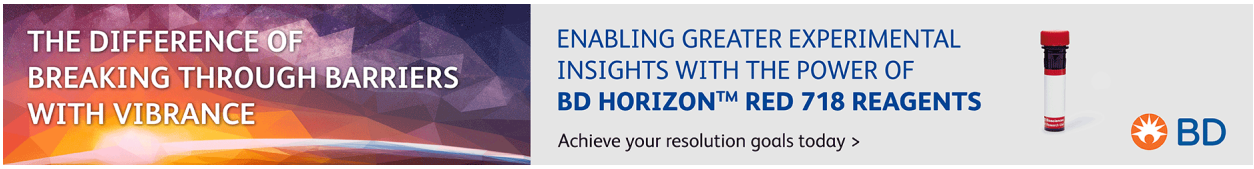

The Journal of Immunology

Analysis of the Chinese Alligator TCRα/δ Loci Reveals the Evolutionary Pattern of Atypical TCRδ/TCRμ in Tetrapods

Xifeng Wang, Jinwei Huang, Peng Wang, Renping Wang, Chaolin Wang, Di Yu, Cuncun Ke, Tian Huang, Yu Song, Jianhui Bai, Kongpan Li, Liming Ren, Robert D. Miller, Haitang Han, Xin Zhou and Yaofeng Zhao

J Immunol published online 26 June 2020
http://www.jimmunol.org/content/early/2020/06/25/jimmunol.2000257

This information is current as of June 27, 2020.

Supplementary Material **http://www.jimmunol.org/content/suppl/2020/06/25/jimmunol.2000257.DCSupplemental**

Why *The JI*? Submit online.

- **Rapid Reviews! 30 days*** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

**average*

Subscription Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

Permissions Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852
Copyright © 2020 by The American Association of Immunologists, Inc. All rights reserved.
Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

Published June 26, 2020, doi:10.4049/jimmunol.2000257

Analysis of the Chinese Alligator TCRα/δ Loci Reveals the Evolutionary Pattern of Atypical TCRδ/TCRμ in Tetrapods

Xifeng Wang,*[,1] Jinwei Huang,[†,1] Peng Wang,[†,1] Renping Wang,[‡] Chaolin Wang,[‡] Di Yu,[†] Cuncun Ke,[†] Tian Huang,[§] Yu Song,[†] Jianhui Bai,[†] Kongpan Li,[†] Liming Ren,[†] Robert D. Miller,[¶] Haitang Han,[†] Xin Zhou,* and Yaofeng Zhao[†]

Atypical TCRδ found in sharks, amphibians, birds, and monotremes and TCRμ found in monotremes and marsupials are TCR chains that use Ig or BCR-like variable domains (VHδ/Vμ) rather than conventional TCR V domains. These unconventional TCR are consistent with a scenario in which TCR and BCR, although having diverged from each other more than 400 million years ago, continue to exchange variable gene segments in generating diversity for Ag recognition. However, the process underlying this exchange and leading to the evolution of these atypical TCR receptor genes remains elusive. In this study, we identified two TCRα/δ gene loci in the Chinese alligator (*Alligator sinensis*). In total, there were 144 V, 154 Jα, nine Jδ, eight Dδ, two Cα, and five Cδ gene segments in the TCRα/δ loci of the Chinese alligator, representing the most complicated TCRα/δ gene system in both genomic structure and gene content in any tetrapod examined so far. A pool of 32 VHδ genes divided into 18 subfamilies was found to be scattered over the two loci. Phylogenetic analyses revealed that these VHδ genes could be related to bird VHδ genes, VHδ/Vμ genes in platypus or opossum, or alligator VH genes. Based on these findings, a model explaining the evolutionary pattern of atypical TCRδ/TCRμ genes in tetrapods is proposed. This study sheds new light on the evolution of TCR and BCR genes, two of the most essential components of adaptive immunity. ***The Journal of Immunology*****, 2020, 205: 000–000.**

T cell receptors play a significant role in the specific recognition of pathogen-associated epitopes (1). Four TCR chains, called α, β, γ, and δ, are expressed as heterodimers either as α with β on αβ T cells or γ with δ on γδ T cells. These have been found in all jawed vertebrates, with only the squamate reptiles appearing to have lost the γ/δ TCR (2, 3). These polypeptide chains are encoded by the TCRα, TCRβ, TCRγ, and TCRδ genes, respectively (4, 5). Generally, the TCRβ and TCRγ are encoded by their own loci present on separate chromosomes. The TCRδ locus, in contrast, is nested within the TCRα locus at a separate chromosomal location (5, 6).

Like Igs, TCRs generate diverse binding specificities to Ags by undergoing somatic recombination of V(D)J segments (7, 8). The TCR and Ig genes share a common origin in jawed vertebrates; the conventional variable domain genes (TCRV) are estimated to have diverged from IgV genes ∼400 million years ago, and the two are readily distinguishable (9, 10). Interestingly, V genes, which are indistinguishable from IgH V genes (VH), have recently been identified within TCR gene loci of several vertebrates and are used in generating the TCR repertoire (10–16).

Such atypical TCRV genes were first identified in an additional TCR locus, named TCRμ, found, so far, only in marsupials and monotremes (10, 12, 13, 17). TCRμ genes are unlinked to those gene loci that encode conventional TCR chains and, in marsupials, are organized in tandem repeats of a core unit: Vμ–Dμ_n–Jμ–Vμj–Cμ (12, 17). Most notably, the Cμ genes related to Cδ, whereas the Vμ genes are more related to VH genes than to conventional TCRV genes (12, 17). Furthermore, TCRμ chains are predicted to have an extracellular atypical structure, containing three extracellular IgSF domains rather than conventional two due an extra V domain (12). The N-terminal V of marsupial TCRμ (Vμ) is encoded by somatically recombined VDJ segments, whereas the C-proximal V domain (Vμj) is encoded by an exon in which the V and J genes are already prejoined in the germline DNA, and it is relatively invariant (12, 13, 17). In the platypus, the C-proximal V domain still requires somatic recombination of a V and J segment (17).

So far, TCRμ homologs have not been found in nonmammalian species (12, 13, 17). However, in primordial jawed vertebrates, such as sharks and other cartilaginous fish, a TCR chain structurally

*Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Plant Protection, China Agricultural University, Beijing 100193, People's Republic of China; [†]State Key Laboratory of Agrobiotechnology, College of Biological Sciences, National Engineering Laboratory for Animal Breeding, China Agricultural University, Beijing 100193, People's Republic of China; [‡]Administration Bureau of *Alligator sinensis* National Nature Reserve Protection, Anhui 242000, People's Republic of China; [§]Henan Engineering Laboratory for Mammary Bioreactor, School of Life Sciences, Henan University, Kaifeng 475004, Henan, People's Republic of China; and [¶]Department of Biology, Center for Evolutionary and Theoretical Immunology, University of New Mexico, Albuquerque, NM 87131

[1]X.W., J.H., and P.W. contributed equally to this work.

ORCIDs: 0000-0001-8298-0177 (D.Y.); 0000-0002-8479-8819 (R.D.M.); 0000-0002-1407-7952 (X.Z.); 0000-0003-4499-8610 (Y.Z.).

Received for publication March 11, 2020. Accepted for publication May 22, 2020.

This work was supported by the National Natural Science Foundation of China (31472085 and 31772589), the Program of Introducing Talents of Discipline to Universities, Plan 111 (B12008), and funding from the Beijing Advanced Innovation Center for Food Nutrition and Human Health, China Agricultural University. R.D.M. was supported in part by National Institute of General Medical Sciences of the National Institutes of Health Award P30 GM110907 and National Science Foundation Award IOS-13531232.

The sequences presented in this article have been submitted to the National Center for Biotechnology Information GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under accession numbers MT081963 and MT081964.

Address correspondence and reprint requests to Dr. Haitang Han and Prof. Yaofeng Zhao or Prof. Xin Zhou, State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, 2 Yuanmingyuan Western Road, Beijing 100193, People's Republic of China (H.H. and Y.Z.) or Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Plant Protection, China Agricultural University, 2 Yuanmingyuan Western Road, Beijing 100193, People's Republic of China (X.Z.). E-mail addresses: hanhaitang@cau.edu.cn (H.H.) and yaofengzhao@cau.edu.cn (Y.Z.) or xinzhou@cau.edu.cn (X.Z.)

The online version of this article contains supplemental material.

Abbreviations used in this article: BAC, bacterial artificial chromosome; RSS, recombination signal sequence.

Copyright © 2020 by The American Association of Immunologists, Inc. 0022-1767/20/$37.50

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

similar to TCRμ has been identified, called NAR-TCR. NAR-TCR also employs two V domains, each encoded by separate, somatically rearranging VDJ segments (11, 18). The distal V domain of NAR-TCR is more closely related to the V domains of IgNAR than to the V domains of conventional TCR. IgNAR is an H chain homodimer that is not associated with light chains and is found only in cartilaginous fish (19, 20). It can be assumed that the structures of TCRμ and NAR-TCR may represent an ancient receptor system. However, phylogenetic analysis and structure are consistent with NAR-TCR and TCRμ being more likely the results of convergent evolution rather than being related by direct descent (10, 17).

