Immunogenetics (2016) 68:365–379
DOI 10.1007/s00251-016-0904-3

ORIGINAL ARTICLE

Genomic organization of the zebrafish (*Danio rerio*) T cell receptor alpha/delta locus and analysis of expressed products

Stacie L. Seelye[1] · **Patricia L. Chen**[1] · **Thaddeus C. Deiss**[1] · **Michael F. Criscitiello**[1,2]

Received: 11 October 2015 / Accepted: 18 January 2016 / Published online: 26 January 2016

Abstract In testing the hypothesis that all jawed vertebrate classes employ immunoglobulin heavy chain V (IgHV) gene segments in their T cell receptor (TCR)δ encoding loci, we found that some basic characterization was required of zebrafish TCRδ. We began by annotating and characterizing the TCRα/δ locus of *Danio rerio* based on the most recent genome assembly, GRCz10. We identified a total of 141 theoretically functional V segments which we grouped into 41 families based upon 70 % nucleotide identity. This number represents the second greatest count of apparently functional V genes thus far described in an antigen receptor locus with the exception of cattle TCRα/δ. Cloning, relative quantitative PCR, and deep sequencing results corroborate that zebrafish do express TCRδ, but these data suggest only at extremely low levels and in limited diversity in the spleens of the adult fish. While we found no evidence for IgH-TCRδ rearrangements in this fish, by determining the locus organization we were able to suggest how the evolution of the teleost α/δ locus could have lost IgHVs that exist in sharks and frogs. We also found evidence of surprisingly low TCRδ expression and repertoire diversity in this species.

Keywords *Danio rerio* · TCR α · TCR δ · γδ T cells

This work was supported by the National Science Foundation through a grant to MFC (IOS 1257829).

Electronic supplementary material The online version of this article (doi:10.1007/s00251-016-0904-3) contains supplementary material, which is available to authorized users.

✉ Michael F. Criscitiello
mcriscitiello@cvm.tamu.edu

Stacie L. Seelye
sseelye@cvm.tamu.edu

Patricia L. Chen
pchen@cvm.tamu.edu

Thaddeus C. Deiss
tdeiss@cvm.tamu.edu

[1] Comparative Immunogenetics Laboratory, Department of Veterinary Pathobiology, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843, USA

[2] Department of Microbial Pathogenesis and Immunology, School of Medicine, Texas A&M University, Bryan, TX 77807, USA

Introduction

Zebrafish (*Danio rerio*) continues to increase in popularity a vertebrate model species (Iwanami 2014). Zebrafish entered the forefront of animal research in the 1980s due to the ability to perform large-scale genetic screens and production of developmental mutants in the species with studies by George Streisinger (Chakrabarti et al. 1983; Walker and Streisinger 1983). Over time, the use of the fish species was extended to other fields, such as pathology, toxicology, behavior, and evolution (Harper 2011). One significant area zebrafish has contributed is developmental and comparative immunogenetics (Iwanami 2014).

Understanding the organization of the genes that code for zebrafish lymphocyte antigen receptors is integral to our understanding of the immune system of this useful animal model. T cell receptors (TCR), along with immunoglobulin, confer clonal specificity for activation of lymphocytes and are heterodimers of two chains. The chains are typically divided into four classifications, the α/β and γ/δ each forming pairs. T cells bearing the γ/δ heterodimer have many subsets with unique properties and often exhibit features of innate immune responses. They are typically found in epithelial and gastrointestinal tissues and are prevalent in early and fetal

development in many species. Some γ/δ T cells migrate early in development to particular tissues such as the liver, skin, mucosa of the lungs, digestive, and reproductive organs and persist as resident cells (Bonneville et al. 2010). γ/δ T cells form a much larger proportion of the peripheral T cell pool in adult ruminants, rabbits, and chickens than in primates and rodents (Hein and Mackay 1991; Holderness et al. 2013). Relatively little is known about the functional importance and prevalence of γ/δ T cells in teleost fish. We do know that the physiological roles fulfilled by γ/δ T cells in mammals are varied. Some subsets of γ/δ T cells are unique in that they recognize conserved non-peptide antigens that are often upregulated by stressed cells, the expression modalities and distribution of which resemble those of pathogen associated molecular patterns (PAMPs) or danger associated molecular patterns (DAMPs). This is in contrast to α/β T cells which are restricted to recognizing and responding to peptide antigen presented by self MHC molecules (Bonneville et al. 2010). During development of the thymocyte, TCR genes undergo somatic rearrangement of the genetic elements that encode components of the receptor, called the variable (V), diversity (D), and joining (J) gene segments. There is considerable information about V(D)J rearrangement and the facilitating TCR locus organization in many mammalian species but the data available for teleost and lower ectothermic vertebrates is more sparse (Moulana et al. 2014). In most mammalian species the TCRδ locus is imbedded within the TCRα locus and has the following arrangement: Vα/δ-Dδ-Jδ-Cδ-Jα-Cα, with some V's being used by both α and δ (Murphy 2012). The genomic arrangement for some teleost fish has been elucidated, specifically for the Japanese pufferfish (*Takifugu rubripes*) (Wang et al. 2001b) and the green pufferfish (*Tetraodon fluviatilis*) (Fischer et al. 2002). These species show a unique organization for teleost fish of Dδ-Jδ-Cδ-Jα-Cα-Vα/δ, with the Vs in an inverted orientation with respect to the other elements.

The genomic arrangement of the TCR αδ locus of some teleost fish has been elucidated, specifically for the Japanese pufferfish (*T. rubripes*) (Wang et al. 2001a) the green pufferfish (*T. fluviatilis*) (Fischer et al. 2002), and the Atlantic salmon (*Salmo salar*) (Yazawa et al. 2008a). These species show a unique organization for teleost fish of Dδ-Jδ-Cδ-Jα-Cα-Vα/δ, with the Vs in an inverted orientation with respect to the other elements. Additionally, the genes coding for TCRδ and TCRα have been identified but the organization of the genomic locus has yet to be elucidated in channel catfish (*Ictalurus punctatus*) (Moulana et al. 2014). There is also limited information about the TCRγ and TCRδ genes of the mandarin fish (*Siniperca chuatsi*) (Tian et al. 2014), sea bass (*Dicentrarchus labrax*) (Buonocore et al. 2012), common carp (*Cyprinus carpio*) (Shang et al. 2008), and flounder (*Paralichthys olivaceus*) (Nam et al. 2003). To date, little has been published about the genomic organization of the α/δ locus in zebrafish. Various papers have focused on individual aspects of these receptors and have found ample cDNA sequences coding for TCRα yet a very few TCRδ rearrangements (Haire et al. 2000) (Danilova et al. 2004) (Schorpp et al. 2006).

Salmon have 128 potentially functional Vα/δ, human has 57, mouse has 98, and chicken has 70. In zebrafish 148 Vα genes have so far been found on BAC clones containing no apparent defect (Danilova et al. 2004). Previous work identified two Vδ, two Jδ segments, three Dδ segments, as well as one Cδ segment (Schorpp et al. 2006). A cDNA library recovered four related TCRα clones, each with unique V, J, and C sequences and several Jα segments. The cDNA sequence for one Cδ and four Cα rearranged products has been identified (Haire et al. 2000).

An interesting immunogenetic phenomenon concerning TCRδ has been discovered in nurse shark (Criscitiello et al. 2010), *Xenopus* (Parra et al. 2010b), chicken (Parra and Miller 2012), opossum (Parra et al. 2008), platypus (Wang et al. 2011), and most recently in the coelacanth (Amemiya et al. 2013). These vertebrates have the ability to utilize immunoglobulin heavy chain (IgH) V gene sequences to create apparently functional TCR δ (and perhaps α) chains. This process was originally coined transrearrangement. Elucidation of the organization of the TCRδ loci of some model species has shown that many of these species possess V segments located within the α/δ locus that show much higher identity to immunoglobulin heavy chain V sequences than to TCRα/δ V sequences. In originally setting out to determine if such IgHV segments are used in the teleost TCRα/δ loci, we annotated the zebrafish locus and found evidence for low or at least unusual expression of canonical TCRδ.

Methods

C region search

A tBLASTn search was performed using the TCRδC sequence from *P. olivaceus* (accession # BAC65463.1) against version GRC Z10 of the zebrafish genome (as well as other bony fish genomes in our initial interrogation for IgHV gene segments). One match was found on chromosome 2. To verify that this was TCRδC, this sequence was used to perform a tBLASTn search against the nonredundant database. A phylogenetic analysis of various mammalian, amphibian, teleost, and chondrichthyes species TCRδC sequences using MUSCLE for multiple sequence alignment and the neighbor joining method to create a phylogenetic tree with bootstrap values from 1000 iterations was done using the MEGA 6.0 software package (Tamura et al. 2013).

V region search

Using a previously annotated TCRδV sequence from *P. olivaceus* (accession #AB076071.1) as bait sequence, a tBLASTn search was performed on the GRCz10 reference assembly of zebrafish. Genomic sequences were downloaded into the Geneious version 7 (Kearse et al. 2012) software suite for annotation of the locus, multiple sequence alignments and phylogenetic tree analyses. Recombination signal sequences and intron splice signals were identified manually for all sequences and were used to determine the limits of the coding nucleotide sequences of V, D, and J segments. All sequences were trimmed to remove splice signals and recombination signal sequences before a V gene multiple sequence alignment was performed using Clustal W. A phylogenetic tree was created using the neighbor joining method in MEGA 6.0 (Tamura et al. 2013). These V gene sequences were analyzed using a percent identity matrix generated from the multiple sequence alignment. V segments were placed in families based on the rule that sequences that shared 70 % nucleotide identity with at least one other sequence were placed in the same family. Families were then ordered based on their position within the locus. Groupings within families that showed higher percent identity were placed in subgroups. This was represented by the number after the first decimal in the naming protocol. If the sequence did not belong in a subgroup, the second digit of 0 was used to denote no subgroup.

