GENOMIC ORGANIZATION OF ZEBRAFISH (*DANIO RERIO*) T CELL RECEPTOR α/δ LOCUS AND ANALYSIS OF EXPRESSED PRODUCTS

A Thesis

by

STACIE LYNN SEELYE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Michael F Criscitiello
Committee Members,	Vincent Gresham
	Robert Alaniz
Head of Department,	Roger Smith III

May 2016

Major Subject: Biomedical Sciences

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.

DEDICATION

This work is dedicated to someone who was my resident mate, lab mate, and all around wonderful person. Dr. Ashley Peterson, you were taken from us way too soon. Your absence makes the lab animal community, scientific community, and the world a lesser place.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Criscitiello, and my committee members, Dr. Alaniz and Dr. Gresham, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues in the CMP department faculty for making my time at Texas A&M University a great experience. I also want to thank my Comparative Medicine Residency mentors, Dr. Elliott, Dr. Browder, and Dr. Gresham, for their mentorship in this great learning opportunity.

Finally, thanks to my residency mates and family. Dr Klipsic, Dr. McCormack, Dr. Milligan, Dr. Smith, and Dr. Richert. Without your support and friendship I would never have made it through these past three years. I cannot say thank you enough.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION*

Zebrafish (*Danio rerio*) continues to increase in popularity a vertebrate model species (Iwanami 2014). Zebrafish entered the forefront of animal research in the 1980's 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

---

* Document reprinted from "Genomic organization of zebrafish (*Danio rerio*) T cell receptor α/δ locus and analysis of expressed products" by Stacie Seelye, 2016. *Immunogenetics*, 68(5), 365-379, Creative Commons Copyright 2016 by Stacie Seelye.

(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 (*Takifugu rubripes*) (Wang et al. 2001a) the green pufferfish (*Tetraodon 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 mandarinfish (*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, 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 trans-rearrangement. 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.

2. METHODS

2.1 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 2013).

2.2 V region search

Using a previously annotated TCRδV sequence from *P. olivaceus* (accession #AB076071.1) as bait sequence, a tBLASTn search was performed. Genomic sequences were downloaded into the Geneious version 7 (Kearse 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 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.

2.3 Locus annotation

Scaffold version 10 of chromosome 2 was downloaded from NCBI and imported into Geneious version 7. All 149 V 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.

2.4 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 twelve 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 Table 1. Primary PCR was performed using 1ul 10uM dNTP, 10ul 5X Phusion buffer (New England Biolabs, Ipswich MA), 2.5ul 10uM GENERACER 5' primer, 2.5ul 10uM reverse primer MFC527, 2ul 50ng/ul template, 0.5ul 2U/ul High Fidelity Phusion DNA polymerase, and PCR quality water to total volume of 50ul. 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 30s, 3) annealing and elongation 72°C for 30 sec repeat steps 2-3 30 times 4) final elongation 72°C for 5 min. Secondary PCR was performed using 3ul 25mM $MgCl_2$, 1ul 10mM dNTP, 1ul 10uM GENERACER 5' nested primer, 1ul 10uM reverse primer MFC525, 3ul of template from primary PCR, 5ul 10X Buffer, 0.25ul 5U/ul HotFire DNA Polymerase (Solis BioDyne, Tartu, Estonia), and PCR quality water to total volume of 50ul. Thermocycler protocol was as follows: 1) initial denaturation 95°C for 15 min 2) denaturation 95°C for 30 sec 3) annealing 63°C for 30 sec 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 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 1ul 10uM MFC535, MFC536, or MFC537 primers and 1ul 10uM reverse primer MFC525. The amount of template used was 3ul of 50ng/ul 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 sec 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 25mM $MgCL_2$ was increased to 4ul 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 sec and the number of cycles was increased to 35. A positive control using 1ul 10uM 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 25mM $MgCl_2$ was returned to 3ul and the amount of the template was varied so that either 2, 4, 6, or 8 ng/ul 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 sec and 35 cycles were performed. Secondary PCR was performed using the reverse primer MFC525 and 1ul 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.

2.5 Quantitative real-time PCR

Quantitative PCR was performed with 50ng of random hexamer primed cDNA generated with SuperScript III from RNA samples from pooled adult zebrafish spleen immunized with DNP-KLH (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 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 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).

2.6 *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 (Table 1). cDNA was initially denatured at 98°C for two minutes then amplified with Phusion (NEB) high fidelity polymerase for 34 cycles consisting of two steps: 98° C for ten seconds and 72°C for 40 seconds, ending with a final elongation at 72°C for 5 minutes. 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 (1-3kb) 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).

3. 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 (Figure 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 attached EXCEL file supplemental table 1.

3.1 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-161L11 (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 (Figure B-1). The multiple sequence alignment used to create this tree is shown in Figure B-2. No additional potential Cδ sequence locations were found in the zebrafish genome.

3.2 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 (Attached EXCEL file supplemental table 2) created from a Clustal W (Tamura 2013) multiple sequence alignment (Figure B-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 2013) multiple alignment using the Mega 6.0 software (Tamura 2013). This tree (Figure 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 Figure B-3, the amino acid sequence for these V regions is found in Figure B-4.

3.3 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, Steiner, *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 hepatmer, 12/23 spacer and nonamer are shown in Figure B-5.

3.4 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 Figure 3. An alignment of all 8 of the original clones obtained by plasmid transformation and Sanger sequencing is shown in Figure B-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 are shown in Figure B-7.

3.5 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 the four TCR chain transcripts (Figure 4). Not only were both transcripts required for the γδ TCR heterodimer very low compared to TCRα, so was the β chain of the αβ receptor. At least in the spleen of immunized zebrafish the levels of TCRα 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 Figure B-8.

4. DISCUSSION

In testing the hypothesis that all jawed vertebrate classes have integrated immunoglobulin heavy chain V gene segments in 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 3 Dδ, 2Jδ, 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 (Figure 4). Here we completed the α/δ locus annotation, finding no evidence for IgH V segments. We attempted to analyze the expressed TCRδ repertoire in the fish to rule out trans-rearrangements 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).

4.1 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 *T. 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 (Figure 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 trick 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 (Figure 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 (Figure 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 of 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/δVδJ-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 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).

4.2 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 traditional 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 (*Cyprinus 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 Figure B-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 *Danio rerio* TCRδ C does contain the highly conserved cysteine residues.

4.3 V region analysis

Zebrafish appear to have the second largest number of potentially functional α/δ V genes of any species previously studied (Figure 6). The species with the greatest number of V sequences is the cow (Connelley et al. 2014), which curiously has a 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 may not be an adequate representation of the number of V segments at the genomic level.

In comparison to *S. salar*, *D. rerio* appear 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 6 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. nigriviridis* 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 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 (Figure 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. nigriviridis* V4, V7, V10, and V12. *D. rerio* sequence tradv6.0.1 groups with *S. salar* sequence V6.1 and V3.1. In regards 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.

4.4 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. 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 TCRδ, β and γ are expressed at extremely low levels in spleen of untreated 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).

5. 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 20:5032-40

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 496:311-6

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. European journal of immunology 37*:*1621-1633

Bonneville M, O'Brien RL, Born WK (2010) γδ T cell effector functions: a blend of innate programming and acquired plasticity. Nature Reviews Immunology 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*:*109-23

Chambers JM, Freeny, A and Heiberger, R. M. (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 Opinion on Drug Discovery 9*:*725-739

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*:*6950-60

Danilova N, Hohman VS, Sacher F, Ota T, Willett CE, Steiner LA (2004) T cells and the thymus in developing zebrafish. Developmental & Comparative Immunology 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 Press, New York

Hein WR, Mackay CR (1991) Prominence of gamma delta T cells in the ruminant immune system. Immunol Today 12*:*30-4

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 y8 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

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., And A. Drummond (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−ΔΔCT 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*:*545-61

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. The Journal of Immunology 170*:*3081-3090

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, 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 104:9776-81

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. Molecular Biology and Evolution 29:3205-3214

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

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 188:3912-9

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

Parra ZE, Ohta Y, Criscitiello MF, Flajnik MF, Miller RD (2010b) The dynamic TCRδ: TCRδ chains in the amphibian Xenopus tropicalis utilize antibody-like V genes. European journal of immunology 40:2319-2329

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. The Journal of Immunology 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, Austria

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

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

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*:*2463-76

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*:*412-25

Tamura K, Stecher, G., Peterson, D., Filipski, and S. Kumar (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 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*:*174-9

Walker C, Streisinger G (1983) Induction of Mutations by gamma-Rays in Pregonial Germ Cells of Zebrafish Embryos. Genetics 103*:*125-36

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*:*1379-93

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

Wang K, Gan L, Kunisada T, Lee I, Yamagishi H, Hood L (2001b) Characterization of the Japanese pufferfish (Takifugu rubripes) T-cell receptor α 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*:*5246-54

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. Frontiers in Immunology 6*:*Article 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. Molecular immunology 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*:*204-12

APPENDIX A

FIGURES

Figure 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 circles represent V sequences expressed with TCR δ constant region. Blue circles 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.

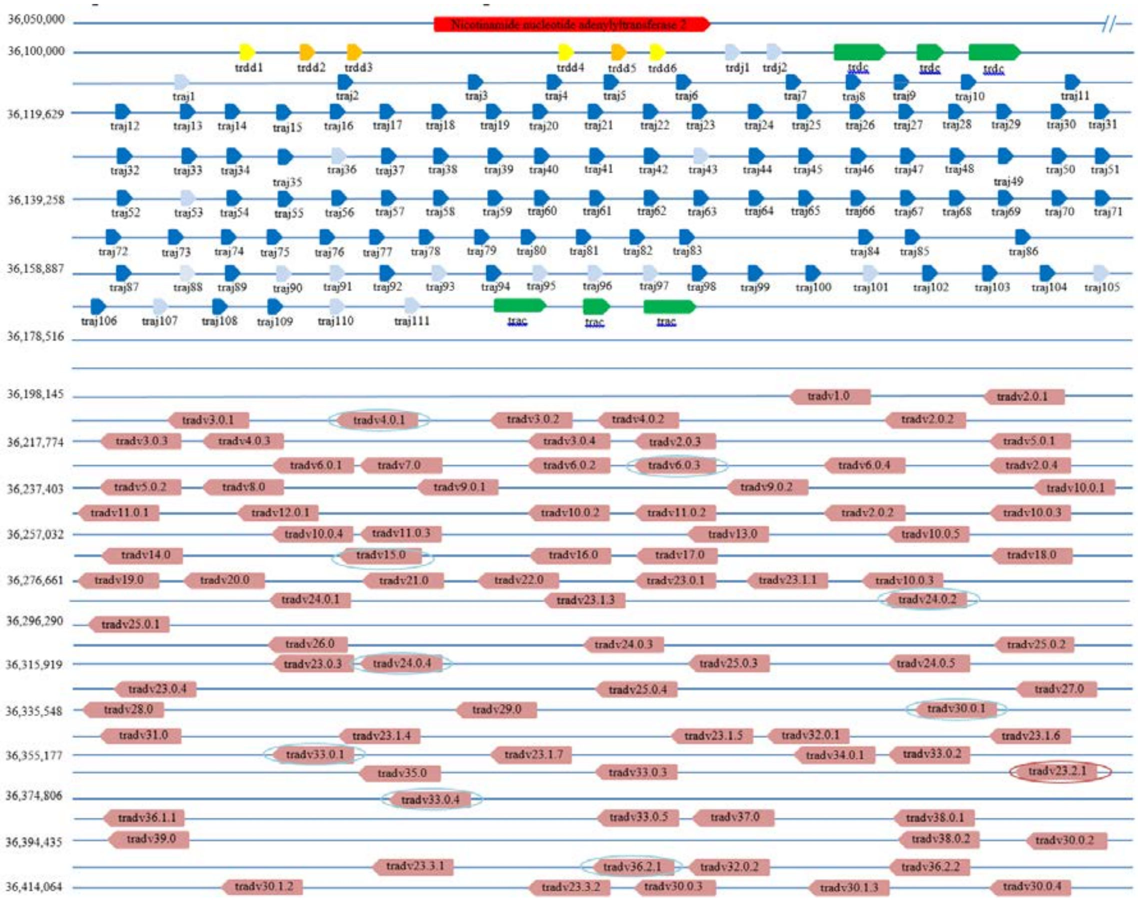


Figure 1 continued

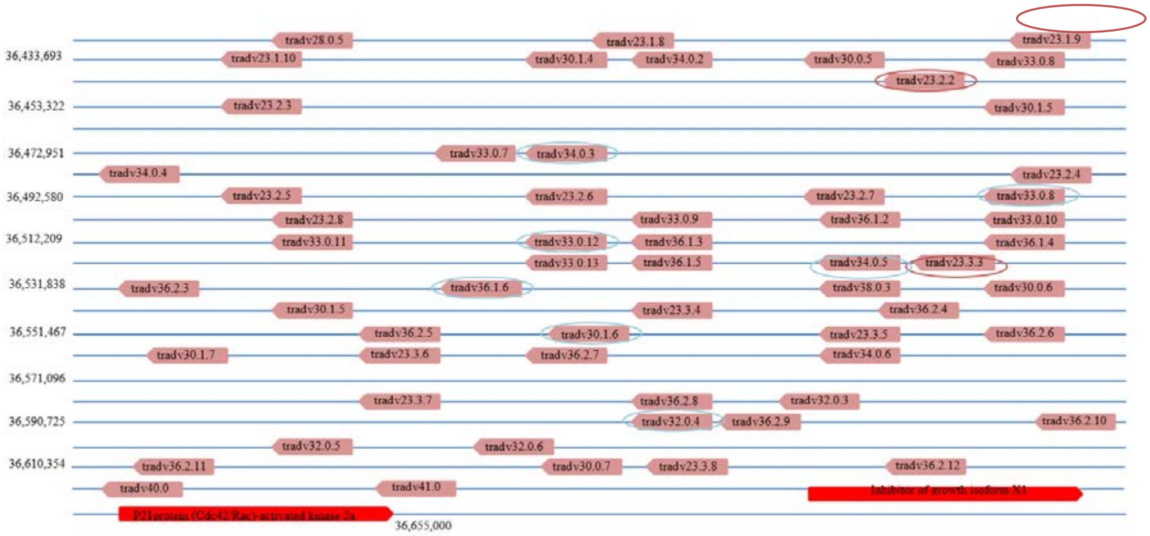


Figure 2. Phylogenetic analysis of all genomic Vα/δ sequences from *D. rerio*. The neighbor joining tree was drawn using MEGA 6.0 and 1,000 bootstrap replications. Colored asterisks represent different phylogentic families. (Some colors were re-used due to limitations in number of colors.) Triangles 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 attached EXCEL file supplemental table 1.

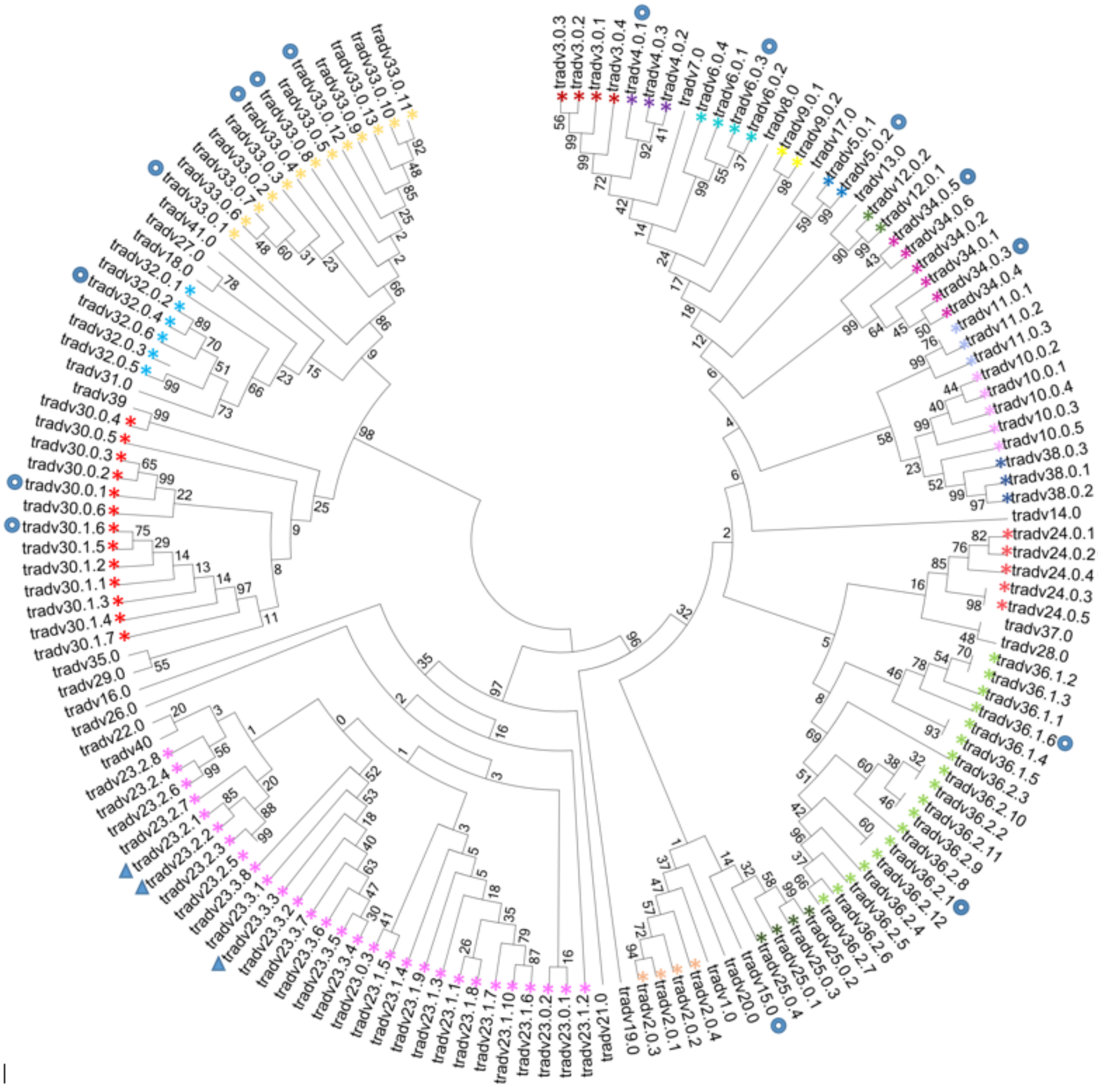


Figure 3. Alignment of all three known expressed zebrafish Vδ sequences. Differences in V and J sequences are highlighted in yellow. Differences in D and N/P nucleotides were not highlighted. Conserved hallmark sequences were highlighted in grey. Blue, green and magenta highlighting marks the nucleotides 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).

```
<                        Leader                    ><V sequence 23.2.2
     M  E  K        Q  L  M  L  I  L  I  L  T  P  G  V  M  T  A  D  Q
C1  ATGGAGAAA------CAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
<                        Leader                    ><V sequence 23.2.1
     M  E  K        Q  L  M  L  I  L  I  L  T  P  G  V  M  T  A  D  Q
B35 ATGGAGAAA------CAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
<                        Leader                    ><V sequence 23.3.3
     M  E  R  G  F  L  L  V  V  V  V  I  M  A  T  G  L  V  F  G  D  N
EST ATGGAAAGAGGCTTTCTACTCGTTGTTGTTGTTATTATGGCGACAGGTTTGGTATTTGGGGACAAT


     I  S  P  N  K  E  A  L  T  V  K  E  E  E  T  V  T  F  S  C  S  Y
C1  ATTAGCCCAAATAAAGAAGCTCTTACTGTAAAGGAAGAGGAGACAGTGACCTTCAGTTGCTCATAT
     I  R  P  N  K  E  A  F  T  V  K  E  E  E  T  V  T  F  S  C  S  Y
B35 ATTAGGCCAAATAAAGAAGCTTTTACTGTAAAGGAAGAGGAGACTGTGACCTTCAGCTGCTCATAT
     I  E  P  E  E  K  D  V  M  T  K  E  R  E  A  V  K  L  A  C  S  Y
EST ATTGAGCCAGAGGAGAAAGATGTTATGACAAAAGAAAGAGAAGCTGTCAAGCTGGCCTGCTCATAC


     D  T  S  S  S  Y  V  R  L  Y  W  Y  R  Q  Y  L  N  G  E  P  Q  Y
C1  GATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAGAACCTCAGTAT
     E  T  S  S  S  Y  V  W  L  Y  W  Y  R  Q  H  L  N  G  E  P  Q  Y
B35 GAAACAAGCAGCAGTTATGTTTGGCTTTACTGGTACAGACAGCATCTTAACGGAGAACCTCAGTAT
     S  T  T  N  N  R  V  R  L  Y  W  Y  R  Q  N  P  N  A  E  L  L  L
EST AGTACAACCAACAATAGAGTTCGGCTTTATTGGTACAGACAGAATCCAAATGCAGAACTTTTGCTT


     L  L  F  K  A  A  R  S  S  S  G  G  G  R  P  D  N  P  R  F  K  S
C1  TTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATAATCCTCGTTTTAAGTCG
     L  I  F  K  P  A  K  S  A  S  V  S  G  D  P  V  D  R  R  F  Q  S
B35 TTAATCTTTAAACCTGCAAAATCAGCTTCTGTAAGTGGAGATCCAGTTGATCGTAGGTTTCAGTCG
     L  T  Y  K  G  A  R  S  L  S  A     K  H  S  S  N  D  R  F  Q  S
EST TTAACATACAAAGGTGCTCGATCTCTGAGTGC---TAAACACTCCTCTAATGATCGGTTTCAATCC
```

Figure 3 continued

```
     T  T  S  D  S  S  T  E  L  T  I  S  G  V  T  L  S  D  S  A  L
C1  ACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAACTCTGTCAGATTCAGCTCTC
     S  T  S  D  S  S  T  E  L  T  I  S  G  A  T  L  S  D  S  A  L
B35 AGCACATCAGACTCATCCACTGAACTCACTATTAGCGGTGCAACTCTGTCAGATTCAGCTCTC
     T  T  S  D  S  S  T  E  L  T  I  T  D  V  R  L  S  D  S  A  L
EST ACAACATCAGACTCATCCACTGAACTCACTATTACTGATGTGCGTCTGTCAGATTCAGCTCTC

      V  sequence        >  <  N/P  ><   D1   > <    D4   >
                                     TACTTGGGAC GATTGGGGTAC
     Y  Y  C  A  L  R  V  G       E                      Y
C1  TATTATTGTGCTCTAAGAGTAGGAG       A                    GTAC
     Y  Y  C  A  L  R  V  G            T  W  D   D  W  G
B35 TATTATTGTGCTCTTAGAGTTGG           TACTTGGGAC GATTGGGGTA
     Y  Y  C  A  L  R  V  G                           V G  V
EST TATTATTGTGCTCTAAGAGTAGGAG                           TTGGGGTAC

  <   D6   > <  N/P  >
  TCTGGACTAC
       D  Y
C1    GACTAC
    T  G      H  A
B35 CTGGAC    ATGCC
              R  V
EST           GGGTT

    <                            J2
        A  T  D  P  L  T  F  G  K  P  I  T  L  T  V  I  P
C1  ---GCTACTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCA-Cδ
    <                            J1
              D  P  L  T  F  G  A  P  I  R  L  T  V  N  P
B35 ---------GACCCACTAACTTTCGGAGCTCCCATCCGTCTCACGGTCAATCCA-Cδ
    <                            J2
     S  A  T  D  P  L  T  F  G  K  P  I  T  L  T  V  I  EST
EST TCAGCTACTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCA-Cδ
```

Figure 4. Splenic TCR-chain expression preference. qPCR for each TCR chain from zebrafish splenic cDNA. Center lines represent mean 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.

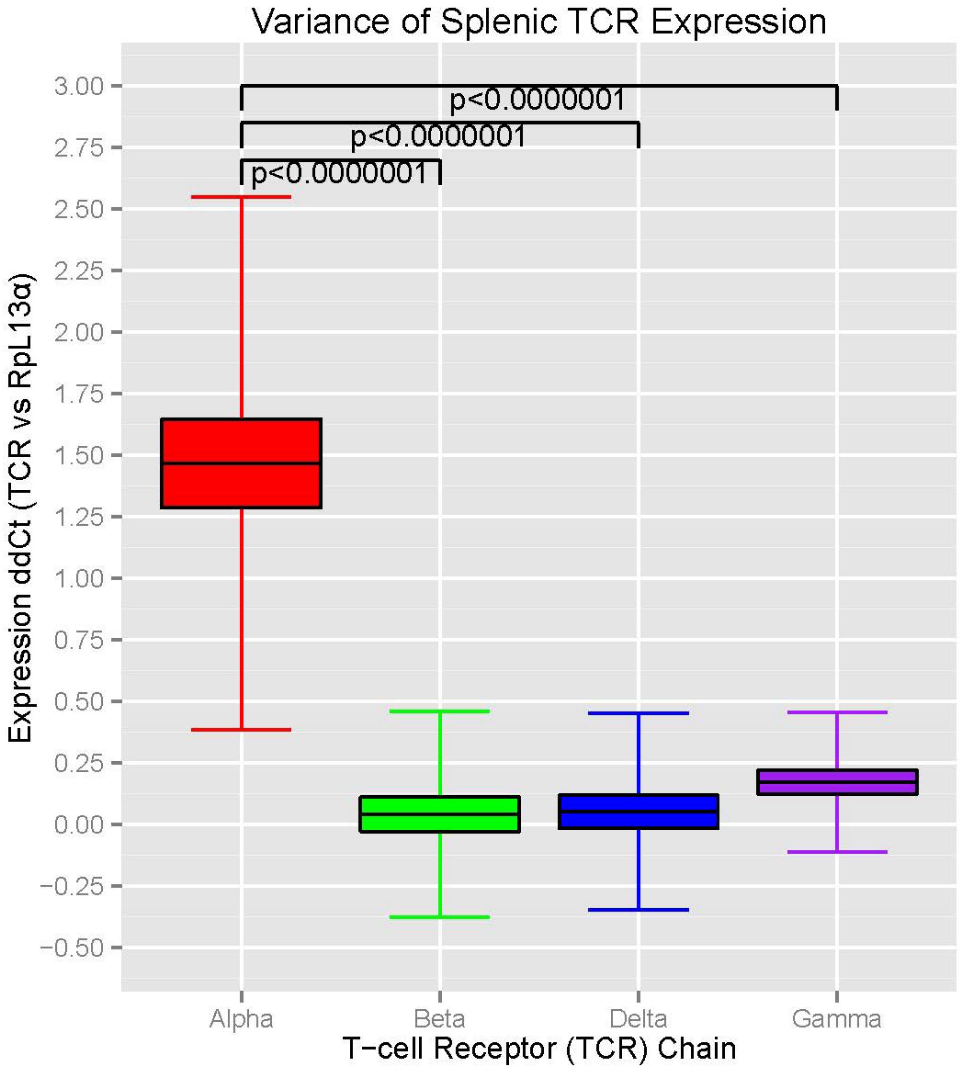


Figure 5. TCR α/δ locus organization has rearrangement consequences and evolutionary insights. A. The zebrafish locus can rearrange deletionally for TCRδ or inversionally for TCRα, yet both of these processes leave open the possibility of subsequent rearrangements for making the other TCR chain. **B.** 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. Grey 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.