The origins of TCRμ remained unclear until the discovery of VHδ genes in the TCRα/δ locus of platypus, which provided a model for the TCRμ origin (10). Such atypical V genes that appear to be indistinguishable from VH genes but are located within the conventional TCRα/δ locus have also been detected in amphibians, the platypus, passeriform birds, coelacanths, and nurse sharks (10, 14–16, 21). The VHδ gene, which is used with Cδ as part of TCRδ repertoire, was first identified as a cluster of gene segments in a reverse transcriptional orientation at the 5′ end of the *Xenopus tropicalis* TCRα/δ locus (14). Later, one additional VHδ gene was identified in the TCRα/δ locus of platypus (10). The phylogenetic relationship of VHδ and Vμ genes in platypus provided a critical link in reconstructing the evolutionary history of TCRμ. Possibly, V gene segments have recurrently translocated between the IgH and TCRδ loci. The TCRδ genes are hypothesized to have translocated out of the TCRα/δ locus early in evolution of mammals, creating the TCRμ locus (10). Supporting this, a second TCRδ locus that uses Ig-like V domains was found in Galliformes, which is unlinked to the conventional TCRα/δ locus. Additionally, flanking genes of the second TCRδ locus cannot be found in the conventional TCRα/δ locus (15). A IgH-like V segment, nestled within the TCRδ locus and expressed with TCRδ (but not IgH) was also found in the nurse shark and termed as TAILVs (16).

VHδ genes exist in TCRα/δ locus perhaps because of direct transfer of VH genes from an IgH locus to the TCRα/δ locus. But it should be pointed out the TAILV/VHδ/μ genes found in the nurse sharks, coelacanths, frogs, birds, and monotremes are not closely related. Phylogenetic evidence did not support TAILV as the progenitor of VHδ (16). Indeed, VHδ and Vμ genes in tetrapods appear to have derived from different, ancient VH clans. Avian VHδ are derived from VH clan I, amphibian VHδ from VH clan II, and platypus VHδ from VH clan III, for example (14, 15, 17). The translocation of VH into the TCRα/δ loci seemed to have occurred independently in different lineages. Strangely, birds are thought to only have VH clan III genes in their IgH locus, yet all avian VHδ groups belong to VH clan I (10, 22). This raises interesting questions regarding the origins of the VHδ and Vμ genes in tetrapods.

In this study, we generated a detailed genomic organization for the TCRα/δ locus in the Chinese alligator, an ancient reptile that represents a key evolutionary position linking amphibians, birds, and mammals (23). Notably, a large number of VHδ genes were found to be dispersed throughout the TCRα/δ locus in the Chinese alligator, and phylogenetical analysis suggested some of them grouped with frog VHδ, some with bird VHδ, and some with VHδ and Vμ genes in the platypus and the opossum. Based on this study, we developed and present a model to explain the evolutionary pattern of atypical VHδ genes in tetrapod.

Materials and Methods

Sample collection, DNA and RNA extraction, and reverse transcription

Blood samples of Chinese alligators were collected from the Beijing Zoo. Genomic DNA was extracted from blood cells according to a routine protocol. Various tissues of Chinese alligator were provided by the Administration Bureau of Chinese Alligator National Nature Reserve Protection of Anhui Province. Total RNA was extracted from the spleen and thymus using a TRIzol kit (TIAGEN Biotech, Beijing, China) according to the manufacturer's instructions. Reverse transcription was conducted using Moloney Murine Leukemia Virus Reverse Transcriptase (Promega, Madison, WI) with an oligo(dT) adapter primer. Our studies were approved by the Animal Care and Use Committee of China Agricultural University.

Bacterial artificial chromosome screening

Based on the sequences of TCR Cα and TCR Cδ derived from chicken (Cα, ABU97797.1; Cδ1, AAD51740.1 [both from GenBank]), we performed tblastn against National Center for Biotechnology Information GenBank. Two scaffolds (scaffold1585_1 and scaffold955_1) were identified. Based on the sequences of Cα and Cδ from scaffold1585_1 and scaffold955_1, positive bacterial artificial chromosome (BAC) clones covering the Chinese alligator TCRα/δ loci were isolated from the BAC library via PCR-based screening. The first positive BAC clone was sequenced from both ends, and the end sequences were used to design primers for the next round of screening to determine the BAC clone overlap. Finally, 17 positive BAC clones were sequenced by shotgun sequencing (the sequence had ~7 × coverage of each BAC) and assembled with a next-generation sequencing platform by Beijing Genomics Institute (Beijing, China).

Genome walking

We mapped the gap between Y251L7 and Y104L3 of Chinese alligator TCRα/δ locus A by a genome walking kit (Clontech Laboratories) according to the manufacturer's instructions. All amplifications were conducted using LA Taq DNA Polymerase (Takara, Dalian, China) with proofreading activity. The PCR products were sequenced.

Gene annotation and characterization

IgBLAST (http://www.ncbi.nlm.nih.gov/igblast/) was used to predict the V gene segments by assessing similarity to homologs from human and mouse. Germline Vα and Vδ genes were grouped into families using the ImMunoGeneTics numbering system (24). If there was more than one member in the subgroup, we named the V gene from 3′ to 5′, with the subgroup number followed by the gene segment number. The recombination signal sequences (RSSs) for the V, D, and J gene segments were analyzed using the online programmer FUZZNUC (http://embossgui.sourceforge.net/demo/fuzznuc.html). Non-TCR genes located in or flanking the TCRα/δ locus were identified using GENSCAN (http://genes.mit.edu/GENSCAN.html), and then BLAST was used on all predicted exon domains against the GenBank database using blastp to identify the genes. DNA and protein sequence editing, alignments, and comparisons were performed with MegAlign software (DNASTAR) (25). Dot blot analyses of the V regions of the TCRα/δ locus were conducted with the dotter programmer (26).

Southern blotting

Restriction endonucleases were used to digest the genomic DNA. Digested genomic DNA was electrophoresed in a 0.9% agarose gel for 6 h and then transferred to a positively charged nylon membrane (Roche, Mannheim, Germany) for hybridization with individual Cα1, Cα2, Cδ1, Cδ2, Cδ4, and Cδ5 probes, which were labeled by a PCR digoxigenin probe synthesis kit (Roche). Hybridization and detection were performed according to the manufacturer's instructions.

Construction of phylogenetic trees

The phylogenetic trees were constructed in MEGA version 6.0 (27) using the neighbor-joining method with 1000 bootstrap replicates. Multiple amino acid/nucleotide alignments for tree construction were performed using ClustalW. Each Vα/δ and VHδ subgroup was represented by one family per species chosen at random. The accession numbers of the sequences used to construct the phylogenetic trees are listed in Supplemental Table I. All other sequences were derived in this study.

Cloning of expressed Chinese alligator TCRα and TCRδ chain genes at the cDNA level

5′ RACE System for RACE (Invitrogen, Carlsbad, CA) was applied to total RNA to obtain the expressed repertoire of TCRα and TCRδ V segments. We designed the gene-specific primers for the TCRα and TCRδ chains. All PCR amplifications were performed using a proofreading

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

enzyme, Pyrobest DNA polymerase (Takara). The PCR products were cloned into the pMD19-T vector (Takara) and sequenced.

Results

Identification of two TCRα/δ loci in the Chinese alligator

To analyze the TCRα/δ locus in the Chinese alligator, we employed a BAC genomic library that was generated by using peripheral blood leukocytes isolated from a Chinese alligator (28). Using a PCR-based approach and shotgun sequencing, 17 overlapping genomic TCRα/δ gene–positive BAC clones (Y12D21, Y135N2, Y144M13, Y251L17, Y104L3, Y257J18, Y179I18, Y107L2, Y13B10, Y183J22, Y132K5, Y75A3, Y360O6, Y173I18, Y236E17, Y96L4, and Y110J22) were identified. Upon filling a gap between Y251L7 and Y104L3 by genome walking, we obtained two contiguous distinct genomic sequences containing the alligator TCRα/δ genes. These spanned ~547 kb and 1013 kb of DNA sequences, respectively, and we labeled them as TCRα/δ locus A and locus B.

Subsequently, we annotated and characterized the genes of the two TCRα/δ loci. As in all tetrapods analyzed so far, the genes encoding the alligator TCRα and TCRδ chains are tightly linked