Locus annotation

Scaffold version 10 of chromosome 2 was downloaded from NCBI and imported into Geneious version 7. All 149V regions from the above BLAST search were annotated as well as previously found D, J, and C regions for TCRα/δ. Some TCRδ D, J, and C segments matched published sequences (Schorpp et al. 2006), previously described Jα sequences were found using a custom annotation database created by IMGT/LIGM-DB (http://www.imgt.org/ligmdb/) and previously described Cα sequence was confirmed on the scaffolded assembly (Haire et al. 2000). Annotated sequences were then manually evaluated and the start and stop codons were identified based on appropriate splice sites. The entire locus was again visually inspected for the presence of additional, potentially functional D and J segments. Additional D segments were analyzed for the presence of the heptamer, spacer, and nonamer sequence on both the 5′ and 3′ end of sequences. Additional J segments were identified based on the presence of the heptamer, spacer, nonamer sequence and the presence of the FGxG motif as a hallmark of J regions as well as the FGxP motif identified in zebrafish Jδ.

Search for evidence of TCRδ transcripts

A 5′ RACE library was created using the GeneRacer Kit (Invitrogen Waltham MA) from spleen RNA originally isolated from 12 outbred zebrafish (kind gift from Matt Young). PCR was performed using a reverse primer to target the TCRδC region and forward primer to the GeneRacer 5′ oligo adapter. All primers used can be found listed in Supplemental Table 1. Primary PCR was performed using 1 μl 10 μM dNTP, 10 μl 5× Phusion buffer (New England Biolabs, Ipswich MA), 2.5 μl 10 μM GENERACER 5′ primer, 2.5 μl 10 μM reverse primer MFC527, 2 μl 50 ng/μl template, 0.5 μl 2 U/μl High Fidelity Phusion DNA polymerase, and PCR quality water to total volume of 50 μl. Thermocycler (Bio-Rad C1000 thermal cycler, Bio-Rad Laboratories, Hercules CA) protocol was as follows: (1) initial denaturation 95 °C for 15 min, (2) denaturation 95 °C 30 s, (3) annealing and elongation 72 °C for 30 s repeat steps 2–3 30 times, (4) final elongation 72 °C for 5 min. Secondary PCR was performed using 3 μl 25 mM $MgCl_2$, 1 μl 10 mM dNTP, 1 μl 10 μM GENERACER 5′ nested primer, 1 μl 10 μM reverse primer MFC525, 3 μl of template from primary PCR, 5 μl 10× Buffer, 0.25 μl 5 U/μl HotFire DNA Polymerase (Solis BioDyne, Tartu, Estonia), and PCR quality water to total volume of 50 μl. Thermocycler protocol was as follows: (1) initial denaturation 95 °C for 15 min, (2) denaturation 95 °C for 30 s, (3) annealing 63 °C for 30 s 4) elongation 72 °C for 1 min repeat steps 2–4 35 times (5) final elongation 72 °C for 5 min.

Amplicon DNA was extracted from the gel slice and ligated into the PCR-II vector (Invitrogen). The plasmid was used to transform chemically competent One Shot Top10 *E. coli* cells (Life Technologies, Waltham MA). Cultures were grown on ampicillin plates coated with X-gal. White and light blue colonies were selected and prepared for sequence analysis using the ZR Plasmid MiniPrep Kit (Zymo Research, Irvine, CA). Sequencing products from the plasmid were amplified using BigDye (Life Technologies) and samples were sequenced by the Gene Technology Laboratory at Texas A&M. Geneious version 7 (Kearse et al. 2012) was used for sequence analysis.

Additional short, minimally degenerate primer PCR (Rast and Litman 1994) was performed targeting the conserved framework sequence encoding the WYRQ motif with the same reverse primers from the 5′ RACE PCR. PCR was first performed according to the same protocol as mentioned above for the secondary PCR, with the primers used being 1 μl 10 μM MFC535, MFC536, or MFC537 primers and 1 μl 10 μM reverse primer MFC525. The amount of template used was 3 μl of 50 ng/μl 5′ RACE library cDNA. The thermocycler protocol was also the same with the exception of using an annealing gradient of 30–54 °C for 30 s at step

three and only 30 cycles were performed. No bands were obtained from this attempt on 0.8 % agarose gel except for positive control amplicons.

A second attempt was performed where the amount of 25 mM $MgCl_2$ was increased to 4 μl and only the primer MFC537 was used. The thermocycler protocol used was the same as previous except with an annealing gradient 47–56 °C for 30 s and the number of cycles was increased to 35. A positive control using 1 μl 10 μM forward and reverse primers for the housekeeping gene RpL13α under the same conditions was included. The thermocycler protocol for the control was the same with the exception of using an annealing temperature of 59 °C. Again, no bands were obtained on 0.8 % agarose gel except for positive control amplicons.

A third attempt was made where the amount of 25 mM $MgCl_2$ was returned to 3 μl and the amount of the template was varied so that either 2, 4, 6, or 8 ng/μl of the 5′ RACE library template cDNA was used. The reverse primer was changed to MFC 527. The remainder of the mixture was the same. The thermocycler protocol used the same temperatures used previously with the exception of the annealing temperature of 52 °C for 30 s and 35 cycles were performed. Secondary PCR was performed using the reverse primer MFC525 and 1 μl of each of the primary PCR products. The remainder of the components remained the same. The same thermocycler protocol was used with the exception of 30 cycles in this run. Again, no bands were obtained on 0.8 % agarose gel.

Quantitative real-time PCR

Quantitative PCR was performed with 50 ng of random hexamer primed cDNA generated with SuperScript III from RNA samples from pooled adult zebrafish spleen immunized with DNP-KLH via intraperitoneal injection (Weir et al. 2015). We used the SYBR Green PCR Master Mix (Roche, Branford CT) following the manufacturer's recommendation. Triplicate wells were assayed in a Roche LightCycler 480, for 45 cycles annealing at 58 °C, followed by a melting curve analysis. Primers for all four TCR constant region genes and the ribosomal protein gene RpL13α are listed in Supplemental Table 1. The $2^{-\Delta\Delta Ct}$ method using RpL13α as the calibrator (Livak and Schmittgen 2001) was used to calculate relative TCR chain constant gene expression comparing unimmunized control fish to immunized fish. Summary statistics were performed in R with the *summarySE* function of Rmisc package (Hope 2013; R Core Team 2014). Statistical analyses of the variance of mean, ANOVA, and Tukey HSD, were performed in R using the base stats package with a 95 % confidence level (Chambers et al. 1992; R Core Team 2014). Visualization of the data performed in R with the *ggplot2* package, statistical significance indicated by *p* values of the Tukey HSD post-hoc test (Wickham 2009).

Pacbio sequencing

The same primers that successfully amplified the zebrafish TCRδ rearrangement above with the 5′ RACE approach were barcoded for Pacbio SMRT deep amplicon analysis (Supplemental Table 1). cDNA was initially denatured at 98 °C for 2 min then amplified with Phusion (NEB) high fidelity polymerase for 34 cycles consisting of two steps: 98 °C for 10 s and 72 °C for 40 s, ending with a final elongation at 72 °C for 5 min. Bands were excised after visualization in a 1 % agarose gel, and extracted using Qiaquick gel extraction columns (Qiagen). Samples were pooled and sent to Duke University Genome Sequencing Center for Pacbio small insert library preparation (1–3 kb) and SMRT sequencing (P6-C4 Chemistry). Initial quality control, read filtering, and Circular Consensus Sequence (CCS) analysis were performed at the Duke University Genome Center. CCSs containing the proper barcoded primers were then annotated within the Geneious R7 Software Suite (Biomatters).

Results

In order to analyze canonical TCRδ use in zebrafish and search for Ig/TCR transrearrangements in the teleosts, the TCR α/δ locus of *D. rerio* was manually annotated using the latest genome assembly and taking into account the scant expression data in the literature, sequence databases, and PCR cloning in our laboratory. This resulted in the first complete map of the locus and the description of many previously undescribed genetic elements (Fig. 1). The general organization of the locus follows that of other teleosts studied: Dδ-Jδ-Cδ-Jα-Cα-Vα/δ. The gene names, functionality, genomic sequences, and deduced amino acid sequences of all V, D, J, and C segments for TCR α and δ are found in Supplemental Data 1.

Constant region

A putative zebrafish TCR delta C gene at 36,107,203-36,108,481 of chromosome 2 showed 100 % identity to the one identified on clone DKEY-161 L11 (accession #BX681417) of zebrafish linkage group 2 (Schorpp et al. 2006), a 56 % identity match with the TCR delta C of carp and 42 % with that of salmon. Based on this homology paired with weaker identity to other TCR chain C genes and expression in transcripts with D and J segments 5′ of this C gene, we annotate this as *TCRDC*. A neighbor joining phylogenetic tree with bootstrap values supporting that this gene is indeed TCRδ (Supplemental Figure 1). The multiple sequence alignment used to create this tree is shown in Supplemental Figure 2. No additional potential Cδ sequence locations were found in the zebrafish genome.