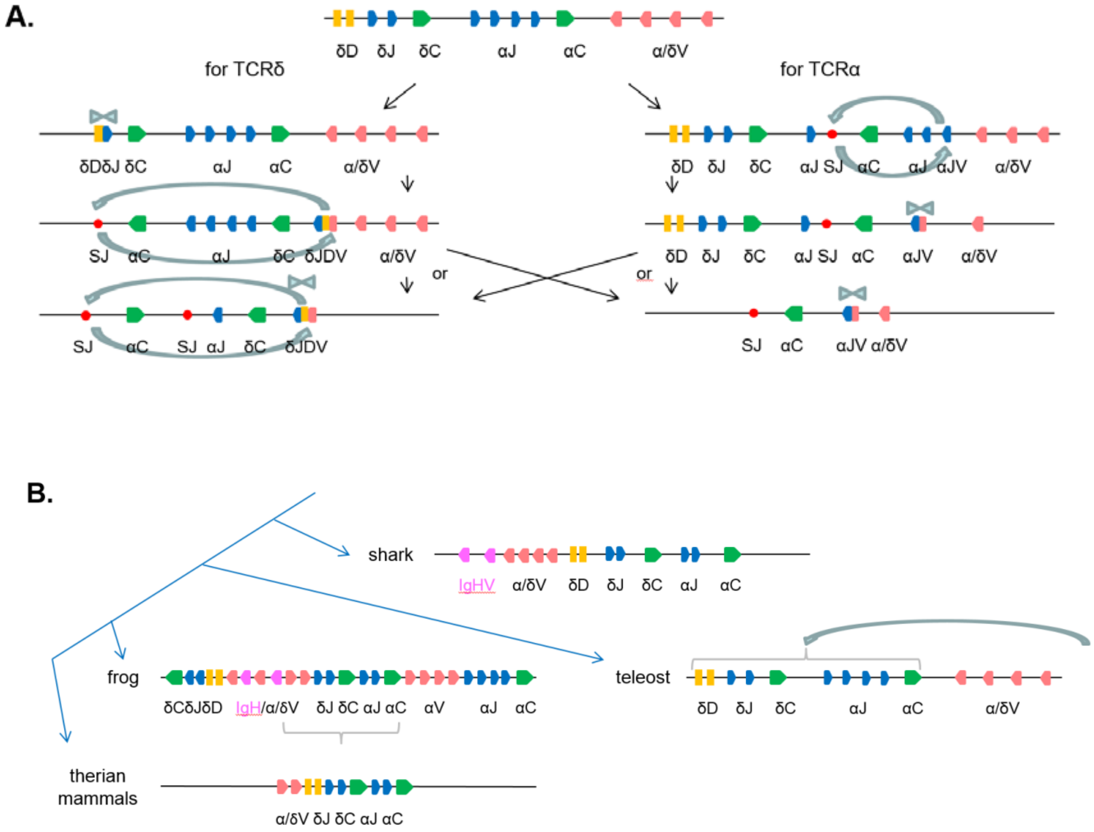


Figure 6. Functional α/δV genes in different vertebrates. Graph compares the number of V α/δ segments found in the germline DNA to the number of those that are thought to be functional and to the number of V segments that have been found to be expressed with α, δ, or both. Cattle information can be found 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 attached EXCEL file supplemental table 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).

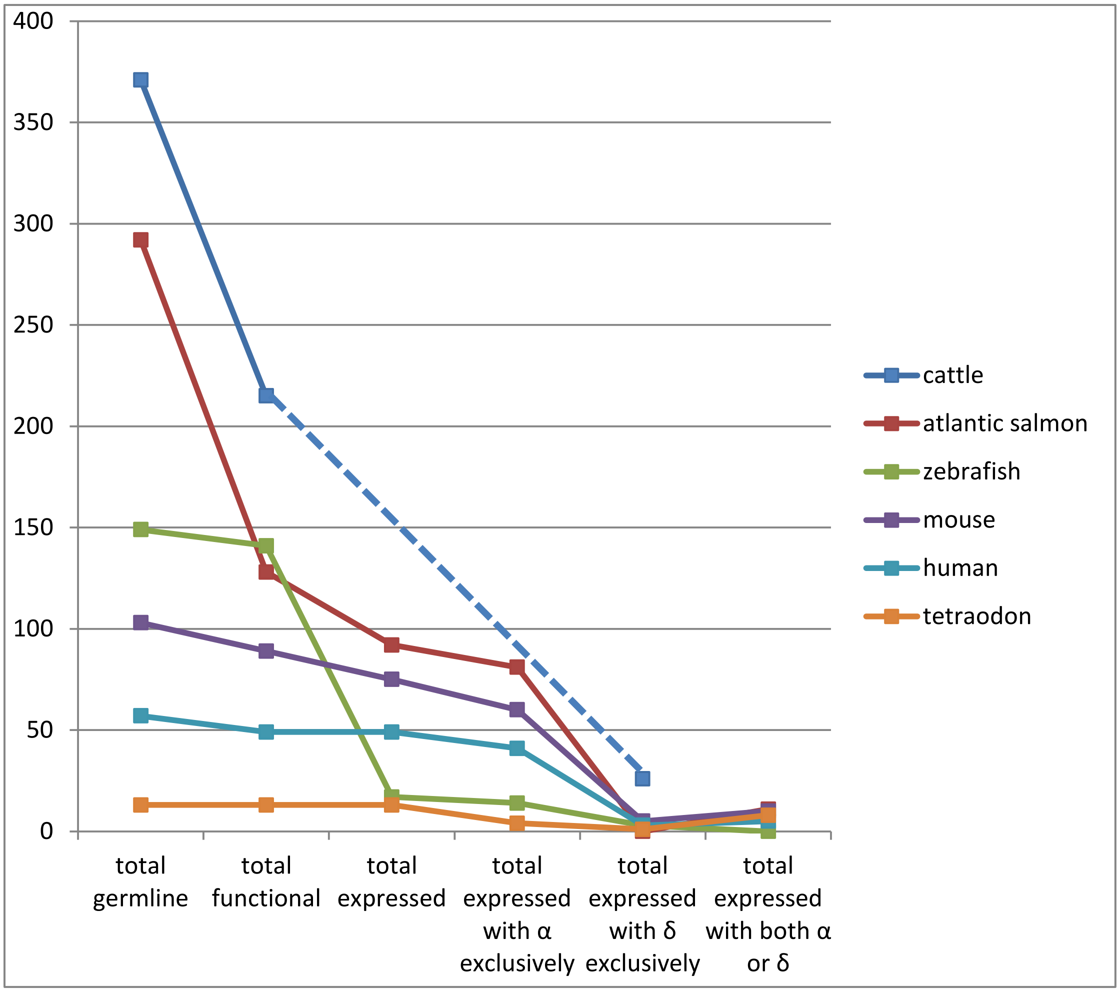


Figure 6 continued

	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

Figure 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 1,000 bootstrap replications. GenBank accession numbers for selected sequences are foud in table 2.

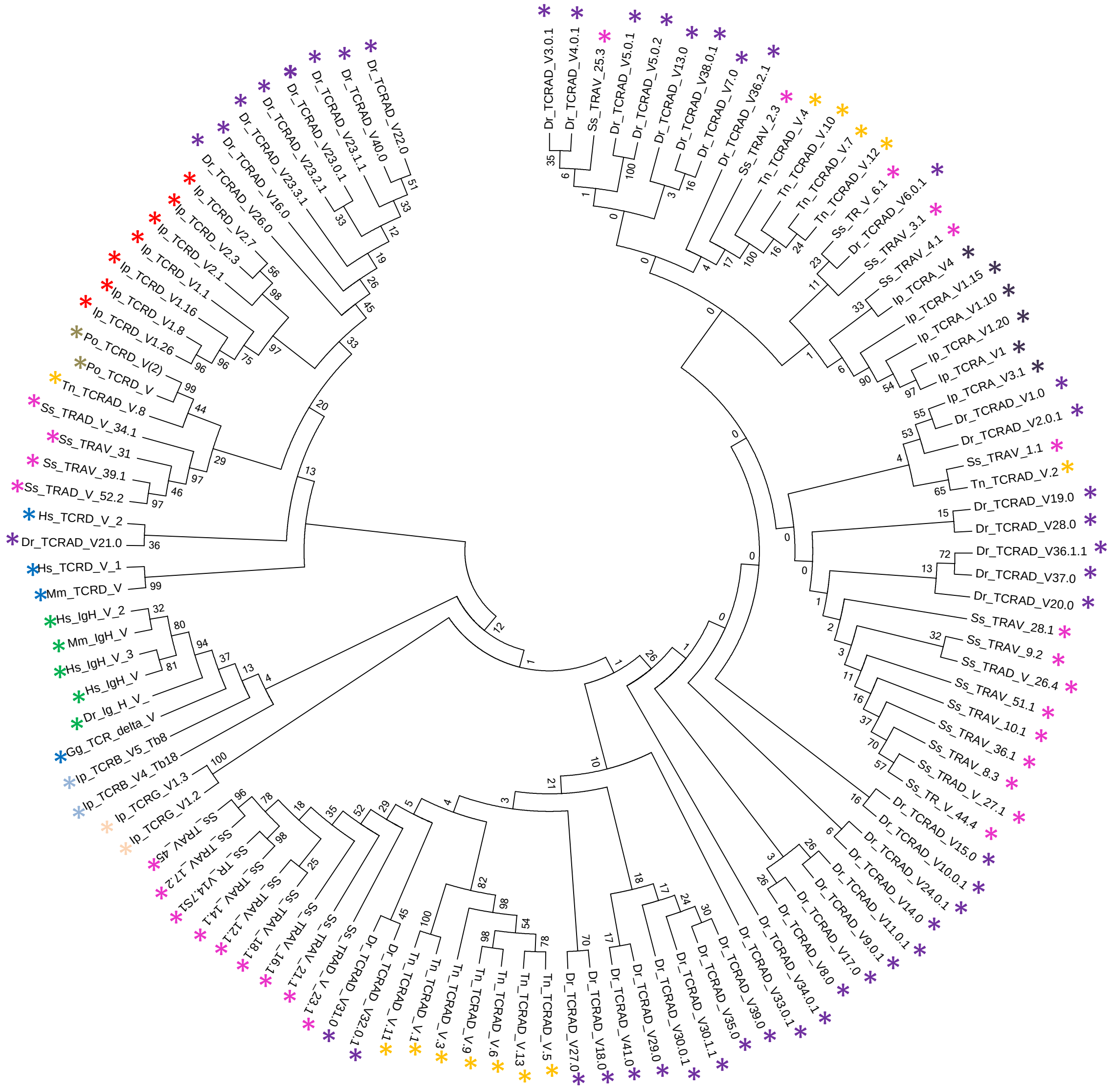


Figure 7 continued

Legend

*Dr	*Danio rerio* TCR A/D sequences
*Ss	*Salmo salar* TCR A/D sequences
*Tn	*Tetraodon nigroviridis* TCR A/D sequences
*Ip	*Ictalurus punctatus* TCR D sequences
*Ip	*Ictalurus punctatus* TCR A sequences
*Po	*Paralichthys olivaceus* TCR D sequences
*Ip	*Ictalurus punctatus* TCR G sequences
*Ip	*Ictalurus puncatatus* TCR B sequences
*Gg,Mm,Hs	TCRD V sequences from *Gallus gallus*, *Mus musculus, Homo sapiens*
*Hs, Mm,Dr	Ig H V sequences from *Homo sapiens*, *Mus musculus, Danio rerio*

APPENDIX B

SUPPLEMENTAL FIGURES

Figure B-1 Phylogenetic tree from multiple sequence alignment of selected C sequences. Created in Mega 6.0 using the Muscle and Neighbor joining tree building methodology. Ambystoma mexicanum TCR delta (AY029365.1); Callorhinchus milii TCR delta (NW_006897485); Cyprinus carpio TCR delta (AB541473.2/BAI83002); Danio rerio TCR alpha (AF246178.1); Danio rerio TCR beta(NC_007128.6); Danio rerio TCR gamma (AY973943.1); Dicentratchus labrax TCR gamma (EU853841); Gallus gallus TCR delta (AF175433); Ginglymostoma cirratum TCR delta(JF507709.1); Heterodontus francisci TCR delta (U22673); Homo sapiens TCR delta (A31326); Ictalurus punctatus TCR delta (HQ913590); Ictalurus punctatus TCR gamma (DQ435303); Mus musculus TCR delta (AAA51274); Ovis aries TCR delta (AJ290095.1); Paralichthys olivaceus TCR alpha (AB053227.1); Paralichthys olivaceus TCR beta (AB081559.1); Paralichthys olivaceus TCR delta (AU050806); Paralichthys olivaceus TCR gamma (AB081560.1); Salmo salar TCR delta (EF467299); Siniperca chuatsi TCR beta (EF596786.1); Takifugu rubripes TCR delta (AB222469); Tetraodon nigroviridis TCR delta (AJ410421.1); Xenopus tropicalis TCR delta (GQ262017.1)

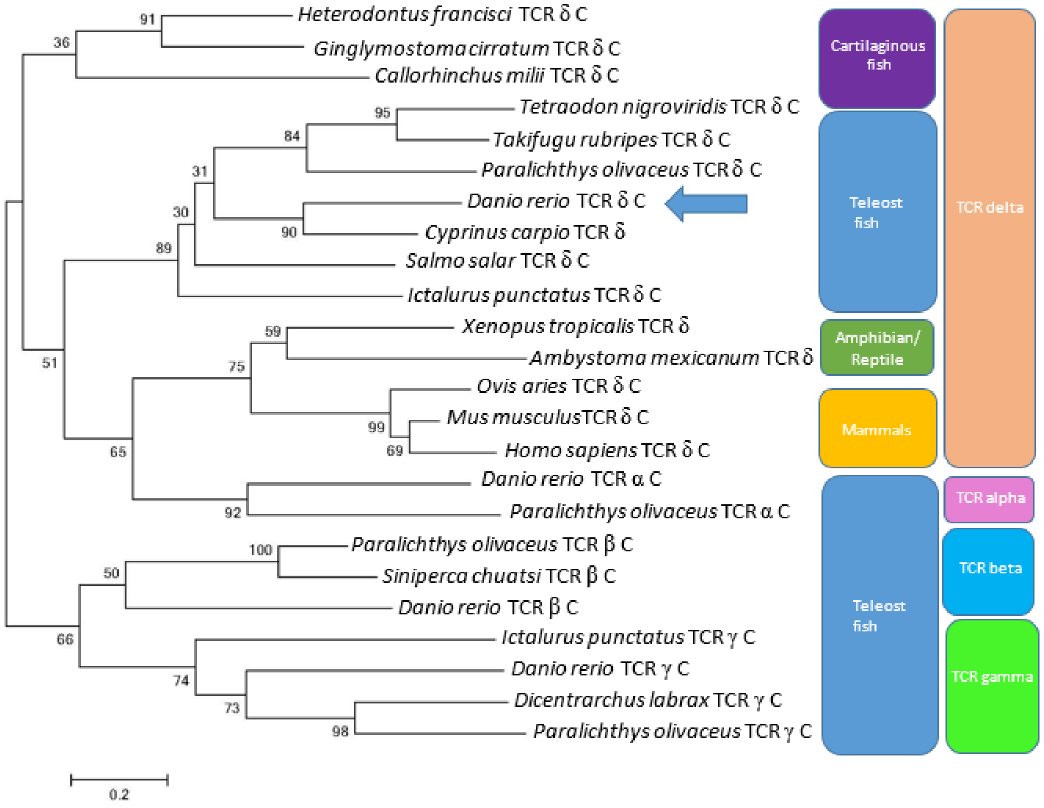


Figure B-2 Multiple sequence alignment of selected C sequences. Created in Mega 6.0 using the Clustal W methodology. *Ambystoma mexicanum* TCR delta (AY029365.1); *Callorhinchus milii* TCR delta (NW_006897485); *Cyprinus carpio* TCR delta (AB541473.2/BAI83002); *Danio rerio* TCR alpha (AF246178.1); *Danio rerio* TCR beta(NC_007128.6); *Danio rerio* TCR gamma (AY973943.1); *Dicentratchus labrax* TCR gamma (EU853841); *Gallus gallus* TCR delta (AF175433); *Ginglymostoma cirratum* TCR delta(JF507709.1); *Heterodontus francisci* TCR delta (U22673); *Homo sapiens* TCR delta (A31326); *Ictalurus punctatus* TCR delta (HQ913590); *Ictalurus punctatus* TCR gamma (DQ435303); *Mus musculus* TCR delta (AAA51274); *Ovis aries* TCR delta (AJ290095.1); *Paralichthys olivaceus* TCR alpha (AB053227.1); *Paralichthys olivaceus* TCR beta (AB081559.1); *Paralichthys olivaceus* TCR delta (AU050806); *Paralichthys olivaceus* TCR gamma (AB081560.1); *Salmo salar* TCR delta (EF467299); *Siniperca chuatsi* TCR beta (EF596786.1); *Takifugu rubripes* TCR delta (AB222469); *Tetraodon nigroviridis* TCR delta (AJ410421.1); *Xenopus tropicalis* TCR delta (GQ262017.1).

```
          <              Ig Domain         *
Am TCRD : THVQPS--VFVLRS-------------DDAAPCLVKDFSLKQ-AEIFM------NSSSVSETGIETSAPVLSKSGKYSA------
Cm TCRD : KPVNPK--VTMFYPAIKS---NP--PASDPAALCLADEYIPKE-LSILV--HD-INDMDLTKSATTSSSIAIIDHKLYHK------
Cc TCRD : -ETAPV--LSILSP-----------FKGDGPDICLAAGFFPQE-KNMTL-----TLEGNEPKSLTSSNAPLFSTTKTYYY------
Dr TCRA : TEVKPN--IYQVGN------------------SCLATDFTKHN--EFNI----------TDSPFLETSAVRYSGQSYYSS------
Dr TCRB : NVTEPE--VKILPPSPKELCSQNTNDKKLSTLVCVATGFYPDH-VSMSW-----KVNDQERKDNISTDHIAIQDKTTLMYNISSRM
Dr TCRD : TVNSPPAFLSVLSP-----------IKGHGSDICVAAGFFPQQ-KTMIL-----TSEDGNTVNQETSNAVLSLSSKNYYY------
Dr TCRG : SVVTPK--LSGYLNK----------ANEKPAALCQAKDMFPDL-VSFKW---E-KKSSSGGWTEVSNDQIVEHSHNEDPV------
Dl TCRG : QIVKPV--VSVYPAASRA---H---LEGKISLLCVASDMFPPL-VRFSW--KRQKDNGPLEKVTSADKKQLKLKESGCTA------
Gg TCRD : RDSIPE--VIVMKSKKHK---E---DGDTGKAACLARNFYKKN-ISLEM---------SSNEVVYEPKAPIVTSNGMYNT------
Gc TCRD : SIVKPK--LSAFYPPKSS---S---KDAVQAAVCLASDFFPKD-ISLQL------AFGDKPKANVTRPSVLKPD-GSYVA------
Hf TCRD : RPENPK--LSIFYPSLAR---SDDLDPDETAAVCLASEFTPKE-IELSV-------VWDINHKSNVTRSSILLNDGYYWS------
Hs TCRD : PHTKPS--VFVMKN--------------GTNVACLVKEFYPKD-IRINL--------VSSKKITEFDPAIVISPSGKYNA------
Ip TCRD : PRSLPT--LSILTSH----------DSNEKLEICLAAGFFPKE-GEVSL-----YTGDNVTPENHTVENAAMSAAGTYYY------
Ip TCRG : RVKEPQ--VYVYPVSTPE-------KGEKSFLLCHARGMFPDL-VRFTW-----QAKDQGGKNVDLRGDERLEQRDEVPEVRITSM
Mm TCRD : PPAKPS--VFIMKN--------------GTNVACLVKDFYPKE-VTISL--------RSSKKIVEFDPAIVISPSGKYSA------
Ov TCRD : PAASPS--VFVMKN--------------GTNVACLVKEFYPKE-VTISL--------QSSKKIIEYEPAIVVSPGGRYSA------
Po TCRA : DEYEPS--YYELE------------DQDTTITACLATGFSKHN-ATLEH-------DLFKNGSTNMTGAARASEESLY--------
Po TCRB : AVKSPI--VKVFRPSSKE-CRNPIDNEREKTLVCVASDFYPDH-VSVYWQIIQLNVTSGVNVTRGVTTDEAALRKDKVYTITSRLK
Po TCRD : SPVKPS--LSVLSP-----------LDPGGPRVCLAAGFRPQE-VAMVL-------NGSNSIDISEEEAVASLRHNSFFF------
Po TCRG : ---------RMYRTASRD---D---LDGKSVLLCVASNMSPPL-VHFSW--KRQKEGGPLEDLTSAQGEQLELRVPERTA------
Re TCRD : DKQEPE--LSIHYPPVHQ---SG--KTDQTATVCLLGPFYPNK-VTVNI----TLDGTTDNGSRTVLTEDVLTDKGYYRT------
Ss TCRD : IAARPT--LSILTPLHKS-------GDDDQPKVCLATGFFPST-KNMNL------ILGNNDTVQLSTTKAALSSTRTYFF------
Sc TCRB : ------------------------KQGKGETIVCVASEFYPDH-VSVSW-----TINGSIVREGVATDNAAL-RVGKYYRITSRLR
Tr TCRD : TVLSPT--LFILTPLL---------SEAPPADVCLAADFRPKE-GEMVV----------NGRVQKTSNAVLSLGHKSFFF------
Tn TCRD : LVLQPT--LFILNPQ----------PPAEPVDVCLAAHFSPRQ-GEMGV----------NGLKQETSGAVMSHRHGSFFF------
Xt TCRD : ---APS--VFVLKP-----------QKGDSPTACLVKDFTPKEDVEIYM-------NSTQSNGSIISASPVVSLNGTYSA------

                          •                             Ig Domain               ><       Tm
Am TCRD : VRVREFPVSD---HVTCAWGHGGSVVMSNDVMETQKENPSVDADISKSTEECEAPTDQLIVRS------SEKVN----ILSMTVMS
Cm TCRD : TSFLTVPQDT—-TGMRCIVQHEGKNVSAEAKYT----------------------------------PPQTS----V-------
Cc TCRD : AGFSDNKKIN-----ECKMDGQTADKTSDTPETTDKPERPTQNKCETNARTSASGFIDTG---------VPKMN----FMNLLVTG
Dr TCRA : FKFSSNDCKE---TGVCSSSGGGTDLER----------------------------------------DEKVN----FLSLGIFW
Dr TCRB : VYVSEWMDPN—-NKFTCTVRFFNGDHYINVSYTTKGQ-------------------------------KVKPK----GLKLVGFG
Dr TCRD : VGSSEKKIQE------CVMDGKTAKADK---IDKPSGNPVEDNPKKTATLECHTNTTTPPSTNND----DPKTN----SMTLLVIG
Dr TCRG : TSMVILNTPE-DNIYRCTVTHEGSKDPQQIEIKKKDEKPSVIKPSGGPDPTCPPSTETPISQESAN---PEKKQ----SLQLFVYG
Dl TCRG : TILQIPQRESSTYKYYCSVQHEGETVEAQIKQEVSTLPPPTTSPSPPPSLPPSVQPSLIPSPPPSLIPAPPASAPSQYQVKLLCLM
Gg TCRD : IKVVKVTKQS---EVTCTARLSNGNFTANSTTPEMKAEEIKPANICNTTDTSAEDVK-----------MEKVN----MLSMAVLG
Gc TCRD : FSFLFNPENE—-ANFVCEAVHENQHSHANITLGTKSGTGPTQSSCAESSKDMMEAGNDPFQE-------RPQAN----LLSLTLMG
Hf TCRD : SGFLPFPKDQKPVNVTCEAKHNGDTIVQQNIKEPTAAPPKPIDCNKSSNGTSAGLNDTDND--------LTEWN----FMSLTVMG
Hs TCRD : VKLGKYEDSN---SVTCSVQHDNKTVHSTDFEVKTDSTDHVKPKETENTKQPSKSCHKPKAIVH-----TEKVN----MMSLTVLG
Ip TCRD : AAFSKDGIKK------CAMKDVSLDKNDVKPTAQIPSCDQNSTLGISTSNNLTSPNKIKISG-------DPKGN----TMLLIVTG
Ip TCRG : LIVGKDKAKN--NNFICTVKHDSSVKDKELPIPREEDTSKSNAGVLNTCPTQKAEEAEEVEEEIMNFGVFEHSR----SLYLFSVT
Mm TCRD : VKLGQYGDSN---SVTCSVQHNSETVHSTDFEPYANSFNNEKLPE-----------------------PENDT----QISEPCYG
Ov TCRD : VKLGQYNDPD---SVTCSVEHNKKTWHSSDFEPKKDISETTPKPTESENTTEIQVPATCYEPQVQ----PGKVN----MMSLSVLG
Po TCRA : SQVALWSDAS-----QCESRTGKNESEVCADVLKK---------------------------------DPAVN----TVSLLVVA
Po TCRB : VSAEDWYKPE—-WNFECIVRFFNGTHDTYHKDSISGPDILTR--------------------------EKYLR----ITRQAKLS
```

Figure B-2 continued

```
Po TCRD : TGFSDGAISS------CEMNDVSSSNDHE----------------------------------------NVRFN----SYLLLLNA
Po TCRG : AIMVVDRDASYTYKYRCNVRHEGGPTERVVPALPPSTPPPSTTPPSTPPPSETSP--------------EPTWS—QQQLKLLCLLY
Re TCRD : SGFREFSHDLQPKTIGCSAKHGDKYIETKKEITISAPGVTPALDCDAGTSSNHTGTD------------MPEVN----FMTLTVMG
Ss TCRD : AGFSKENIKR------CEMDGNVEPPVDDKDDSTDVSNYNTTATTASVNCKDEPLNNVTTTNYTD----YTKMN----FTSLVVYG
Sc TCRB : VAAETWYNPS—-IKFTCIVTFYNVTNTDYPAHIFGETAVGDGMTR------------------------DKYLR----ITQTAKLS
Tr TCRD : AGFNRTIES-------CRLHNSSVPVDREQDTDQDTDQDTDQDTDQDTDQDTDQDTVQDTDQDTDQEWT-------RSREN----L
Tn TCRD : AGFNRTVKD-------CRLQESSDSVDPAGKDQGTD--------------------------------PELWT----VTRRNLLL
Xt TCRD : VHVDRLGTEG----LHCQAKHKG---------------------------------------------------------------


Am TCRD :                ><Cyt
Cm TCRD :LRVLFAKSLAFNLLLSATFFLL---------------------
Cc TCRD :-------------------------------------------
Dr TCRA :LRILLAKCVAVNVMMTVKAFVF---------------------
Dr TCRB :LRILFLKTVVFNVLVTFKAWMS---------------------
Dr TCRD :YILFLSKSFLYALVVSAMFWKLKFSMKTKEIPEE---------
Dr TCRG :LRILLAKCVAVNVMLSIKAFLF---------------------
Dl TCRG :YAVMLMKNVLYFIVVFIVLLKRKAGKKEERS------------
Gg TCRD :YTVLMVKSLVYCCGLSCLMIFRNKGPSTNCTHAD---------
Gc TCRD :LRVLLAKSIAFNTLMSIKLILF---------------------
Hf TCRD :LRFLLFKSISVNLLMTARVWIS---------------------
Hs TCRD :LRVLFFKSVAFNVMMTARCCSFKEFSAMRWIQR----------
Ip TCRD :LRMLFAKTVAVNFLLTAKLFFL---------------------
Ip TCRG :LRLLLAKAVGINILMTIKAFLV---------------------
Mm TCRD :YVLLLVKNVLYFCTVFVLLFKRNPAKI----------------
Ov TCRD :PRV-----TVHTEKVNMMS------------------------
Po TCRA :LRMLFAKSVAVNFLFTAKLF-----------------------
Po TCRB :LRLLFLKTVVFNVLLTLRLWLSHRV------------------
Po TCRD :YSVLIIKSSVYGAFVAFLVWRLQSSTEKQN-------------
Po TCRG :ARVMFTKIVAFTSVLTIRTLLA---------------------
Re TCRD :TVLMVKSLVYCCGLSLINILRNRERPPAAHKRTELPPAASTSH
Ss TCRD :LRILFFKSVAFNVLMTARAWVF---------------------
Sc TCRB :LRVLFAKAVAFNVLFTVKALVF---------------------
Tr TCRD :YGVFIIKSCIYGAFVVFLVWRLQGSCGKQNN------------
Tn TCRD :LLLRMDMLRVVLTKTVSLSTIVTIRALLL--------------
Xt TCRD :LQ-----------------------------------------
-------------------------------------------
```