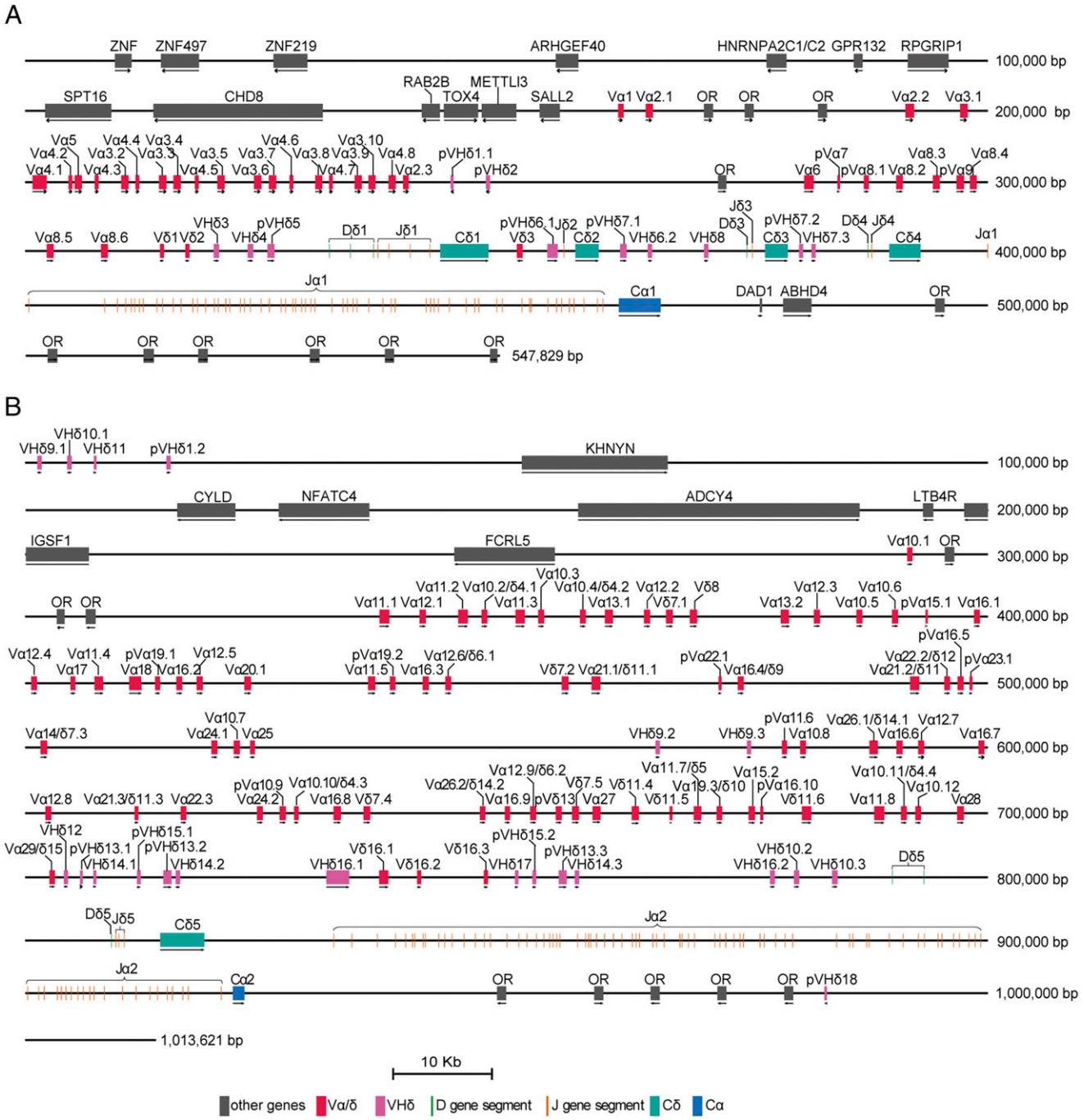

FIGURE 1. Physical map of the two Chinese alligator TCRα/δ loci. Vα (VHδ or Vδ), potentially functional variable gene segments; pVα (VHδ or Vδ), pseudovariable gene segments; D, diversity gene segments; J, joining gene segments; and C, C region gene. V, D, and J gene segments and C regions were numbered based on their position on the locus (5′–3′) and were color coded as follows: Vα/δ, red; D, green; J, orange; Cα, blue; Cδ, dark green; and VHδ, pink. Non-TCR genes located in or flanking the TCRα/δ locus were shaded by light gray. Transcriptional orientation is indicated by the arrow under each segment except for the D and J genes.

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

in both loci, with TCRδ genes nested among the TCRα genes (Fig. 1). Flanking genes *METTL3* and *SALL2* were detected at the 5′ end of locus A, which was similar to that in human, mouse, opossum, *X. tropicalis*, and coelacanth (Fig. 1). We also identified *DAD1* and *ABHD4* at the 3′ end of locus A, as in mammals, chicken, lizard, turtle, and *X. tropicalis* (Fig. 1). Additionally, there was also a single Vδ gene in reverse transcriptional orientation downstream of the Cδ1 gene like in mammalian TCRα/δ loci (Fig. 1). However, genes (*KHNYN*, *CYLD*, *NFATC4*, *ADCY4*, *LTB4R*, *IGSF1*, and *FCRL5*) located at the 5′ end of locus B were not found in TCRα/δ loci of any other vertebrates, There were, however, several olfactory receptors interspersed among the Vα genes (Figs. 1, 2).

Genomic structures of the two TCRα/δ loci

Based on the existence of RSSs, we have annotated a pool of 46 V genes positioned at the 5′ of Dδ–Jδ–Cδ clusters, which were followed by a large number of Jα ($n = 63$) segments and a Cα region in locus A (Fig. 1). The striking feature of the locus A is the presence of four tandemly repeated Dδ–Jδ–Cδ clusters. There are three Dδ segments, four Jδ segments, and one Cδ in the first Dδ–Jδ–Cδ cluster. Only one Jδ segment and one Cδ was found in the second cluster. The Cδ2 appeared to be a pseudogene because of a frameshift mutation at its first exon. Both the third and fourth Dδ–Jδ–Cδ clusters had a single Dδ segment, a single Jδ segment, and one Cδ (Fig. 1). The genomic structure of locus B was simpler than the locus A. There are 98 V genes and one Dδ–Jδ–Cδ cluster containing three Dδ, three Jδ, and one Cδ. A Jα–Cα cluster containing 91 Jα segments and a single Cα is positioned downstream of the Dδ–Jδ–Cδ cluster (Fig. 1). All Dδ and Jδ/Jα gene segments are flanked by canonical RSS (Supplemental Table II).

To determine whether the TCRα/δ loci in another alligator species is also organized as the similar structure to that in Chinese alligator, we identified the V, D, J, and C gene segments on two scaffolds (GenBank: NW_017707656.1 and NW_017710963.1) in the current genome assembly of American alligator (*Alligator mississipiensis*). All genes of the TCRα/δ locus in American alligator are syntenic to those in Chinese alligator.

Characterization of the Cα and Cδ genes in Chinese alligator

Two Cα were detected, one from locus A and one from locus B; they share an amino acid identity of 64.8% (Figs. 1, 2). Southern blot analyses with the first exons of Cα1 and Cα2 as probes were conducted to verify the copy numbers of the Cα genes, and the two probes each detected a single hybridizing fragment (Supplemental Fig. 1A). An amino acid sequence alignment of the two Cα with those of other vertebrates revealed conserved cysteine distribution (Supplemental Fig. 2A). The Cα of Chinese alligator shared <35% sequence identity with Cα from other species (Supplemental Fig. 2A). Moreover, a phylogenetic analysis indicated that the divergence of Cα1/Cα2 genes in alligators (including the Chinese alligator and American alligator) occurred in a common ancestor (Fig. 3).

We also identified five Cδ genes in the Chinese alligator (Fig. 1). Pairwise comparisons revealed no more than 40% aa sequence identity between Chinese alligator Cδ and that of other vertebrates (Supplemental Fig. 2B). The identical cysteine distribution in the Cδ is conserved between Chinese alligator and Cδ from other vertebrate species. Southern blot analysis corroborated the presence of five Cδ genes in the genome (Supplemental Fig. 1B). We observed only a single band each when using the first exons of Cδ1, Cδ4, and Cδ5 as probes (Supplemental Fig. 1B, 1C). When the first exon of Cδ2 was used, two bands were detected in *Xba* I–, *Bam*H I–, and *Pst* I–digested genomic DNA (Supplemental Fig. 1B), most likely because the nucleotide sequence identity between exon 1 in Cδ2 and Cδ3 is to 90.2%.

Based on a phylogenetic analysis, Cδ genes from the reptilian lineage, including birds and nonavian reptiles, clustered together (Fig. 4). They formed three subclades: Cδ I subclade comprised all avian Cδ genes, alligator Cδ1 and Cδ5 and turtle Cδ1; Cδ II

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

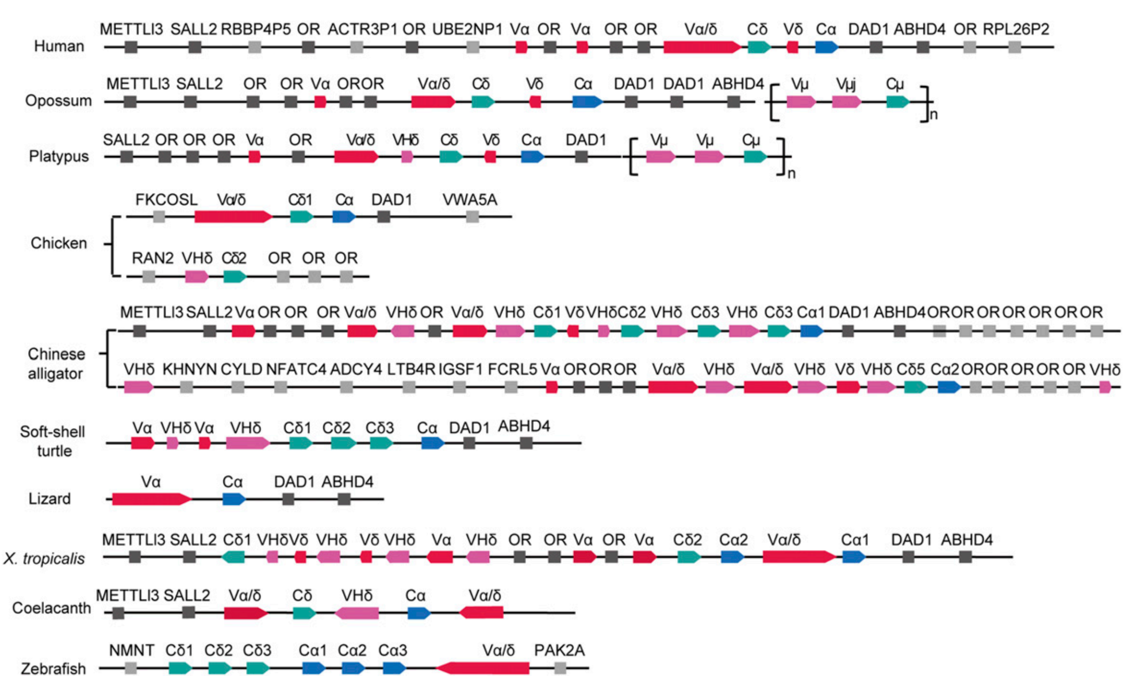

FIGURE 2. Simplified representations of the TCRα/δ and TCRμ loci indistinct vertebrate lineages. Vα/δ genes are marked in red, VHδ and Vμ genes are marked in pink, Cδ and Cμ genes are represented in dark green, and Cα genes are shown in blue. Syntenic genes shown in dark gray are those conserved with other tetrapod TCRα/δ locus: methyl-transferase like 3 (*METTL3*), spalt-like transcription factor 2 (*SALL2*), olfactory receptors (*OR*), defender against cell death gene 1 (*DAD1*), and abhydrolase domain containing 4 (*ABHD4*). Other genes are marked by light gray.