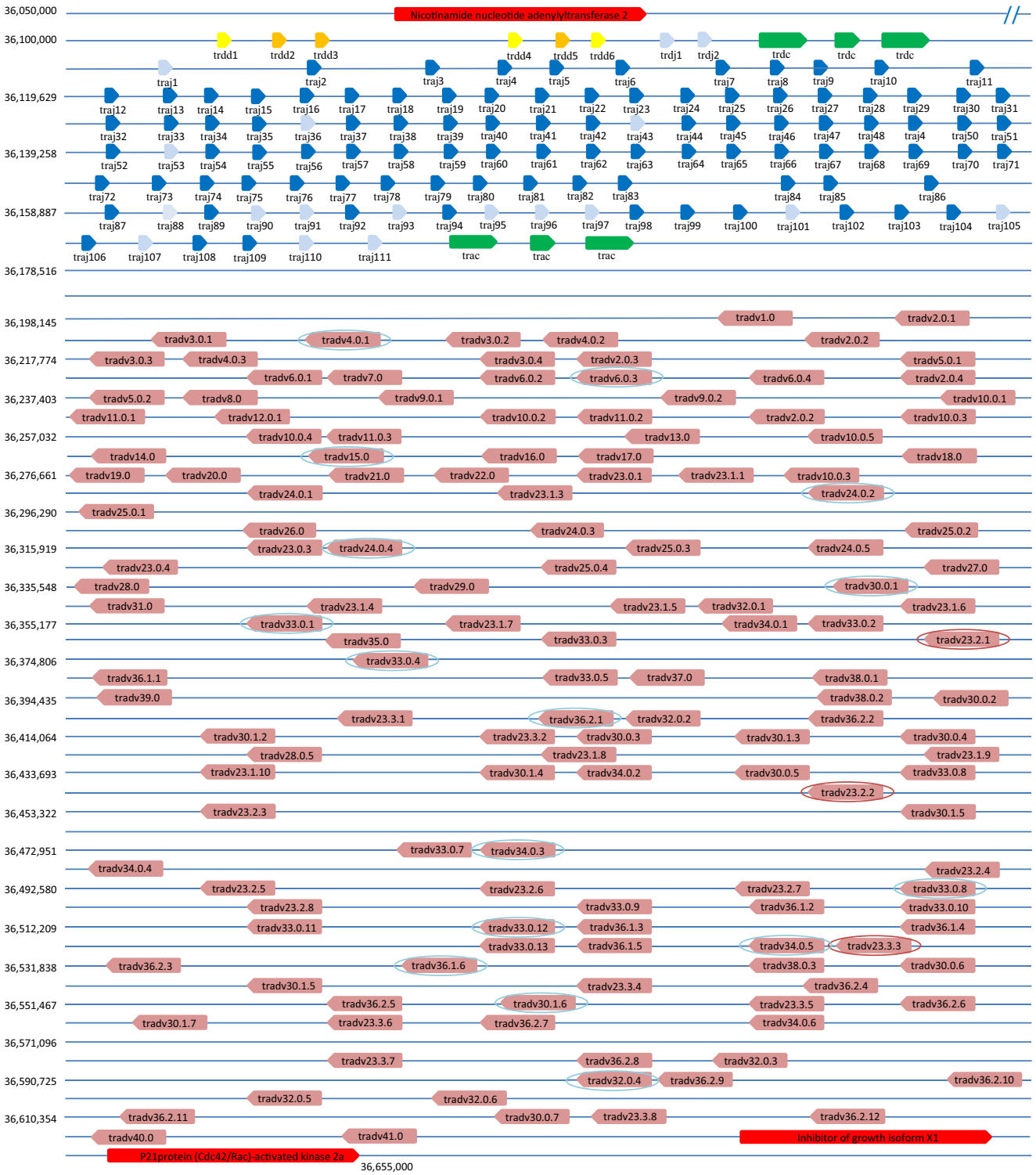


Fig. 1 Annotation of TCR α/δ locus. This annotation is based on the version 10 assembly for Chromosome 2, accession number NC_007113, released on September 24, 2014. *Yellow arrows* in δ D region represent expressed sequences and orange represent potential segments. *Light blue arrows* in α J region represent expressed sequences and *dark blue arrows* represent potential segments. Red ovals represent V sequences expressed with TCR δ constant region. *Blue ovals* represent V sequences expressed with TCR α constant region. *Arrows* represent transcriptional orientation. The numbers represent the nucleotide designation at the beginning of each row. The last annotation represents the last nucleotide for the P21protein (Cdc42/Rac)-activated kinase 2a

V regions

Initial tBLASTn searches with flounder Vδ sequence revealed 159 sequence matches on chromosome 2. Of these 147 sequences were located downstream of the putative Cα region ranging from nucleotides 36,200,000 to 36,600,000. Twelve were discounted as being too many megabases away to be used in this locus. Five sequences (tradv23.2.8, tradv30.1.5, tradv36.2.8, tradv36.2.9, and tradv36.2.11) were incomplete and after visually inspecting the genomic sequence, it was determined that they should be classified as pseudogenes as they did not contain appropriate splice sites. Two additional sequences were removed from consideration due to one being an overlapping duplicate and the second sequence showing only minimal sequence homology to the other V sequences. Three contained stop codons (tradv12.0.2, tradv12.0.1, tradv23.2.6). Four additional V sequences (not of the original 159 identified by BLAST) were found using a custom annotation database created by IMGT/LIGM-DB (http://www.imgt.org/ligmdb/). This gives a total of 141 theoretically functional Vα/δ segments and 8 pseudogenes in this locus. All 141 of these sequences have the canonical sequence structure; the conserved cysteine residues necessary for intradomain disulfide bonds, conserved WYXQ motif in the FR2 region, the YYCA motif in FR3, as well as the RSS located at the 3′ end of each coding segment.

Next, an analysis of these V region genes was performed. Based on a percent identity matrix (Supplemental Table 1) created from a Clustal W (Tamura et al. 2013) multiple sequence alignment (Supplemental Figure 3), these 141 V sequences were placed into 41 different families. Of these, 23 represented single gene families. Three of these families (tradv23, tradv30, and tradv36) were further annotated into subfamilies. The sequences in subfamily tradv23.1 and tradv23.3 all have at least 70 % nucleotide identity between each other sequence of these respective subfamilies (Pascual and Capra 1991). There are two members of subfamilies tradv23.2 and tradv23.3 that share greater than 70 % identity and six sequences in subfamilies tradv23.1 and tradv23.2 whose identity is above that threshold. There are no sequences between tradv23.1 and tradv23.3 that share 70 % identity but they are linked by their similarity with subfamily tradv23.2, hence they were all placed in the same family. Sequences from subfamily 30.1 share 70 % nucleotide sequence identity with each additional sequence in the subfamily. Sequences with designation 30.0 have 70 % identity to only some of the other sequences in the subfamilies 30.0 and 30.1, but not each sequence. Within subfamilies tradv36.1 and tradv36.2, each gene again has 70 % nucleotide sequence identity to each additional sequence within these respective subfamilies, but between these two subfamilies, there are four sequences that share 70 % or greater identity, justifying their placement in the same family.

A neighbor joining tree was created from a Clustal W (Tamura et al. 2013) multiple alignment using the Mega 6.0 software (Tamura et al. 2013). This tree (Fig. 2) confirmed our placing of sequences into families and subfamilies from the percent identity matrix. While the tree is based on the nucleotide sequence alignment in Supplemental Figure 3, the amino acid sequence for these V regions is found in Supplemental Figure 4.

D and J segments

There were three new Dδ and no new potentially functional Jδ segments identified to add to the three Dδ and two Jδ segments already identified (Schorpp et al. 2006). There were an additional 94 Jα segments located, bringing the total to 111 when including the 17 Jα sequences found using a custom annotation database created by IMGT/LIGM-DB (http://www.imgt.org/ligmdb/) which references previously unpublished work by Hohman et al. and submitted to NCBI Genbank in 2001 (Genbank accession numbers AF424544, AF424545, AF424546, AF424547, AF424548, AF425590, and AAL29405.1) and Hammond from the Wellcome Trust Sanger Institute (Genbank accession number AL591399). All of these segments had the canonically accepted sequence structure where the D sequences can be read in all three frames and RSS at both the 5′ and 3′ end. The Jα sequences contain the conserved FGxG motif and a 5′ RSS were also found for each sequence. These sequences and their respective RSSs consisting of a conserved heptamer, 12/23 spacer and nonamer are shown in Supplemental Figure 5.

PCR cloning of TCRδ cDNA products

Multiple PCR strategies from cDNA of multiple fish only yielded one TCRδ rearrangement. An alignment of the one select clone and two additional sequences previously published is shown in Fig. 3. An alignment of all 8 of the original clones obtained by plasmid transformation and Sanger sequencing is shown in Supplemental Figure 6.

No successful PCR amplification was obtained from minimally degenerate primers to avoid any inefficiencies in RACE RNA adaptor ligation targeting the conserved V framework sequence. PacBio sequencing revealed an additional 440 clones. All clones had identical V sequences tradv23.2.2 as well as identical CDR3 regions. An alignment of all 440 clones is shown in Supplemental Figure 7.