Figure B-3 Multiple sequence alignment of nucleotide sequences used for Figure 2 tree. Mega 6.0 program was used and Clustal W methodology was employed for alignment. Highlighted nucleotide sequences represent coding sequences for the conserved C, WYRQ, YxC sequences respectively.

```
V12.0.1  ---------ATTATACTCATGGAGAAGGAATAGCTCCTCT------CTCTTCATCTGAATTGGTCATCAATGGTCACTCCACAACTTTGTCCTGTAATTACAATGGC
V12.0.2  ---------ATGATACTAGTGGAGAAGGAATAGCTCATCA------CTCTTCATCTGAATTTTTCATCAATGGTCACTCCACAACTTTGTCCAGCAATTACAATGGC
V13.0    ---------GTGATGCCTTTGGTGAGGGAATTGCACCTGT------CTCGGGTTACAAAACCGTCCCAGAGGAGAAAACAACGACCCTGTCCTGCATTTATAATGGC
V5.0.1   ---------GAGACTCTCTTGGTGACATTATAACATCTCG------TGGTTCTGAAAAAGCAGCTCAAGATGGTACCAATGTCACTCTGTCCTGCAACTACACTGGT
V5.0.2   ---------GAGAAGTAATAGGTGAAGTGATAACATCTCG------TGGTTCTAAAAAAGAAGCTCAAGATGGTACCAAGGTTACTCTGTCCTGCAACTACACAGGA
V17.0    ---GCACCAGCAGACAAGACAGCATTACTCCAACATCTCC------TGCTGTCCATGTAAAAGAAAAGGAAGCAGCCAGAATTTCTTGTGAATACAAGTACACAGTG
V3.0.1   ---------GGAATGCTGGTAAGGAAGTGATATCATCACT------GTCCACAGAACAGATTGTTTTGGCTGGAAAAAATGCCACACTTTCCTGCAAATACAATAG-
V3.0.3   ---------AGACTGCTGGTGAGGAAGTGATATCATCACT------GTCCACAGAACAGATTGTTTTGGCCGGAAAAAATGCCACACTTTCCTGCAAATACAATGG-
V3.0.2   ---------GGGATGCTGGTGAGGAAGTGATATCATCACT------GTCCACAGAACAGATTGTTTTGGCTGGAAAAAATGCCACACTTTCCTGCAAATACAATGG-
V3.0.4   ---------AGAATTCGGGTGAAGAAGTGATATCAACACT------ACCCACAAACCAATATATTATTGAAGAAAAAAATGCCACACTTTCCTGCAATTACAGTGG-
V4.0.1   -----------AATGCTGGTGAAGAAGTGATAAGACCACT------CTTCAAAGAGAAACATGTTCTGGAAGGAGATGTTGTTACACTGATCTGCAGCTACTCTGG-
V4.0.3   ---------AGAATGCTGGTGAAGAAGTGTTAAGACCACT------CTTCAAAGAAAAACATGTTCTGGAAGGAGATGTTGTTACACTGATCTGCAGCTACTCTGG-
V4.0.2   -------------TCATTACAGTGGA-CATTATTTATACT------GTTCAAAGAGAAACATGTTCTGGAAGGAGATGTTGTTACACTGATCTGCAGCTACTCTGG-
V9.0.1   --------AAAGCGTTTTTTTCAAATGCAATAAGACCTCT------CTCTGCTGAACAGCAGGTTTTAGAAGGAGGAAAGGTGATCCTTTCATGTGCCTACAATGGA
V9.0.2   --------TAGGCATTTTTG-CAAATTCAGTAACTCCTTG------GTCAACTAAACAGCAGGTTTTGGAAGGAGAAAATGTGACTCTTTCATGCAACTACAGTGGA
V8.0     ---------AGATTGCTGATGCATATACAATAGAATCTCT------CTCAGCTGAAAAGCAGGTTTCAGTTGGAGGAAATGCTGTGATTTCTTGTAAATACAGTGG-
V6.0.1   ---------AACACACAGAGGCAGATGTAATAAAACCTTT------ATCCCTGGATGAAAATAGCCATGTTGGTGATGCTGTTACTCTGTCCTGCAGTTACAAGGAT
V6.0.3   ------CACAACACACAGAGGCAGATGTAATAAAACCTTT------ATCCCTGGATGAAAATAAACATGTTGATGACACCTTTACTCTGTCCTGCAGTTATAAAGAT
V6.0.2   ---------AACACACAGAGGCAGATGTAATAAAACCTTT------ATCCCTGGATAAAAACTGTCATGTTGGTGACGCTGTTACTCTGTCCTGCAGTTATAAAGAT
V6.0.4   ---------AACACACAGAGGCAGATGTAATAAAACCTTT------ATCCCTGGATGAAAACAGTCATGTTGGTGATGCTGTTACTCTGTCCTGCAGTTATAAAGAT
V7.0     -----AATGTTGATGCAGATGGACATATGATAAATCCTTT------GAGTTTAGAGACACACGGCCCAGTAGGACACAATGTGACCCTCTCCTGCAATTACAGTGGT
V34.0.1  ---------GAATGTCTTTTTCAGATGAAATACATCCCCT------GTCTTCAGATAAACATGTAGCTGAAGGTCAAAATGTGACTCTATGTTGTAATTACAGAGCC
V34.0.3  ---------GAATTTCAATTGCAGATGAAATACATCCCCT------GTCTACAGACAAACATGTAGCTGAAGGTCAAAATGTGACTCTGTCTTGTAATTACAGCACC
V34.0.4  ---------GAACTGCAATTGCAGATGAAATACAGCCCCT------GTCTTCAGACAAACGTGTAGCTAAAGGCGAAAATGTGACTCTGTCCTGTAATTACAGCACC
V34.0.2  ---------GAACATCTTCTGCTGATGAAATACATCCCCT------GTTTTTAGATAAACAGGTAGCCGAAGGTCAAAATGTGACTCTGTCTTGTAATTACAGCATC
V34.0.5  ---------GAAATTTATTTGGTGATACAATCGAACCGCT------GTTTCCAGACAGGCATGGTTCTGAGGGTGAAAATGTGACTTTCTCCTGTAACTACAGTCTC
V34.0.6  ---------GAAATTCTTTTGGGGATGAAATTCAATCCCT------GGTTAAACACGAACATGGTTCTGAGAATGAAAATGTGACTTTCTCCTGTAAATACCGTCTA
V10.0.2  ------------ATACACTGGCACAGTCAATAAGACCACT------GCAGAGCGATGTTCATATCACTGAGAAGACTCCAGTCAGGCTCTCATGCAAATATGATGG-
V10.0.3  ------------ATACACTGGCACAGTCAATAAGACCACT------GCAAGGCGATGTTCATATCACGGAGAGGACTCCAGTCAGGCTCTCATGCAAATATGATGG-
V10.0.1  ------------ATACACTGGCACAGTCAATAAGACCACT------GCAGAGCGATGTTTATATCACTGAGAGGACTCCAGTCAAGCTCTCATGCAAATATGATGG-
V10.0.4  ------------ATACACTGGCACAGTCAATAAGACCACT------GCAGAGCGATGTTTATATCACTGAGAGGACTCCAGTCAAGCTCTCATGCAAATATGATGG-
V10.0.5  --------ACGAATGC-TTTGCGCAAAATATCACACCACT------GCAGGGCACGACACAAGTTACTGAGGGAAAGACAGTCACTCTGTCATGTAAATATGATGT-
V38.0.1  ---------AGGAGACGTTTGCACAATCAATAACACCACT------GGAAAACAAGACACTTAAAAGTGAGGGAGAAACGGTCACCCTGTCCTGTAAATATGAAGG-
V38.0.2  ---------AAGAGACGTTTGCACAATCAATAACACCACT------GGAAAACAAGACACTTAAAAGTGAGGGAGAAACGGTCACCCTGTCCTGTAAATATGAAGG-
V38.0.3  ---------GTGACACATTTGCACAGACAATAACACCGCT------GGAAAACAAGATAGTTAAATCTGAGGGAGAAACTGCCTCTCTGTCCTGTAAATATGAAGG-
V11.0.1  ---------ATCAATCATTTGGCCAGTCAATAACAGCACT------GCAGAGCACAAAACAGGCCGCTGAGAGCGAAACGGTCATTCTGTCATGCAAATATGATAC-
V11.0.2  ---------ATCAATCATTTGGTCAGTCAATAACAGCACT------GCAGAGCACAAAACAGGCTGCTGAGAGCGAAACGGTCAGTCTGTCCTGCAAATATGATAC-
V11.0.3  ---------ATCAATCATTTGGCCAGTCAATAACAGCACT------GCAAACCACAAAACAGGCTGTTGAGAGCGAAACGGTCATTTTGTCCTGCAAATATGATAC-
V14.0    ---------GTATCATCCAGACAAACATGATAACATCCCA------TGACCCACAAAAACAAGCATTTGAAGGGGAATATGTAAAACTTGCATGTAACTACAGTGG-
V36.2.8  -----ACCGGCTGTTT--TTGGAAATGTCATCAAACCAAA------GCAAACTGATGTTTTTGCAGATGAAGAATCAAATGTGACATTATCCTGTAGTTTTACAGCC
V36.2.10  -------CGGCTGTTT--TTGGAAATGTCATCAAACCAAT------CAACACTGATGTTTTTGCTGAGGAAGGATCAAATGTGACATTATCCTGCAGTTTTACAGC
V36.2.2  -------CTGCTGTTT--TTGGAAATGTCATCAAACCAAT------CAACACAGATGTTTTTACCGAGGAAAGATCAAATGTGACATTATCCTGTAGTTTTACAGCC
V36.2.11 TTTTCAGCTGCTGTTT--TTGGAAATGTCATCAGACCAAT------CAACACTGATGTTTTTGCTGAGGAAGGATCAAATGTGACATTATCCTGTAGTTTTACAGCC
V36.2.9  --ACACCTCACTAAAAAATCAAATGTGTCTTGGATCTACTGGTATGTAAAACAGATGTTTTTGTTGAGGAAGGATCAAATGTGACATTATCCTGTAGTTTTACAGCC
V36.2.1  -------CTTCTGTTT--ATGGAAATGTCATCATACCAGA------TGAAACAGAAGCTGTTCATGCGAAGGGTTCAGATGTTAAACTATCCTGCAGTTTTTTAGCC
V36.2.12 -------CAACTGTCT--TTGGAAATACCATCAAACCAGA------TGAAACAAATGTAGTTGCTCAGAAGGATTCAGATGTTAAATTGTCCTGTAGTTATTCA---
V36.2.6  -------CTTCAGTCT--TTGGAAGTGATATCAAATCAAA------CATCACAGATGTTGTTTATAAGGAGGGTTCAGATGGTACAATATTCTGCAGTCATTAT---
V36.2.7  -------CTTCAGTCT--TTGGAAATGATATCAAACCAAA------CTCAACAGATGTT---TATGAGGAGGGTTCAGATGGTACAATATTCTGCAGTCATTCT---
V36.2.5  -------CATCTGTTC--TTGGAAATGACATCAAGCCAAA------CATCACAGATGTTGTTTATGAGGAGGGTTCAGATGGTACAATATTCTGCAGTCATTCT---
V36.2.4  -------CTTCAGTCT--TTGGAAATGATATCAAACCAAA------CTCAACAGAAGTTGTTTTAAAGGAGGGTTCAGATGGTACAATATTCTGCAGTCATTCT---
V36.2.3  -------TTTCGGTTT--CTGGAAACCTCATCAAACCAAT------CCAAACTGATGTTTTTGCTGATGAGGGTTCAAATGTGACATTATCCTGCAGTTATTCA---
V25.0.2  -------GTGTCAGTT--TTGGGAATGTCATCACATCGGT------TGAAACCCGTCTGCTTCATAGCGAAGGAGACGTTATAAGTGTCTCCTGTAGCTACTCC---
V25.0.3  -------GTGTCAGTT--TTGGGAATGTCATCACATCGGT------TGAAACCCGTCTGATTCATACAGAAGGAGACGTTATAAGTGTCTCCTGCAGCTACTCC---
V25.0.1  -------GTGATGTTT--TAGGAAATGTCATCATACCGGT------CCACTCTGAAGTATATAAAACAGAAAAGGAAACTATATCTCTGTCCTGCAACTACTCT---
V25.0.4  -------GTGTTACTT--TTGGGAATGACATCACACCAGT------CCAGACCGAGGTGTTTGGGGCAGAGATGGACAATATAACAGTGTCCTGCAAATACTCT---
V15.0    -------GTGTCAGCT--GTGGCGATGTCATCACTCCAGC------CCAAACTGAGGTGTATGGGAAAAGTGGGAATAAAGTCACGCTGTCGTGCAACTACTCC---
V19.0    -------GTGGAAAGT--CTGATAGTCCCATCACATCAGA------TCAGACTGAGGTCTTTCCTAGGGAGAATGCCAATGTCACTTTGTCCTGCAGTTACTCA---
V20.0    -------ATGTCACTC--ATGCAGATGCAATTACCCCAGA------CCGCCCTGAAGTGAATGTTAATGAGACCGAAAGCATAACGTTATCCTGTCAATATTCC---
V37.0    -------GTGTCATTT--TTGGAGATGAAATCAGTCCTGA------CAGCAGTGAAGTGTTTGCTGCAGTGGCTTCAAAAGTAACTTTATCCTGCAGGTATTCC---
V36.1.2  -------ATGCAACAT--ATGCAGATGAAATTAAACCAGT------CAAAACTGAGGTGTTTGCTGATGAGGGCTCAAATGTTACTTTATCCTGCACTTATTCA---
V36.1.3  -------ATGCAATAT--ATGCAGATGAAATTAAACCAGT------CAAATCAGAGGAGTTTGCTGATGGGGGCTCAAATGTTACTTTATCCTGCACTTATTCA---
V36.1.1  -------ATGCAACAT--ATGGAGATGAAATTAAACCAGT------AAAAACAGAGAAGTTTGCTGACAAGGGTTCAAGCGTTAAATTATCCTGCACTTATTCA---
V36.1.6  -------GTGTTGCAT--ATGGAAATGAAATTAAACCAGT------CAAAACAGAGGAGTTTGCTAAGGAGGATTCAAGTGTTACTTTATCCTGCAGTTATTCA---
V36.1.4  -------ATGCCAGTA--ACGGAGATACTATTACACCAGA------CAAAACAGAGGAGTTTGCTGATGAGGATTCAAATCTTACTTTATCCTGCAGTTATTCC---
V36.1.5  -------GTGCCAGTA--GTGAAGATACTATTACAGCAGA------CAAAACAGAGGAGTTTGCTGATGATGGTTCAAATGTTACTTTATCCTGCACTTTTTCA---
V24.0.1  ---------GTAACCTTGCTAAAGATGTGATCACACCATA------CAGTGATGAAATTTTTGCCAGCGAGACTGAGAGTGTGAAATTCTCCTGTAACTACACTGGT
V24.0.2  ---------GTAACCGTGCTAAAGATGCGATCACTCCATA------TAGGGATGCAGTTTTTGCCAGTGAGAGTGAGAGAGTGAAACTCTCCTGTAACTACACTGGT
V24.0.4  ---------ATTCCAGTGCTACAGAAGCCATCACACCATA------CAGGGCTTATGAATTTGCTTTTAAGGATGACAGAGTGATCCTCTCCTGTAACTACACTGGT
V24.0.3  ---------GTTCCAGAGCCTCGGAAGTCATCACTCCTAA------AACTGACAGAGAGTTTGCTGTTGAAGGTGACAATGTGACTCTCTCTTGTAACTACAGTGGA
V24.0.5  ---------GTTCCAGAGCTTTAGAAGCCATCGATCCAAA------AACTGACAGAGAGTTTGCTGTTGAAGGTGACAATGTGACTCTCTCTTGTAACTACAGTGGA
V28.0    ------GTGTTTCTTCAGCTGAAGAAATCATATTCCCTAA------CAATAAAACAGTAATTGCCAAAGCAGATGAAAAAGTTCTCCTATCCTGCAGCTACACTGGC
V2.0.1   ---------GTGTAGTAAATGGGAATGTCATCACACCAGC------GAGAACATCTGTAGTTTTAACTGAAGGAAGCAACATCACTCTGTCTTGTACATTTAATACA
V2.0.3   ---------AGGACTCAAATGGGAATGTCATCAGACCAGA------GAAAACATCTGTAGTTTTAACTGAAGGAAGCAACATCACTCTGTCTTGTACATTTAATATA
V2.0.4   ---------GAAACGCTAATGAGAACTCTATCAGACCAGA------TGAAGAGTTTGTGTCTTTAAACGAAGGCAGTAACACCACCCTGTCCTGTACATTTGATGGA
V2.0.2   ---------GTGACTCTAATGGTGATATTATTAGACCAGA------GCAAACATCAGTAGTTGTAACTGACGGAAGCAGCATCACTCTGTCCTGCACATATGATGAA
V1.0     ---------GTGAAACAACTGGGGATAAAATTAAACCAGA------TAACGACACCGCACAGTTAATGGAGAATTCAAATATAACACTCTCATGCACATATGAAGGC
V23.2.6  -------GTGTGACGAATGCA--GATCAGATTGGACCAAATAAGGGTGAAAAT---TCCATCAGTAAAGAGGGAGAGACTG-------------------------
V23.2.4  -------GTGTGACGAATGCA--GATCAGATTGGACCAAATAAGGGTGAAAAT---GCCATCAGTAAAGAGGGAGAGACTGCTACAATGAGCTGCTCGTATGATACA
V23.2.8  -----AGGTGTGACGAATGCA--GATCAGATTGGATCAAATATTGGTGAAAAT---GTCTTCAGTATAGAAGGAGAGACTGCTACATTGAACTGTTCATATGCTACA
```

Sequences continued:

```
V12.0.1  TCT---TACAGCAGTGACT-------CTCTCATGTGGTA--TCGACAATACAG---CAGCTCCAAACCACAGTTTCTGTATTTGGTTAGTGAAGCC------AAGCT
V12.0.2  TCT---TACAGCAGTGACT-------CTCTCTTGTGGTA--TTGACAATACAG---CAGCTCCAAACCACAGTTTCTGTATTTGGTTAGTGAAGCA------AAGCT
V13.0    TCA---GCAGCTACAGATG-------CTCTCCAGTGGTA--TCGTCAGTACCC---AAGATCCAGACCAGACTTCCTTTTTCTAGTTAACGAAGCT------GCATT
V5.0.1   GC---------GGTTTATA-------ATCTGTTATGGTA--TCGTCAGTATCA---GAGATTCAAACCTGAACTCCTTCTCTCCATCACTGAATCT------GGAGA
V5.0.2   GT---------GGTTTATA-------ATCTGTTATGGTA--TCGTCAGTATCA---GAGATTCAAACCTGAACTCCTTCTCTCCATCACTGAATCT------GGAGA
V17.0    TC---------CATGAATA-------ATCTGCAGTGGTA--TCGCCAATATCC---CAATGCTAAGCCAGAATTCCTGGTTCTAATCATGGAGTCA------GGACA
V3.0.1   --------CAATATTCGCA-------ATTTGCAGTGGTA--TCGACAGTATCC---CGGATCTAAACCAGAGAACATCATTTTTCACACTGAGAGC------AACAA
V3.0.3   --------CAATGTTTACA-------GCTTGCAGTGGTA--TCGACAGTATCC---CGGATCCAAACCTGAGAATATCATTTATCACACTGAGAGC------AACAA
V3.0.2   --------CAATGTTTACA-------ATTTGCAGTGGTA--TCGACAGTATCC---CGGATCCAAACCTGAGAATATCATTTTTCACACTGAGAGC------AAAAA
V3.0.4   --------CTATGTTCAAA-------ATTTGCAGTGGTA--TCGTCAGTATCC---TGGATCCAAACCTGAGAATATTATTTTTTACACTGAGAGC------AACAA
V4.0.1   --------AAGTGTTCAGA-------ATCTGCAGTGGTA--CCATCAGTTTCC---CGGATCTAAACCTGAAAATATATTTCTTTACTTTGAGAAC------AA---
V4.0.3   --------AAATGTTCTGA-------ATTTGCAGTTGTA--CCATCAGTTTCC---CGGATCTAAACCTGAAAATATATTTCTTTACTTTGAGAAC------AA---
V4.0.2   --------AAATGTTCAGA-------ATTTGCAGTGGTA--CCATCAGTTTCC---CGGATCCAAACCTGAAAATATATTTCTTTACTTTGAGAAC------AT---
V9.0.1   AG------TAATATAAAGA-------GTTTACAGTGGTA--CAAACAGCATC---ACAGCCCTGCACCCGAATATCTCCTGCAAACTTTTGAAA---------ATGC
V9.0.2   AG------TAATATTCAAA-------GTTTGCAGTGGTA—CAGAAAGTATCCAAGCAGCACAGCTCCTGAATATATCCTGCAAACCTTTGAAA---------GTGCG
V8.0     --------TATTGGGCCAA-------GCTTCCAGTGGTA--TCGCCAATATCCA---GGATCTAGACCAGAGTACCTTATTTTAATACAGAGACT-------GGTGC
V6.0.1   TATACGGGTAATATGCAGA-------ATTTACAATGGTA--CCGGCAATTTCC---CAGATCCAAACCAGAGTTCATCCTTTACATCTTTGAAAGT------GGAAC
V6.0.3   TATACGGACACTGTACGGA-------ATTTACAATGGTA--CCGGCAATTACC---CAGATCCAAACCAGAGTTCATCCTTTACATCTTTGAAAGC------GGAAC
V6.0.2   TATACGGGTACTGTACAGA-------ATTTACAATGGTA--CCGGCAATTTCC---CAGATCCAAACCAGAGTTCATCCTTTACATCTTTGAAAGC------GGAAC
V6.0.4   TATACGGGTACTGTTAATG-------GTTTACAATGGTA--CCGACAATTACC---CAGATCCAAACCAGAGTTCATCCTTTACATCTTTGAAAGT------GGAAC
V7.0     CGT---------GTTCATA-------CTCTGCAATGGTT--TCGACAGTATCC---AGGATCTCGAATAGAGTTCCTCATTTTTGCCACTG--AGC------TCAAT
V34.0.1  AGTACTGGCAATGTTGACT-------CTCTCCTGTGGTA--TCGTCAGTATCC---CGGAGCAAACCCAGAGTATTTACTGCTGCTCAATGAATAT------TCAAA
V34.0.3  AGTACTGGCACTGTTAACT-------CTCTCCAGTGGTA--TCGCCAGTATCC---AGGAGCAAAACCAGAGTTTTTACTTTCTCTAAATGAATAT------TCAAA
V34.0.4  A---CTGTCACTGTTGACT-------CTCTCCAGTGGTA--TCGTCAGTATTC---CGGAGCCAAACCTGAGTTTTTACTTTCTCTAAATGAATAT------TCAAA
V34.0.2  AGTATTGGCATTCTTAGGT-------CTCTCCAGTGGTA--CCGCCAGTATTC---CGGAGCAAAACCTGAATTTTTACTGATGCTCAATGAATAT------TATAA
V34.0.5  ACTAGTGGCACCAGTGTTA-------ATCTTCTGTGGTA--TCGCCAGTATTC---TGGAGCAAAACCAGAGTATTTACTTTCTGTCACTGAATAT------TCAAA
V34.0.6  ACTAGTGGGACTAATGTCA-------ATTTCCAGAGGTA--TCGCCAGTATTC---CGGAGCAAAACCAGAGTATTTACTTTCTGTCACTGAATAT------TCAAA
V10.0.2  --------CAGTGCAAACA-------GTCTGCATTGGTA--TAGACAGCACCT---CGGATCAAAACCTGAATTTCTGCTTCTGGTAATGGAGTCAGTCATCAGACA
V10.0.3  --------CAGTGCATACA-------GTCTGCATTGGTA--TCGACAGTATCC---AGGAGCAAAACCTGAATTTCTGCTTCTTGTAAGGGAATCAGCTAAAAAATC
V10.0.1  --------CAGTGTATACA-------GTCTGCATTGGTA--TCCACAGTATCC---AGGAGCAAAACCTGAATTTCTGCTTCTTGTATTGAAGTCAGCTACAAAATC
V10.0.4  --------CAGTGTATACA-------GTCTGCATTGGTA--TCGCCAGTATTC---AGGAGCAAAACCTGAATTTCTGCTTCTTGTAACAGAGTCAACCAGCAAAAT
V10.0.5  --------CAGTGTCCAGA-------GTTTGCTTTGGTA--TCGACAGTATCC---TGGATCAGGACTGGAATTTCTGCTGCTTGTCATCGAGTCGCCAACAAAGAC
V38.0.1  --------CGTTGTGAATA-------ACCTGCACTGGTA--TCGTCAGTATCC---TGGATTCAGACCAAAGTTCCT-----------CGCATACATTTATGAGCAT
V38.0.2  --------CGTTGTGAATA-------ACCTGCACTGGTA--TCGTCAGTATCC---TGGATTCAGACCAGAGTTCCT-----------TGCATACATTTATGAGCAT
V38.0.3  --------GTCAGTAACTA-------ACTTGCACTGGTA--TCGTCAGTATCC---TGGATCCAGACCAGAGTTCCT-----------CGCATACATTTATGAGCAT
V11.0.1  --------CCGAGCAGACA-------ATTTACAGTGGTA--TAGACATTATCC---AGGATCTACACCAGAATATCTGCTGCTTGTTTATGAATCA------AACAC
V11.0.2  --------TCAAGCAAACA-------ATTTACAGTGGTA--TCGACAATATCC---AGGATCTACACCAGAATATCTGCTGCTTGTTTATGAATCA------AACAC
V11.0.3  --------TCAAGCAGACA-------GTTTACACTGGTA--TCGACAATATCC---AGGATCTACACCAGAATATCTGCTGCTTGTTTATGAATCA------AACAC
V14.0    --------AGCACCTGATA-------ATCTTCACTGGTA--CCGTCAATATCC---CGGATCGACACCAATGTTCCTCCTTTACATCTACGAAAGC------GGACT
V36.2.8  TCTGG---AGGGTCAGATT-------ATCTCCACTGGTA--CCGTCAGTATGG---AAGATCAAAACCTGAATTTCTTGTACTCACCTACAGAACAGCAACAAAAGC
V36.2.10 CTCGGG---TGTCACAGATT-------ATCTCCACTGGTA--CCGTCAGTATGG---AAGGTCAAAACCTGAATTTCTTGTACCACCTACAGCTCTGCAACAAAAGC
V36.2.2  TCTGG---AGGGTCAGATT-------ATCTCCACTGGTA--CCGTCAGTATGG---AGGATCAAAACCTGAATTTCTTGTCCTCATCTATGGGTCTGCAACAAAAGC
V36.2.11 TCTGG---AGGTACAGATT-------ATCTCCACTGGTA--CCGTCAGTATGG---AGGATCAAAACCCGAATTTCTTGTACTCACCTACAGCAGTGCAAAAAAATC
V36.2.9  TCTGG---AGGGTCAGATG-------ATCTCCACTGGTA--CCGTCAGTATGG---AGGATCAAAACCTGAATTTCTTGTACTCACCTACAGAACAGCAAAAGAAGC
V36.2.1  TCTGG---ATATACAGACT-------ATGTCTACTGGTA--CCGTCAGTATAG---AAGATCAAAACCTGAATATCTTGTACAAATCATTAGTAATTCAAATAAAGC
V36.2.12 TCTAC---AACTACAGATA-------CTCTCTACTGGTA--CCGTCAGTATGA---AAGATCAAAACCAGAATTTCTTGTACTCATCTACAGTAGTGCAAAAGCTGC
V36.2.6  CCTGG---AACTGGAGATA-------ATATCTTCTGGTA--CCGCCAATATGG---AGGATCAAAACCTGAATTTTTGCTTGCTACCTATGGCAGTGCAAAAGACAC
V36.2.7  CCTGG---AGCTGGAGATT-------ATATCTACTGGTA--CCGTCAGTATGG---AGGATCAAAACCTGAATTTCTTGTATTCACTTTCGGCAGTGTAAAAGACGC
V36.2.5  CCTGG---AGCTGGAGATT-------ATATTTACTGGTA--CCGTCAGTATGG---AGGATCAAAACCTGAATTTCTTGTACTAATTTTTGGCAGTACGAAAGACAC
V36.2.4  CCTGG---AACTGGAGATT-------ATATTCACTGGTA--CCGTCAGTATGG---AGGATCAAAACCTGAATTTCTAGTTCTCATCTATGGGACTGCAAACAAAGG
V36.2.3  TCTG---------CAGATT-------ATCTATTCTGGTA--TCGTCAGTATCC---CAGAGGAGCTCCTGAATTTCTTGTGTCCATCTTTGAGGGTGAAAGAAATAG
V25.0.2  TCTG---------CAAGAA-------CTTTATTCTGGTA--TCGCCAAAATCC---TGGAACGGTTCCTCAATTTCTTTTTGTCATTCTAGAATCCAATGGAGATGT
V25.0.3  TCAG---------CAAGAA-------CTTTATTCTGGTA--TTTCCAAAATCC---TGGATCGGTTCCTCAATTTCTTTTCATCATTCTCCATTCCAATGGAGAAGT
V25.0.1  TCAG---------CATCGA-------GTCTCTACTGGTA--TCGTCAATATCC---TAAATCAACACCTGAATTTCTTTTTGTCATTCTACATGCTACAGGGAAAGT
V25.0.4  TCAG---------CAGTTA-------GTCTGCAGTGGTA--TCGCCAGTATCC---TGGTTCAGCACCTGACTTTCTTTTGACAAT---ACATAGGACTGGAAGTGT
V15.0    ACAG---------CACTTT-------CTTTATACTGGTA--TCGTCAATATCC---AGGATCAAATCCACATTGCCTTCTTCTGATACTA------ACAAAAAGTG-
V19.0    TCAG---------CAATCA-------GTCTGCATTGGTA--CCGCCATTTACC---CAGATCAGCACCTGAATTTCTTCTATTAGTTCTGCAAACCACTGGACGGGT
V20.0    TCAG---------CATTCA-------GTCTCCACTGGTA--TCGACAGCATGA---TGCATCCGGTCCTCAGTTCCTCATGCTCATCCAACAAAGCACTGGACAAAT
V37.0    TCAG---------CATATA-------ATCTTCAATGGTA--CCGTCACTATCC---TGGATCAGCTCCTCAAAATCTTGTCCTCATATCAGAGGGTCTAACAGAGAC
V36.1.2  TCTG---------CAAATA-------ATCTTCAATGGTA--CCGTCAGTATCC---TGGATCAGCTCCACAATATCTTGTGCTCATTCCAGACAGTTCTAAAGAAGA
V36.1.3  TCTG---------CATATT-------ATCTTTACTGGTA--CAGACAGTATCC---TGGATCAGCTCCTCAACATATTGTGCTTATTCCAGAGAGTAGAAAAGATGA
V36.1.1  TCTG---------CAAATA-------CTCTTTTCTGGTA--CCGTCAGTATCC---CGGATCAGCTCCTCAATATCTTGTGCTCATTCCAGAAGGTACTAAAGATGA
V36.1.6  TCTG---------CACGGA-------CTTTATTCTGGTA--CCGTCAGTATCC---TGGATTAGCTCCTCAATTCCTTGTGTTCTTTATTCATGGTACAAGTGAAGC
V36.1.4  TCTG---------CCTGGA-------ATTTACTTTGGTA--CCGCCAGTATCC---TGGATCAGCTCTTGAATATCTTGTGCTTACTGTAGATACTGTAAACGAAGC
V36.1.5  TCTG---------CATGGT-------CTTTACTTTGGTA--CCGTCAGTATCC---TGGATCAGCTCCCGAGTTCCTTGTAAACATTGCAGAAAGGTCTAAAAAGGT
V24.0.1  TCTG---------TTGACG-------GTCTTCACTGGTA--TCGGCAGTATCC---AGATTCACCACCACACTTTCTCATACTGGATTATTATGGGTCAATTACTCA
V24.0.2  TCTG---------TTGATA-------GTCTTCACTGGTA--TCGACAGTATCC---CGGTTCACGACCACACTTCCTCATACTGGATTACTCTGGATTTGTTATT--
V24.0.4  TCTG---------CTCGAG-------GTCTGCACTGGTA--TCGACAGTATCC---AGGTTTACCACCCCATTTCCTCATACTGGATTACTATGGGGCAGTTACTCA
V24.0.3  TCAG---------TTCGTA-------GTGTTCACTGGTA--TCGCAAATATTC---AGGTTCACCTCCACAGTTCCTCATACTGGAGTATTCAGGAGTTATTACTCC
V24.0.5  TCTG---------TTCGTG-------GCGTTCACTGGTA--TCGACAGTACGC---AGCTTCGCCACCACAATTCCTCATACTAGAAGATTCAGGAGTAACAACTCC
V28.0    ACGG---------TTTATG-------ATTTATATTGGTA--CCGTCAATCCCC---TGGATCAGCACCTCAGTTCCTCATAATGGAATCGTCTGGAGTTATAATACA
V2.0.1   TCGC---------CTTACA-------GGCTCCTCTGGTA--CCGACAGACGCC---TGGATCAAGACCAGAGTTTCTGCTGCTGATTGTGAGGTCTACTAAAACTAC
V2.0.3   ACAC---------CTTACA-------GCCTCCACTGGTA--CCGCCAGATGCC—TGGATCAAGACCAGAGTTTCTGCTACTGAGGTTGTGAGGTCTACTAAAACTAA
V2.0.4   TCTG---------CTTATC-------GCCTCCACTGGTA--TCGACAGAAAAC---TGGATCGTTGTTAGAGTTTCTAGTGCTGATTGATAAATCTACCAAAGTTGT
V2.0.2   TCAG---------CCTACA-------GTCTCCACTGGTA--CCGACAGACACC---TGGATCAAGACCAGAGTTTCTGCTGCTGATAGATGAATCTTCAAAGTACGT
V1.0     TCCA---------TTGACA-------GCTTGTACTGGTT--CCAACAAAAAAT---TGGATCAAGACTGGAGTTTCTCATTATGATTGATGAAGCTTCTGAATACGT
V23.2.6  --------------ATATG----TTAGCCTTTACTGGTATACAGACAATATAC---TAACAAAGAACTTCAATATTTAATATGGAAAGAGGCT--CGATCATGGAGT
V23.2.4  ------AGCAGCAGATATG----TTAGGCTTTACTGGTA--CAGACAATATAC---TAACAAAGAACTTCAATATTTAATATGGAAAGATGCT--CGATCATGGAGT
V23.2.8  ------AGCAGCAGGAATG----TTAGACTTTATTGGTT--CAGACAATATCA---TAATAAAGAACCTCAATATTTAATATGGAAAGATTGT--AGATCATATAGT
```

