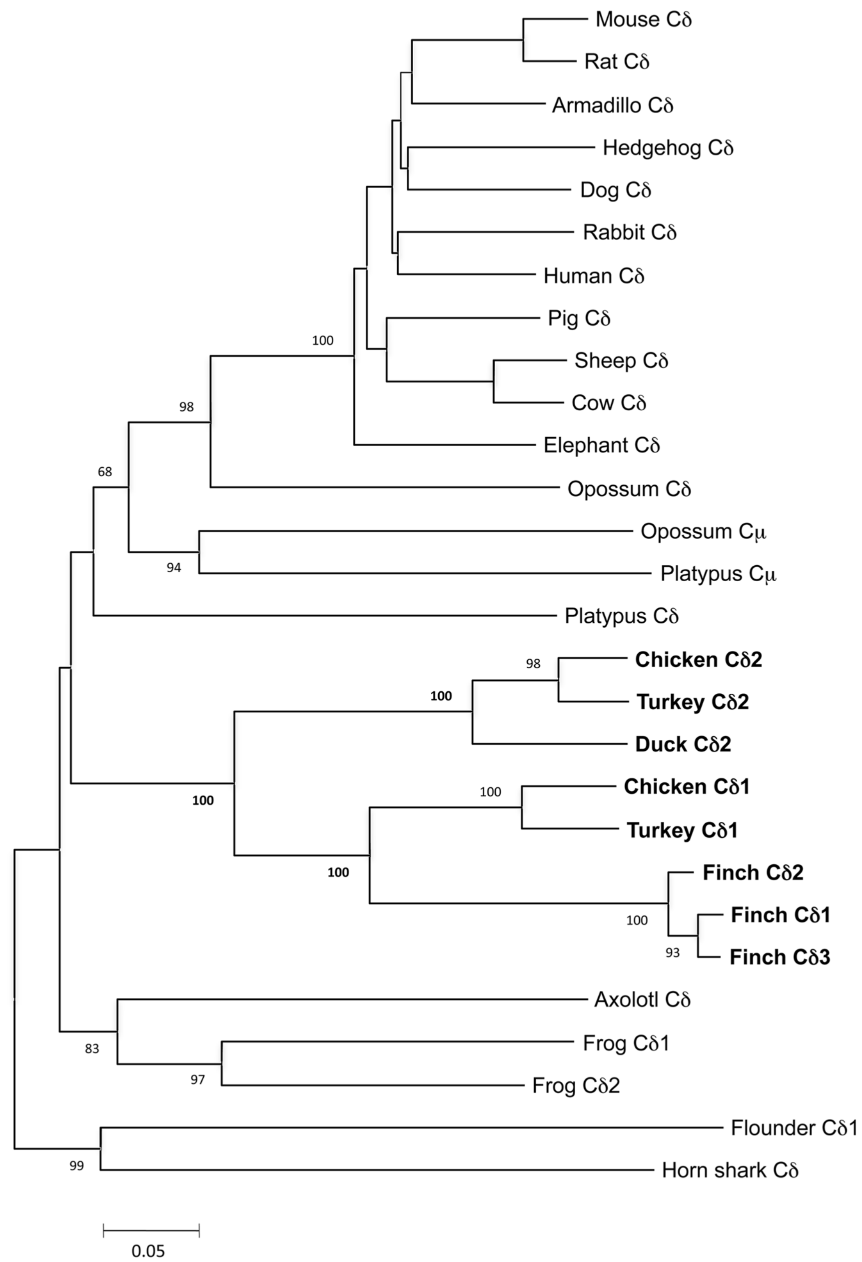

Figure 6.
Phylogenetic tree of TCR Cδ and Cμ regions. The tree was constructed using the Ig-C domains and analyzed using the minimum evolution distance method. Bootstrap values are indicated per 1000 replicates. Avian sequences are shown in bold. A distance bar is shown below the tree.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