Quantitative real-time PCR of four TCR chain gene transcripts

The paucity of cloned functional TCRδ transcripts by traditional PCR prompted us to use quantitative real-time PCR to analyze the levels of change upon immunostimulation of the

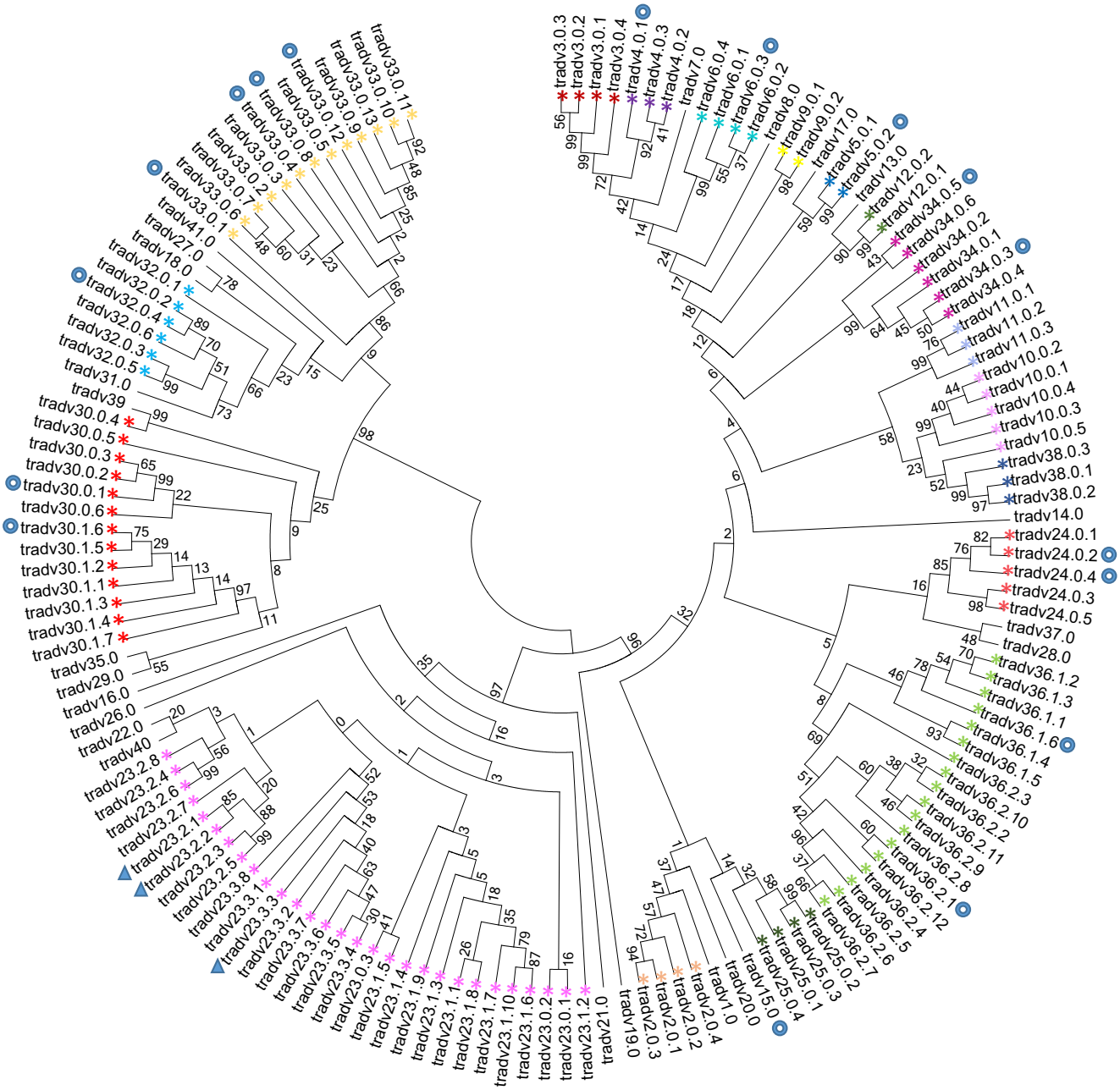


Fig. 2 Phylogenetic analysis of all genomic Vα/δ sequences from *D. rerio*. The neighbor joining tree was drawn using MEGA 6.0 and 1000 bootstrap replications. *Colored asterisks* represent different phylogentic families (some colors were re-used due to limitations in number of colors.) *Triangles* represent those sequences known to be expressed with Cδ. *Circles* represent those Vs known to be used with Cα. Sequences, expression data, and nucleotide locations for start and stop codons are found in Supplemental Data 1

four TCR chain transcripts (Fig. 4). Not only were the relative increases in both transcripts required for the γδ TCR heterodimer very low compared to TCRα, so was that of the β chain of the αβ receptor. At least in the spleen of immunized zebrafish the levels of TCRα upregulation at the mRNA level appears much higher than the other three TCR chains and may explain the lack of TCRδ expressed gene rearrangements we found, as both TCRγ and TCRδ transcripts appear to be limiting. Cloning of the quantified amplicons for sequence confirmation is shown in Supplemental Figure 8.

Discussion

In testing the hypothesis that all jawed vertebrate classes have integrated immunoglobulin heavy chain V gene segments in

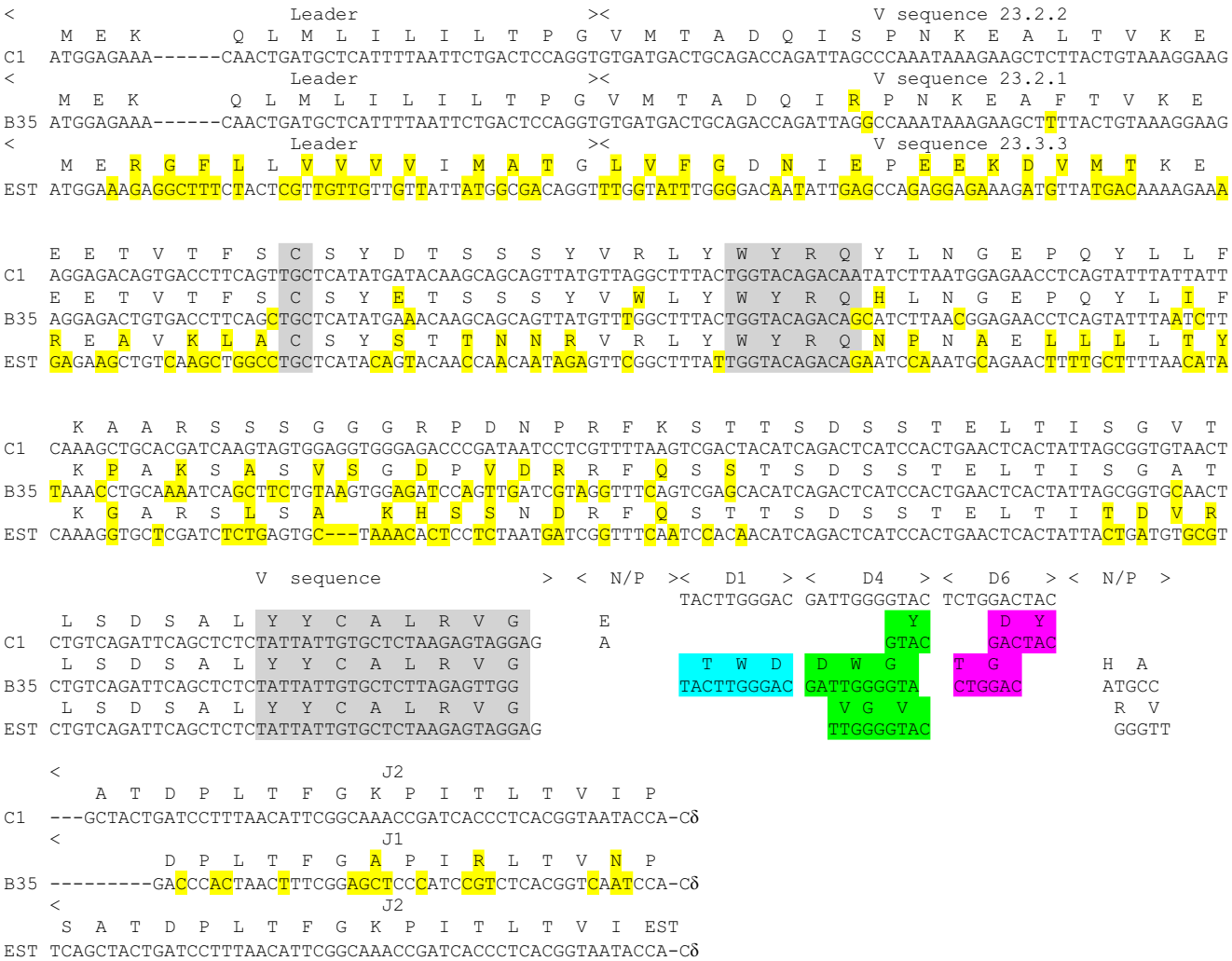


Fig. 3 Alignment of all three known expressed zebrafish Vδ sequences. Differences in V and J sequences are highlighted in *yellow*. Differences in D and N/P nucleotides were not highlighted. Conserved hallmark sequences were highlighted in *gray*. *Blue*, *green*, and *magenta* highlighting marks the nucleotides and amino acids encoded by D1, D4 and D6, respectively. C1 was selected sequence from clones obtained from zebrafish spleen in this work. Sequences were from database search in NCBI yielding the EST (#DT064263.1) and B35 was from the literature (Schorpp et al. 2006)

their TCRδ encoding loci and TCRδ repertoires in a dominant teleost model species, we found that much basic characterization was required of zebrafish TCRδ and the genomic and expressed mRNA levels of this gene. Despite the growing popularity of *D. rerio* as an animal model, there was a surprising scarcity of information detailing the genetic organization and expression data surrounding their use of γδ T cells. Seminal early work described TCRα products of the zebrafish α/δ locus (Haire et al. 2000), and described 8 Vα families that are highly expressed (Danilova et al. 2004). This later publication refers to unpublished work by T. Ota and the Sanger Center identifying at least 148 Vα sequences that have been grouped into 87 families. These genomic annotations were submitted to NCBI Genbank with ascension numbers clone101L20 (Genbank Accession Number AL591481.5), clone 71H18 (AL596128.9), clone 18F12 (AL592550.11), clone 172D23 (AL591399.3), and clone 40G1 (AL591674.3). However, no further information was provided about the criteria used for these family groupings or their position on the genome assembly. A third formative paper provided the first look into the genetics of TCRδ (Schorpp et al. 2006). This paper provided the genomic coding sequence for three Dδ, two Jδ, and one Cδ gene segments from a BAC library (GenBank accession number BX681417.10). Interestingly the Jδ genes use an FGxP motif instead of the more common FGxG amino acid motif, where there is a proline substituted in place of a second glycine in the di-glycine bulge. These works only reported one complete TCRδ V rearrangement from zebrafish (Schorpp et al. 2006), and one more was found in the NCBI database (Fig. 4). Here, we completed the α/δ locus annotation, finding no evidence for IgH V segments. We attempted to analyze the expressed