Sequences continued:

```
V12.0.1  GATCCAACCAGCTGATCCT----------CCAATTTCTGGACTATCTGCT-AAATTAAA--TGAGGAGAATAAC---CTTGTGTATCTGGAGATCTCCTCTGCTGCA
V12.0.2  TGAACAACCAGCTGATCCT----------CCAATTCCTGGACTACCTGCT-AAATTAAA--TGAGGAGAATAAC---CTTGTGTATCTGGAGATCTCCTCTGCTGCA
V13.0    TAAACAACCTGCTAATCCT----------CCCATTCCTGGAATATCTGCT-AGATTAAA--CGAGGAGAAAAAT---TGTGTGTACCTGGACATCACCAGCACTGCA
V5.0.1   AGCAGTTAAAGCAGACCCA----------TCAGCCCACTGGCTATCAGCT-ACAGTCCA--CAAACAAAGCAAA---CTTACAGAACTGGAAATCAGCCCTGCTAAA
V5.0.2   AGCAGTTAAAGCAGACCCA----------TCATCCCACTGGCTATCAGCT-ACAGTCCA--CAAACAAAGCAAA---CTCACAGAACTGGAAATCAGCCCTGCTAAA
V17.0    AAATCAAACTGCGACACCG----------CCTCACCCCCGACTGACCACC-TCAGTTAA--TAAAGACAAAAAG---CTGGTGAATCTGGAAATCTTCCCCGCGGAT
V3.0.1   TCAGTCTTACTCCAG-------------------GCTTCGTCTGCTCGCT-GTGGCTGA--AAAAGAAATCAAA---CGTATGAATTTGAGCATATTTCATACAGAG
V3.0.3   GCAGTCTCACTCCAG-------------------GCTTCGTCTGCTCGCT-GTGACTGA--GAAAGGAATGAAA---CGTATGAATTTGAGCATATTTCATACAGAG
V3.0.2   TCAGTCAGACTTCAA-------------------GCTTCGCCTGCTTGCT-GTGGGTGA--GAAAGGAATGAAA---CGTATGAATTTGAGCATATTTCATACAGAG
V3.0.4   TCAGTCAGATTCCAA-------------------GCTTCGTCTGCTCGCT-GTGGCTGA--TAAGGGACAGAAA---CTTATGAATTTAAGCATGTATTTTACAAAA
V4.0.1   ---GCCACAAACTGA-------------------GGGCCGCATGACCGCT-GCATCTGA--TAAAGCAGTCAAA---CACATGAATCTGACCATCACTTCTACACAA
V4.0.3   ---GCCACAAACTGA-------------------GGGCCGCATGACTGCT-GCATCTGA--TAAAGCAGTCAAA---CACATGAATCTGACCATCACTTCTACAAAA
V4.0.2   ---GTTTACAACTAA-------------------GAACCGCATGATTGCT-GCGTCTGA--TAAAGCAGTCAAA---CACATGAATCTGACCATCAGTTCTGCAGAA
V9.0.1   GGCACCAGAGCAGAA-------------------GGACCGTCTACTTGCT-AAAGCACA--AAAAGACAAAAAA---CAACTGGATCTGGAGATCTCCAAAGCTGAA
V9.0.2   GAACCAGCGCAACA--------------------GGGCCGACATCTTGCT-AAAGCACA--AAAGAACAAAAAA---CAACTGGATCTGCTCATCTCCAAAGCTGAA
V8.0     TGGATCTGAACCCAC-------------------TTTACGTCTTACATCG-ATGGCCAA--AAAAGCACTCACA---CAAGTGGATCTGGAGATCTCCTCTACTGAA
V6.0.1   TATGAGTAAAGACAAACC----------------AGACAGATTTTCTGCT-CAAGTTAA--TAAAGGCACCAAA---CAAGTGGATCTGAACATCTCCCCTGCTGTA
V6.0.3   TATGAGTAAAGACAAACC----------------AGACAGATTTTCTGCT-CAAGTTAA--TAAAGACAGCAAA---CAAGTGGATCTGAACATCTCCTCTGCTTTA
V6.0.2   TATGAGTAAAGACAAACC----------------AGACAGATTTTCTGCT-CAAGTTAA--TAAAGACACCAAA---CAAGTGGATCTGAACATCTCCTCTGCTGTA
V6.0.4   TTTGAGTGAAGACAAACC----------------AGACAGATTTTCTGCT-CAAGTTAA--TAAAGACACCAAA---CAAGTGGATCTGAACATCTCCTCTGCTGTA
V7.0     GATCAATCAA-ACGCTGC----------------GCTTCGTATTTCAAAT-GTGGTCGA--TAAACA-AAAAAG---CA--TGAGTCTGAGCATTTTCCATGCAGAA
V34.0.1  CAGCACTAAGTCTTA-------------------TCTTCGTCTATTTTCA-AACGCTGT--AAAAGCACTGAAG---CGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V34.0.3  CAGCTCTAAACCTGA-------------------TCTTCGTCTGGTTTCA-AACGCTGT--AAAAGCACTGAAG---CGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V34.0.4  CAGCTCTAAGCTTGA-------------------TCTTCGTCTGGTTTCA-AACGCTGT--AAAAGCACTGAAG---CGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V34.0.2  CAAATCCAAGCTTGA-------------------TCTTCGTCTGTTTTCA-AACGCTGT--AAAAGCACTGAAG---CGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V34.0.5  CAGCTCTGAGTCTGA-------------------TCTTCGTCTGTTTTCA-AACGCTGT--AAAAGCACTGAAG---CGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V34.0.6  CAGCTCTGAGTCTGA-------------------TCTTCGTCTGTTTTCA-AACGCTGT--AAAAGCACTGAAG---CGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V10.0.2  TGTAACATATGCAGATCCT----------AAAATCCCTGGAATGGATGGA-GAAATGAG--CATGAGTGAAAAA---CAAGTGGATCTGGAGATCTTCTCTGCTACA
V10.0.3  TGTAACATATGCAGATCCT----------AAAATCCCCGGAATGGATGGA-GAAATGAG--CATGAGTGAAAAA---CAAGTGGATCTGGAGATCTTCTCTGCCGCA
V10.0.1  TGTAACATATGCAGATCCT----------AAAATCCCCGGAATGGATGGG-GAAATGAG--TATAAGTGAACAA---CAAGTAGATCTGGAGATCTCCTCTGCTGCA
V10.0.4  TGTGACCTATGCAGATCCC----------AAAATTCCTGGAATGGATGGA-GAAATGAG--GATGAGTGAAAAA---CAAGTGGATCTGGAGATCTCCTCTGCTGCA
V10.0.5  AGTTATACATGCAAAACCT----------CCCATTCCTAGAGTGGATGGA-GAAATGAG--TATGAGTGATAAA---CGTGTGAATCTGGAGATCTCCTCTGCTGAA
V38.0.1  GGAGCTACCAGTAAACCTC----------TTCCTCCACGACTTTTACCTA-AAGTTGA---TAAAAACAGCAAA---CGTGTGAGTCTGGAGATCTCTGAAGCTGAA
V38.0.2  GGAGCTACCAGTAAACCTC----------TTCCTCCACGACTTTTACCTA-AAGTTGA---TAAAAACAGCAAA---CGTGTGAGTCTGGAGATCTCTGAAGCTGAA
V38.0.3  GGAGCCACAAGTAAACCTC----------TTCCTCCTCGACTTTTACCAA-AAGTTGA---TAAAAACAGCAAA---CGTGTGAGTCTGGAGATCTCTGAAGCTGAA
V11.0.1  AGTGTTTCATGCTGAACCT----------CCTTTCCCAAGACTGAACGCC-AGTGTGAA--TAAAGTGGAGAAG---CAGGTGAATCTGACCATCTCTTCTGCTGCA
V11.0.2  AGTGTTTCATGCTGAACCT----------CCTTTCCCAAGACTGGACGCC-AGTGTGAA--TAAAGTGGAGAAG---CAGGTGAATCTGACCATCTCTTCTGCTGCA
V11.0.3  CGTGTTTCATGCTGAACCT----------CCTTTCCCAAGACTGAACGCC-AGTGTGAA--TAAAGTGGAGAAG---CAGGTGAATCTGACCATCTCTTCTGCTGCA
V14.0    GAAGAGTGACAACATACC----------------AGAGCGCTTAACGCCG-AAAATCAA--CAAAACTAAGAAA---GTTGTGGATCTGGAGATCTCCTCTGCTACA
V36.2.8  TG---CACAGTCCGAT-------ATAGA------CACAAGATTCACTACT-AATGTTAC--AAAAAAGGAG------CATGTGGATCTGCTGCTCTCCTCTGCTGCA
V36.2.10 TG---CTCAGTCCGAT-------ATAGA------CACAAGATTCACTACT-AATATTAC--AAAAAAGGAG------CATGTGGATCTGCTGCTCTCCTCTGCTGCA
V36.2.2  TG---CACAGTCTGAT-------ATAGA------TACAAGATTCACTGCT-AATATTAC--TAAAAAGGAG------CATGTGGATCTGCTGCTCTCCTCTGCTGCA
V36.2.11 TA---AGGAGTCTGAT-------ATAGA------TCCAAAATTCAGTGCT-AATGTCTC--AAAAAGTGAA------CATGTGGATCTGCTGCTCTCTTCTGCTGCA
V36.2.9  TG---AGAAGTCTAAA-------GTAGA------TCCAAGATTTATTGCC-AATGTCCC--AAAAAGTGAG------CATGTGGATCTGCTGCTCTCCTCTGCTGCA
V36.2.1  AC---AGCAATCTCTA-------GTAGA------TCCAAGATTCACTGCT-TACATTGA--AAACAGAGACAAAATTCATGTGGATCTGCTGCTCTCCTCTGCTGCA
V36.2.12 CG---AGAAATCTAAA-------GTAGA------TCCAAGATTCTTTGTT-AAAGTTGA--AAAAAGACAACAAATCCATGTGGATCTTCTGCTCTCCTCTACTGCA
V36.2.6  TG---AAACATCAAAA-------GTAGA------TCCAAGATTCTCAGTT-AAAGTTGA--AAAAAAGGAACGAAACCATGCAGATCTGAAGATCTCCTCTGCTGCA
V36.2.7  TG---AAACATCAAAA-------GTAGA------CCCAAGATTCTCTGTT-AAAATTGA--AAAAAAGGAACAAAACCATGCAGATCTGAAGATCTCCTCTGCTGCA
V36.2.5  TG---AAACATCAAAA-------GTAGA------TCCAAGATTCTCTGTT-ACAGTTGA--GAAAAAGGAACGAAACCATGCAGATCTGCTGCTCTCCTCTGCCGCA
V36.2.4  TG---AAACATCAAAA-------GTAGA------TCCAAGATTCTCTGTT-AAAGTTGA--AAAAAAGGAACGAAACCATGCAGATCTGAAGATCTCCTCTGCTGCA
V36.2.3  CA---AATCATCTGAT-------GTAGA------TCCAAGACTCTCTGCT-AAACTCAC--TAAAAGTGAGCATAACCATGCAGATTTGATGCTCTCCTGTGCTGCA
V25.0.2  TTTACAAAAATCTGAA----GCTGTGAA---AGATCCTCGATTCTCTGCA-AAAGTGAA--TGAAAAGAAAACC---CATGTGTTTCTGGAGATCTCCTCTGCTAAA
V25.0.3  TTTACAAAAATCTGAA----GCTGTGAA---AGATCCTCGATTCTCTGCA-AAAGTGAA--TGAAAAGAAAACC---CATGTGTTTCTGGAGATCTCCTCTGCTAAA
V25.0.1  TTTACAGAAATCTGAT----ATTGTGGATCAAGATCCTCGATTCTCTGGA-AAACTGAA--TGAGGAGAAAACC---CACACGAGCCTTAATATCTCCTCTGCTAAA
V25.0.4  TTCCCGAAGTTCTACA----ATTGTGAATGAAGATCCTTGGTACTATGGA-AAACTGAA--TAAAGAGAAAACC---CATGTGTTTCTGGAGATCTCCACTGCTAAA
V15.0    -----AAACCTCTGAA----AACATGCA------GCCTGGAATGTCAGTG-AGATTACA--TAAAGAGAAAGCC---TCTATGGATCTGAAGATCTCCTCTGCCAAA
V19.0    C---CAGACTGCACAG----AATCTGGA------TTCTCGGTTTGAGGCA-ATGCTGAA--TGAAAAGAAAAAC---AAAGTTCATCTGATCATCTCCTCTGTCAAA
V20.0    GATACAGCGTTCAGAC----GACA---------------GGCTGTCAGGA-AAACTTAA--CAAAGAGAAAAAC---CGTGTAGATCTGCAAATCTCCTCTTCCACC
V37.0    TA---AGATTTCTGAAGTAGATTCCAGGTTTACAACTAAGCTGAGAAAAG-AAAATCAAGCTGGAAAAGAAATCAAGCGTGTGGATCTGCTGCTCTCCTCTGCTGCA
V36.1.2  TT---GGAAGTCTAAT-------GTTGA------TGACAGATTTACAGCC-AGAGTCAC--TAAAAACAAAGAAAAGCGTGTGGATCTGCTGCTCTCCTCTGCTGCT
V36.1.3  TA---CAAAGTCTGAT-------GTAGA------TCACAGATTCATAGCC-AGAGTCAC--TAAAGATAAAGAAAAGCGTGTGGATCTGCTGCTCTCCTCTGCTGCT
V36.1.1  TT---GGAAGTCAGAT-------GTTGA------TAAAAGATTTACAGCC-AGAATTAC--TAAAGACAAAGAGAAGCATGTGGATCTGCTGCTCTCCTCTGCTGCT
V36.1.6  TA---AAGAGTCTGGT-------GTAGA------TCCCAGATTCACTGCT-AAACCCAA--TGGAGAAAAACAAAGCCATGTGGATCTGCTCCTCTCCTCTGCTGCT
V36.1.4  TC---GAACGTCTGAT-------GTAGA------TGTCAGATTTGCTACT-AAAACCAG--AAAAGAAAATGCAAACCATGTGGATCTGGAGATCTCCTCTGCTGCA
V36.1.5  AG---AGAGGTCGAAT-------ACTGA------TCACCGATTTACTACA-AAGATCAG--GATAGAGAAAGAAAAGCATGTGGATCTGCTCTTGTCCCCTGCTGCA
V24.0.1  TGC-GAATCCACCTAT------------------TGCTGGAATCTCCATT-AAACACAG--AAAAGACAACAGC---AGTGTGGATCTGGAGATCTCCTCTGTTGCA
V24.0.2  ACTCATGC-TAATCCTCC----------------TGATGGAATCTCCATT-AAACACAG--AAAAGAAAACAGC---AGTGTGGATCTGCTGATCTCCTCTGCTGCA
V24.0.4  TGC-AAAACCTCCGAT------------------TGCTGGAATCTCCATC-AAACACAG--AAAAGACGACAGC---AGTGTGGATCTGGAGATCTCCTCTGCTGCA
V24.0.3  AGC-GGATCCACCAGT------------------TCCAGGAATCTCCATC-AATCACAG--AAAAAAGCAAAGC---CATGTGGATCTGCTGCTCTCCTCTGCTGCA
V24.0.5  AGC-AAATCCACCAGT------------------TCCTGGAATCTCCATC-AATCACAG--AAAAGCAGAAAAA---CATGTGGATCTGCTGCTCTCCTCTGCTGCA
V28.0    TGC-TACTCCTCAGGT------------------TCCAGGGATAAAAATT-ATCAATAA--AAAAAACACCAAG---CATGTGGAGCTGGAGATCTTCAGTGCATCA
V2.0.1   AATTGAAGCTTCTCCA-------CCACA------GCCACATATGTCCATC-AACCTCAG--TG---AGAAA------CTTGTGGATCTGGAGATCTCCTCTGTTGCA
V2.0.3   AGTAGCAGCTTCTCCA-------CCTCA------GCCACATATGTCCATC-AACCTCAG--TG---AGAAA------CGTGTGGATCTGGAGATCTCCTCTGCTGCA
V2.0.4   AGTGACCGCTTCTCCT-------CCGCA------TCCGCACGTGTCTATT-AATCTCCA--TG---ATAAC------CGTGTGGATTTGGAGATCACTCCTGCTGCA
V2.0.2   CACCAAAGCTGAACAA-------CCTCA------CCCACATATGTCCATCCAACTTCAGAACA---ACAAA------CGTGTAAATCTGATCCTCTCATTCGCTGCT
V1.0     GACAAATGCCATTCCA-------CCTCA------TCCACGCCTGTCAGTCCAACTTCACAAAGCACAAAAA------CGTGTTAGTCTAGAGATGTCGTCTGCAGCC
V23.2.6  G-CCACTGGGACACCTGAT-------------GATCCTCGCATTCGGGCT-TCAACAT--------CAGAAACC---ACCACTGAACTCATTATTACTGGTGTAACT
V23.2.4  G-GTGTGGCGACTCCTGAT-------------GATCCTCGCATTCGGTCT-ACAACAT--------CAGAAACC---ACCACTGAACTCATTATTACTGGTGTAACT
V23.2.8  A-ACACTGTGACCTCCGCT-------------GATCCTCGCGTTGAGCCC-AAAACAT--------CAGAAACC---ACTGCTGAACTCATTATTCAGAGTGTAACT
```

Sequences continued:

```
V12.0.1  -TATCAGACTCTG--CTGTGTATTA-CTGCGCTCTGAAGCC----------------
V12.0.2  CTATCGGACTCTG--CTGTGTATTA-CTGCGCTCTGACGCC----------------
V13.0    ATATCAGACTCTG--CGGTGTATTA-CTGTGCTCTGCAGCC----------------
V5.0.1   GTGACTGACTCTG--CTGTGTATTA-CTGTGCCCTGAAGCC----------------
V5.0.2   GTGACTGACTCTG--CTGTGTATTA-CTGTGCCCTGAAGCC----------------
V17.0    GTGAAGGATTCTG--CTGTGTATTA-CTGTGCACTGCAGCC----------------
V3.0.1   ATGCAGGATTCTG--CTCTTTACTA-CTGTGCCCTGGAGCC----------------
V3.0.3   ATGGAGGATTCTG--CTCTTTACTA-CTGTGCCCTGGAGCC----------------
V3.0.2   ATGCAGGATTCTG--CTCTTTACTA-CTGTGCCCTGCAGCC----------------
V3.0.4   GTGGAGGATTCTG--CTCTTTACTA-CTGTGCCCTGCAGCC----------------
V4.0.1   TTGACAGACTCTG--CTGTGTACTT-TTGTGCCCTGGTGCC----------------
V4.0.3   GTGAAAGACTCTG--CTGTGTACTA-CTGTGCCCTGCAGCC----------------
V4.0.2   ATTACAGACTCTG--CTGTGTACTA-CTGTGCCCTGCAGCC----------------
V9.0.1   ATGTCAGACTCTG--CTGTGTATTA-CTGTGCACTGGTGCC----------------
V9.0.2   ATGACAGACTCTG--CGATGTATTA-CTGTGCACTAGTGCC----------------
V8.0     GTGAAAGACTCTG--CTATGTATTA-CTGTGCCCTGCAGCC----------------
V6.0.1   GAAACAGATTCTG--CCATGTATTA-CTGCGCCCTGGTGCC----------------
V6.0.3   GAAACAGATTCTG--CCATGTATTA-CTGTGCCCTGGTGCC----------------
V6.0.2   GAAACAGATTCTG--CCATGTATTA-CTGTGCCCTGGTGCC----------------
V6.0.4   GAAACAGATTCTG--CCATGTATTA-CTGTGCCCTGGTGCC----------------
V7.0     ATTCAAGATTCAG--CGACGTACTA-CTGTGCTCTGGAGAC----------------
V34.0.1  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V34.0.3  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V34.0.4  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V34.0.2  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V34.0.5  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V34.0.6  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V10.0.2  ATATCAGACTCTG--CAGTGTATTA-CTGTGCCCTGCAGCC----------------
V10.0.3  GTATCAGACTCTG--CTGTGTATTA-TTGTGCCCTGGAGCC----------------
V10.0.1  GTATCAGACTCTG--CAGTGTATTA-CTGTGCCCTGCAGCC----------------
V10.0.4  GTATCAGACTCTG--CGGTGTATTA-TTGTGCCCTGCAGCC----------------
V10.0.5  GTGACTGACTCTG--CTCTGTACTA-CTGTGCTCTGGTGCC----------------
V38.0.1  GTGACTGACTCTG--CGCTGTATTA-CTGTGCACTGACGCC----------------
V38.0.2  GTGACTGACTCTG--CGCTGTATTA-CTGTGCACTGACGCC----------------
V38.0.3  GTGACTGACTCTG--CGCTGTATTA-CTGTGCACTGACGCC----------------
V11.0.1  GTCACAGATTCAG--CTGTTTATTA-CTGTGCTCTGCAGCC----------------
V11.0.2  GTCACAGATTCAG--CTGTTTATTA-CTGTGCTCTGCAGCC----------------
V11.0.3  GTCACAGATTCAG--CTCTGTATTA-CTGTGCTCTGCAGCC----------------
V14.0    GTATCAGACTCTG--CTGTGTATTA-CTGTGCCCTGAAGCC----------------
V36.2.8  GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGGC----------------
V36.2.10 GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V36.2.2  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V36.2.11 GTATCAGACTCTGGTCCGTGTATCATCCGTTCATTTGATTATTCCTTTTATTACGGA
V36.2.9  GTATCAGACTCTG--CTGTATATTA-CTGCGCTCTGAGGCC----------------
V36.2.1  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V36.2.12 GTATCAGACTCTG--CTGTGTATTA-CTGCGCTCTGCAGCC----------------
V36.2.6  GTATCAGACTCTG--CTGTGTATTA-CTGCGCTCTGAGGCC----------------
V36.2.7  GTATCAGACTCTG--CTGTGTATTA-CTGCGCTCTGAGGCC----------------
V36.2.5  GTATCAGACTCTG--CTGTGTATTA-CTGCGCTCTGCAGCC----------------
V36.2.4  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V36.2.3  GTATCAGACTCCG--CTGTGTATTA-CTGCGCTCTGAGGCC----------------
V25.0.2  GTCACAGATTCAG--CCATTTACTA-CTGTGCCTTGGAGCC----------------
V25.0.3  GTCACAGATTCAG--CAATTTACTA-CTGTGCCTTGGAGCC----------------
V25.0.1  GTGACAGACTCAG--CAATGTATTT-CTGTGCCATGGAGCC----------------
V25.0.4  GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V15.0    CTAACAGACTCGG--CAGTTTACTA-CTGTGCGTTGAAGCC----------------
V19.0    GTGACAGATTCAG--CACTGTATTA-CTGTGCTATGGAGCC----------------
V20.0    ATAAAAGACTCTG--CGCTGTATTT-CTGTGCTCTGCAGCC----------------
V37.0    GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V36.1.2  GTATCGGACTCTG--CTGTGTATTA-CTGTGCTCTGGAGCC----------------
V36.1.3  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGGAGCC----------------
V36.1.1  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V36.1.6  GTATCAGACTCTG--CTCTTTACTA-CTGTGCCATGGAGCC----------------
V36.1.4  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAAGC----------------
V36.1.5  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGACGTCCATGAC----------
V24.0.1  GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V24.0.2  GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V24.0.4  GTATCAGACTCTG--CTCTTTACTA-CTGCGCACTGGAGAC----------------
V24.0.3  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V24.0.5  GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V28.0    GTGGAACATTCAG--CTGTGTATTA-TTGCGCTGCGCGGCC----------------
V2.0.1   GTGTCAGACTCGG--CTGTGTACTA-CTGTGCCCTGGAGCC----------------
V2.0.3   GTGTCAGACTCGG--CTGTGTACTA-CTGTGCCCTGGAGCC----------------
V2.0.4   GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V2.0.2   GTATCAGATTCTG--CCGTTTACTA-CTGTGCCCTGCAGCC----------------
V1.0     GTTACGGATTCTG--CTCTTTACTT-TTGTGCTTTGAGGCC----------------
V23.2.6  CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTGATAGTTGGAG-----------
V23.2.4  CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTGAGAGTTGGAG-----------
V23.2.8  CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTACGA------------------
```