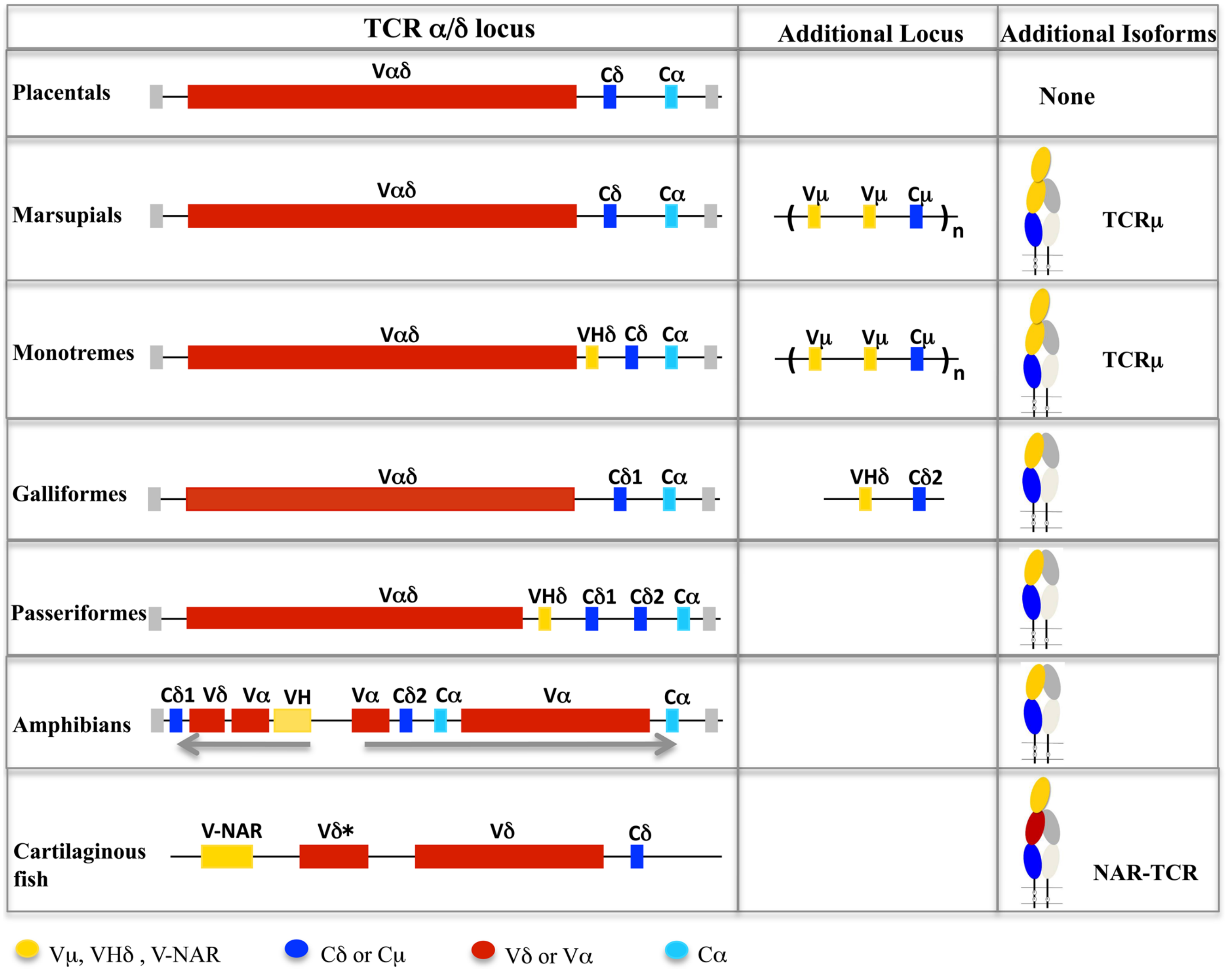

Figure 7.

Simplified representation of the *TCRα/δ* locus, atypical TCR and atypical isoforms in distinct vertebrate lineages. Genes are color coded as indicated at the bottom. Extended red and yellow rectangles in the locus maps represent the presence of several gene segments. Vδ* represent Vδ genes in cartilaginous fish that have lost the sequence encoding the leader peptide and that are only used as supportive Vδ on NAR-TCR.