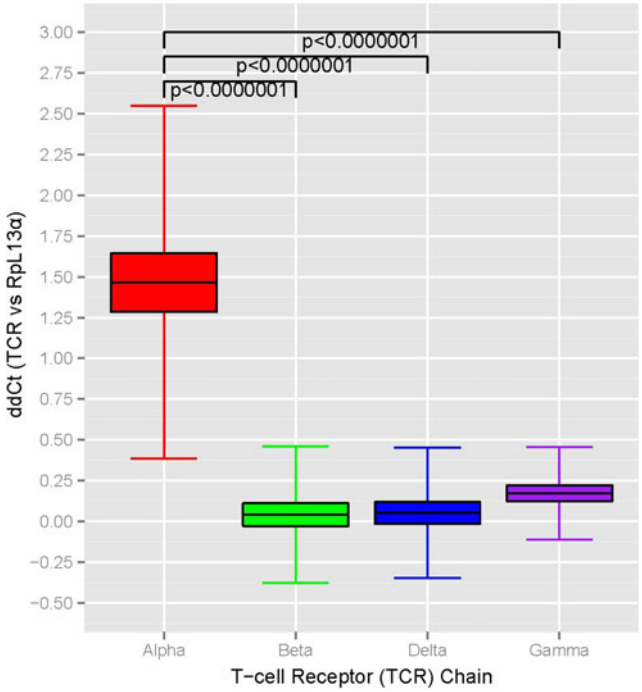


Fig. 4 Splenic TCR chain relative expression preference. qPCR for each TCR chain from zebrafish splenic cDNA. Center lines represent mean relative change in expression, boxes represent a 95 % confidence interval centered on the mean, and whiskers represent standard deviation from the mean. Statistically significant variances in chain expression were determined through ANOVA and post-hoc Tukey HSD

TCRδ repertoire in the fish to rule out transrearrangements to distant IgH V segments, however we found isolation of even canonical (TCRδV-TCRδD-TCRδJ-TCRδC) TCRδ transcripts a challenge, only finding one (and no IgHV-TCRδD-TCRδJ-TCRδC products).

Locus organization

The unusual locus organization with α/δV genes in an inverted transcriptional orientation 3′ of TCRαC we found in *D. rerio* has also been observed in *Tetraodon nigroviridis* (Fischer et al. 2002), *T. rubripes* (Wang et al. 2001b), and *S. salar* (Yazawa et al. 2008b). There is no locus organization data available regarding other teleost fish, only expression data based on cDNA analysis. This may be important evolutionarily as analysis of the skate TCR α and δ loci shows evidence of a larger linkage distance than seen in mammals (Rast et al. 1997). The difference seen in transcriptional orientation for the various segments, suggests that teleosts use more inversional recombination to generate their T cell α and δ functional V encoding exons than the deletional recombination that is commonly seen in mammals (Fig. 5a). When sequences are in opposite orientation, recombination results in inversion of the gene segments instead of deletion (Agard and Lewis 2000). Importantly, this organization does not delete the δ locus at first functional α rearrangement, as is the case in most vertebrates. Thus other mechanisms (possibly greater Erk influence) must control ultimate γδ versus αβ lineage commitment in teleosts.

Further exploration of the locus organization of additional teleost species as well as other cartilaginous and bony fish is warranted as this organization may give additional insight in the evolution of the αδ T cell receptor locus and T lineages. The use of IgHV on TCRδ seemed to be an immunogenetic device evolved in shark and maintained in many vertebrate groups (Criscitiello 2014), yet apparently teleosts discarded it as did several endotherm lineages. It seems possible that an incomplete recombination of the V array in an ancestral teleost to the other side of the D-J-C exons could have produced the downstream V's absent the IgHVs seen in many other vertebrates (Fig. 5b). Duplications within this hypothesized ancestral shark organization could explain the distinct organization in the amphibian *Xenopus* (Parra et al. 2010a). In considering the use of IgHV segments in the TCRδ repertoire of sharks and frogs yet so far not fish, is it possible that this hypothesized inversional locus reorganization in an ancestral teleost is responsible for the loss of the IgHV-TCRδ chimeric receptors in fish? More comparative loci analysis is needed and in the meantime we suggest such a model (Fig. 5b). Starting with the locus organization as we understand it in cartilaginous fish, recombination moving δD-δJ-δC-αJ-αC to a location to the other side of the αδV array could have disrupted the synteny of the IgHV segments with the Vs of the locus yielding the organization and lack of IgHV seen in zebrafish. From the shark organization tandem duplication of many elements and movement of duplicated blocks could yield the much more complex locus seen in *Xenopus*. Deletion of the IgHV/α/δV-δJ-δC-αJ-αC center of the locus organization seen in the anuran amphibian would result in the genomic organization of the TCRαδ locus seen in most mammals.

There also appears to be only a single TCRδ locus in the zebrafish. This is based on our BLAST search results that revealed only one matching genomic sequence in the *D. rerio* genome. The 5′, δD end of the locus is flanked by the nicotinamide nucleotide adenylyltransferase 2 gene in zebrafish as it is in salmon. This is similar to what is thus far found in cartilaginous fish, and higher vertebrates. There are a few exceptions to this rule however. For instance, the *P. olivaceus* was found to possess a second Cδ sequence that existed within the Cγ gene locus (Nam et al. 2003). Additionally, the occurrence of a second TCR δ locus is seen in the Galliforms, such as chickens and turkeys (Parra et al. 2012b). In addition to the conventional TCR α/δ locus, they have a second TCRδ lineage that is unusual in that the V genes are more related to IgHV genes than to TCRV genes. There is evidence that both loci can be active as there is evidence of expression of traditional TCRδ receptors as well as those that utilize the IgHV gene with the TCRδ constant region, similar to the transrearrangement phenomenon seen in sharks and the

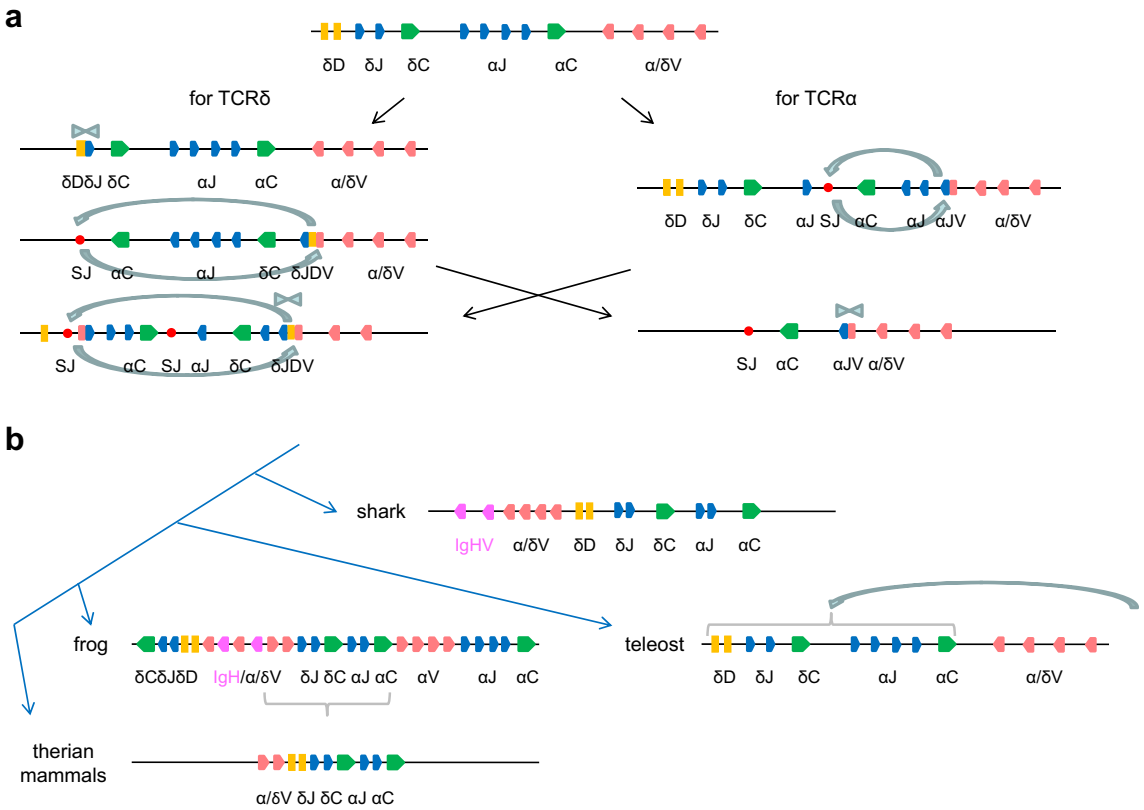


Fig. 5 TCR α/δ locus organization has rearrangement consequences and evolutionary insights. **a** The zebrafish locus can rearrange with an initial V-D or V-J inversion for TCRδ or TCRα respectfully, yet both of these processes leave open the possibility of subsequent rearrangements for making the other TCR chain, and even both. **b** With what we now know about teleost, frog, endotherms, and preliminary data in elephant shark (Venkatesh et al. 2014) and nurse shark (data not shown), it is possible to hypothesize the inversions, duplications and deletions that could have shaped the TCR α/δ locus organization in different vertebrate classes. Pointed ends of the pentagons representing V, J, and C gene segments denote transcriptional orientation. *Red circles* are signal joints left in genome by inversional V(D)J recombination. *Gray arrows and brackets* denote inversions and block movements, *bowties* mark deletions. *Placental mammals* refers to Infraclass Placentalia, monotremes such as the platypus do have IgHV in their TCR α/δ locus (Parra et al. 2012a) and use a choriovitelline placenta that provides nutrients primarily from the yolk sac

chimeric receptors in frogs. This second locus is not found elsewhere, including in other avian lineages such as the Passeriformes (Parra and Miller 2012). There is strong evidence (Parra et al. 2012a) that the IgHV genes used in TCRδ loci and the plasticity in TCRδV use facilitated the evolution in monotremes and marsupial mammals of an additional fifth TCR chain (TCRμ) that is distantly related to TCRδ (Parra et al. 2007; Wang et al. 2011).