Next set of sequences:

```
V23.2.7  -------GTTTGATGAGTGTG--GATAAAATTGGGCCAGATAAAGACACAAAT---GTCATCAGTAAAGAAGGAGAGACTGTGACACTAAGCTGTTCATATGATACA
V23.2.1  -------GTGTGATGACTGCA--GACCAGATTAGGCCAAATAAAGAAGC---T---TTTACTGTAAAGGAAGAGGAGACTGTGACCTTCAGTTGCTCATATGAAACA
V23.2.2  -------GTGTGATGACTGCA--GACCAGATTAGCCCAAATAAAGAAGC---T---CTTACTGTAAAGGAAGAGGAGACAGTGACCTTCAGTTGCTCATATGATACA
V23.2.3  -------GTGTGATGACTGCA--GACCAGATTAGGCCGAATAAAGATGT---C---ACAGTCATAAAGGAAGAGGAGACTGTGACCCTCAGCTGCTCATATGA---A
V23.2.5  -------GTGTGATGACTGCA--GACCAGATTAGGCCAAATAAAGATTC---C---ACAGTCATAAAGGAAGAGGAGACTGTGACCCTCAGCTGCTCGTATGA---A
V40.0    -------ATGTGACGGCTGTG--GACAAAATTGAGCCAAACAGTGGGAGCAAA---GTCATCAGAACAGAACGAGAGTCTGTAAAGCTGAGCTGCTCTTATGATTCA
V23.3.4  -------GTTTGGTGTGTGAA--GACAGCATTGAGCCAGATAAAGGGACAGAA---AAAACATCTGAAGAAACAGAAAATATTAAGCTAAGCTGCTTATATACTACA
V23.3.7  -------GTTTGGTGTGTGGA--GACAGCATTGAGCCAGATAAAGGGACAGAA---AAAAACTCTGAAGAAACAGAAAATGTCAAGCTGAGCTGCTCATATAGTACA
V23.3.6  -------GTTTGGTGTGTGGA--GACAGCATTGGGCCAGATAAAGGAACAGAA---AAAA---CTGAAGAAACTGAAACTGTCAAGTTGAGATGCTCATATAGTACA
V23.3.5  -------GTTTGGTGTGTGGA--GACAGCATTGAGCCAGATAAAGGGACAAAA---AAAAAGTCTGAAGAAACAGAAAGTGTCAAGTTGAGCTGCTCTTACAGTACA
V23.3.1  -------GTTTGGTGTGTGGA--GACAGCATTGAACCAGATAAAGGGAAAGAA---AAAACATCTGAAGAAATAGAAAATGTCAAGTTGAGCTGCTCTTACAGTACA
V23.3.2  -------GTTTGGTGTGTGGA--GACAGCATTGGACCAGATGAAAGGACAGAA---AACACATCTGAAGAAACAAAAACGGTCAAGTTGAGCTGCTCATATAGTGCA
V23.3.8  -------GTTTGGTGTGTGGA--GACAGTATAGAGCCAAATAAAGGAATGGAA---AAAACAATTAAAGAAACAGAACCTGTTCAGCTGAGCTGTGCGTATACTACA
V23.3.3  -------GTTTGGTGTTTGGG--GACAATATTGAGCCAGA---GGAGAAAGAT---GTTATGACAAAAGAAAGAGAAGCTGTCAAGCTGGCCTGCTCATACAGTACA
V22.0    -------GTGCAATGTGTCAA--GATCAGATTGGACCAAGCAGAACCAG---T---GTCACGAGTACAGAAGGAGATTCGGTCAATCTGGAGTGTTCATACGACTCA
V23.0.1  ----------GTGGATATACT--GATAGGATTGGACCAGGGGAAGGAGACACAAATATAATCAGTATGGAAGGAGAGTCTGTGACACTGAGCTGCACATATGAAACA
V23.0.3  ----------GTGCACTTTCC--GACAGCATTGGACCAAATGAAAAAGAAACTAATGTGATGAGTAGAGAAGGAGAGTCTATGAGATTGAGCTGCTCATTTGAATCA
V26.0    ----------GTGTGTTTGCT--GACAGCATTGCACCAAAAAAGGAGGGTAGAAATAAAACCAGAAAAGAAGGAGAGTCTGTGACACTGAGCTGCACATTTAACCGC
V23.1.3  ----------GTGTGTTTGCT--GACAGCATTAAACCAAAAGGCAAAGATAAA---ATAATCAGAAGCGAAGGAGAGTCTGTGACACTGAGCTGCGAATATGAAACA
V23.1.5  ----------GTGTGTTTTCT--GACAGCATTGGACCAAAAGATGAAGACAAACATATAATCAGGAGAGAAGGAGAGTCGGTGACATTGAGTTGCTCATATGAAACA
V23.1.6  ----------GTGTATTTGCT--GACAGCATTCTTCCAGACAAAGATAGTAAT---ATGACGAGTAATGAAGGAGAGTCTGTGACACTGAAATGCTCTTATGATACG
V23.1.10 ----------GTGTATTTGCT--GACAGCATTCGTCCAGACAAAGATACTA------TGACGAGTTATGAAGGAGAGTCTGTGACACTAAAATGCTCTTATGATACG
V23.1.7  ----------GTGTATTTGCT--GACAGCATTCGTCCAGACAAGGACACTAAT---ATAATCAGAAGTGAAGGAGAGTCTGTGACACTGAGATGCTCGTATGAAACA
V23.1.8  ----------GTGTGTTTCCT--AACAACATTTGGCCAAAAGACAAAGAAGTG---ATAGTCAGTAATGAAGGAGAGTCTGTGACACTGACCTGCTCATACAGTTCA
V23.1.1  ----------GTGTGTTTGCT--GACAGCCTTCAACCGAAGGACAAGGACAAA---ATAATCAG---CGAAGGACAGTCTGTGACACTGAGCTGCACATATGATACA
V23.1.4  ----------GTGTGTTTGCT--GACAGCATTGGACCAAATGACAAAGAAACTAGTGTGAGCAGTAATGAAGGAGAGTCTGTGACACTGAGCTGCTCATATGATGCA
V23.1.9  ----------GTGTGTTACCA--GACAGCATTGGACCAACTGACAAAGATGCAATTAAAACCAGGAAAGAAGGAGAATCTGTAACACTGAGTTGCTCATATGATACA
V23.1.2  ----------GTGTTTTGTCT--GACAGCATTCAACCAACAGACAAGAAAACTAATACTGACACAAAAGAAGAAGAGCATGTAACACTGAGCTGCTCATATGATACA
V23.0.2  ----------------GCACT--GATAGAATTGGACCAGAAGGAGAAACTAAT---ATACTCCAAGAAGAAGGACAGTCTGTGACACTGAGCTGCACATATGAAACA
V16.0    ----------TTGCTTTTGGT--GATACCATTGGACCAGATGAGCCTTTA------ATTTTCAGAAGTGAAACTGACCGGGTGACACTGAGTTGTTCGTACCAAACA
V21.0    -------AAATCACTTTCTGT--GCTCATATTGCACCAAACACAAAATTC------GTCATTGGTAGAGAGGGGGAATCTGCAACTCTCTCATGCCTGTATGAATCC
V30.1.1  ------GATGTGATTCTCAG---GACAAAGTTGATCAGGACACAAAGACTCAA---TCTGCTTTTGAAGGTGA---TACTGTAACCATTGACTGCACATATCAGACT
V30.1.3  ------GATGTGATTCTCAG---GATAAAGTTGATCAGCACACAAAGACTCAA---TCTGCTTTTGAAGGTGA---TACTGTAACCATTGACTGCACATATCAGACT
V30.1.4  ------GATGTGATTCTCAG---GACAAAGTTGATCAGCACACAAAGACTCAA---TCTGCTTTTGAAGGTGA---TACTGTAACCATTGACTGCACATATCAGACT
V30.1.2  ------GATGTGATTCTCAG---GACAAAGTTGATCAGCACACAAAGACTCAG---TCTGCTTTTGAAGGTGA---TACTGTAACCATTGACTGCACATATCAGACT
V30.1.6  ------GATGTGATTCTCAG---GATAAAGTTGATCAGCACACAAAGACTCAA---TCTGCTTTTGAAGGTGA---TACTGTAACCACTGACTGCACATATCAGACT
V30.1.7  ------GATGTGATTCTCAG---GATAAAGTTGATCAGCACATAAAGACTCAA---TCTGCTTTTGAAGGTGA---TACTGTAACCATTGACTGCACATATCAGACT
V30.1.5  ----AGGATGTGATTCTCAG---GACAAAGTTGATCAGCACACAAAAACTCAA---TCTGCTTTTGAAGGTGA---TACTGTAACCATTGACTGCACATATCAGACT
V30.0.2  ------GCTGTGATTGTGCT---GAGAGTGTTGATCAGAATACAAGAGTTGAA---ACTTCTGTTGAAGAAGG---ATCTGTGAGCATTAGCTGTACATATCAAACC
V30.0.3  ------GCTGTGATTGTGCT---GAGAGTGTTGATCAGAATACAAGAGTTGTA---ACTTTTGTTGAAGAAGG---ATCTGTGACCATTAACTGTACATATCAAACC
V30.0.1  ------GCTGTGATTGTGCT---GAGAGTGTTGATCAGAATACAAGAGTTGAA---ACTGCTGTTGAAGGAGG---ATCGGTAACCATTAACTGTACATATCAAACC
V35.0    ------GGTGTAAATGTAAG---GATTCTGTCCACCAGAA---GAGAGTTGAA---ACTGCTGTTGAAGGCAC---CGATTTATTCATCAGTTGCTCATATACAAGC
V30.0.4  ------GTCGTGTTTCTCTG---GATAAAGTTGAACAAAACTCAAGAGTTAAA---ACTGCTGTTGAAGGTGA---AGATGTAATAATCAACTGTCAGTATGAAACC
V39.0    ------GTCGTGTTTCTCTG---GATAAAGTTGAACAAAACTCAAGAGTTGAA---ACTGTTGTTGAAGATGA---AGATGTAATAATCAACTGTCAGTATGAAACC
V30.0.5  ------GATTTGTTTCCCCG---GACAGTGTTGAACAGAAAACAAGAGTTCAA---ACTGCTGTAGAAGGTGA---AAGTGTAACAATCAACTGCACTTATTCAACC
V30.0.6  ------GAATAGAGTCGCAG---GACAATGTTGAGCAGAAAGCAAGAGTTCAA---ACTGCTGTTGAAGGTGA---AGCTGTTGAGATCGACTGCACATATCAAACC
V29.0    ------AATGCAGAGCACAA---GATAGTGTTACTCAGACAGCTACAGTTCAA---ACTGCTACTGAAGGTCA---AGCAGTCACTATCGAATGCACATACAAAAC-
V41.0    ------GATGCAAAGCACAG---GACAATGTAATACAGCCCAAACAAGTTATG---ACAGCTTATGAAGATGG---AACAATCACTCTTGAGTGCACGTACAAAGCG
V32.0.2  ------GGTGCAGATCAGAA---GACAGAGTTGATCAGCCGGACAAACAACTG---ACTGAAGCTGAAGGAGG---ATCAGTGACATTACAGTGCAAATATAAAACA
V32.0.4  ------GTTGCAGATCAGAG---GACAGAGTTGATCAGCCGGACAAACAACTG---ACTGAAGCTGAAGGAGG---ATCAGTGACATTACAGTGCAAATATAAAACA
V32.0.60 ------AGTTCAGATCAGAG---GACAGAGTTGATCAACCGGAAAAGGAAATA---ACTGTAGCTGAAGGAGG---ATCAGTGACATTACAGTGCAAATATGAATCA
V32.0.3  ------AGTGCAGATCAGAA---GAGAAAGTTGATCAGCCAGACAAACACATG---ACTGAAGTTGAAGGAAA---GTCAGTGACGTTACGATGTACATTTACAACA
V32.0.5  ------AGTGCAGATCAGAA---GAGAAAGTTGATCAGCCAGACAAACACATG---ACTGAAGTTGAAGGAAA---GTCAGTGACGTTACGATGTACATTTACAACA
V32.0.1  ------AATGCAGACCAGA------TAGTGTTAATCAGCCAGACACACACGTA---ACTAAATCTGAAGATGG---AGCAGTGACATTACAATGCAAATACACTGCA
V31.0    ------GGTGTCATTCAGAT---GACACGGTTGATCAGCCACTTAAACATGTT---ACCGCATTAGAGGGAAT---CTCAGTAACACTGGACTGCAAATATACCACA
V33.0.6  ------TGTGCAGGGGGCAA---GACAAAGTTCAACAGCCTTCAGGAGAAAAG---ACTGAAAATGAAGGAGA---TCAAGTCAATTTAATGTGTAATTACA---CT
V33.0.7  ------TGTGCAGGGGGCAA---GACAAAGTTCAACAGCCTTCAGGAGAAATG---ACTGAAAATGAAGGAGA---TCAAGTCAGTTTAATGTGTAATTACA---CT
V33.0.2  ------TGTGTAGGGCGCAA---GACAGAGTTGAACAGCCTACAAGAAAAAAG---ACTGAAAATGAAGGAGA---TCAAATCACTTTAATGTGTAATTACA---CT
V33.0.5  ------TGTGCAGAGGACAG---GACAGAGTTGAACAGCCTTTAAGAAGCATA---AGTGCAAATGAAAGAGG---TCAAGTTATAATAAGCTGTAGCTACA---GC
V33.0.8  ------TGTGCAGAGGTCAA---GACAGAGTTGAACAGCCTTTAAGAGACATA---ACTGCAAATGAAGGAGA---TCAAGTTATTATACCCTGTAATTACA---GC
V33.0.9  ------AGTGCAGAGGACAA---GACAGAGTTGAACAGCCATCAATAGAAATG---ACTGAAAGTGAAGTGGA---AGAAGTTATTATAAGATGTAATTACA---GT
V33.0.13 ------TGTGCAGAGGACAA---GACCGAGTTGAACAGCCTATAAGAGAAATG---ACAGAAAGTATAGGAGA---TAAAGTTGTTATCAGCTGTATTTATA---GC
V33.0.10 ------TGTGCAGAGGACAA---GACAGAGTGGAACAGGTTGAAGGAGAAATG---AGTGACAGTGAAGGAAA---TCATGTTATTATAAGCTGTAATTACA---GC
V33.0.11 ------TGTGCAGAGGACAA---GACAGAGTGGAACAGGCTGAGGGAGAAATG---ACTGAAAGTGAACGAAA---TCATGTTATTATAAGCTGTAATTACA---TT
V33.0.12 ------TGTGCAGAGGACAA---GACAAAGTTGAACAGCCTATAAGAGAAATA---ACTGAAGGTGAAGGAGC---TCATGTTTTTGTCCAGTGTAATTATA-----
V33.0.3  ------TGTGCCGGGGATCATCAGACAATGTTGAACAACCTTCAGGACAAATT---ACAGCAAATGAAGCAGA---TTCAGTTCATCTACAGTGCAGTTACT---CT
V33.0.4  ------TGTGCAGGGGACAG---AACAGAGTTGAACAACCTTCAAGAGAAATG---ACGGCAGATAAAGGAGA---TACAGTTACTCTGCCGTGCAACTATT---CT
V33.0.1  ------TATGTGGGGAACAA---GACAGTGTTGATCAGCCGGACAGAGAACTG---ACTTCTGCTGAAGGAAA---TGAAGTGATTTTATTGTGTAATTATAAACTG
V18.0    ------AATGCAGATGTGAA---GACGCTGTACGTCAGCCAACAGCTTCTGTA---TCTGCTTTCACAGAAGA---AACCATAACCTTGAGCTGTGAATATTATACA
V27.0    ------AATCCAGAGCACAA---GACTCAGTTTATCAGCCAAAAACCCTCCAA---TCAGCAAATGAAGAGGATGCAGCAGTAACCTTGGATTGTAAATACAGTACA
```