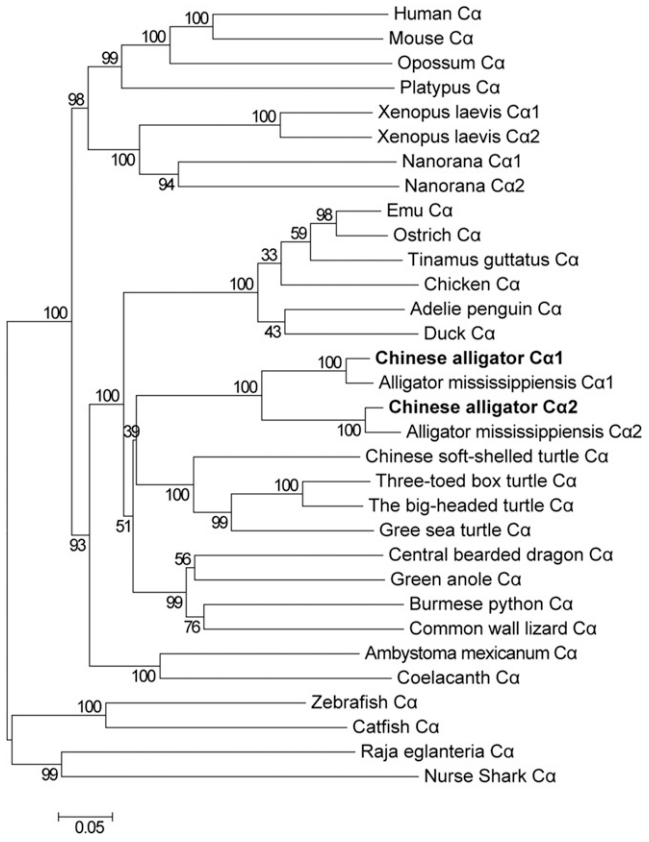

FIGURE 3. Phylogeny of the TCRα constant regions in jawed vertebrates. The phylogenetic tree was constructed using TCR Cα domains in MEGA 6.0 (27) with the neighbor-joining method (42). Chinese alligator sequences are shown in bold. Bootstrap percentage values are shown at the major interior branch nodes. The horizontal length is proportional to the distance score generated by the computer program. Similar topology was generated using the maximum likelihood method.

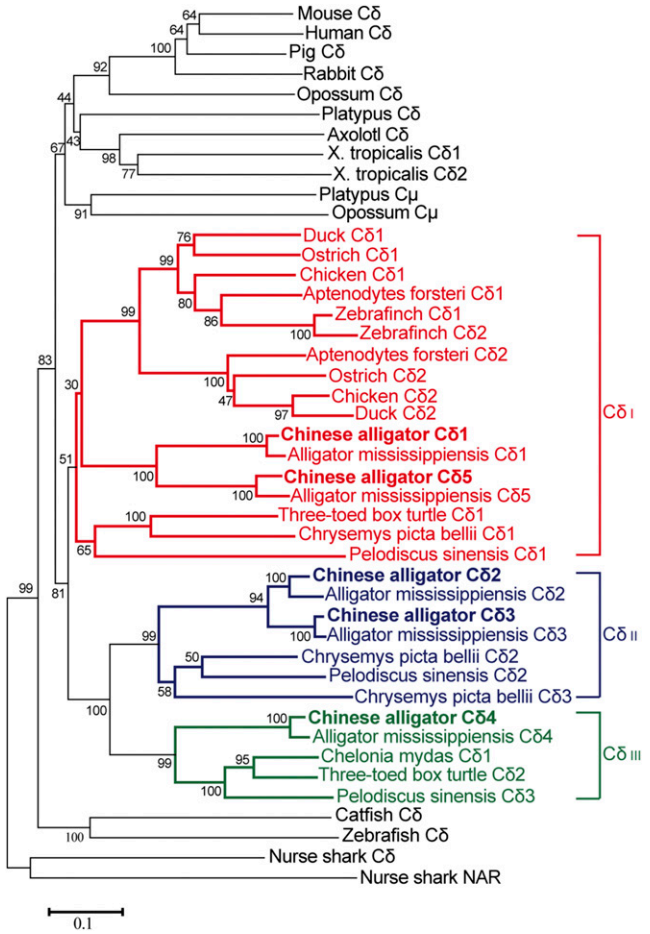

FIGURE 4. Phylogenetic analyses of Chinese alligator TCRδ constant regions. The tree was constructed using the Ig-C domains and analyzed using the neighbor-joining method. Bootstrap values are indicated per 1000 replicates. Chinese alligator sequences are shown in boldface. A distance bar is shown below the tree. In the figure, all bird Cδ genes, alligator Cδ1 and Cδ5 and turtle Cδ1 as Cδ I in red; turtle Cδ2, alligator Cδ2 and Cδ3 as Cδ II in blue; and turtle Cδ3 and alligator Cδ4 as Cδ III in green. Similar topology was generated using the maximum likelihood method.

subclade consisted of turtle Cδ2 and alligator Cδ2 and Cδ3; and Cδ III subclade was composed of turtle Cδ3 and alligator Cδ4 (Fig. 4).

A large number of unconventional VHδ gene segments dispersed in the two TCRα/δ loci

As noted above, a total of 144 V gene segments combined were detected in two TCRα/δ loci (Fig. 1). One hundred and twelve conventional Vα/δ genes could be classified into thirty-four families using the rule that sequences sharing 75% nucleotide identity are in the same family. Based on the ImMunoGeneTics rules described previously (29), 97 Vα/δ (~86%) genes were predicted to be functional, and 15 appeared to be pseudogenes because of the presence of unusual RSS or in-frame stop codons, frameshifts, and/or mutations in the sequences (Supplemental Table II). The derived amino acid sequences of these Vα/δ families contained the canonical Vα/δ structure, which included the conservative sites Cys^{23}, Trp^{41}, and Cys^{104} (Supplemental Fig. 2C).

Thirty-two unconventional V genes were found to share greater identity with VH genes than conventional Vδ genes and were designated as VHδ genes, following the nomenclature established for *X. tropicalis* (14). Indeed, compared with Ig and TCR V genes from other vertebrate species, these VHδ genes were found to be most similar to VH. Phylogenetic analysis supported that the VHδ genes of Chinese alligator grouped within the same clade as the VH genes rather than the Vα/δ genes, suggesting that the VHδ genes were more closely related to the VH genes (Fig. 5). These VHδ genes were segregated into 18 families based on 75% or greater nucleotide identity (Fig. 1, Table I, Supplemental Fig. 2D).

In locus A, two pseudo-VHδ genes were identified among the Vα genes. The other VHδ genes were nested within the TCRα locus as part of four VHδ–Dδ–Jδ–Cδ clusters. In locus B, four VHδ genes clustered at the 5′ end of the locus. Nine were dispersed among the Vα/δ genes, and seven were located immediately upstream of the Dδ–Jδ–Cδ cluster. Finally, a single VHδ (pVHδ18) was identified at the 3′ end the locus (Fig. 1).

Eighteen VHδ gene segments located on two genomic scaffolds from the American alligator were also identified and could be classified into 12 families. The position of the VHδ genes in Chinese and American alligators were highly conserved (data not shown).

VHδ genes identified in the Chinese alligator fell into three different ancient VH clans

Ig VH genes from mammals and other tetrapods have been shown to cluster into three ancient clans (22, 30). Previous studies indicated that the *X. tropicalis* VHδ all clustered with VH clan II genes (14), whereas all VHδ genes identified in coelacanth clustered into VH