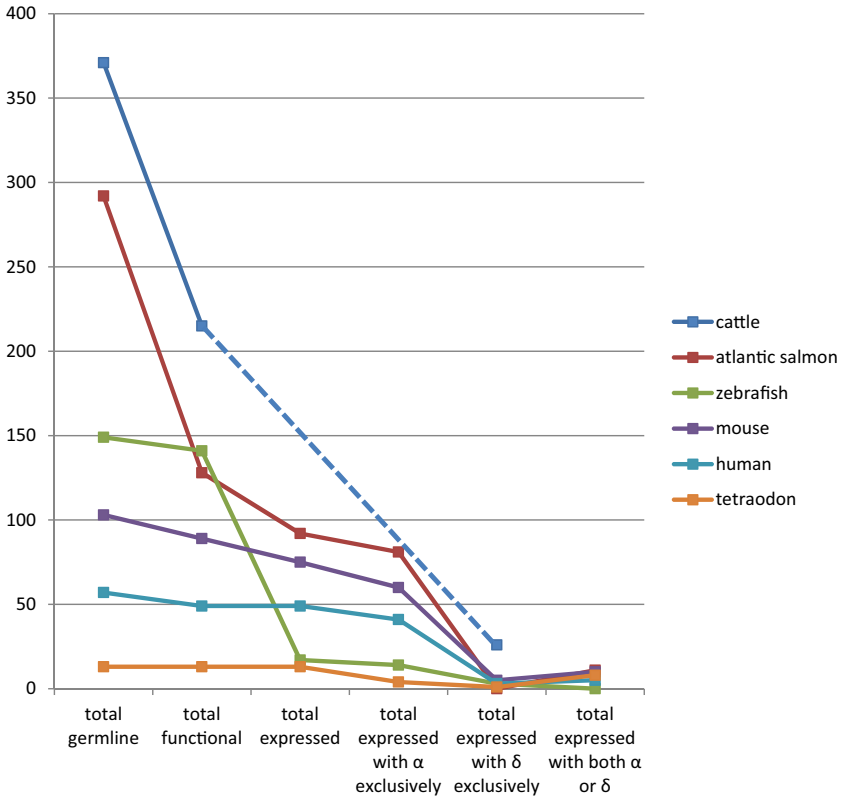
D, J, and C segment analysis

While analyzing and identifying the J segments for TCRδ, it was noticed that these J sequences did not contain the hallmark FGxG sequence; instead, the second glycine is replaced with a proline. This same substitution was also found in the catfish (Moulana et al. 2014). The Atlantic salmon contains one functional Jδ that utilizes a FGKA sequence (Yazawa et al. 2008b) while *Tetraodon* utilizes the canonical FGxG motif (Fischer et al. 2002).

The complete TCRδ C protein sequence found in zebrafish shows high sequence homology to the TCRδ C sequences of other teleost fish, particularly with that of the common carp (*C. carpio*) and *S. salar*, with 56 and 42 % amino acid identity respectively. TCRδ C sequences have also been found for other teleost species including the channel catfish, Japanese flounder, puffer, and fugu. Based on the phylogenetic tree in Supplemental Figure 1, all of these sequences do show high homology with each other and usually to TCRδ C of other classes than to any of the other TCR chains (TCRα of closely related fish being the exception). In addition to its high homology to other TCRδ C, the *D.rerio* TCRδ C does contain the highly conserved cysteine residues.

V region analysis

Zebrafish appear to have the second largest number of potentially functional α/δ V genes of any species previously studied (Fig. 6). The species with the greatest number of V sequences is the cow (Connelley et al. 2014), which curiously has a



	total germline	total functional	total expressed	total expressed with α exclusively	total expressed with δ exclusively	total expressed with both α or δ
cattle	371	215			26	
atlantic salmon	292	128	92	81	0	11
zebrafish	149	141	17	14	3	0
mouse	103	89	75	60	5	10
human	57	49	49	41	3	5
tetraodon	13	13	13	4	1	8

Fig. 6 Functional α/δV genes in different vertebrates. Graph compares the number of V α/δ segments found in the germline DNA to the number of those that are thought to be functional and to the number of V segments that have been found to be expressed with α, δ, or both. Cattle information can be found in http://www.imgt.org/IMGTrepertoire/Proteins/index.php#C and (Connelley et al. (2014), Atlantic salmon in (Yazawa et al. (2008b), zebrafish information can be in found in Supplemental Data 1 of this manuscript, mouse and human information can be found in http://www.imgt.org/IMGTrepertoire/Proteins/index.php#C, and *Tetraodon* in (Fischer et al. 2002)

greatly restricted IgHV repertoire but diversifies some ultralong CDR3H into an additional diverse microdomain (Wang et al. 2013). The species with the next highest number of TCR α/δ V segments after zebrafish is *S. salar* which possess 128 potentially functional out of a total of 292 V genes, the remainder being pseudogenes (Yazawa et al. 2008b). This salmon study goes on to compare the number of α/δ V genes in *S. salar* to the number in chicken (*Gallus gallus*) (70), human (*Homo sapiens*) (57), and mice (*Mus musculus*) (98). An exhaustive search for the number of V segments has not been conducted in other teleost species to make a valid comparison of these numbers. *T. nigroviridis* has only 13 Vα/δ segments (Fischer et al. 2002), but this species has a condensed genome, so this is not surprising. A total of 21 distinct V sequences have been found by cDNA sequencing in *I. punctatus* (Moulana et al. 2014) but this has not been traced back to the number of segments at the locus so this may not be an adequate representation of the number of V segments at the genomic level.

In comparison to *S. salar*, *D. rerio* appears to possess a substantially lower number of pseudogenes (9 versus 164). However, when comparing the number of pseudogenes in pufferfish (0), mice (9), and humans (11) (Yazawa et al. 2008b), it does appear that *S. salar* are the outlier with an unusually high number of pseudogenes.

The 149 α/δ V sequences that were located could be placed in 41 families based on 70 % nucleotide identity. Previous unpublished work by T. Ota placed 148 V genes into 87 families, but only 58 of these are in Genbank and no further information was available on their annotation or assignment to families (Danilova et al. 2004). Based on the percent identity matrix, there were some sequences that showed 70 % identity to some other sequences in a group but not to all. For the sake of clarity, we defined the family to include all

sequences that share 70 % nucleotide identity with at least one other sequence in that family. We chose 70 % as our cutoff based on the guidelines put forth by IMGT (http://www.imgt.org). Numerous papers have utilized various methods for classifying V sequences into families. Inconsistency in naming and classifying V sequences may potentially compound our ability to compare the number of families and characteristics of these families across various species of teleost fish. Yazawa et al. utilized 70 % nucleotide identity to classify the 292 V segments of *S. salar* into 62 families. But this is far from standardized. For example, one paper characterizing TCRδ and γ of *I. punctatus* utilized 75 % nucleotide identity in conjunction with a pairwise alignment to define their groups (Moulana et al. 2014). In *T. nigroviridis*, the 13 V segments have been placed into six families based on 75 % nucleotide identity (Fischer et al. 2002). Another study characterizing the TCRα chains in the rainbow trout (*Oncorhynchus mykiss*) used 75 % amino acid identity to classify 9 Vα segments into 6 groups and one pseudogene that they were unable to classify (Partula et al. 1995). In *T. rupribes*, there are 17 complete V sequences that were placed in 4 subfamilies based on sequence similarity, but it was unclear what percentage was used (Wang et al. 2001b). Further complicating the family analysis is the phenomenon that TCR α and δ typically share a common pool of V segments. The most common way to classify a V segment as either α or δ is based on expression data, however, it is not reasonable to say that a certain V is only an α or δ, just that it has been found to be expressed with one or the other TCR chain or both. For this reason, for example, the V segments of *T. nigroviridis* are classified as Vα/δ since they were identified at the genomic level and expression data was not obtained. In contrast, the V segments of *I. punctatus* and *S. salar* are classified as either α or δ or both because those sequences were obtained from cDNA expression data. This does not mean that the TCRα sequences are not expressed with TCRδ or vice versa, just that we do not have exhaustive data.