Sequences continued:

```
V23.2.7  ------CAGAGCAGCTATG----TTAGACTTTACTGGTA--CAGACAATATGC---TGACAGAGAACCGCAGTATTTAATTTGGAAAGAGGCT--CGATCATTGAGG
V23.2.1  ------AGCAGCAGCTATG----TTTGGCTTTACTGGTA--CAGACAGCATCT---TAACGGAGAACCTCAGTATTTAATCTTTAAACCTGCA--AAATCAGCTTCT
V23.2.2  ------AGCAGCAGTTATG----TTAGGCTTTACTGGTA--CAGACAATATCT---TAATGGAGAACCTCAGTATTTATTATTCAAAGCTGCA--CGATCAAGTAGT
V23.2.3  ------AGCAGCAGTACTG----TTAGGCTTTACTGGTA--CAGACAGTATCT---CAATGGAGAACCTCAGTTTTTAGTATCT---TCTGCA--CAATCAAGTAGT
V23.2.5  ------AGCAGCAATTATG----TTAGGCTTTACTGGTA--CAGACAGTATCT---CAATGGAGAACCTCAGTATTTAATACAT---TCTGCA--GGTTCAAATATT
V40.0    ------AATAATGAATATG----TACGTCTATATTGGTA--CAGAAGACATTC---AAGTGGAGAACTTCAATATTTATTATATAAAGGAGCT--CGATCAAACTCA
V23.3.4  ------AGCAGTGATTATG----TTTGGCTTTACTGGTT--CAGACAATATCT---CAAAGGAGAACCTCAGTATTTATTATACAAAGGTGCA--CGATCAT---AT
V23.3.7  ------AGCAGTGAATATG----TTTATCTTTACTGGTA--CAGACAGTATCT---CACTGGAGAACTGCAGTATCTACTACGCAAAGGTGCA--CAATCAT---AT
V23.3.6  ------AATAGTGAATATG----TTTGGCTATATTGGTA--CAAACGGTCTCT---CAATAAAGAACCTCAGTATTTATTGTTCGAAAGTGCA--CGATCGG---CT
V23.3.5  ------AGTGGTGAAGATG----TTCTGCTTTACTGGTA--CAGACAGTATCT---CAATGGTGAAACTCAGTATTTACTGCGCAGATATGCT--CGATCATTGGGT
V23.3.1  ------AGCAGCAGTTATG----TTTGCCTCTACTGGTA--CAGACAGTATCT---CAATGGAGAACCTCAGTATTTAGTGCACAGATATGCT--CGATCATGGGGT
V23.3.2  ------AGTATTGAATATG----TTGTGCTTTACTGGTA--CAGACAGTATAA---ACATGGAGAACCTCAGTATTTACTTCGCAGATATGCT--CGATCATATAAT
V23.3.8  GCCACAAGTAGAAGTTATG----TTTATCTTTACTGGTT--CAGACAATCTTT---CAATGGAGAGCCACAGTATTTGTTAAACAAAAATGCA--CAAACAG---CC
V23.3.3  ------ACCAACAATAGAG----TTCGGCTTTATTGGTA--CAGACAGAATCC---AAATGCAGAACTTTTGCTTTTAACATACAAAGGTGCT--CGATCTC---TG
```

```
V22.0    ------AGCATCAACAATG----TCTTCCTATATTGGTA--CAGAAAATACTC---TAATGCACAGCCACAATATATACTTCTGAAAGGCGGC--CGATCACAGACA
V23.0.1  ------AACAGTAATGACA----TCTGGCTTTATTGGTA--CAGACAGTTTTC---TAACAGAGAACCAGAATATATACTGTTTAAAGGTGCT--CGATCAAGGAGC
V23.0.3  ------AGCAGCAATTATG----TTTTGCTTTACTGGTT--CAGACAGAACCC---TGATGGAGAGCTGCAATATTTACTATATAAAGGTGCT--CGATCGTGGAGC
V26.0    ------AGCGATACTGCGA----TTTCACTTTACTGGTA--CATGCAGTATCC---AAACAGAGAACTTCAGTATTTACTGTATAAAGGTCAT--CGATCGTCTGAG
V23.1.3  ------AGCAGCAATAATG----TAATGCTATACTGGTA--CAGACAATATAC---TAACAGAGAACCTCAGTTTTTGCTGCGTAAAGGAGCT--CGATCAAGGAAT
V23.1.5  ------CAGAGCAATAATG----TTTGGCTTTACTGGTA--CAGACAATATAC---TAACAAAGAACCTCAGTATTTACTTTTGAAAGGAGCT--CGATCAAATAGC
V23.1.6  ------AGCAGCAATAATG----TTTTGCTTTACTGGTA--CAGACAGTATGC---TAATACTGAACCTCAGTATTTACTGTGGAAAGGAGCT--CGATCTAACAGC
V23.1.10 ------AGCAGCAATAATG----TTTTGCTTTACTTGTA--CAAACAGAATGT---TAATACTGAACCTCAATATTTACTGTGGAAAAAAGCT--CGATCATACAAT
V23.1.7  ------AACAGCAACTATA----TTAGACTTTACTGGTA--CAGACAGTATGT---TAATACCGAACCTCAGTATTTACTGTGGAAAGGAGCT--CGATCAGACAGC
V23.1.8  ------GGCAGCAATGATG----TTTGGATTTACTGGTA--CAAATATCTTTC---TAACACAGAACCTAAGTTTTTACTGTGGAAAGGTGCT--CGAGCACAGAAC
V23.1.1  ------AGCAGCAGTTATG----TTTATCTCTATTGGTA--CAAACATAACAT---AAAGGGAGAACCTAAATTCGTACTGTGTAAAGGTGCC--CGATCCCGAAGT
V23.1.4  ------CAGAGCAATAATA----TTAGACTATACTGGTA--CAGACAAAACAG---TAATCGAGAACCTGAGTATTTACTGTGGAAAGGAGCT--CGATCATATGGA
V23.1.9  ------CAGAGCAATAATG----TTTATCTTTTCTGGTA--CAGACAATTTAC---TAACAGAGAACCTGAGTATTTACTAATTAAAGGAGCT--CGATCATATGGA
V23.1.2  ------AGGAGCGAGTATG----TCTATCTCTACTGGTA--CAGACAATATCC---TAATAGAGAACCAGAATATTTACTGCGTAAAGGTGCC--CGTTCGTGGAGC
V23.0.2  ------AACAGCAACACCA----TTTATCTTTACTGGTA--CAGACAGTATCC---AAAGACAAAACCTGAATTTATACTGTATAAAGGTGCT--CGATCATGGAGC
V16.0    ------AATAGTGAACATG----TTGATCTTTACTGGTA--CAGACACTATCC---TAAAAGTGAACCACAGTATTTACTATGGAAGGGAGCC--AGAGCAAGGGAT
V21.0    ------AGCAGCTCCCATG----TGTTGCTCTTCTGGGG--CAGGATACTCCC---AAAACAGGCACCTGAGTTTTTACTGCTCAAGGGTGCA--CGATCCGAGACA
V30.1.1  TCA----AC---AAACCCA------ACTCTATTCTGGTA--CCAGCAGAAAGT---CAAAGGAGTCCCCAAACACATGCTAAATAAAATCACCAACACAACTGGACA
V30.1.3  TCTT-ATAC---AACACCG------ACTCTATTCTGGTA--CCAGCAGAAAGT---CAACGGAGTCCCCAAACACATGCTTAATAAAATCTCT---ACATCTGGACA
V30.1.4  TCTT-CAAC---AGGCCCG------ACTCTATTCTGGTA--CCAGCAGAAAGT---CAATGGTGTCCCAAAATACATGCTGAATAAAATCACT---ACGACTGGACA
V30.1.2  TCTT-ATAC---AGGCCCA------ACTCTATTCTGGTA--CCAGCAGAAAGT---CAACGGAGTCCCCAAATACATGCTGAATAAAATCACT---ACAACTGGACA
V30.1.6  TCTT-CTAC---AGGCCCA------ACTCTATTCTGGTA--CCAGCAGAAAGT---CAGTGGAGTCCCCAAACACATGCTGAATAAATTCTCT---ACAACTGGAAA
V30.1.7  TCTT-ATAC---AACACCG------ACTCTATTCTGGTA--CCAGCAGAAAGT---CAATGAAGTCCCCAAACACATGCTGAATAAATTCTCT---ACAACTGGAAA
V30.1.5  TCTT-ATAC---AGGCCCG------ACTCTATTCTGGTA--CCAGCAGAAAAT---CAA----GTCCCCAGATATATGCTGAATAAAATCACTAACACATCAGGACA
V30.0.2  AGTG-ATCC---TTCCCCA------AATCTCTTTTGGTA--CCAACAAAAGCC---GAACACAATCCCCAAACACATGATGACGATAATCACT---ACAACTGTACA
V30.0.3  AGCG-ATCC---TTCCCCA------TATCTGTTCTGGTA--CCAACAAAAGCC---GAACACAATCCCCAAACACATGATGACAATAATCACT---ACAACTGTACA
V30.0.1  AGTG-ATCC---TTCCCCA------TATCTGTTCTGGTA--CCAACAAAAGCC---AAACACAATCCCTAAATACATGATGATGATATTTGCT---ACAACTGTACA
V35.0    ACTG-ACCC---AATACCA------ACGCTTTTCTGGTA--CCAGCAAAAAGT---AAATGAGTTCCCTAAGTACATGCTAT---------CT---ACTTTTGAAAG
V30.0.4  ACTG-ATAC---ATCGCCA------TATCTCTTCTGGTA--CCAGCAAAAAGT---CAACAGTGTCCCAGAATACATGCTGAACAGATTC------AGTTCTGATGA
V39.0    ACTG-ATAC---ATTGCCA------TATCTCTTCTGGTA--CCAGCAAAAAGT---CAACAGATTCCCAGAATACATGCTGAACAGATTT------AGTTCTGATGA
V30.0.5  ACTG-ATAC---ATTTCCA------TATCTCTTCTGGTA--CCAGCAACAAGT---AAACGGAATCCCTCAATATATGCTGAAGAAATTTCCT---GGATCCGGTGA
V30.0.6  --CA-GCTC---ACCCCGA------C-TCTCTTCTGGTA--CCAGCAAAAGAC---AAACACAATCCCAAAATATATGCTCAACAGAGTTGGA---AAGACTGGAGG
V29.0    --CA-ACGC---ATTCCCA------ACGCTTTACTGGTA--CCAGTATAAAGC---AAATCAGTTCCCAAAGTATATGCTGAAGAGATAT---GCTGGATCAAGCGA
V41.0    TCCA-GCACTCAATATCCA------AACCTCCTCTGGTA--CCAACAGATGAC---GAATGGATCCCCAAAATACATGCTGATGAGACTT---TCTGGATCTAGTGA
V32.0.2  GATT-CAGT-GCAAGA--A------AATTTATTCTGGTA--CATCCAGAGAAC---AAATGACCTACCAAAATACATCCTGAGGAGAAGCAAATACGGAG---GAGA
V32.0.4  GATT-CAAC-ACAAGA--A------GATTTATTCTGGTA--CATCCAGAGAAT---AAATGACCAACCAAAATACATCCTGAGGAGAAGCAAATATGGAG---GAGA
V32.0.60 TATT-CTGT-GCAAAA--A------GATATATTCTGGTA--CGTGCAGAGAAC---AAATGACTTCCCAAAATACATCCTGAGGAGAAGTAACTTTGGAG---GAGA
V32.0.3  ACTT-CAAA-CCAACC--A------GACCTCTTCTGGTA--CATCCAGAGAGC---AAATGATCATCCCAAATACATCCTGAGGAGGGACAAGTATGGAGCTGGAGA
V32.0.5  ACTT-CAAA-CCAACC--A------GACCTCTTCTGGTA--CATCCAGAAAGC---AAATGATCATCCCAAATACATCCTGAGGAGGGACAAGTATGGAGCTGGAGA
V32.0.1  ACTT-CCACAGCAACACCA------GAACTGTACTGGTA--CATCCAGAGAGA---ACATGAAATTCCCAAATACATCCTGAGGAAAAACAAATATGAAG---GAGA
V31.0    TCCT-CAAC-ATCACA--A------GAGCTTTTCTGGTA--CATTCAGAGAGC---AGATGGATCTCTAAAA------CTTCAAAGAAATTCATACGGCGGAGGAGT
V33.0.6  ACTA-CAGCTTCAGACACAAATATTTATCTCTTCTGGTA--TAAACAGCTGCC---AAACAGATCACCAACATTCATCCTGAATAAATCTCCATTTAGTGAAGGAAC
V33.0.7  ACTA-CAGCTTCAGATGGAAACATTTTTCTCTTCTGGTA--CAAACAGCTGTC---AAACAGATCCCCAACATTCATCCTGAATAAAATTCCATTTAGTGAAGGAAC
V33.0.2  ACTA-CTTCTACAAATGCAGATGTTTATCTCTTCTGGTA--CAAACAGCTGCC---AGACAGATCCCCAACATTCATCCTGAATAAATTTCCATTTAGTGAAGGAAC
V33.0.5  ACCA-CTGATATAAATATA------TACCTCTTCTGGTA--CAAACAGCTTCC---AAACAGATCACCAACATTCATTCTGAATAAACAAACATATGGTGAAGGAAT
V33.0.8  ACCA-CTGATACAAATCCA------TATCTCTTCTGGTA--CAAACAGCTGCC---AAACAGATCACCAACATTCATTCTGAATAAACTCACTTATGGTGAAGGAAC
V33.0.9  ACTC-TTTCTACAAGTCCA------GATCTCTTCTGGTA--CAAACAGCTGCC---AAACAGATCACCAACATACATTCTGAATCAATACAG---------------
V33.0.13 ACCA-CTTATACAAGTCTG------ACTCTCTTCTGGTA--CAAAAAGCTGCC---AAACAGATCACCAACATACATTCTGAATCAATACAG---------------
V33.0.10 AC----TTCACAAACTGCA------TATCTCTTCTGGTA--CAAACAGCTGCC---AAACAGATCACCAACATTCATTCTGAGTCAATACAG---------------
V33.0.11 ACAA-CTTCAACAACAGTG------TATCTCTTCTGGTA--CAAACAGCTGCC---AAACAGATCACCAACATACATTCTGAGTCAATACAG---------------
V33.0.12 -----CTTCATCAAATCCA------ACACTCTTCTGGTA--CAAACAGCTGCC---AAACAGATCACCAACATTAATTCTGAGTCAATACAG---------------
V33.0.3  ACAA-CTTTCACAAACGCT------TATCTCTTCTGGTA--CAAACAACTGCC---AAACAGATCACCGACTTTCATCCTGAGTGAATTTACCATTGGTGAAGGAAC
V33.0.4  GTCA-CGACCACAAATGCT------TATCTCTTCTGGTA--CAAACAACTTCC---AAACAGATCGCCAACATTCATGCTAAGTGCATTTACAGTTGGCAAAGGAAC
V33.0.1  ATGA-GCGCTGCTAATGCA------TATCTCTACTGGTA--CAAACAACTACC---AGACACATCACCAGCATTCATTCTGAGTGAGTTTTCCATCGGGAAAGGAAA
V18.0    ACAG-ACTC--GATGCCTG-------CTCTGCTATGGTA--CCGGCAAACAAA---GAATTTCCCACCACAATATATTTTAGTAAGGCACAATTTGGGAAATGGAAG
V27.0    TCAA-GCTC--AAACCCTA-------CTCTACTCTGGTA--CAGACAAGATGA---TTATGGCTCTCCGAAATACATTCTACTAAGATATAAATTTGGAGATGGAGA
```

```
Sequences continued:
```

```
V23.2.7  G-GTGCTGGAAACTCTTTT-------------GACAGCCGGTTTCAGTCG-ACCACAT--------CTGAAAGC---TCCACTGAACTCACTGTTACTGGTGTATCT
V23.2.1  G-TAAGTGGAGATCCAGTT-------------GATCGTAGGTTTCAGTCG-AGCACAT--------CAGACTCA---TCCACTGAACTCACTATTAGCGGTGCAACT
V23.2.2  G-GAGGTGGGAGACCCGAT-------------AATCCTCGTTTTAAGTCG-ACTACAT--------CAGACTCA---TCCACTGAACTCACTATTAGCGGTGTAACT
V23.2.3  G-GAAA----------------------------------ATTTCAAACA-GCTAAAT--------CACGCTCA---TCTACTGAACTCACTATTAGCGGTGTAACT
V23.2.5  G-GAAG----------------------------------GTTTCAAATC-GCTAAAT--------CGGACTCA---TCTACTGAACTCACTATTAGCGGTGTAACT
V40.0    G-CCAAAGACACCTCTGAT-------------G---CCCGATTTGAGTCA-AGTACAT--------CTCGCATA---TCCACTGACCTCATTATTAAGACTGCAGCT
V23.3.4  A-GTACTCATCACAAGTCT-------------GATGATCGGTTTGAGTCG-GCTACAT--------CACGTTCA---TCCACTGAACTCACTATTACTAATTTGCGT
V23.3.7  C-TTACTCATCACACTTCT-------------GATAAACGGTTTGAGTCG-GCTACAT--------CACGTTCA---TCCACTCAACTCACTATTAATGATGTGAGT
V23.3.6  T-CTGATTCACACTCCTCT-------------GATGATCGGTTTCACTCG-GCTACAT--------CACGTTCA---TCCACTGAACTTACTATTAATAATGTGCGT
V23.3.5  A-ATAATGAATACAAGTCT-------------GATGATCGGTTTCACTCG-GCTACAT--------CAAGTTCA---TCCACTGAACTCACTATTACTAATGTGCGT
V23.3.1  A-ATGATGAATACAAGTCT-------------GATGATCGGTTTCAGTCT-GCAACAT--------CAGAGTCA---TCCACTGAACTCACTATTACTAATGTGCAT
V23.3.2  G-ATTATGAAGCCACTTCT-------------GATGATCGCTTTCAATCA-GATTCAT--------CATATACA---TCCACTGAACTCACTATTAAAGATGCACGT
V23.3.8  A-GTGATTCACACACGTCG-------------GATGATCGGTTTCAGACG-ACTATAA--------CACAGACA---GCCACTGAACTCACTATTACTGATGTACGA
V23.3.3  A-GTGCTAAACACTCCTCT-------------AATGATCGGTTTCAATCC-ACAACAT--------CAGACTCA---TCCACTGAACTCACTATTACTGATGTGCGT
V22.0    T-ATGAGGACATACCAGAC-------------CGGGCTAAATTTGGATCA-GCTACGT--------CGCAGAAA---TCTACTACACTGACCATTAAAAGTGTGACT
V23.0.1  ------GAGGAGGATT-TA-------------CCTGACGGTTTTCAGTCA-AGGACAT--------CAGACTCA---ACCACTGAAATCACTATTAACAGTGTGCGT
V23.0.3  TCTAAAGAAGATCATT-CA-------------GAAGATCGTTTTCAGTCA-ATTGCAG--------CATATTCC---TCCACTGAACTCAATACTGCTGAAGCGGCT
V26.0    ----GAGGGGATTGTAACA-------------GACAGTCGTTATAAATCA-ACTGTCT--------CACCTACG---TCCACTTCACTCACCATTAACTCTGTGAAT
V23.1.3  G-GACTTGATGACATACCA-------------GACCGTCAGTTTGATTCG-ACTTCAT--------CCCAAACC---TCCACTGAACTCACTAATACTGGTGTAAGC
V23.1.5  A-GATCTGAGGACAGACCA-------------GATCATCACTTTCAATCA-TCTACAG--------GCCAAACC---ACCACAGAACTAACAATTGATAAAGCAGCT
V23.1.6  A-ACTATAAGAAAATACCA-------------GGCAGTCAGTTCGATTCC-ATTACAT--------CCCAGACG---TCCACTGAACTCATTATTAAGAGTGTATCT
V23.1.10 A-GAGATGAAGACATACCT-------------GACAGTCGGTTTGATTCG-ATTACAT--------CCCAAACC---TCTACTGAACTCATTATAAAGAGTGTATCT
V23.1.7  G-ACTATAAGAATATACCA-------------GACAGTCGATTTGATTCA-ATTACAT--------CCCAAACC---TCCACTGAACTCATTATTAAGAGTGTATCT
V23.1.8  A-TGCA---GCATATCCCA-------------GATCAGCGCTTTGAATCT-ACTACAT--------ACCAAACA---TCCACTAACCTCTTTATTAAGAGTGTATCT
V23.1.1  G-ATAGACAGTATATACCA-------------AACCGTCAGTTTGATTCA-ATTACAT--------CCCAAACC---TCCACTGACCTCATTATTAAAAGTGTAACA
V23.1.4  G-GAGA---GGATAAAACA-------------GACTCTCGGTTTAATTCG-AAAACAT--------CCCGAACT---GCAACTGAACTCACTATTTCTGCAGTAACT
V23.1.9  G-GAGA---GGATAAAATA-------------GATTCACGGTTTCGTCCG-ACTACCT--------CCCAAACC---TCCACTGAACTCACTATTCCTGGTGTAACT
V23.1.2  A-GTTACAATAATATGCCA-------------GATCCTCAGTTTCAGGCG-ACTTCAT--------CCCAAACC---TCCACTGAACTCACTATTACTGCTGTAACT
V23.0.2  A----GTGAGGATATTCCA-------------GATAGTTA---TGAATCA-GCTACAT--------CACATACA---TCCACTAAACTCATTATTAAGAGTGTATCT
V16.0    G-GTGAGAAGCATATGCCA-------------GATTCTCATTTTGAATCT-TCTACAA--------GCAGAACG---TCCACTGAGCTGATTATTAAAGTAGCGGCT
```

```
V21.0    C----AATTTCAAAGCGTA-------------GATCCTCGATACGATTCT-GCCACTT--------CAAAGACC---GCCACTTTTCTGATCATTAACAATCTAACA
V30.1.1  CATTGAGGAAGAATTCAAT-------------GAG---AGATTTCACGCA-ACTCTCAG--CAAAACT---------TCAGTTCCTCTGACAATCAAGGATCTACGA
V30.1.3  CACTGAGGAAGAATTTAAA-------------GAG---AGATTTCACGCA-ACTCTCAG--CAAAACT---------TCAGTTCCTCTGACAATCAAGGATCTACGA
V30.1.4  CACTGAGGAAGAATTCAAG-------------AAA---AGATTTAACGCA-ACTCTCAG--CAAAACT---------TCAGTTCCTCTGACAATCAAGGATCTGCGT
V30.1.2  CACTGAGGAAGAATTCAAA-------------GAG---AGATTTCACGCA-ACTCTAAG--CAAAACT---------TCAGTTCCTCTGACAATCCAGGATGTGCGA
V30.1.6  TAGTGAGGAAGAATTAAAA-------------AAAAAAAGATTTCATGCA-ACTCTCAG--CAAAACT---------TCAGTTCCTCTGACAATTCAGAATCTGCGA
V30.1.7  TAGTGAGGAAGAATTCAAG-------------AAA---AGATTTAATGCA-GCTCTCAG--CAAAACA---------TCAGTTCCTCTGACAATCAAGGATCTGCAT
V30.1.5  CACTGAGGATGAATTAAAG-------------GAG---AGATTTCATGCA-GCTCTCAG--CAAAACT---------GACATGACTTTTATACTTCATTACACTTAA
V30.0.2  AAATGACAAAGACTTTGAG-------------GAGGAGAGATTTGGTGCA-AAACACGA--CAAAACTTTAAAA---TCAGTTCCTCTGCTGATCCAGGATCTGCGT
V30.0.3  AAATGACAAAGACTTTGAG-------------GAGGAGAGATTTGGTGCA-AAACACGA--CAAAACTTTAAAA---TCAGTTCCTCTGCTGATCCAGGATCTGCGT
V30.0.1  AAATGACAAAGACTTTGAG-------------GAGGAGAGATTTAGTGCA-AAACACGA--CAAAACTTTAAAA---TCAGTTCCTCTGCTGATCCAGGATCTGCGT
V35.0    CAAAGACGACGAATTCAAA-------------AA---GAGATTTAATGCA-CGACTGAA--CACATCTTCAACT---TCAGCTCCTCTGCTGATCCAGGATCTGCGA
V30.0.4  GG------AGCAATTCAAG-------------GG---TAGATTTAGCAGC-GAACTCAA--CACAAGTTCAAAA---TCATTTCCACTGAAGATCCAGGATGTGCGT
V39.0    AG------AGCAATTCAAG-------------GA---TAGATTTAGCAGC-GAACTCAA--CACAAGTTCAAAA---TCATTTCCACTGAAGATCCAGGATGTGCGT
V30.0.5  GATTGAGAATCAATTTAAA-------------AA---GAGATTTGATGCC-GATCTCGA--CTCCTCTACAACA---TCAGTCCTGCTGAGGATCAAAGAGGTGCAT
V30.0.6  TGAAGACAAAGAATTCAAA-------------GA---AAGATTTACTGCA-AATCTCAA--CACATTTTCAAAA---TCAGTTCCTCTGATGATCAGAGATCTACAT
V29.0    CGAAGATGACAATTTTAAG-------------GAC---AGATTTAATGCA-AATCTCAA--CAGCTCTTCTACA---TCAGTTCCTCTGACAATAACAGATGTGCGA
V41.0    CACTAATGAAGAGTTTAAG-------------CAG---AGATTTAACGCA-GAACTCAA--CACATCTGCATTA---TCAATTCCTCTGACGATCAAGGATCTGCGT
V32.0.2  AAATGGCACAGAGTTTCAG-------------GAG---AGATTTCACTCT-GAACTTT--------CAACTAAT---TCAGTTCCTCTGACGATCAGAGATCTGCAT
V32.0.4  AAATGGCACAGAGTTTCAG-------------GAG---AGATTTTACTCT-GAACTTT--------CAACTAAT---TTTGTTCCTCTGACGATCAGAGATCTGCAT
V32.0.6  AAATGGCACAGAGTTTCAG-------------GAG---AGATTTCACTCT-GAACTGT--------CTTTAGAT---TCAGTTCCTCTGCTGATCAGAGATCTGCGT
V32.0.3  AAATGGCACCGAATTTCAG-------------GAG---AGATTTCACTCT-GAATTTT--------TAACAGAT---TCAGTTCCTCTGCTGATCAGAGATCTGCAT
V32.0.5  AAATGGCACAGAGTTTCAG-------------GAG---AGATTTCACTCT-GAACTTT--------TAACAGAT---TCAGTTCCTCTGCTGATCAGAGATCTGCAT
V32.0.1  AAATGGCACAGAGTTTCAC-------------AAG---AGATTTCATGTT-GAATTAA--------CAACGGAT---TCAGTTCCTCTGCTGATCAAAGATCTGCGT
V31.0    GAATGGAACAGAGTTCAAG-------------GAA---AGATTTTACTCT-GAGGTGAA--AA---CATCAAAA---TCAGTTCCTCTGCTGATTCAGGATACGTGT
V33.0.6  TATTGAGCCTGAATTTAAC-------------AAG---AGACTTTCAGCA-ACACTGAA--TTCAACATCAAAA---ACAGTTCGTCTGCTGATCCAGGATCTGCGA
V33.0.7  TATTGAGCCTGAGTTTAAC-------------AAG---AGATTTACAGCA-AAACTGAA--TTCAACATTAAGA---ACAGTTCCTCTGCTGATCCAGGATCTGCGA
V33.0.2  TATTGAGCCTGAGTTTAAC-------------AAG---AGATTTTCAGCA-AAACTGAA--TTCAACATTAAGA---ACAGTTCCTCTGCTGATCCAGGATCTGCGA
V33.0.5  TACTGCGGCTGATTTTAAG-------------GAG---AGATTTTCTGCA-ACACTGAA--CTCCACTGCAAAA---TCAGTTCCTCTGATGATCCGGGATCTGCGA
V33.0.8  TGTTGAGGCTGATTTTAAG-------------GAT---AGATTTTCTGCA-ACACTGAA--CTCCACAGCAAAA---TCGGTTCCTCTGATGATCAAGGATGCACAA
V33.0.9  CACAGAGGCTGACTTTAAA-------------GAA---AGATTTTCTGCA-ACACTGGA--CTCCAAAGAAAAA---TCATTTCCACTGACAATCAAGAATCTGCGA
V33.0.13 CACTGAGGCTGACTTTAAG-------------GAG---AGATTTTCTGCA-ACACTGGA--CTCCACAGAAAAA---TCATTTCCACTGACAATCAAGAATCTGCAT
V33.0.10 CGCTGAGGCTGATTTTACA-------------AAG---AGATTTTCTGCA-ACACTGGA--CTCCAAAGAAAAA---TCATTTCCACTGACAATCAAGAATCTGCAA
V33.0.11 CGCTGAGGCTGATTTTAAA-------------GAG---AGATTTTCTGCA-ACACTGGA--CTCCAAAGAAAAA---TCATTTCCACTGACAATCAAGAATCTGCGA
V33.0.12 CGCTGAGGCTGA---TAAG-------------AAG---AGATTCTCAGCA-ACACTGAA--CTCCAAAGAAAAA---TCATTTCCACTGCTGATCCAGGATCTACAA
V33.0.3  TACTGAGGATGAATTTAAA-------------GGG---AGATTTTCAGCA-ACACTGAA--TTCAACATTAAAA---TCTGTTCCTCTGCTGATCCAGGATCTGCGA
V33.0.4  TACTGAGGATGAGTTTAAA-------------GAG---AGATTTAACGCA-ACACTGGA--TAAAACATTAAGA---ACGGTTCCTCTGCTGATCCAGGATCTGCAA
V33.0.1  AGCAGAGCCTGAATTTAAG-------------GAC---AGATTCTCAGCA-AAGTTGGA--CTCCAAATTAGAA---ACAGTTCCTCTGACAATCAAGGATGTGCGT
V18.0    TAATTCCAAAGACTTCCCC-------------AACAACAGATATCAGTCC-CGGTTGGA--TTCTGCTATGGGA---CAAGTTACTCTGACTGTACGAAACCTTCTT
V27.0    TGAAGCTACTGAGTT---C-------------AAGGGACGATTTGGGTCC-AGACTCGA--TTCTGCTTCAAAA---TCAGTTCCTCTGACGATTCAACATCCGCAT

Sequences continued:

V23.2.7 CTGTCAGTTTCAG--CTCTTTATTA-CTGTGCTCTATTTGTTGGAG-----------
V23.2.1 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTTGGAG-----------
V23.2.2 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGGAG-----------
V23.2.3 CTGTCAGATTCAG--CTCTTTATTA-TTGTGCTCTAAGAGTTTTAG-----------
V23.2.5 CTGTCAGATTCAG--CTCTTTATTA-TTGTGCTCTACGAGTTGGTG-----------
V40.0   CTGTCAGATTCAG--CTCTTTATTA-TTGTGCTCTAAGAGTTGAAG-----------
V23.3.4 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAAGAG-----------
V23.3.7 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGCGTAGGAT-----------
V23.3.6 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGGAG-----------
V23.3.5 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAATAG-----------
V23.3.1 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAAGAG-----------
V23.3.2 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAATAGTAGGAT-----------
V23.3.8 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGGAG-----------
V23.3.3 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGGAG-----------
V22.0   GTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGGAG-----------
V23.0.1 CTGTCAGATTCAG--CTCTCTATTA-CTGCGCTCTTGA-AGTAG-------------
V23.0.3 CTGTCAGATTCAG--CTCTCTATTA-CTGTGCTGTGAAGAGCCGTCTGTACAG----
V26.0   CTGTCAGATTCAG--CCCTGTATTA-CTGTGCTCTTAAAAAAGGAC-----------
V23.1.3 CTGTCGGATTCAG--CTCTCTATTA-TTGTGCAATAAGAGTTGGAG-----------
V23.1.5 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTTGGAG-----------
V23.1.6 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGGAG-----------
V23.1.10CTGTCAGATTCAG--CTCTCTATCA-TTGTGCTCTAACAGTAGGAG-----------
V23.1.7 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAAGAG-----------
V23.1.8 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTAGTAG-----------
V23.1.1 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTATTAGTTGAAG-----------
V23.1.4 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTTTTAG-----------
V23.1.9 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTTGTAG-----------
V23.1.2 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCTCTAAGAGTTGAAG-----------
V23.0.2 CTGTCAGATTCAG--CTCTCTATTA-TTGTGCCCTATTAGTTAGAG-----------
V16.0   CTGACAGATTCAG--CTCTCTATTA-TTGTGCTCTTTGGAAAAG-------------
V21.0   CTGGAGGACAAGG--CTTCCTACTA-CTGTGTGCTGGAAGAA---------------
V30.1.1 GTGTCAGACT--G--CTATGTATTA-CTGTGCTCTGAGGCC----------------
V30.1.3 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V30.1.4 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V30.1.2 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V30.1.6 GTTCAAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAACC----------------
V30.1.7 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAGCC----------------
V30.1.5 GAAACAAACT-TG--CCACAAAATT-ATTCCTACAGAATCC----------------
V30.0.2 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V30.0.3 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V30.0.1 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V35.0   GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAGCC----------------
V30.0.4 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V39.0   GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V30.0.5 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAGCC----------------
V30.0.6 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
```

```
V29.0   GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V41.0   GCATCAGATTCTG--CTGTGTATTA-CTGTGCCCTGAAGCC----------------
V32.0.2 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAAGCC----------------
V32.0.4 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V32.0.6 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTTTGAGGCC----------------
V32.0.3 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V32.0.5 GTGTCTGATTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V32.0.1 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAGCC----------------
V31.0   GTGTGTGACTCTG--CTGTTTACTA-CTGCGCTCTGAGGCC----------------
V33.0.6 GTGTCAGACTCGG--CTGTGTTTTA-CTGTGCTCTGAGGCC----------------
V33.0.7 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTTTGAAGCC----------------
V33.0.2 GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V33.0.5 GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V33.0.8 TTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V33.0.9 GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAAGCC----------------
V33.0.13GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V33.0.10GTATCAGACTCTG--CTGTGTATTA-CTGTGCTTTGAGGCC----------------
V33.0.11GTATCAGACTCTG--CTGTGTATTA-CTGTGCTCTGAGGCC----------------
V33.0.12GTATCAGACTCTG--CTGTGTATTA-CTGTGCCCTGGAGCC----------------
V33.0.3 GCGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTAAGGCC----------------
V33.0.4 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTTTGCAGCC----------------
V33.0.1 GTGTCAGACTCTG--CTGTGTATTA-CTGTGCTCTGCAGCC----------------
V18.0   CTGTCTGATTCTT--CAGTTTACTT-CTGTGCTTTGAGGCC----------------
V27.0   GCGTCTGATTCTG--CTGTTTACTT-CTGTGCTCTGCTGCC----------------
```