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

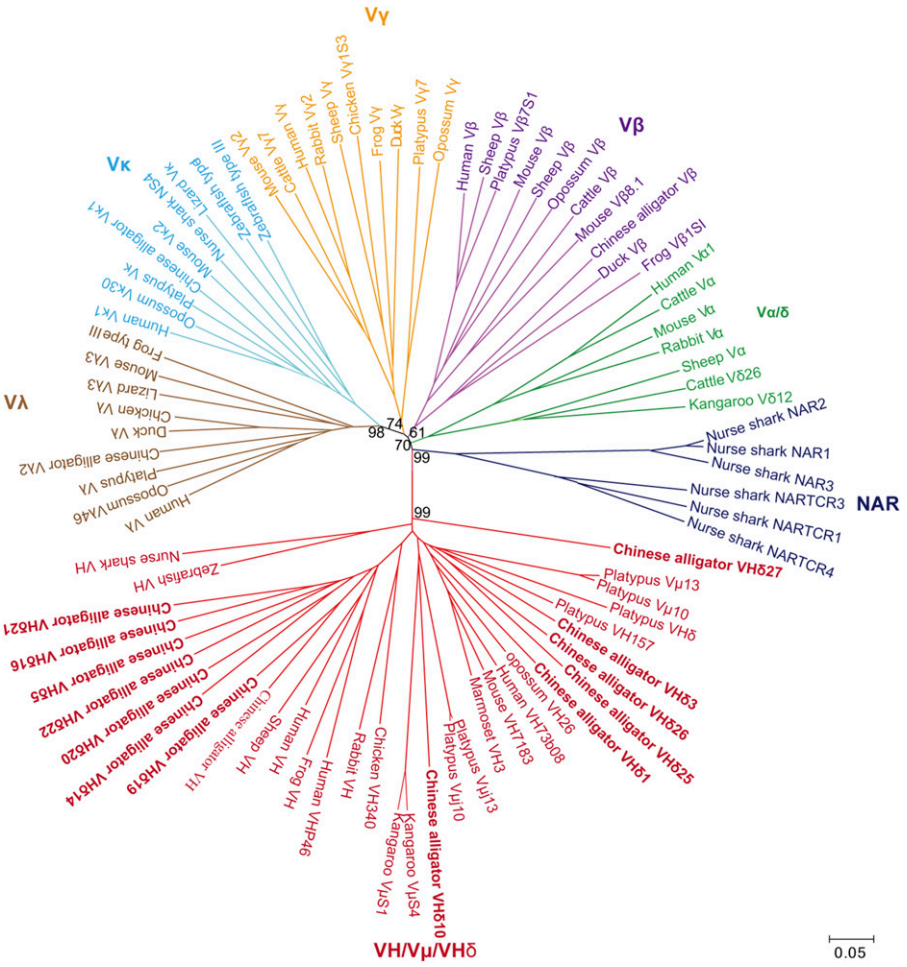

FIGURE 5. Phylogenetic analysis of VHδ regions from four conventional TCRs, TCRμ, and Ig. The trees shown were from nucleotide alignments constructed by MEGA 6.0 (27) and analyzed by using the neighbor-joining method. The numbers are the percentage of bootstrap support for each node based on 1000 replicate trees. Vκ and Vλ are respectively colored in brown and light blue. Vα/δ, Vγ, Vβ, and NAR-TCRV are in blue, orange, green, and blue, respectively. VHδ cluster with Vμ and VH gene together were colored in red. VHδ gene from Chinese alligator were marked in bolded font, and they fall into the VH clade rather than Vα/δ clade. Similar topology was generated using the maximum likelihood method.

clan III (21). Avian VHδ genes, in contrast, are closely related to VH clan I (15). Finally, the VH-related Vμ and VHδ in platypus both clustered within VH clan III (10). These observations are consistent with VHδ having evolved independently from VH multiple times.

To analyze phylogeny of the VHδ genes in Chinese alligator, a phylogenetic tree using VH genes from different tetrapods, Vμ of opossum and platypus, and VHδ of platypus, birds, alligators, soft-shell turtles, and *X. tropicalis* was constructed (Fig. 6, Table II). The VHδ genes from alligators fell into three different ancient clans of VH genes: eight VHδ subgroups (VHδ6–9, VHδ12–13, VHδ15, and VHδ17) fell into VH clan III, among which VHδ12 clustered with platypus VHδ in the same clade, suggesting that they share a common origin (Fig. 6). In contrast, nine subgroups, including VHδ1–5, VHδ10–11, VHδ14, and VHδ16, were located within VH clan I together with VHδ genes of birds (Fig. 6), whereas the VHδ18 fell into the VH clan II (Fig. 6).

In addition, Chinese alligator pVHδ1.1, VHδ2, VHδ9.3, and VHδ15 were clustered with their own VH8, VH17, VH32, and VH10, respectively (Fig. 6). Indeed, these VHδ genes in TCRα/δ share a high degree of identity (63.4–91.1%) with their corresponding VH genes in the IgH locus.

Expression of VHδ with Cδ1, Cδ3, Cδ4, and Cδ5

To identify that VHδ genes are found used in Chinese alligator TCRδ transcripts, 5′ RACE PCR was performed. We sequenced and analyzed 296 cDNA clones (including 74 TCRδ1, 20 TCRδ3, 25 TCRδ4, and 177 TCRδ5) with unique VJ junctions. Four different Cδ had different preferences for usage of V gene segments. Vδ3 was the most frequently used in TCRδ1 chains, which accounted for roughly half of the expressed TCRδ repertoire (45.95%) and only one TCRδ1 clone (clone D2-2-6) contained VHδ3 gene. In TCRδ5 chains, Vα14/δ7 and Vα21/δ11 were preferentially used, accounting for 37.85 and 33.90%, respectively, and there were 18 (10.17%) TCRδ5 cDNA clones containing VHδ genes. But in all TCRδ3 and TCRδ4 transcripts obtained, the Cδ were only found coexpressed with VHδ genes.

The four functional Cδ genes coexpressed with VHδ genes detected in Chinese alligator encoded single V and C domains, similar to the structure of atypical TCRδ chain in *X. tropicalis* and birds (14, 31). To analyze the sequence diversity of the TCRδ transcripts with VHδ, we compared VDJ junctions in sequenced transcripts. A biased usage of VHδ genes segments was observed in these clones (Fig. 7). Only a single TCRδ1 chain containing VHδ3 gene segment was found. The 18 TCRδ5 clones using VHδ genes showed a preferential VHδ usage with most clones using VHδ12 and VHδ16 (six clones and eight clones, respectively), with VHδ10 (three clones) being the third most frequently used (Fig. 7). In the TCRδ3 chains, which are only coexpressed with VHδ gene segments VHδ6.2 and VHδ8 were found to be nearly equally used, whereas most TCRδ4 clones, also only coexpressed with VHδ, preferentially used VHδ7.3 (92%). Notably, all the VHδ gene segments tended to be used by the nearest Cδ.

From the sequence analyses, it appears that the Cδ transcripts use the Dδ gene segments and Jδ gene segments located within the same cluster as the Cδ (Figs. 7, 8). Most of the TCRδ1 and TCRδ5 appear to use more than one D gene (VDDJ) because of the presence of the asymmetrical RSS leading to incorporation of

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

Table I. Summary of the germline VHδ subgroups in Chinese alligator TCRα/δ loci

VHδ Subgroup	VHδ Gene	Functional[a]	Total
VHδ1	pVHδ[b]1.1 and pVHδ1.2	0	2
VHδ2	VHδ2	1	1
VHδ3	VHδ3	1	1
VHδ4	VHδ4	1	1
VHδ5	pVHδ5	0	1
VHδ6	pVHδ6.1 and VHδ6.2	1	2
VHδ7	pVHδ7.1, pVHδ7.2, and VHδ7.3	1	3
VHδ8	VHδ8	1	1
VHδ9	VHδ9.1, VHδ9.2, and VHδ9.3	3	3
VHδ10	VHδ10.1, pVHδ10.2, and VHδ10.3	2	3
VHδ11	VHδ11	1	1
VHδ12	VHδ12	1	1
VHδ13	pVHδ13.1, pVHδ13.2, and pVHδ13.2	0	3
VHδ14	pVHδ14.1, VHδ14.2, and VHδ14.3	2	3
VHδ15	VHδ15.1 and VHδ15.2	2	2
VHδ16	pVHδ16.1 and VHδ16.2	1	2
VHδ17	pVHδ17	0	1
VHδ18	pVHδ18	0	1
Total		18	32

[a]Functionally mean potentially functional variable gene segments with open reading frames.

[b]pVHδ represented pseudovariable gene segments that have defect in splicing sites, RSS, and/or regulatory elements and/or changing the conserved amino acids has been suggested to lead to incorrect folding.

multiple D genes in recombination (32, 33). Therefore, the average length of CDR3 in unconventional TCRδ1 and TCRδ5 was longer than that in TCRδ3 and TCRδ4 (13 ± 2.5 versus 7 ± 2.1 aa).

Rare usage of VHδ genes in TCRα chains

The use of Vα gene segments was also analyzed. A total of 313 cDNA fragments, including 123 TCRα1 and 188 TCRα2, all with unique V–J junctions, were obtained. We analyzed the usage of V and J gene segments and observed a biased usage of Vα and Jα in these clones. Cα1 preferred to use the Vα genes of locus A (75/123, 60.97%), and Cα2 had a biased usage pattern of Vα in locus B (147/188, 78.19%). Additionally, Vδ and VHδ genes, albeit rare, were also found to be used in TCRα chains.

Discussion

In this study, we systematically characterized the genomic organization, phylogeny, and expression of the TCRα/δ locus in Chinese alligator. The results provide a better and deeper understanding of the origin and evolution of VHδ genes, as well as TCRα/δ locus as a whole in tetrapods.

The Chinese alligator TCRα/δ locus exhibits a complicated genomic organization

The genomic region containing TCRα/δ has been relatively conserved over the 400 million years of evolution of vertebrates because the genes flanking the TCRα/δ gene locus in coelacanth, which is considered to be phylogenetically mostly close to the ancestor of tetrapod, are the same as in *X. tropicalis*, alligators, birds, and mammals (Fig. 1) (5, 14, 21, 31). Chinese alligator TCRα/δ locus A was also flanked by the *METTL3*, *SALL2*, *DAD1*, and the *ABHD4* genes, as seen in other tetrapods (6, 14, 34). However, flanking the locus B were *KHNYN*, *CYLD*, *NFATC4*, *ADCY4*, *LTB4R*, *IGSF1*, and *FCRL5* genes, none of which could be found in the conventional TCRα/δ locus in amphibians, birds,