A phylogenetic analysis was performed using one select V sequence from each family of *D. rerio*, *S. salar*, and *I. punctatus*. In addition, selected outgroup sequences

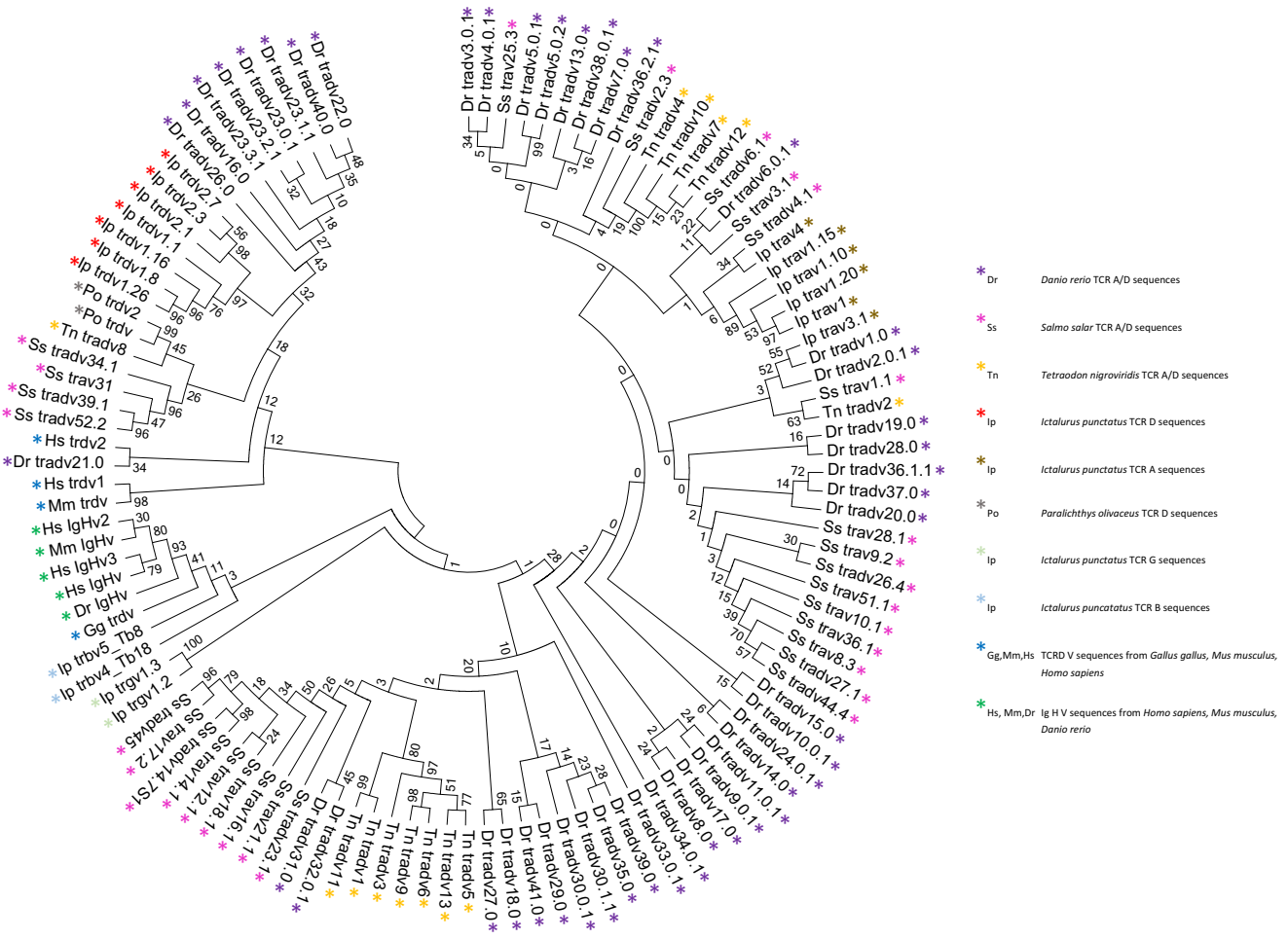


Fig. 7 Vertebrate TCRαδV phylogeny. Phylogenetic analysis of selected Vα/δ sequences from *D. rerio* as well as selected Vδ, Vα, Vγ, Vβ, and IgH V segments from various species of teleost fish and mammals for comparison. The neighbor joining tree was drawn using MEGA 6.0 and 1000 bootstrap replications. GenBank accession numbers for selected sequences are found in Supplemental Data 3

representing TCRα, β, and γ V sequences of other teleosts, as well as TCRδ from select mammalian species, and IgH V from select teleost and mammalian species were included for comparisons (Fig. 7). As expected, these sequences showed grouping first along receptor or isotypes lines. Additionally, these V sequences appeared mostly to cluster by species and not along V family lines. There are a few exceptions. The *D. rerio* sequences tradv3.0.1 and tradv4.0.1 group with the *S. salar* 25.3 sequence. *D. rerio* sequence tradv36.2.1 cluster with *S. salar* V2.3, *T. nigroviridis* V4, V7, V10, and V12. *D. rerio* sequence tradv6.0.1 groups with *S. salar* sequence V6.1 and V3.1. In regard to the subfamilies of *D. rerio* tradv23, tradv30, and tradv36; families tradv23 and tradv30 cluster together on the phylogenetic tree while tradv36.1.1 and tradv36.2.1 do not. Importantly, the bootstrap support of many of these bifurcations is low.

Expression data

Three of the identified Dδ sequences (Dδ1, Dδ4, Dδ6) were found to be expressed in the three transcripts analyzed. Both Jδ1 and Jδ2 were used in these transcripts as well, and Vδ sequences tradv23.2.1, tradv23.2.2, and tradv23.3.3 were found to be expressed. Additionally, N/P nucleotides were present at the coding joint of the V-D and D-J sequences. Through our attempts at amplifying the repertoire of sequences expressed in the zebrafish spleen, we were only able to obtain one unique clone. This single clone was further supported by PacBio sequencing which produced 440 identical sequences. No evidence was found for constant domain allelic polymorphism, which we have seen in some teleost fish TCR α and β including zebrafish (Criscitiello et al. 2004a, b; Kamper and McKinney 2002). Because of the same CDR3 sequence and two V(D)J coding joints, it is most likely that this sequence represents an individual clone amplified from only one of the zebrafish in the pool and not a homogenous population of γδ thymocytes with greatly restricted (fixed) TCRδ diversity that hone to the spleen or peripheral blood. This is supported by our qPCR data showing that relative upregulation of TCRδ, β and γ is low compared to alpha in spleen of immunized adult zebrafish, and this was the only TCRδ product we isolated. TCRδ expression data from the Atlantic salmon showed a higher diversity in their expression repertoire, utilizing 13 of the available Vα/δ segments to produce diverse TCRδ receptors (Yazawa et al. 2008b). However, this case of species-specific highly restricted diversity has been seen before in the axolotl (André et al. 2007) and in mouse mucosal epithelia (Itohara et al. 1990).

Conclusion

In this paper, we have provided an annotation for the complete TCR α/δ locus of the zebrafish. We found no evidence for IgHV in this locus, but did find the V array to be inverted 3′ of the αC as in other teleosts offering a possible explanation for the loss of IgH use in teleosts that appears in (at least some) sharks and amphibians. We had difficulty obtaining diverse canonical expression data for TCRδ from zebrafish spleen, suggesting that γδ T cell numbers may be especially low in the circulating periphery of normal adult zebrafish. Perhaps certain states of immunostimulation, particular tissues, or specific developmental stages will reveal more TCRδ expression and diversity. It is hoped that this study will represent an important first step in defining the curious expression of TCRδ in zebrafish and the annotation and phylogenetic analysis of the locus will provide a useful resource to investigators using this model.

References

Agard EA, Lewis SM (2000) Postcleavage sequence specificity in V(D)J recombination. Mol Cell Biol 20:5032–5040

Amemiya CT, Alfoldi J, Lee AP, Fan S, Philippe H, Maccallum I, Braasch I, Manousaki T, Schneider I, Rohner N, Organ C, Chalopin D, Smith JJ, Robinson M, Dorrington RA, Gerdol M, Aken B, Biscotti MA, Barucca M, Baurain D, Berlin AM, Blatch GL, Buonocore F, Burmester T, Campbell MS, Canapa A, Cannon JP, Christoffels A, De Moro G, Edkins AL, Fan L, Fausto AM, Feiner N, Forconi M, Gamieldien J, Gnerre S, Gnirke A, Goldstone JV, Haerty W, Hahn ME, Hesse U, Hoffmann S, Johnson J, Karchner SI, Kuraku S, Lara M, Levin JZ, Litman GW, Mauceli E, Miyake T, Mueller MG, Nelson DR, Nitsche A, Olmo E, Ota T, Pallavicini A, Panji S, Picone B, Ponting CP, Prohaska SJ, Przybylski D, Saha NR, Ravi V, Ribeiro FJ, Sauka-Spengler T, Scapigliati G, Searle SM, Sharpe T, Simakov O, Stadler PF, Stegeman JJ, Sumiyama K, Tabbaa D, Tafer H, Turner-Maier J, van Heusden P, White S, Williams L, Yandell M, Brinkmann H, Volff JN, Tabin CJ, Shubin N, Schartl M, Jaffe DB, Postlethwait JH, Venkatesh B, Di Palma F, Lander ES, Meyer A, Lindblad-Toh K (2013) The African coelacanth genome provides insights into tetrapod evolution. Nature 496:311–316

André S, Kerfourn F, Affaticati P, Guerci A, Ravassard P, Fellah JS (2007) Highly restricted diversity of TCR delta chains of the amphibian Mexican axolotl (Ambystomamexicanum) in peripheral tissues. Eur J Immunol 37:1621–1633