Figure B-4 Multiple Sequence Alignment of T cell receptor alpha/delta V sequences from *Danio rerio*. Figure was constructed using Mega 6.0 software. Clustal W methodology used for alignment. * represent hallmark sequences

```
                                       *               ****
V30.1.3  ---CDSQDKVDQH--TKTQSAFEGDTVTIDCTYQTSYTT---PTLFWYQQKVN-GVPKHMLNKIS-TSG-----HTEEEFKE-RFHATLSKTS--------VPLTIKD-
V30.1.4  ---CDSQDKVDQH--TKTQSAFEGDTVTIDCTYQTSSTG---PTLFWYQQKVN-GVPKYMLNKIT-TTG-----HTEEEFKK-RFNATLSKTS--------VPLTIKD-
V30.1.1  ---CDSQDKVDQD--TKTQSAFEGDTVTIDCTYQTS-TN---PTLFWYQQKVK-GVPKHMLNKITNTTG-----HIEEEFNE-RFHATLSKTS--------VPLTIKD-
V30.1.6  ---CDSQDKVDQH--TKTQSAFEGDTVTTDCTYQTSSTG---PTLFWYQQKVS-GVPKHMLNKFS-TTG-----NSEEELKKKRFHATLSKTS--------VPLTIQN-
V30.1.7  ---CDSQDKVDQH--IKTQSAFEGDTVTIDCTYQTSYTT---PTLFWYQQKVN-EVPKHMLNKFS-TTG-----NSEEEFKK-RFNAALSKTS--------VPLTIKD-
V30.1.5  --GCDSQDKVDQH--TKTQSAFEGDTVTIDCTYQTSYTG---PTLFWYQQKIK-SPDIC*IKSLT-HQD-----TLRMN*RR-DFMQLSAKLT--------*LLYFIT-
V10.0.1  ----TLAQSIRPL--QSDVYITERTPVKLSC--KYDGS---VYSLHWYPQYPG-AKPEFLLLVLKSAT-KS---VTYADPKIPGMDGEMSISE------QQVDLEISS-
V10.0.4  ----TLAQSIRPL--QSDVYITERTPVKLSC--KYDGS---VYSLHWYRQYSG-AKPEFLLLVTESTS-KI---VTYADPKIPGMDGEMRMSE------KQVDLEISS-
V10.0.3  ----TLAQSIRPL--QGDVHITERTPVRLSC--KYDGS---AYSLHWYRQYPG-AKPEFLLLVRESAK-KS---VTYADPKIPGMDGEMSMSE------KQVDLEIFS-
V10.0.2  ----TLAQSIRPL--QSDVHITEKTPVRLSC--KYDGS---ANSLHWYRQHLG-SKPEFLLLVMESVI-RH---VTYADPKIPGMDGEMSMSE------KQVDLEIFS-
V10.0.5  ---ECFAQNITPL--QGTTQVTEGKTVTLSC--KYDVS---VQSLLWYRQYPG-SGLEFLLLVIESPT-KT---VIHAKPPIPRVDGEMSMSD------KRVNLEISS-
V11.0.1  ---QSFGQSITAL--QSTKQAAESETVILSC--KYDTR---ADNLQWYRHYPG-STPEYLLLVYES---NT---VFHAEPPFPRLNASVNKVE------KQVNLTISS-
V11.0.2  ---QSFGQSITAL--QSTKQAAESETVSLSC--KYDTQ---ANNLQWYRQYPG-STPEYLLLVYES---NT---VFHAEPPFPRLDASVNKVE------KQVNLTISS-
V11.0.3  ---QSFGQSITAL--QTTKQAVESETVILSC--KYDTQ---ADSLHWYRQYPG-STPEYLLLVYES---NT---VFHAEPPFPRLNASVNKVE------KQVNLTISS-
V17.0    ---TSRQDSITPT--SPAVHVKEKEAARISCEYKYTVS---MNNLQWYRQYPN-AKPEFLVLIMES---GQ---NQTATPPHPRLTTSVNKDK------KLVNLEIFP-
V38.0.1  ---ETFAQSITPL--ENKTLKSEGETVTLSC--KYEG---VVNNLHWYRQYPG-FRPKFLAYIYEHG--------ATSKPLPPRLLPKVDKNS------KRVSLEISE-
V38.0.2  ---ETFAQSITPL--ENKTLKSEGETVTLSC--KYEG---VVNNLHWYRQYPG-FRPEFLAYIYEHG--------ATSKPLPPRLLPKVDKNS------KRVSLEISE-
V38.0.3  ---DTFAQTITPL--ENKIVKSEGETASLSC--KYEG---SVTNLHWYRQYPG-SRPEFLAYIYEHG--------ATSKPLPPRLLPKVDKNS------KRVSLEISE-
V14.0    ---IIQTNMITSH--DPQKQAFEGEYVKLAC--NYSG---APDNLHWYRQYPG-STPMFLLYIYESG--------LKSDNIPERLTPKINKTK------KVVDLEISS-
V5.0.1   ---DSLGDIITSR--GSEKAAQDGTNVTLSC--NYTG---AVYNLLWYRQYQR-FKPELLLSITESG----E--AVKADPSAHWLSATVHKQS------KLTELEISP-
V5.0.2   ---EVIGEVITSR--GSKKEAQDGTKVTLSC--NYTG---VVYNLLWYRQYQR-FKPELLLSITESG----E--AVKADPSSHWLSATVHKQS------KLTELEISP-
V3.0.2   ---DAGEEVISSL--STEQIVLAGKNATLSC--KYNG---NVYNLQWYRQYPG-SKPENIIFHTESK--------NQSDFK-LRLLAVGEKGM------KRMNLSIFH-
V3.0.3   ---TAGEEVISSL--STEQIVLAGKNATLSC--KYNG---NVYSLQWYRQYPG-SKPENIIYHTESN--------KQSHSR-LRLLAVTEKGM------KRMNLSIFH-
V3.0.1   ---NAGKEVISSL--STEQIVLAGKNATLSC--KYNS---NIRNLQWYRQYPG-SKPENIIFHTESN--------NQSYSR-LRLLAVAEKEI------KRMNLSIFH-
V3.0.4   ---NSGEEVISTL--PTNQYIIEEKNATLSC--NYSG---YVQNLQWYRQYPG-SKPENIIFYTESN--------NQSDSK-LRLLAVADKGQ------KLMNLSMYF-
V7.0     -NVDADGHMINPL--SLETHGPVGHNVTLSC--NYSG---RVHTLQWFRQYPG-SRIEFLIFATELN--------DQSNAA-LRISNVVDK-Q------KSMSLSIFH-
V8.0     ---IADAYTIESL--SAEKQVSVGGNAVISC--KYSG---IGPSFQWYRQYPG-SRPEYLIFNTETG--------AGSEPT-LRLTSMAKKAL------TQVDLEISS-
V34.0.3  ---ISIADEIHPL--STDKHVAEGQNVTLSC--NYSTSTGTVNSLQWYRQYPG-AKPEFLLSLNEYS--------NSSKPD-LRLVSNAVKAL------KRVDLLLSS-
V34.0.4  ---TAIADEIQPL--SSDKRVAKGENVTLSC--NYST-TVTVDSLQWYRQYSG-AKPEFLLSLNEYS--------NSSKLD-LRLVSNAVKAL------KRVDLLLSS-
V34.0.2  ---TSSADEIHPL--FLDKQVAEGQNVTLSC--NYSISIGILRSLQWYRQYSG-AKPEFLLMLNEYY--------NKSKLD-LRLFSNAVKAL------KRVDLLLSS-
V34.0.1  ---MSFSDEIHPL--SSDKHVAEGQNVTLCC--NYRASTGNVDSLLWYRQYPG-ANPEYLLLLNEYS--------NSTKSY-LRLFSNAVKAL------KRVDLLLSS-
V34.0.5  ---NLFGDTIEPL--FPDRHGSEGENVTFSC--NYSLTSGTSVNLLWYRQYSG-AKPEYLLSVTEYS--------NSSESD-LRLFSNAVKAL------KRVDLLLSS-
V34.0.6  ---NSFGDEIQSL--VKHEHGSENENVTFSC--KYRLTSGTNVNFQRYRQYSG-AKPEYLLSVTEYS--------NSSESD-LRLFSNAVKAL------KRVDLLLSS-
V4.0.1   ---NAGEEVIRPL--FKEKHVLEGDVVTLIC--SYSG---SVQNLQWYHQFPG-SKPENIFLYFEN---------KPQTE--GRMTAASDKAV------KHMNLTITS-
V4.0.3   ---NAGEEVLRPL--FKEKHVLEGDVVTLIC--SYSG---NVLNLQLYHQFPG-SKPENIFLYFEN---------KPQTE--GRMTAASDKAV------KHMNLTITS-
V4.0.2   ----SLQWTLFIL--FKEKHVLEGDVVTLIC--SYSG---NVQNLQWYHQFPG-SKPENIFLYFEN---------MFTTK--NRMIAASDKAV------KHMNLTISS-
V9.0.1   --KAFFSNAIRPL--SAEQQVLEGGKVILSC--AYNGS--NIKSLQWYKQHHS-PAPEYLLQTFENA--------APEQK--DRLLAKAQKDK------KQLDLEISK-
V9.0.2   ---GIFANSVTPW--STKQQVLEGENVTLSC--NYSGS--NIQSLQWYRKYPSSTAPEYILQTFESA--------EPAQQ--GRHLAKAQKNK------KQLDLLISK-
V6.0.1   ---HTEADVIKPL--SLDENSHVGDAVTLSC--SYKDYTGNMQNLQWYRQFPR-SKPEFILYIFESG--------TMSKDKPDRFSAQVNKGT------KQVDLNISP-
V6.0.2   ---HTEADVIKPL--SLDKNCHVGDAVTLSC--SYKDYTGTVQNLQWYRQFPR-SKPEFILYIFESG--------TMSKDKPDRFSAQVNKDT------KQVDLNISS-
V6.0.4   ---HTEADVIKPL--SLDENSHVGDAVTLSC--SYKDYTGTVNGLQWYRQLPR-SKPEFILYIFESG--------TLSEDKPDRFSAQVNKDT------KQVDLNISS-
V6.0.3   --QHTEADVIKPL--SLDENKHVDDTFTLSC--SYKDYTDTVRNLQWYRQLPR-SKPEFILYIFESG--------TMSKDKPDRFSAQVNKDS------KQVDLNISS-
V12.0.2  ---DTSGEGIAHH--SSSEFFINGHSTTLSS--NYNGS-YSSDSLLWY*QYSS-SKPQFLYLVSEA----KL--EQPADPPIPGLPAKLNEEN------NLVYLEISS-
V13.0    ---DAFGEGIAPV--SGYKTVPEEKTTTLSC--IYNGS-AATDALQWYRQYPR-SRPDFLFLVNEA----AF--KQPANPPIPGISARLNEEK------NCVYLDITS-
V2.0.1   ---VVNGNVITPA--RTSVVLTEGSNITLSC--TFNTS-PYR--LLWYRQTPG-SRPEFLLLIVRST---KT--TIEASPPQPHMSINLSE--------KLVDLEISS-
V2.0.3   ---DSNGNVIRPE--KTSVVLTEGSNITLSC--TFNIT-PYS--LHWYRQMPG-SRPEFLLLIVRST---KT--KVAASPPQPHMSINLSE--------KRVDLEISS-
V2.0.4   ---NANENSIRPD--EEFVSLNEGSNTTLSC--TFDGS-AYR--LHWYRQKTG-SLLEFLVLIDKST---KV--VVTASPPHPHVSINLHD--------NRVDLEITP-
V2.0.2   ---DSNGDIIRPE--QTSVVVTDGSSITLSC--TYDES-AYS--LHWYRQTPG-SRPEFLLLIDESS---KY--VTKAEQPHPHMSIQLQNN-------KRVNLILSF-
V1.0     ---ETTGDKIKPD--NDTAQLMENSNITLSC--TYEGS-IDS--LYWFQQKIG-SRLEFLIMIDEAS---EY--VTNAIPPHPRLSVQLHKAQ------KRVSLEMSS-
V24.0.1  ---NLAKDVITPY--SDEIFASETESVKFSC--NYTGS---VDGLHWYRQYPD-SPPHFLILDYYGS-------ITHANPPIAGISIKHRKDN------SSVDLEISS-
V24.0.2  ---NRAKDAITPY--RDAVFASESERVKLSC--NYTGS---VDSLHWYRQYPG-SRPHFLILDYSGF-------VTHANPPFDGISIKHRKEN------SSVDLLISS-
V24.0.4  ---SSATEAITPY--RAYEFAFKDDRVILSC--NYTGS---ARGLHWYRQYPG-LPPHFLILDYYGA-------VTHAKPPIAGISIKHRKDD------SSVDLEISS-
V24.0.3  ---SRASEVITPK--TDREFAVEGDNVTLSC--NYSGS---VRSVHWYRKYSG-SPPQFLILEYSGV-------ITPADPPVPGISINHRKKQ------SHVDLLLSS-
V24.0.5  ---SRALEAIDPK--TDREFAVEGDNVTLSC--NYSGS---VRGVHWYRQYAA-SPPQFLILEDSGV-------TTPANPPVPGISINHRKAE------KHVDLLLSS-
V28.0    --VSSAEEIIFPN--NKTVIAKADEKVLLSC--SYTGT---VYDLYWYRQSPG-SAPQFLIMESSGV-------IIHATPQVPGIKIINKKNT------KHVELEIFS-
V36.2.10 ---AVFGNVIKPI--NTDVFAEEGSNVTLSC--SFTASG-VTDYLHWYRQYGR-SKPEFLVLTYSSAT------KAAQSDIDTRFTTNITKKE-------HVDLLLSS-
V36.2.2  ---AVFGNVIKPI--NTDVFTEERSNVTLSC--SFTASG-GSDYLHWYRQYGG-SKPEFLVLIYGSAT------KAAQSDIDTRFTANITKKE-------HVDLLLSS-
V36.2.8  --PAVFGNVIKPK--QTDVFADEESNVTLSC--SFTASG-GSDYLHWYRQYGR-SKPEFLVLTYRTAT------KAAQSDIDTRFTTNVTKKE-------HVDLLLSS-
V36.2.11 FSAAVFGNVIRPI--NTDVFAEEGSNVTLSC--SFTASG-GTDYLHWYRQYGG-SKPEFLVLTYSSAK------KSKESDIDPKFSANVSKSE-------HVDLLLSS-
V36.2.9  TPH*KIKCVLDLLVCKTDVFVEEGSNVTLSC--SFTASG-GSDDLHWYRQYGG-SKPEFLVLTYRTAK------EAEKSKVDPRFIANVPKSE-------HVDLLLSS-
V36.2.1  ---SVYGNVIIPD--ETEAVHAKGSDVKLSC--SFLASG-YTDYVYWYRQYRR-SKPEYLVQIISNSN------KAQQSLVDPRFTAYIENRDK-----IHVDLLLSS-
V36.2.12 ---TVFGNTIKPD--ETNVVAQKDSDVKLSC--SYSST--TTDTLYWYRQYER-SKPEFLVLIYSSAK------AAEKSKVDPRFFVKVEKRQQ-----IHVDLLLSS-
V36.2.5  ---SVLGNDIKPN--ITDVVYEEGSDGTIFC--SHSPG--AGDYIYWYRQYGG-SKPEFLVLIFGSTK------DTETSKVDPRFSVTVEKKER-----NHADLLLSS-
V36.2.7  ---SVFGNDIKPN--STD-VYEEGSDGTIFC--SHSPG--AGDYIYWYRQYGG-SKPEFLVFTFGSVK------DAETSKVDPRFSVKIEKKEQ-----NHADLKISS-
V36.2.6  ---SVFGSDIKSN--ITDVVYKEGSDGTIFC--SHYPG--TGDNIFWYRQYGG-SKPEFLLATYGSAK------DTETSKVDPRFSVKVEKKER-----NHADLKISS-
V36.2.4  ---SVFGNDIKPN--STEVVLKEGSDGTIFC--SHSPG--TGDYIHWYRQYGG-SKPEFLVLIYGTAN------KGETSKVDPRFSVKVEKKER-----NHADLKISS-
```

```
Sequences continued:

                        ***
V30.1.3  -LRVSDS------AVYYCALR----
V30.1.4  -LRVSDS------AVYYCALR----
V30.1.1  -LRVSDC------YVLLCSEA----
V30.1.6  -LRVQDS------AVYYCALK----
V30.1.7  -LHVSDS------AVYYCALK----
V30.1.5  -LKKQTC------HKIIPTES----
V10.0.1  -AAVSDS------AVYYCALQ----
V10.0.4  -AAVSDS------AVYYCALQ----
V10.0.3  -AAVSDS------AVYYCALE----
```

```
V10.0.2  -ATISDS------AVYYCALQ----
V10.0.5  -AEVTDS------ALYYCALV----
V11.0.1  -AAVTDS------AVYYCALQ----
V11.0.2  -AAVTDS------AVYYCALQ----
V11.0.3  -AAVTDS------ALYYCALQ----
V17.0    -ADVKDS------AVYYCALQ----
V38.0.1  -AEVTDS------ALYYCALT----
V38.0.2  -AEVTDS------ALYYCALT----
V38.0.3  -AEVTDS------ALYYCALT----
V14.0    -ATVSDS------AVYYCALK----
V5.0.1   -AKVTDS------AVYYCALK----
V5.0.2   -AKVTDS------AVYYCALK----
V3.0.2   -TEMQDS------ALYYCALQ----
V3.0.3   -TEMEDS------ALYYCALE----
V3.0.1   -TEMQDS------ALYYCALE----
V3.0.4   -TKVEDS------ALYYCALQ----
V7.0     -AEIQDS------ATYYCALE----
V8.0     -TEVKDS------AMYYCALQ----
V34.0.3  -AAVSDS------AVYYCALQ----
V34.0.4  -AAVSDS------AVYYCALQ----
V34.0.2  -AAVSDS------AVYYCALQ----
V34.0.1  -AAVSDS------AVYYCALR----
V34.0.5  -AAVSDS------AVYYCALR----
V34.0.6  -AAVSDS------AVYYCALQ----
V4.0.1   -TQLTDS------AVYFCALV----
V4.0.3   -TKVKDS------AVYYCALQ----
V4.0.2   -AEITDS------AVYYCALQ----
V9.0.1   -AEMSDS------AVYYCALV----
V9.0.2   -AEMTDS------AMYYCALV----
V6.0.1   -AVETDS------AMYYCALV----
V6.0.2   -AVETDS------AMYYCALV----
V6.0.4   -AVETDS------AMYYCALV----
V6.0.3   -ALETDS------AMYYCALV----
V12.0.2  -AALSDS------AVYYCALT----
V13.0    -TAISDS------AVYYCALQ----
V2.0.1   -VAVSDS------AVYYCALE----
V2.0.3   -AAVSDS------AVYYCALE----
V2.0.4   -AAVSDS------AVYYCALQ----
V2.0.2   -AAVSDS------AVYYCALQ----
V1.0     -AAVTDS------ALYFCALR----
V24.0.1  -VAVSDS------AVYYCALQ----
V24.0.2  -AAVSDS------AVYYCALR----
V24.0.4  -AAVSDS------ALYYCALE----
V24.0.3  -AAVSDS------AVYYCALQ----
V24.0.5  -AAVSDS------AVYYCALR----
V28.0    -ASVEHS------AVYYCAAR----
V36.2.10 -AAVSDS------AVYYCALR----
V36.2.2  -AAVSDS------AVYYCALQ----
V36.2.8  -AAVSDS------AVYYCALQ----
V36.2.11 -AAVSDSGPCIIRSFDYSFYYG---
V36.2.9  -AAVSDS------AVYYCALR----
V36.2.1  -AAVSDS------AVYYCALR----
V36.2.12 -TAVSDS------AVYYCALQ----
V36.2.5  -AAVSDS------AVYYCALQ----
V36.2.7  -AAVSDS------AVYYCALR----
V36.2.6  -AAVSDS------AVYYCALR----
V36.2.4  -AAVSDS------AVYYCALQ----

Next set of sequences:

                                       *               ****
V36.2.3  ---SVSGNLIKPI--QTDVFADEGSNVTLSC--SYSSA----DYLFWYRQYPR-GAPEFLVSIFEGER------NSKSSDVDPRLSAKLTKSEH-----NHADLMLSC-
V36.1.2  ---ATYADEIKPV--KTEVFADEGSNVTLSC--TYSSA----NNLQWYRQYPG-SAPQYLVLIPDSSK------EDWKSNVDDRFTARVTKNKE-----KRVDLLLSS-
V36.1.3  ---AIYADEIKPV--KSEEFADGGSNVTLSC--TYSSA----YYLYWYRQYPG-SAPQHIVLIPESRK------DDTKSDVDHRFIARVTKDKE-----KRVDLLLSS-
V36.1.1  ---ATYGDEIKPV--KTEKFADKGSSVKLSC--TYSSA----NTLFWYRQYPG-SAPQYLVLIPEGTK------DDWKSDVDKRFTARITKDKE-----KHVDLLLSS-
V36.1.6  ---VAYGNEIKPV--KTEEFAKEDSSVTLSC--SYSSA----RTLFWYRQYPG-LAPQFLVFFIHGTS------EAKESGVDPRFTAKPNGEKQ-----SHVDLLLSS-
V37.0    ---VIFGDEISPD--SSEVFAAVASKVTLSC--RYSSA----YNLQWYRHYPG-SAPQNLVLISEGLT------ETKISEVDSRFTTKLRKENQAGKEIKRVDLLLSS-
V36.1.4  ---ASNGDTITPD--KTEEFADEDSNLTLSC--SYSSA----WNLLWYRQYPG-SALEYLVLTVDTVN------EARTSDVDVRFATKTRKENA-----NHVDLEISS-
V36.1.5  ---ASSEDTITAD--KTEEFADDGSNVTLSC--TFSSA----WSLLWYRQYPG-SAPEFLVNIAERSK------KVERSNTDHRFTTKIRIEKE-----KHVDLLLSP-
V19.0    ---GKSDSPITSD--QTEVFPRENANVTLSC--SYSSA----ISLHWYRHLPR-SAPEFLLLVLQTTGRV-----QTAQNLDSRFEAMLNEKKN------KVHLIISS-
V20.0    ---VTHADAITPD--RPEVNVNETESITLSC--QYSSA----FSLHWYRQHDA-SGPQFLMLIQQSTG-------QMIQRSDDRLSGKLNKEKN------RVDLQISS-
V25.1.2  ---VSFGNVITSV--ETRLLHSEGDVISVSC--SYSSA----RTLFWYRQNPG-TVPQFLFVILESNGDVLQ--KSEAVK-DPRFSAKVNEKKT------HVFLEISS-
V25.1.3  ---VSFGNVITSV--ETRLIHTEGDVISVSC--SYSSA----RTLFWYFQNPG-SVPQFLFIILHSNGEVLQ--KSEAVK-DPRFSAKVNEKKT------HVFLEISS-
V25.1.1  ---DVLGNVIIPV--HSEVYKTEKETISLSC--NYSSA----SSLYWYRQYPK-STPEFLFVILHATGKVLQ--KSDIVDQDPRFSGKLNEEKT------HTSLNISS-
V25.1.4  ---VTFGNDITPV--QTEVFGAEMDNITVSC--KYSSA----VSLQWYRQYPG-SAPDFLLTIHR-TGSVSR--SSTIVNEDPWYYGKLNKEKT------HVFLEIST-
V15.0    ---VSCGDVITPA--QTEVYGKSGNKVTLSC--NYSTA----LSLYWYRQYPG-SNPHCLLLILT------K--SETSENMQPGMSVRLHKEKA------SMDLKISS-
V23.2.8  --GVTNADQIGS-NIGENVFSIEGETATLNCSYATSSR---NVRLYWFRQYHN-KEPQYLIWKDCRSYS------NTVTSADPRVEPKTSETT--------AELIIQS-
V23.2.4  ---VTNADQIGP-NKGENAISKEGETATMSCSYDTSSR---YVRLYWYRQYTN-KELQYLIWKDARSWS------GVATPDDPRIRSTTSETT--------TELIITG-
V23.2.6  ---V*RMQIRLD-QIRVKIPSVKRE--RLIC*-----------PLLVYRQYTN-KELQYLIWKEARSWS------ATGTPDDPRIRASTSETT--------TELIITG-
V23.2.1  ---VMTADQIRP-NKEAFTVKEE-ETVTFSCSYETSSS---YVWLYWYRQHLN-GEPQYLIFKPAKSAS------VSGDPVDRRFQSSTSDSS--------TELTISG-
V23.2.2  ---VMTADQISP-NKEALTVKEE-ETVTFSCSYDTSSS---YVRLYWYRQYLN-GEPQYLLFKAARSSS------GGGRPDNPRFKSTTSDSS--------TELTISG-
V23.2.3  ---VMTADQIRP-NKDVTVIKEE-ETVTLSCSYE-SSS---TVRLYWYRQYLN-GEPQFLVSSAQSSSG--------------KFQTAKSRSS--------TELTISG-
V23.2.5  ---VMTADQIRP-NKDSTVIKEE-ETVTLSCSYE-SSN---YVRLYWYRQYLN-GEPQYLIHSAGSNIG--------------RFQIAKSDSS--------TELTISG-
V23.2.7  ---MMSVDKIGP-DKDTNVISKEGETVTLSCSYDTQSS---YVRLYWYRQYAD-REPQYLIWKEARSLR------GAGNSFDSRFQSTTSESS--------TELTVTG-
V23.3.1  ---MVCGDSIEP-DKGKEKTSEEIENVKLSCSYSTSSS---YVCLYWYRQYLN-GEPQYLVHRYARSWG------NDEYKSDDRFQSATSESS--------TELTITN-
V23.3.5  ---MVCGDSIEP-DKGTKKKSEETESVKLSCSYSTSGE---DVLLYWYRQYLN-GETQYLLRRYARSLG------NNEYKSDDRFHSATSSSS--------TELTITN-
V23.3.4  ---MVCEDSIEP-DKGTEKTSEETENIKLSCLYTTSSD---YVWLYWFRQYLK-GEPQYLLYKGARSY-------STHHKSDDRFESATSRSS--------TELTITN-
V23.3.7  ---MVCGDSIEP-DKGTEKNSEETENVKLSCSYSTSSE---YVYLYWYRQYLT-GELQYLLRKGAQSY-------LTHHTSDKRFESATSRSS--------TQLTIND-
V23.3.6  ---MVCGDSIGP-DKGTEKT-EETETVKLRCSYSTNSE---YVWLYWYKRSLN-KEPQYLLFESARSA-------SDSHSSDDRFHSATSRSS--------TELTINN-
```

```
                                       *               ****
V23.3.2  ---MVCGDSIGP-DERTENTSEETKTVKLSCSYSASIE---YVVLYWYRQYKH-GEPQYLLRRYARSYN------DYEATSDDRFQSDSSYTS--------TELTIKD-
V23.3.8  ---MVCGDSIEP-NKGMEKTIKETEPVQLSCAYTTATSR-SYVYLYWFRQSFN-GEPQYLLNKNAQTA-------SDSHTSDDRFQTTITQTA--------TELTITD-
V23.3.3  ---MVFGDNIEP-EE-KDVMTKEREAVKLACSYSTTNN---RVRLYWYRQNPN-AELLLLTYKGARSL-------SAKHSSNDRFQSTTSDSS--------TELTITD-
V40.0    ---VTAVDKIEP-NSGSKVIRTERESVKLSCSYDSNNE---YVRLYWYRRHSS-GELQYLLYKGARSN-------SAKDTSDARFESSTSRIS--------TDLIIKT-
V23.0.1  ----GYTDRIGPGEGDTNIISMEGESVTLSCTYETNSN---DIWLYWYRQFSN-REPEYILFKGARSRS------E-EDLP-DGFQSRTSDST--------TEITINS-
V23.0.2  ------TDRIGP-EGETNILQEEGQSVTLSCTYETNSN---TIYLYWYRQYPK-TKPEFILYKGARSWS------S-EDIP-DSYESATSHTS--------TKLIIKS-
V23.0.3  ----ALSDSIGPNEKETNVMSREGESMRLSCSFESSSN---YVLLYWFRQNPD-GELQYLLYKGARSWS------SKEDHSEDRFQSIAAYSS--------TELNTAE-
V23.1.10 ----VFADSIRP-DKDT-MTSYEGESVTLKCSYDTSSN---NVLLYLYKQNVN-TEPQYLLWKKARSYN------RDEDIPDSRFDSITSQTS--------TELIIKS-
V23.1.6  ----VFADSILP-DKDSNMTSNEGESVTLKCSYDTSSN---NVLLYWYRQYAN-TEPQYLLWKGARSNS------NYKKIPGSQFDSITSQTS--------TELIIKS-
V23.1.7  ----VFADSIRP-DKDTNIIRSEGESVTLRCSYETNSN---YIRLYWYRQYVN-TEPQYLLWKGARSDS------DYKNIPDSRFDSITSQTS--------TELIIKS-
V23.1.3  ----VFADSIKPKGKDK-IIRSEGESVTLSCEYETSSN---NVMLYWYRQYTN-REPQFLLRKGARSRN------GLDDIPDRQFDSTSSQTS--------TELTNTG-
V23.1.5  ----VFSDSIGPKDEDKHIIRREGESVTLSCSYETQSN---NVWLYWYRQYTN-KEPQYLLLKGARSNS------RSEDRPDHHFQSSTGQTT--------TELTIDK-
V23.1.1  ----VFADSLQPKDKDK--IISEGQSVTLSCTYDTSSS---YVYLYWYKHNIK-GEPKFVLCKGARSRS------DRQYIPNRQFDSITSQTS--------TDLIIKS-
V23.1.8  ----VFPNNIWPKDKEV-IVSNEGESVTLTCSYSSGSN---DVWIYWYKYLSN-TEPKFLLWKGARAQN-------MQHIPDQRFESTTYQTS--------TNLFIKS-
V23.1.4  ----VFADSIGPNDKETSVSSNEGESVTLSCSYDAQSN---NIRLYWYRQNSN-REPEYLLWKGARSYG------GEDKT-DSRFNSKTSRTA--------TELTISA-
V23.1.9  ----VLPDSIGPTDKDAIKTRKEGESVTLSCSYDTQSN---NVYLFWYRQFTN-REPEYLLIKGARSYG------GEDKI-DSRFRPTTSQTS--------TELTIPG-
V23.1.2  ----VLSDSIQPTDKKTNTDTKEEEHVTLSCSYDTRSE---YVYLYWYRQYPN-REPEYLLRKGARSWS------SYNNMPDPQFQATSSQTS--------TELTITA-
V26.0    ----VFADSIAPKKEGRNKTRKEGESVTLSCTFNRSDT---AISLYWYMQYPN-RELQYLLYKGHRSSE------E-GIVTDSRYKSTVSPTS--------TSLTINS-
V16.0    ----AFGDTIGP--DEPLIFRSETDRVTLSCSYQTNSE---HVDLYWYRHYPK-SEPQYLLWKGARARD------GEKHMPDSHFESSTSRTS--------TELIIKV-
V22.0    ---AMCQDQIGP--SRTSVTSTEGDSVNLECSYDSSIN---NVFLYWYRKYSN-AQPQYILLKGGRSQT------YEDIPDRAKFGSATSQKS--------TTLTIKS-
V21.0    ---ITFCAHIAP--NTKFVIGREGESATLSCLYESSSS---HVLLFWGRILPK-QAPEFLLLKGARSET------QFQS-VDPRYDSATSKTA--------TFLIINN-
V27.0    ---SRAQ-DSVYQPKTLQSANEEDAAVTLDCKYSTSSSN---PTLLWYRQDDY-GSPKYILLRYKFGDG-----DEATEF-KGRFGSRLDSAS------KSVPLTIQH-
V41.0    ---CKAQ-DNVIQPKQVMTAY-EDGTITLECTYKASSTQY--PNLLWYQQMTN-GSPKYMLMRLS-GSS-----DTNEEF-KQRFNAELNTSA------LSIPLTIKD-
V32.0.2  ---CRSE-DRVDQPDKQLTEA-EGGSVTLQCKYKTDSVQ---ENLFWYIQRTN-DLPKYILRRSKYG-G-----ENGTEF-QERFHSELST--------NSVPLTIRD-
V32.0.4  ---CRSE-DRVDQPDKQLTEA-EGGSVTLQCKYKTDSTQ---EDLFWYIQRIN-DQPKYILRRSKYG-G-----ENGTEF-QERFYSELST--------NFVPLTIRD-
V32.0.6  ---FRSE-DRVDQPEKEITVA-EGGSVTLQCKYESYSVQ---KDIFWYVQRTN-DFPKYILRRSNFG-G-----ENGTEF-QERFHSELSL--------DSVPLLIRD-
V32.0.3  ---CRSE-EKVDQPDKHMTEV-EGKSVTLRCTFTTTSNQ---PDLFWYIQRAN-DHPKYILRRDKYGAG-----ENGTEF-QERFHSEFLT--------DSVPLLIRD-
V32.0.5  ---CRSE-EKVDQPDKHMTEV-EGKSVTLRCTFTTTSNQ---PDLFWYIQKAN-DHPKYILRRDKYGAG-----ENGTEF-QERFHSELLT--------DSVPLLIRD-
V31.0    ---CHSD-DTVDQPLKHVTAL-EGISVTLDCKYTTSSTS---QELFWYIQRAD-GSLK--LQRNSYGGG-----VNGTEF-KERFYSEVKTS-------KSVPLLIQD-
V32.0.1  ---CRP--DSVNQPDTHVTKS-EDGAVTLQCKYTATSTAT--PELYWYIQREH-EIPKYILRKNKYE-G-----ENGTEF-HKRFHVELTT--------DSVPLLIKD-
V29.0    ---CRAQ-DSVTQTATVQTAT-EGQAVTIECTYK-TNAF---PTLYWYQYKAN-QFPKYMLKRYAGSSD-----EDDNFK--DRFNANLNSSS------TSVPLTITD-
V30.0.6  ---IESQ-DNVEQKARVQTAV-EGEAVEIDCTYQ-TQLT---PTLFWYQQKTN-TIPKYMLNRVGKTGG-----EDKEFK--ERFTANLNTFS------KSVPLMIRD-
V35.0    ---CKCK-DSVHQKR-VETAV-EGTDLFISCSYTSTDPI---PTLFWYQQKVN-EFPKYMLSTFE---S-----KDDEFK--KRFNARLNTSS------TSAPLLIQD-
V30.0.5  ---FVSP-DSVEQKTRVQTAV-EGESVTINCTYSTTDTF---PYLFWYQQQVN-GIPQYMLKKFPGSGE-----IENQFK--KRFDADLDSST------TSVLLRIKE-
V39.0    ---RVSL-DKVEQNSRVETVV-EDEDVIINCQYETTDTL---PYLFWYQQKVN-RFPEYMLNRFSSD--------EEQFK--DRFSSELNTSS------KSFPLKIQD-
V30.0.2  ---CDCA-ESVDQNTRVETSV-EEGSVSISCTYQTSDPS---PNLFWYQQKPN-TIPKHMMTIITTTVQ-----NDKDFE-EERFGAKHDKTL------KSVPLLIQD-
V30.1.2  ---CDCA-ESVDQNTRVETSV-EEGSVSISCTYQTSDPS---PNLFWYQQKPN-TIPKHMMTIITTTVQ-----NDKDFE-EERFGAKHDKTL------KSVPLLIQD-
V30.0.3  ---CDCA-ESVDQNTRVVTFV-EEGSVTINCTYQTSDPS---PYLFWYQQKPN-TIPKHMMTIITTTVQ-----NDKDFE-EERFGAKHDKTL------KSVPLLIQD-
V30.0.1  ---CDCA-ESVDQNTRVETAV-EGGSVTINCTYQTSDPS---PYLFWYQQKPN-TIPKYMMMIFATTVQ-----NDKDFE-EERFSAKHDKTL------KSVPLLIQD-
V33.0.10 ---CRGQ-DRVEQVEGEMSDS-EGNHVIISCNY-STSQT---AYLFWYKQLPN-RSPTFILSQY----------SAEADF-TKRFSATLDSKE------KSFPLTIKN-
V33.0.11 ---CRGQ-DRVEQAEGEMTES-ERNHVIISCNYITTSTT---VYLFWYKQLPN-RSPTYILSQY----------SAEADF-KERFSATLDSKE------KSFPLTIKN-
V33.0.13 ---CRGQ-DRVEQPIREMTES-IGDKVVISCIYSTTYTS---LTLFWYKKLPN-RSPTYILNQY----------STEADF-KERFSATLDSTE------KSFPLTIKN-
V33.0.9  ---CRGQ-DRVEQPSIEMTES-EVEEVIIRCNYSTLSTS---PDLFWYKQLPN-RSPTYILNQY----------STEADF-KERFSATLDSKE------KSFPLTIKN-
V33.0.12 ---CRGQ-DKVEQPIREITEG-EGAHVFVQCNY--TSSN---PTLFWYKQLPN-RSPTLILSQY----------SAEAD--KKRFSATLNSKE------KSFPLLIQD-
V33.0.6  ---CRGQ-DKVQQPSGEKTEN-EGDQVNLMCNYTTTASDT-NIYLFWYKQLPN-RSPTFILNKSPFSEG-----TIEPEF-NKRLSATLNSTS------KTVRLLIQD-
V33.0.7  ---CRGQ-DKVQQPSGEMTEN-EGDQVSLMCNYTTTASDG-NIFLFWYKQLSN-RSPTFILNKIPFSEG-----TIEPEF-NKRFTAKLNSTL------RTVPLLIQD-
V33.0.2  ---CRAQ-DRVEQPTRKKTEN-EGDQITLMCNYTTTSTNA-DVYLFWYKQLPD-RSPTFILNKFPFSEG-----TIEPEF-NKRFSAKLNSTL------RTVPLLIQD-
V33.0.5  ---CRGQ-DRVEQPLRSISAN-ERGQVIISCSYSTTDIN---IYLFWYKQLPN-RSPTFILNKQTYGEG-----ITAADF-KERFSATLNSTA------KSVPLMIRD-
V33.0.8  ---CRGQ-DRVEQPLRDITAN-EGDQVIIPCNYSTTDTN---PYLFWYKQLPN-RSPTFILNKLTYGEG-----TVEADF-KDRFSATLNSTA------KSVPLMIKD-
V33.0.3  ---CRGSSDNVEQPSGQITAN-EADSVHLQCSYSTTFTN---AYLFWYKQLPN-RSPTFILSEFTIGEG-----TTEDEF-KGRFSATLNSTL------KSVPLLIQD-
V33.0.4  ---CRGQ-NRVEQPSREMTAD-KGDTVTLPCNYSVTTTN---AYLFWYKQLPN-RSPTFMLSAFTVGKG-----TTEDEF-KERFNATLDKTL------RTVPLLIQD-
V33.0.1  ---CGEQ-DSVDQPDRELTSA-EGNEVILLCNYKLMSAAN--AYLYWYKQLPD-TSPAFILSEFSIGKG-----KAEPEF-KDRFSAKLDSKL------ETVPLTIKD-
V18.0    ---CRCE-DAVRQPTASVSAF-TEETITLSCEYYTTDSM---PALLWYRQTKN-FPPQYILVRHNLGNG-----SNSKDFPNNRYQSRLDSAM------GQVTLTVRN-
V30.1.1  --DVILRTKLIRT--QRLNLLLKVIL*PLTAHIRLQQT-----QLYSGTSRKSKESPNTC*IKSPTQLD------TLRKNSMRDFTQLSAKLQ---------FL*QSR-
V12.0.1  ----LYSWRRNS----SSLFI*IGHQWSLHN---FVL*------LQWLLQQ*LSHVVSTIQQLQTTVSVFG**SQADPTS*SSNFWTIC*IK*------GE*PCVSGDL
```