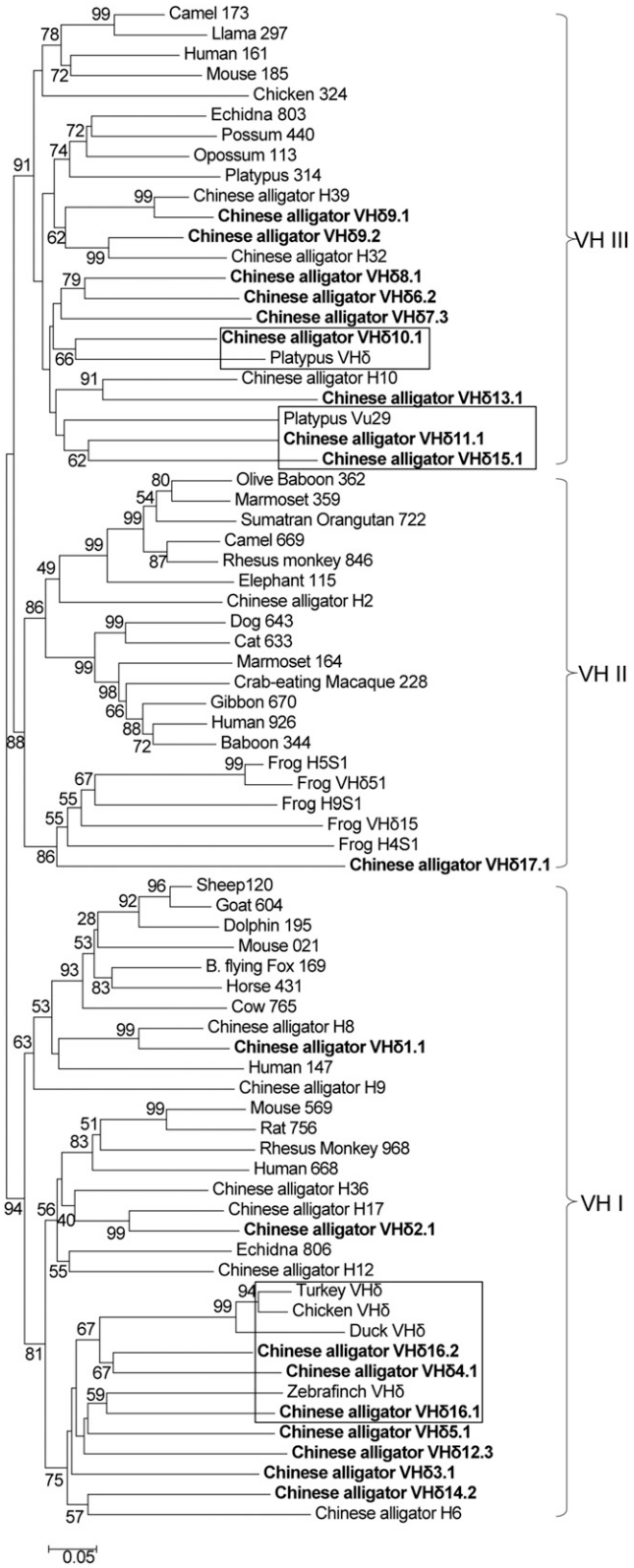

FIGURE 6. Phylogenetic tree of Chinese alligator VHδ genes compared with atypical V gene, including Vμ, VHδ, and VH genes from Chinese alligator, birds, mammals, and amphibians. All Chinese alligator VHδ are in boldface. The tree shown was generated using the neighbor-joining method. Bootstrap values 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 on each node. The orthologous of VHδ/Vμ genes of birds and platypus were showed in boxes. A scale bar representing proportional evolutionary distance is shown below the tree. Similar topology was generated using the maximum likelihood method.

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

Table II. Numbers of V, D, J, and C gene segments in TCRα/δ loci of selected tetrapods

Species	Vα/Vδ[a]	VHδ	Jα	Jδ	Dδ	Cα	Cδ	Total[b]	Reference
Human	57 (44)[c]	0	61	4	3	1	1	127	(35)
Mouse	104 (28)	0	60	2	2	1	1	170	(35, 36)
Cattle	>400 (>44)	0	—[d]	5	3	1	1	>410	(37, 38)
Opossum	74 (47)	1 (1)	53	6	2	1	1	138	(34)
Platypus	99 (19)	2 (1)	32	7	2	1	1	144	(10)
Chicken	96 (3)	1 (1)	48	2	2	1	2	152	(31, 39)
Zebra finch	>14 (7)	1 (1)	14	3	2	1	2	>37	(15)
Duck	~79 (8)	1 (1)	68	2	2	1	2	155	(40)
X. tropicalis	57 (31)	14 (5)	77	5	2	2	2	159	(14, 15)
Chinese alligator	**112 (34)**	**32 (18)**	**154**	**9**	**8**	**2**	**5**	**322**	This study

All genes of TCRα/δ in Chinese alligator appear in bold.
[a]The Vα/Vδ segments include the Vα in either TCRα and/or TCRδ locus.
[b]Numbers of all elements in different species.
[c]Numbers preceding parentheses are numbers of V gene segments, and numbers in parentheses are numbers of V subgroups.
[d]Number not confirmed.

and mammals (Figs. 1, 2). These results are consistent with locus A being the ancient and locus B being derived from locus A.

The genomic structure of the two Chinese alligator TCRα/δ loci revealed that the genes encoding TCRδ were nested within the TCRα genes, like in other tetrapods (14, 31, 34). What was atypical was four tandemly repeated Dδ–Jδ–Cδ clusters were found to be present in locus A (Fig. 1). More surprising was the discovery of a large number of VH-related genes dispersed within both Chinese alligator TCRα/δ loci, which has not been found previously in any other tetrapod (Figs. 1, 5, 6). In total, there were 144 V (including 32 VHδ), 154 Jα, 9 Jδ, eight Dδ, 2 Cα, and 5 Cδ gene segments in the TCRα/δ loci of Chinese alligator (Fig. 1). The results are consistent with Chinese alligator having a few more Jδ segments and Dδ segments than other tetrapod species and had twice the number of Jα gene segments than in other vertebrates because of the presence of the additional locus (10, 14, 15, 31, 34–40). Furthermore, a large number of V gene segments were also found in the TCRα/δ loci of Chinese alligator, second only to that found in cattle (Table II) (32, 33).

In brief, although the organization of the Chinese alligator TCRα/δ locus has many features typical of tetrapods (6, 10, 14, 34, 35), it appears that the overall gene content and organization of the Chinese alligator TCRα/δ locus is the most complex among all tetrapod examined thus far (Fig. 2).

Transfer of VH genes into the TCRα/δ locus occurred frequently in vertebrates after the divergence of the IgH and TCR genes

The presence of TCR chains that use Ig-like V domains, such as atypical TCRδ using VHδ (termed as TAILVs in the nurse shark), NAR-TCR, or TCRμ, have now been widely found in vertebrates, with the exception of the bony fish and eutherian mammals (10–12, 14–16, 21). NAR-TCR uses an N-terminal V domain related to IgNAR, which is unique to cartilaginous fish (19). The VDJ cluster of IgNAR may have recombined with Vδ genes sometime early in the evolution of modern sharks to create the NAR-TCR cluster (11).

The nurse shark TCRα/δ locus houses a complex of V segments from multiple lineages, including an IgH-like V segment coexpressed with TCRδ rearrangements. Phylogenetic analyses suggest the proximal IgH clusters could be the ancient precursors to TAILVs that illustrate that IgH and TCRδ have freely exchanged gene segments over the course of the nurse shark evolution (16). *X. tropicalis* possesses a great number of VHδ genes, all of which cluster with clan II VH gene segments (14). The VHδ5 genes in the TCRα/δ locus share a surprisingly high degree of nucleotide identity (91–96%) with VH5 genes in the *Xenopus* IgH locus, consistent with the VHδ5 genes being paralogous to VH5 genes in *X. tropicalis* (14). In Chinese alligator, we also identified 32 VHδ gene segments that can be segregated into 18 subgroups (Fig. 1, Table I). Based on nucleotide identity, four of them were found to be more similar to VH genes of Chinese alligator itself as supported by the phylogenetic analysis (Fig. 5). It is highly likely that these VHδ genes were translocated from the IgH locus into the TCRα/δ locus during the alligator speciation.

Therefore, a transfer of VH genes from the IgH locus to TCRα/δ locus has been occurring repeatedly after the divergence of the IgH and TCR genes in vertebrates (10, 14, 15). This transfer appears to be unidirectional, as thus far, no TCRV genes have been reported in the Ig loci (14).

Analysis of the Chinese alligator TCRα/δ locus reveals a possible evolutionary route of VHδ genes in tetrapod

Previous studies suggested that the translocation of VH genes from IgH to TCR locus appeared to have occurred independently in different species of tetrapods, as the VHδ genes in frogs, birds, and monotremes belonged to different VH clans (10, 14, 15). However, in this study, numerous VHδ genes were identified in Chinese alligators that were found to be related to three distinct VH gene clans (Figs. 1, 5). Most interestingly, as supported by the phylogenetic tree, all the avian VHδ and the platypus VHδ and Vμ, identified previously, are associated with distinct VH clans and could be perfectly matched to their phylogenetic counterparts of VHδ genes in the alligator (Fig. 6). This strongly suggested that all VHδ and Vμ genes identified in birds and mammals had a common origin with the VHδ genes found in alligators. Based on this study and previous efforts by other teams, we propose a model

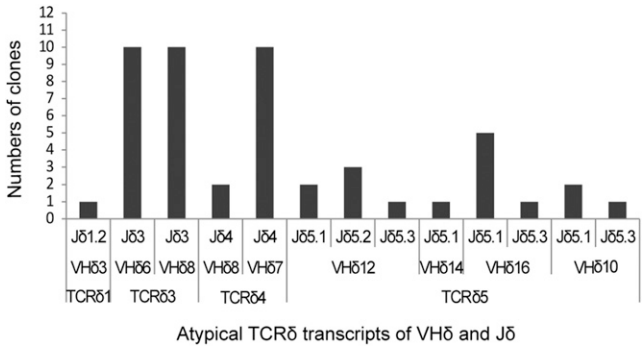

FIGURE 7. Usage frequency of VHδ and Jδ genes in Chinese alligator. In horizontal axis, TCRδ1, TCRδ3, TCRδ4, and TCRδ5 indicate transcripts amplified from different Cδ, the number behind the VHδ means the number of families in Chinese alligator, and the number behind the J represents the number of allele. The vertical scale represents the number of the clone.