Bonneville M, O'Brien RL, Born WK (2010) γδ T cell effector functions: a blend of innate programming and acquired plasticity. Nat Rev Immunol 10:467–478

Buonocore F, Castro R, Randelli E, Lefranc MP, Six A, Kuhl H, Reinhardt R, Facchiano A, Boudinot P, Scapigliati G (2012) Diversity, molecular characterization and expression of T cell receptor gamma in a teleost fish, the sea bass (*Dicentrarchus labrax*, L). PLoS One 7, e47957

Chakrabarti S, Streisinger G, Singer F, Walker C (1983) Frequency of gamma-Ray induced specific locus and recessive lethal mutations in mature germ cells of the zebrafish, *Brachydanio rerio*. Genetics 103: 109–123

Chambers JM, Freeny A, Heiberger RM (1992) Analysis of variance; designed experiments. In Hastie JMCaTJ (ed.) Statistical models in S. Wadsworth & Brooks/Cole

Connelley TK, Degnan K, Longhi CW, Morrison WI (2014) Genomic analysis offers insights into the evolution of the bovine TRA/TRD locus. BMC Genomics 15:994

Criscitiello MF (2014) What the shark immune system can and cannot provide for the expanding design landscape of immunotherapy. Expert Opin Drug Disc 9:725–739

Criscitiello MF, Kamper SM, McKinney EC (2004a) Allelic polymorphism of TCRalpha chain constant domain genes in the bicolor damselfish. Dev Comp Immunol 28:781–792

Criscitiello MF, Wermenstam NE, Pilstrom L, McKinney EC (2004b) Allelic polymorphism of T-cell receptor constant domains is widespread in fishes. Immunogenetics 55(12):818–824

Criscitiello MF, Ohta Y, Saltis M, McKinney EC, Flajnik MF (2010) Evolutionarily conserved TCR binding sites, identification of T cells in primary lymphoid tissues, and surprising trans-rearrangements in nurse shark. J Immunol 184:6950–6960

Danilova N, Hohman VS, Sacher F, Ota T, Willett CE, Steiner LA (2004) T cells and the thymus in developing zebrafish. Dev Comp Immunol 28:755–767

Fischer C, Bouneau L, Ozouf-Costaz C, Crnogorac-Jurcevic T, Weissenbach J, Bernot A (2002) Conservation of the T-cell receptor α/δ linkage in the teleost fish *Tetraodon nigroviridis*. Genomics 79: 241–248

Haire RN, Rast JP, Litman RT, Litman GW (2000) Characterization of three isotypes of immunoglobulin light chains and T-cell antigen receptor α in zebrafish. Immunogenetics 51:915–923

Harper CCL (2011) The laboratory zebrafish. CRC, New York

Hein WR, Mackay CR (1991) Prominence of gamma delta T cells in the ruminant immune system. Immunol Today 12:30–34

Holderness J, Hedges JF, Ramstead A, Jutila MA (2013) Comparative biology of gammadelta T cell function in humans, mice, and domestic animals. Annu Rev Anim Biosci 1:99–124

Hope RM (2013) Rmisc: Rmisc: Ryan Miscellaneous

Itohara S, Farr AG, Lafaille JJ, Bonneville M, Takagaki Y, Haas W, Tonegawa S (1990) Homing of a γδ thymocyte subset with homogeneous T-cell receptors to mucosal epithelia. Nature 343:754–757

Iwanami N (2014) Zebrafish as a model for understanding the evolution of the vertebrate immune system and human primary immunodeficiency. Exp Hematol 42:697–706

Kamper SM, McKinney CE (2002) Polymorphism and evolution in the constant region of the T-cell receptor beta chain in an advanced teleost fish. Immunogenetics 53:1047–1054

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowits S, Duran C, Thierer T, Ashton B, Mentijies P, Drummond AA (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402–408

Moulana M, Taylor EB, Edholm ES, Quiniou SM, Wilson M, Bengten E (2014) Identification and characterization of TCRgamma and TCRdelta chains in channel catfish, Ictalurus punctatus. Immunogenetics 66:545–561

Murphy K (2012) Janeway's immunobiology. Garland Science, New York

Nam B-H, Hirono I, Aoki T (2003) The four TCR genes of teleost fish: the cDNA and genomic DNA analysis of Japanese flounder (Paralichthys olivaceus) TCR α-, β-, γ-, and δ-chains. J Immunol 170:3081–3090

Parra ZE, Miller RD (2012) Comparative analysis of the chicken TCR alpha/delta locus. Immunogenetics 64:641–645

Parra ZE, Baker ML, Schwarz RS, Deakin JE, Lindblad-Toh K, Miller RD (2007) A unique T cell receptor discovered in marsupials. Proc Natl Acad Sci U S A 104:9776–9781

Parra ZE, Baker ML, Hathaway J, Lopez AM, Trujillo J, Sharp A, Miller RD (2008) Comparative genomic analysis and evolution of the T cell receptor loci in the opossum *Monodelphis domestica*. BMC Genomics 9:111

Parra ZE, Ohta Y, Criscitiello MF, Flajnik MF, Miller RD (2010) The dynamic TCRdelta: TCRdelta chains in the amphibian Xenopus tropicalis utilize antibody-like V genes. Eur J Immunol 40(8): 2319–2329

Parra ZE, Lillie M, Miller RD (2012a) A model for the evolution of the mammalian T-cell Receptor alpha/delta and mu Loci based on evidence from the duckbill platypus. Mol Biol Evol 29:3205–3214

Parra ZE, Mitchell K, Dalloul RA, Miller RD (2012b) A second TCRdelta locus in Galliformes uses antibody-like V domains: insight into the evolution of TCRdelta and TCRmu genes in tetrapods. J Immunol 188:3912–3919

Partula S, De Guerra A, Fellah JS, Charlemagne J (1995) Structure and diversity of the T cell antigen receptor beta-chain in a teleost fish. J Immunol 155:699–706

Pascual V, Capra JD (1991) Human immunoglobulin heavy-chain variable region genes: organization, polymorphism, and expression. Adv Immunol 49:1–74

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rast JP, Litman GW (1994) T-cell receptor gene homologs are present in the most primitive jawed vertebrates. Proc Natl Acad Sci U S A 91: 9248–9252

Rast JP, Anderson MK, Strong SJ, Luer C, Litman RT, Litman GW (1997) α, β, γ, and δ T cell antigen receptor genes arose early in vertebrate phylogeny. Immunity 6:1–11

Schorpp M, Bialecki M, Diekhoff D, Walderich B, Odenthal J, Maischein HM, Zapata AG, Boehm T (2006) Conserved functions of Ikaros in vertebrate lymphocyte development: genetic evidence for distinct larval and adult phases of T cell development and two lineages of B cells in zebrafish. J Immunol 177:2463–2476

Shang N, Sun XF, Hu W, Wang YP, Guo QL (2008) Molecular cloning and characterization of common carp (*Cyprinus carpio* L.) TCRgamma and CD3gamma/delta chains. Fish Shellfish Immunol 24:412–425

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Tian JY, Qi ZT, Wu N, Chang MX, Nie P (2014) Complementary DNA sequences of the constant regions of T-cell antigen receptors alpha, beta and gamma in mandarin fish, *Siniperca chuatsi* Basilewsky, and their transcriptional changes after stimulation with *Flavobacterium columnare*. J Fish Dis 37:89–101

Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, Hoon S, Gangu V, Roy SW, Irimia M, Korzh V, Kondrychyn I, Lim ZW, Tay BH, Tohari S, Kong KW, Ho S, Lorente-Galdos B, Quilez J, Marques-Bonet T, Raney BJ, Ingham PW, Tay A, Hillier LW, Minx P, Boehm T, Wilson RK, Brenner S, Warren WC (2014) Elephant shark genome provides unique insights into gnathostome evolution. Nature 505:174–179

Walker C, Streisinger G (1983) Induction of mutations by gamma-Rays in pregonial germ cells of zebrafish embryos. Genetics 103:125–136

Wang K, Gan L, Kunisada T, Lee I, Yamagishi H, Hood L (2001) Characterization of the Japanese pufferfish (*Takifugu rubripes*) T-cell receptor alpha locus reveals a unique genomic organization. Immunogenetics 53:31–42

Wang X, Parra ZE, Miller RD (2011) Platypus TCRmu provides insight into the origins and evolution of a uniquely mammalian TCR locus. J Immunol 187:5246–5254

Wang F, Ekiert DC, Ahmad I, Yu W, Zhang Y, Bazirgan O, Torkamani A, Raudsepp T, Mwangi W, Criscitiello MF, Wilson IA, Schultz PG, Smider VV (2013) Reshaping antibody diversity. Cell 153:1379–1393

Weir H, Chen PL, Deiss TC, Jacobs N, Nabity MB, Young MH, Criscitiello MF (2015) DNP-KLH Yields changes in leukocyte populations and immunoglobulin isotype use with different immunization routes in zebrafish. Front Immunol 6, 606

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Yazawa R, Cooper GA, Beetz-Sargent M, Robb A, McKinnel L, Davidson WS, Koop BF (2008a) Functional adaptive diversity of the Atlantic salmon T-cell receptor gamma locus. Mol Immunol 45: 2150–2157

Yazawa R, Cooper GA, Hunt P, Beetz-Sargent M, Robb A, Conrad M, McKinnel L, So S, Jantzen S, Phillips RB, Davidson WS, Koop BF (2008b) Striking antigen recognition diversity in the Atlantic salmon T-cell receptor alpha/delta locus. Dev Comp Immunol 32:204–212