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

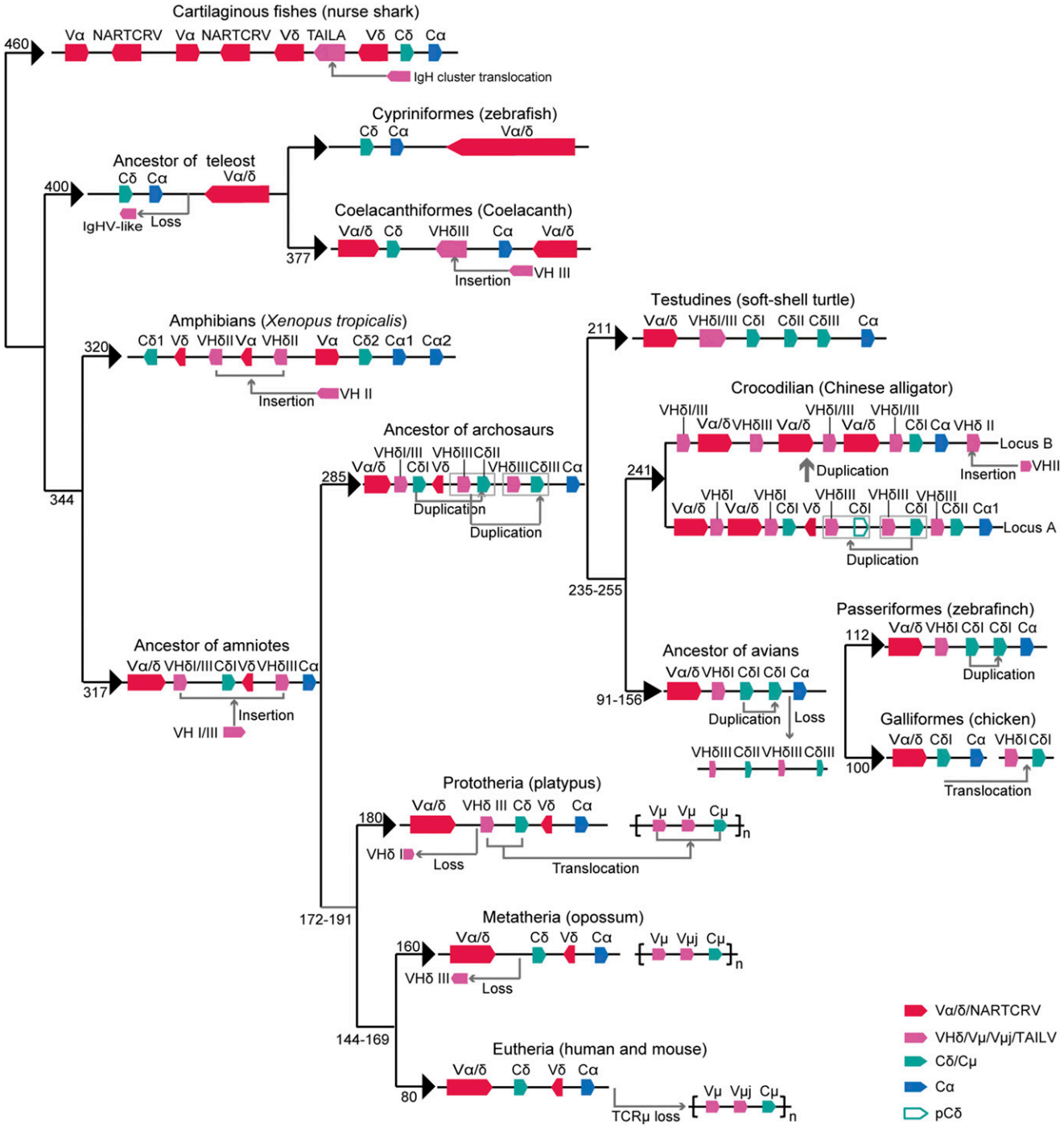

FIGURE 8. A model of the origin and evolution of VHδ gene in TCRα/δ loci of different tetrapod. Genes are color coded as indicated at the right: Vα/δ, red; VHδ and Vμ, pink; Cδ and Cμ, dark green; and Cα, blue. In *X. tropicalis*, VH clan II gene or genes were translocated from the IgH locus and inserted into the TCRα/δ locus, then VHδ II genes expanded (14). In cartilaginous fish, V gene segments of the proximal IgH locus were translocated into the TCRδ locus as TAILVs (16). In the evolution of TCRα/δ locus in an ancestral teleost, IgHV-like segments lost because of an incomplete recombination of the V array to the other side of the D–J–C cluster (41). In coelacanth, the TCRα/δ locus contained VHδ genes (VH clan III) derived from IgH gene locus, most likely during the speciation of coelacanth (21). In an amniotic ancestor, VH clan I and VH clan II gene translocated and inserted into TCRα/δ locus. In the ancestor of archosaurs, duplication of Cδ I led to the presence of Cδ I and Cδ II. Then, VHδ III–Cδ II duplicated to a new VHδ III–Cδ III cluster. In alligator, duplication of whole locus A generated locus B. Then V gene expensed and lost Cδ in TCRα/δ locus A of alligator, whereas in TCRα/δ locus B, additional tandem duplication of the VHδ III–Cδ cluster occurred. In birds, VHδ III genes and Cδ II and Cδ III genes were lost, and another Cδ I gene was duplicated in the ancestor of avian. But in the Passeriformes lineage, one Cδ I gene was lost, and the remaining Cδ I was duplicated after species divergence. In contrast, in the Galliformes lineage, the VHδ–Dδ–Jδ–Cδ cluster was translocated out of the TCRα/δ locus and currently resides on another chromosome (15). In platypus, VHδ I gene was lost, and the VHδ–Dδ–Jδ–Cδ cluster was translocated out of the TCRα/δ locus, giving rise to TCRμ; mammals retained the VHδ III genes (10, 12). The Metatheria lineage lost VHδ from the TCRα/δ locus, remaining a TCRμ locus only. In eutherian mammals, the TCRμ locus was also lost (10, 34).

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

to explain the evolution of TCRα/δ genes in tetrapods, as described below.

In cartilaginous fish, the TCRδ locus is also nested within the TCRα locus and contains the TAILVs segments (IgHV-like) that are the result of a proximal IgH translocation into the TCRδ locus (16). Phylogenetic evidence supports that IgHV segments in proximal IgH clusters of TCRδ locus could be the ancient precursors to TAILVs (16).

In the evolution of TCRα/δ locus in an ancestral teleost fish, the IgHV-like segments may have been lost because of an incomplete recombination of the V array to the opposite side of the D–J–C cluster (41). This rearrangement could have produced the unusual locus organization with Vα/δ genes in an inverted transcriptional orientation 3′ of Cα (41). In coelacanth, which is closely related to the ancestor of terrestrial vertebrates, the TCRα/δ locus contained VHδ genes (VH clan III) derived from IgH gene locus most likely during the speciation of this lineage (21) (Fig. 8).

In the *Xenopus* lineage, a duplication of the Dδ–Jδ–Cδ cluster occurred, resulting in the presence of two Cδ, each with its own set of Dδ and Jδ segments (10, 14). Then, VH clan II gene or genes from the IgH locus were translocated into the TCRα/δ locus (Fig. 8) because *Xenopus* VHδ is all clustered with its own VH II gene segments in the phylogenetic tree (14). Subsequently, the region of the VHδ II–Dδ–Jδ–Cδ cluster was inverted and VHδ II genes expanded, leading to the number of VH genes present in frog (10, 14). In frogs, IgH and TCRα/δ loci are also closely linked, which may have facilitated translocation of VH into the TCRα/δ locus (14).

In a common ancestor of the amniotes, the VHδ I and VHδ III genes were already present in the TCRα/δ locus as a result of VH clan I and VH clan III gene translocation (Figs. 1, 8). Then, in the ancestor of birds, testudines, and crocodilians, duplication of Cδ I led to the presence of Cδ I and Cδ II with its own set of Dδ and Jδ segments (Figs. 1, 4, 8). Subsequently, VHδ III–Cδ II cluster also duplicated in the ancestor of birds, testudines, and crocodilians, generating a new VHδ III–Cδ III cluster (Figs. 1, 4, 8).

After alligators diverged from archosaurs, locus B was duplicated from locus A in alligator, an event supported by phylogenetic analyses of the Cα genes and Cδ genes (Figs. 3, 4, 8). However, the two loci have undergone different evolutionary paths following duplication. In locus A, the VHδ III–Cδ II cluster duplicated again, leading to presence of the fourth Cδ gene (Figs. 1, 8). In locus B, a V gene expansion and loss of selected Cδ genes have shaped the present genomic structure, which contains 98 V gene segments and only one Cδ (Fig. 1). Additionally, VH clan II translocated into locus B, allowing a single VHδ gene located at the 3′ end of the locus B (Figs. 1, 8).

During the evolution of bird TCRα/δ locus, the VHδ III genes and Cδ II and Cδ III genes were lost, but the Cδ I gene was duplicated in the common ancestor of birds (Fig. 8). In the Galliformes lineage (such as chicken and turkey), the VHδ–Dδ–Jδ–Cδ cluster was translocated out of the TCRα/δ locus, creating an additional locus on another chromosome (15). This can explain why all avian VHδ grouped with VH clan I even though no VH clan I genes are found in the avian IgH locus today (15). While in the Passeriformes lineage (such as zebra finch), one ancestral Cδ I gene was lost, and the other remaining Cδ I was later duplicated again after zebra finch separated (Figs. 4, 8).

In contrast to birds, the common ancestor of mammals presumably lost the VHδ I genes, and then the VHδ–Dδ–Jδ–Cδ cluster was translocated out of the TCRα/δ locus, giving rise to TCRμ (10). Therefore, they retained the VHδ III genes (10, 12, 17). However, in the Metatheria lineage, including opossum, VHδ III was lost in the TCRα/δ locus, whereas the TCRμ locus is still maintained in their genomes (12). Later in evolution of eutherian mammals, the TCRμ locus was also lost (Fig. 8).

In summary, this study identified a distinct genomic organization of the TCRα/δ locus in Chinese alligator, which appears to be the most complicated TCRα/δ locus among all tetrapods examined so far. Our study further provides significant insights into the evolution of atypical TCRδ or TCRμ genes in tetrapods and the dynamic history of the TCRα/δ locus.

Acknowledgments

We thank Drs. Li Ma and Yonghe Ma for technical help in the BAC screening and genome walking.

Disclosures

The authors have no financial conflicts of interest.

References

1. Reddy, S. T. 2017. Immunology: the patterns of T-cell target recognition. *Nature* 547: 36–38.
2. Olivieri, D. N., B. von Haeften, C. Sánchez-Espinel, J. Faro, and F. Gambón-Deza. 2014. Genomic V exons from whole genome shotgun data in reptiles. *Immunogenetics* 66: 479–492.
3. Gambón-Deza, F., and D. N. Olivieri. 2018. Immunoglobulin and T cell receptor genes in Chinese crocodile lizard *Shinisaurus crocodilurus*. *Mol. Immunol.* 101: 160–166.
4. Kronenberg, M., G. Siu, L. E. Hood, and N. Shastri. 1986. The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev. Immunol.* 4: 529–591.
5. Rast, J. P., M. K. Anderson, S. J. Strong, C. Luer, R. T. Litman, and G. W. Litman. 1997. Alpha, beta, gamma, and delta T cell antigen receptor genes arose early in vertebrate phylogeny. *Immunity* 6: 1–11.
6. Satyanarayana, K., S. Hata, P. Devlin, M. G. Roncarolo, J. E. De Vries, H. Spits, J. L. Strominger, and M. S. Krangel. 1988. Genomic organization of the human T-cell antigen-receptor α/δ locus. *Proc. Natl. Acad. Sci. USA* 85: 8166–8170.
7. Agrawal, A., Q. M. Eastman, and D. G. Schatz. 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* 394: 744–751.
8. Bassing, C. H., W. Swat, and F. W. Alt. 2002. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* 109: S45–S55.
9. Flajnik, M. F., and M. Kasahara. 2010. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* 11: 47–59.
10. Parra, Z. E., M. Lillie, and R. D. Miller. 2012. A model for the evolution of the mammalian t-cell receptor α/δ and μ loci based on evidence from the duckbill Platypus. *Mol. Biol. Evol.* 29: 3205–3214.
11. Criscitiello, M. F., M. Saltis, and M. F. Flajnik. 2006. An evolutionarily mobile antigen receptor variable region gene: doubly rearranging NAR-TcR genes in sharks. *Proc. Natl. Acad. Sci. USA* 103: 5036–5041.
12. Parra, Z. E., M. L. Baker, R. S. Schwarz, J. E. Deakin, K. Lindblad-Toh, and R. D. Miller. 2007. A unique T cell receptor discovered in marsupials. *Proc. Natl. Acad. Sci. USA* 104: 9776–9781.
13. Parra, Z. E., M. L. Baker, A. M. Lopez, J. Trujillo, J. M. Volpe, and R. D. Miller. 2009. TCR mu recombination and transcription relative to the conventional TCR during postnatal development in opossums. *J. Immunol.* 182: 154–163.
14. Parra, Z. E., Y. Ohta, M. F. Criscitiello, M. F. Flajnik, and R. D. Miller. 2010. The dynamic TCRδ: TCRδ chains in the amphibian *Xenopus tropicalis* utilize antibody-like V genes. *Eur. J. Immunol.* 40: 2319–2329.
15. Parra, Z. E., K. Mitchell, R. A. Dalloul, and R. D. Miller. 2012. A second TCRδ locus in Galliformes uses antibody-like V domains: insight into the evolution of TCRδ and TCRμ genes in tetrapods. *J. Immunol.* 188: 3912–3919.
16. Deiss, T. C., B. Breaux, J. A. Ott, R. A. Daniel, P. L. Chen, C. D. Castro, Y. Ohta, M. F. Flajnik, and M. F. Criscitiello. 2019. Ancient use of Ig variable domains contributes significantly to the TCRδ repertoire. *J. Immunol.* 203: 1265–1275.
17. Wang, X., Z. E. Parra, and R. D. Miller. 2011. Platypus TCRμ provides insight into the origins and evolution of a uniquely mammalian TCR locus. *J. Immunol.* 187: 5246–5254.
18. Flajnik, M. F., N. Deschacht, and S. Muyldermans. 2011. A case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? *PLoS Biol.* 9: e1001120.
19. Stanfield, R. L., H. Dooley, M. F. Flajnik, and I. A. Wilson. 2004. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science* 305: 1770–1773.
20. Dooley, H., and M. F. Flajnik. 2005. Shark immunity bites back: affinity maturation and memory response in the nurse shark, *Ginglymostoma cirratum*. *Eur. J. Immunol.* 35: 936–945.
21. Saha, N. R., T. Ota, G. W. Litman, J. Hansen, Z. Parra, E. Hsu, F. Buonocore, A. Canapa, J. F. Cheng, and C. T. Amemiya. 2014. Genome complexity in the coelacanth is reflected in its adaptive immune system. *J. Exp. Zool. B Mol. Dev. Evol.* 322: 438–463.
22. Ota, T., and M. Nei. 1994. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin VH gene family. *Mol. Biol. Evol.* 11: 469–482.

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020

23. Janke, A., and U. Arnason. 1997. The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent archosauria (birds and crocodiles). *Mol. Biol. Evol.* 14: 1266–1272.
24. Lefranc, M. P., C. Pommié, M. Ruiz, V. Giudicelli, E. Foulquier, L. Truong, V. Thouvenin-Contet, and G. Lefranc. 2003. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* 27: 55–77.
25. Burland, T. G. 2000. DNASTAR's lasergene sequence analysis software. *Methods Mol. Biol.* 132: 71–91.
26. Sonnhammer, E. L., and R. Durbin. 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* 167: GC1–GC10.
27. Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.
28. Cheng, G., Y. Gao, T. Wang, Y. Sun, Z. Wei, L. Li, L. Ren, Y. Guo, X. Hu, Y. Lu, et al. 2013. Extensive diversification of IgH subclass-encoding genes and IgM subclass switching in crocodilians. *Nat. Commun.* 4: 1337.
29. Lefranc, M. P. 2014. Immunoglobulin and T cell receptor genes: IMGT(®) and the birth and rise of immunoinformatics. *Front. Immunol.* 5: 22.
30. Tutter, A., and R. Riblet. 1989. Conservation of an immunoglobulin variable-region gene family indicates a specific, noncoding function. *Proc. Natl. Acad. Sci. USA* 86: 7460–7464.
31. Parra, Z. E., and R. D. Miller. 2012. Comparative analysis of the chicken TCRα/δ locus. *Immunogenetics* 64: 641–645.
32. Carroll, A. M., J. K. Slack, and X. Mu. 1993. V(D)J recombination generates a high frequency of nonstandard TCR D delta-associated rearrangements in thymocytes. *J. Immunol.* 150: 2222–2230.
33. Bassing, C. H., F. W. Alt, M. M. Hughes, M. D'Auteuil, T. D. Wehrly, B. B. Woodman, F. Gärtner, J. M. White, L. Davidson, and B. P. Sleckman. 2000. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature* 405: 583–586.
34. Parra, Z. E., M. L. Baker, J. Hathaway, A. M. Lopez, J. Trujillo, A. Sharp, and R. D. Miller. 2008. Comparative genomic analysis and evolution of the T cell receptor loci in the opossum *Monodelphis domestica*. *BMC Genomics* 9: 111.
35. Giudicelli, V., D. Chaume, and M. P. Lefranc. 2005. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* 33: D256–D261.
36. Bosc, N., and M. P. Lefranc. 2003. The mouse (*Mus musculus*) T cell receptor alpha (TRA) and delta (TRD) variable genes. *Dev. Comp. Immunol.* 27: 465–497.
37. Herzig, C. T., S. L. Blumerman, and C. L. Baldwin. 2006. Identification of three new bovine T-cell receptor delta variable gene subgroups expressed by peripheral blood T cells. *Immunogenetics* 58: 746–757.
38. Reinink, P., and I. Van Rhijn. 2009. The bovine T cell receptor alpha/delta locus contains over 400 V genes and encodes V genes without CDR2. *Immunogenetics* 61: 541–549.
39. Kubota, T., J. Wang, T. W. Göbel, R. D. Hockett, M. D. Cooper, and C. H. Chen. 1999. Characterization of an avian (Gallus gallus domesticus) TCR alpha delta gene locus. *J. Immunol.* 163: 3858–3866.
40. Yang, Z., Y. Sun, Y. Ma, Z. Li, Y. Zhao, L. Ren, H. Han, Y. Jiang, and Y. Zhao. 2017. A comprehensive analysis of the germline and expressed TCR repertoire in White Peking duck. *Sci. Rep.* 7: 41426.
41. Seelye, S. L., P. L. Chen, T. C. Deiss, and M. F. Criscitiello. 2016. Genomic organization of the zebrafish (*Danio rerio*) T cell receptor alpha/delta locus and analysis of expressed products. *Immunogenetics* 68: 365–379.
42. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–425.

Downloaded from http://www.jimmunol.org/ at Univ of Exeter on June 27, 2020