Sequences continued:

```
                        ***
V36.2.3  -AAVSDS------AVYYCALR----
V36.1.2  -AAVSDS------AVYYCALE----
V36.1.3  -AAVSDS------AVYYCALE----
V36.1.1  -AAVSDS------AVYYCALR----
V36.1.6  -AAVSDS------ALYYCAME----
V37.0    -AAVSDS------AVYYCALR----
V36.1.4  -AAVSDS------AVYYCALK----
V36.1.5  -AAVSDS------AVYYCALTSM--
V19.0    -VKVTDS------ALYYCAME----
V20.0    -STIKDS------ALYFCALQ----
V25.1.2  -AKVTDS------AIYYCALE----
V25.1.3  -AKVTDS------AIYYCALE----
V25.1.1  -AKVTDS------AMYFCAME----
V25.1.4  -AKVSDS------AVYYCALQ----
V15.0    -AKLTDS------AVYYCALK----
V23.2.8  -VTLSDS------ALYYCALR----
V23.2.4  -VTLSDS------ALYYCALRVG--
V23.2.6  -VTLSDS------ALYYCALIVG--
V23.2.1  -ATLSDS------ALYYCALRVG--
V23.2.2  -VTLSDS------ALYYCALRVG--
V23.2.3  -VTLSDS------ALYYCALRVL--
V23.2.5  -VTLSDS------ALYYCALRVG--
V23.2.7  -VSLSVS------ALYYCALFVG--
V23.3.1  -VHLSDS------ALYYCALRVR--
V23.3.5  -VRLSDS------ALYYCALRVI--
V23.3.4  -LRLSDS------ALYYCALRVR--
V23.3.7  -VSLSDS------ALYYCALSVG--
V23.3.6  -VRLSDS------ALYYCALRVG--
V23.3.2  -ARLSDS------ALYYCALIVG--
```

```
V23.3.8  -VRLSDS------ALYYCALRVG--
V23.3.3  -VRLSDS------ALYYCALRVG--
V40.0    -AALSDS------ALYYCALRVE--
V23.0.1  -VRLSDS------ALYYCALEV---
V23.0.2  -VSLSDS------ALYYCALLVR--
V23.0.3  -AALSDS------ALYYCAVKSRLY
V23.1.10 -VSLSDS------ALYHCALTVG--
V23.1.6  -VSLSDS------ALYYCALRVG--
V23.1.7  -VSLSDS------ALYYCALRVR--
V23.1.3  -VSLSDS------ALYYCAIRVG--
V23.1.5  -AALSDS------ALYYCALRVG--
V23.1.1  -VTLSDS------ALYYCALLVE--
V23.1.8  -VSLSDS------ALYYCALRVV--
V23.1.4  -VTLSDS------ALYYCALRVL--
V23.1.9  -VTLSDS------ALYYCALRVV--
V23.1.2  -VTLSDS------ALYYCALRVE--
V26.0    -VNLSDS------ALYYCALKKG--
V16.0    -AALTDS------ALYYCALWK---
V22.0    -VTVSDS------ALYYCALRVG--
V21.0    -LTLEDK------ASYYCVLEE---
V27.0    -PHASDS------AVYFCALL----
V41.0    -LRASDS------AVYYCALK----
V32.0.2  -LHVSDS------AVYYCALK----
V32.0.4  -LHVSDS------AVYYCALR----
V32.0.6  -LRVSDS------AVYYCALR----
V32.0.3  -LHVSDS------AVYYCALR----
V32.0.5  -LHVSDS------AVYYCALR----
V31.0    -TCVCDS------AVYYCALR----
V32.0.1  -LRVSDS------AVYYCALK----
V29.0    -VRVSDS------AVYYCALQ----
V30.0.6  -LHVSDS------AVYYCALQ----
V35.0    -LRVSDS------AVYYCALK----
V30.0.5  -VHVSDS------AVYYCALK----
V39.0    -VRVSDS------AVYYCALR----
V30.0.2  -LRVSDS------AVYYCALR----
V30.1.2  -LRVSDS------AVYYCALR----
V30.0.3  -LRVSDS------AVYYCALR----
V30.0.1  -LRVSDS------AVYYCALR----
V33.0.10 -LQVSDS------AVYYCALR----
V33.0.11 -LRVSDS------AVYYCALR----
V33.0.13 -LHVSDS------AVYYCALR----
V33.0.9  -LRVSDS------AVYYCALK----
V33.0.12 -LQVSDS------AVYYCALE----
V33.0.6  -LRVSDS------AVFYCALR----
V33.0.7  -LRVSDS------AVYYCALK----
V33.0.2  -LRVSDS------AVYYCALR----
V33.0.5  -LRVSDS------AVYYCALR----
V33.0.8  -AQLSDS------AVYYCALR----
V33.0.3  -LRASDS------AVYYCALR----
V33.0.4  -LQVSDS------AVYYCALQ----
V33.0.1  -VRVSDS------AVYYCALQ----
V18.0    -LLLSDS------SVYFCALR----
V30.1.1  -IYECQT------AMYYCALR----
V12.0.1  LCCISDS------AVYYCALK----
```

Figure B-5 TCR α and δ D and J sequences. Sequences aligned by RSS and conserved FGxG(P) sequences. Spaces were introduced only to align sequences at these conserved motifs. Terminal GT is underined.

```
6 D delta sequences
                                   <D SEQUENCE>
trdd1  CGTTTTTGCTACAGAGTATTTCATTGTGTACTTGGGAC-CACATTGAGGCAATATGCGACAAGCCCCTACAAAAACTC
trdd2  CTGTCATTTAATATAAGTTCAGTTTGTGAACACATT---CACAGTCTTGACCTGACAGCTCATAGCACACTACCACAA
trdd3  AACGAAACATTGGATCCTTTTCTTTGTGAGTTCAGACTTCACAAGTTTATTGTGGAGAACAAACTCCCGAGGACATGT
trdd4  GGTTTTTGTTTCGAGTTGTGTCAGTGTGGATTGGGGTACCACAGTGATTCTTTTCAGCTTTTGTCTCGTACAAAAACC
trdd5  AGATTCTCCAGATGATTTTCTGTCTGTGAATTACTG---CACAACTGCAGAGGGTTTTTGTATTGAAGTTTTTCATTG
trdd6  GGTTTTTGTATTGAAGTTTTTCATTGTGTCTGGACTAC-CACAGTGATACATTCATTATTAGGCGCTGTACAAAATCT
```

```
2 J delta sequences
         nonomer    spacer   haptomer                     F  G  x  P
trdj1  GGATATTGTCTGAGGGTTGAATGAAGTG----------TGCCGACCCACTAACTTTCGGAGCTCCCATCCGTCTCACGGTCAATCCAAGT
trdj2  TGTTTTTGTAGGGGTTACGCTAGACGTGCATTTTCAGCTACTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAGT
```

```
111 J alpha sequences
        nonamer    spacer   heptamer                                      F  G  x  G
traj1   GTTTTATGTCTGGGAATGAAGCACTGTG--GTCATCCGGTTCGAACTTCAAGCTTTTAATGGGATCTGGAACCAAGTTAAATGTGCAGCAAAGT
traj2   GGTTTATGTATGTGTGTGTAGCCATGTG-TGAATACTGGCAGCCTTGGAAAAATAATCTTTGGTCAAGGAACTAAACTTTTCATCCATTCAAGT
traj3   AGTATTTGATACAGGTTGTAGTGCTGTG----TGGATGCAAATACTAACAAGATGATTTTTGGCAGTGGAACAAAGCTGTTTGTTGAACCAGGCAAGT
traj4   TATTTGTTTCTGCACAAAGCACTGTGTGTCTGGTAGTGGAGCCGCAGGGAAGATTATTTTTGGAAAAGGAGTGAAACTCACAGTAAAAATTCAGGT
traj5   GGTATTTGGTATAGGCTGTAGCTCTGTG---TGAATACAAGGAAATTATAAGATTATATTTGGTGGCGGAACAAAACTCATCGTTGAACAAGGCAAGT
traj6   GGTTTTTGTTTTGTGCTCTGACATTGTG-TGTCTACTGGAATGGCATCAAAGATCTTGTTTGGAACAGGAACAAAACTGATAATCCAAAGCAGT
traj7   GGTTTTAGTTATGTGTTGTCGTGATGTG-TGTCTACTGGAGGAGGTTATAAGTTGATCTTTGGATCAGGAACACGATTGATAATCGACGCAAGT
traj8   AGTTATTGTGATGGTAAATTGCATTGTG-TGACTTCATCTGGGCTCTACCAGATCTATTTTGGCTCTGGAACCAAACTCATAGTTGGAACAAGT
traj9   TTTTTGTTTAGTGATCTGGTATTGTGTGACTAGTGGAGCTGGAGGATTCAAGGTTATCTTTGGGAAAGGCACACAGTTAATAGTTGACAAAAGT
traj10  AGTTTCTGTAGAGACAAAATTTAATGTG-TGACTCAAAGTGGAGGCTTGAAAATCATTTTTGGCTCTGGAACCAAACTGACTGTTGAAAAAAGT
traj11  GGATTTTGATGTGTGCTGTGACATTGTG-TGAATACTGGAGGTGCAAGAAAGCTTGTTTTTGGGAAAGGCACAACATTAACAGTTGGTTCAAGT
traj12  AGTTTTAGTTATTTGGTTTTATAGTGTG-TGAATAGAGCCGGCATGGAAAAGCTTATATTTGGGAAAGGCACACAATTACTCATTGAAACAAGT
traj13  TGTTTTTGTGTAGTGATTTTACTGTGTG-TGACTGGAACTGGAGTAAACAAAGTTATATTTGGGAAAGGCACAAAACTCATCGTAAATACAGGT
traj14  TTTTTGTTTGGTGCTTTTGCATTGTGTGACTACTGGAGGTGGAGGATACAAGATTATCTTTGGTACAGGCACAAAACTGATAGTTGAAACAAGT
traj15  AGTTTTTGCTGTTTGGTTTTGTAATGTG-TGACTAGAGATGGAACCAATAAGCTTCTTTTTGGCAAAGGTACAAATCTGATTGTTGAAACAAGT
traj16  TGTTTTTGAGTGGATGTTTAACTGTGTG-TGACTGGAACTGGACTCCAGAAAGTTATATTTGGGAAAGGCACAAAACTCATAATAAATGCAGGT
traj17  TTTTTGTTTGGTGGTCTTGCATTGTGTGACTGGTTCAAGTGGAGGCTACAAGCTTATATTTGGGACCGGTACAAAGTTAATAGTTGAAACAAGT
traj18  AGTTTTTGCTGTTTGGTTTTGTAGTGTG-TGACTGGAAGTGGATTCAATAAGATTATATTTGGCAAAGGCACAAATCTGATTATTGAAACAAGT
traj19  TGTTTTTGTGTAGCTTTGTAATGCTGTG-TGACTGACACTGGAGGGAGGAAAGTAATATTTGGGAAAGGCACGAGGCTCTTTGTAGATTCAGGT
traj20  AGTTAGAGTTAGTGTTTCATACTGTGTGAATGATGGAAGTGTTTATGGAAAGCTGGTCTTTGGCTCTGGCACCAAATTATATATTAATGAAAGT
traj21  AGTTTTTGTCGTTTGGTTTAGTAGTGTG-TGAATAGAGGGGGAGTCGATAAGCTCATATTTGGAAAAGGCACAAAACTGCTCGTTGAAGCACGT
traj22  TGTTTATGTATAGGTGTTTAACTTTGTG-TGAATCAGGCTGGATTTCAGAAACTAATATTTGGAAAAGGCACAAAGCTCACTGTAAATACAGGT
traj23  AGTTTTTGCTTAGTGGTCTTGCATAGTG----TGACTGGTGGAGTGAACAAGATTATCTTTGGGACAGGAACAAAGTTGGTAGTTGAATCAAGT
traj24  TGTTTTTGTTTAGCTTTATAACTTTGTG-TGACTGATGTCGGAGTAAGAAAATTTATATTTGGGAAAGGCACAAGACTCTTTGTAAATTCAGGT
traj25  TTAAAGTTATTGTGTTTTGTGCTGTGTGAATGATGGAAGTGGATATGGAAAGCTGATCTTTGGTTCTGGCACCAAATTATATATTAATAAAAGT
traj26  AGTTTTTGCTGTTTGGTTTAGTAGTGTG-CGAATGCAGGAGGACTCAGTAAGCTTATGTTTGGGAAAGGCACAAAGTTGCTCATTGAAACAAGT
traj27  TGTTTTTGTGTGCCTATTTTTCTGTGTG-TGACTGGAACTGGAGTACGTAAAATTATATTTGGCAAAGGCACAAAACTCACAGTAAATACAGGT
traj28  TGTTTTTGTATAGTAATATTACTTAGTG-TGACTCAGGGAGGAGCTCAGAAAATAATATTTGGGAAAGGCACAAAGCTCACTGTAATTACAGGT
traj29  TTTTTGTTTGGTAATCTTACATTGTGTGACTTATGGAGGTGGAGGTACCAAACTAATCTTTGGGACAGGCACAATGTTGGTAGTTAAAACAAGT
traj30  TTAGAGTTATGATGTTTCGTACTGTGTGATTGATGGAAGTGCATTTGGAAAGCTGATCTTTGGTTCTGGCACAAAATTATATGTTAATGAAAGT
traj31  AGTTTTTGCTGTTAGGTTAAGTAGTGTG-TGAATACAGGAGGACTCAATAAGCTTATATTTGGGAAAGGAACAAAGTTGCTCATTGAAACAAGT
traj32  AGTTTTTGATTGATAGTTTTGCAATGTG-------GCACTGGATTCAACAAGGTTATCTTTGGAACAGGCACAAAGTTGATAGTTGAAACAAGT
traj33  TGTTTTTGTGTGACTGTAAAACTGTGTG-TGACTCAGACGGGATTACAAAAAATTCTATTCGGCAAAGGCACAAAGCTCACTGTAAATTCAGGT
traj34  TTTTTGACTAGTACTCTTGCATTGTGTGAATTATGGAAGTGGAAACTACAAGCTTATCTTTGGGACAGGCACAAAGCTGATAGTTGAAACAAGT
traj35  CTAGTTTTTGCTGTTTGCTTTTGTAGTG--TGTGACTGGAGGTTTTGACAAGCTTATATTTGGGAAAGGCACAAATCTGATTATTGAAACACGT
traj36  TGTTTTTGTTTAGCTTTATAACTGTGTG-TGACTAACACTGGAGCAGGGAAACTCATATTCGGGAGTGGCACACAGCTTTTTGTAAATTCAGGT
traj37  GGTTTTAGTGTTGTAGATTTATGCTGTG----TGACTGGAAGTGGATATAAGATTATCTTTGGGACGGGTACAAACCTTATAATCGAAACTAGT
traj38  GGTTATTGGTTTGATGAAATGCAATGTG-TGACTTCATCTGGTCTAAGGCAGATCTATTTTGGCTTAGGAACCAAACTCGTTATTGAATCAAGT
traj39  TTTTCGATTAGTGGTCTGACATTGTGTGACTTCTGGAAGTGGTGCTTACAAGCTTTTCTTTGGAACAGGCACAAAGTTGATAGTTGAAACAAGT
traj40  TTTTTGCTATGTGGTCCTACATTGTGTGACATCTGGAAGTGGACTTTACAAGGTTATTTTTGGGAAAGGCACAAAGTTGTTTGTTGAAACAAGT
traj41  AGTTAGTGTAGAGATGAAATATAGTGTG-TGACTCAGTCTGGAAGCTGGAAAATCCATTTTGGCTCTGGAACCAAACTGACCGTTGAAAGCAGT
traj42  AGATTTTGCTGTGTTGTTTCATACTGTG-------TGAATGCTGGACAAAAGCTTATTTTTGGAAGAGGCACAAAGCTAGAAATTCAAACAAGT
traj43  TTTTTGCTGTGTGGCTAAATGCTGTGTGAATAGTGGAGTTAATATCAACAAAATATTCTTCGGGACAGGAACAAAATTGATCATCAGAACAAGT
traj44  TGTTTTTGTACTGCTGTATAATCCTGTG-TGACTAACACTGGAGGATTTAAAATGATTTTTGGCTCTGGAACAAAACTCACAGTTCTGTCAGGT
traj45  ATAGTTACAGTTATGAGGTTTAACAGTG--TGTGGATGCTGGATTTGGAAAGCTCATCTTTGGATCTGGCACAAAATTGCATGTTTATTCGAGT
traj46  TGTTTTTGTGCATCGATTTAACTATGTG-TGACTACCAGTGGAGGCGTTAAAATTATATTTGGGACTGGAACACAACTTACAGTACTGTCAGGT
traj47  AGTAAATGTTTTGGAGTTCCGTGCTGTG-TGTCTAGTAGCTATGGAGGGAAACTCATTTTCGGTCAAGGAACAAATCTGTTTGTTGACAAAAGT
traj48  TTCTCGTTATCAGGTTTCACATTGTGTGAATGCTGGAAGTTCATATGGAAAGCTCATCTTTGGGTCCGGCACAAAATTAGAAGTACACACAAGT
traj49  TTCTAGTTATCAGATTTAACACTGTGTGACTACTGGAAGTGGATATGGAAAGATAATCTTTGGTTTTGGCACAAAATTAGAAGTTAATACAAGT
traj50  AGTAATTGTTGAGGGAAAATGCGGTGTG-TGACTGGATCTGGAAACTGGAAAATCATTTTTGGCTCTGGAACCAGACTTACTGTTGAAACAAGT
traj51  CTTTACTGTCATGGTTGCTAACACTGTG-------TGTCATATTCTGAAAAGATCATCTTTGCACAGGGAACAAAGGTTACCATAAATTCAAGT
traj52  ATTTTGCTATGCAGTCTAATAATGTGTGACGTATGGAAGTGGAGTTAATAAAATCATCTTTGGATCAGGCACAAAATTAATCATTGAAACAAGT
traj53  TGTTTTAGTGTTGTCATGTATCTGTGTG-TGACTGATAGCGGAGGATGGAAAGTAATATTTGGAAAAGGAACCAAACTGACCATCGAGTCAGGT
traj54  GGTAATGTATATGGTTGCTAACACTGTG-------TATCATATTCTGATAAGATGTACTTTGGTGAAGGATCAAGGCTTATTATAAATTCAAGT
traj55  GTTTTTTGTTGTGAGATTTTTCATTGTG-TGAAAACTGGTGGATACGGAAAGATAATTTTTGGAGCTGGAACTAAATTAATAGTTGCAGCACGT
traj56  TAAGTTTATGTAAAGCCACGTAGCTGTG--TGTGACTAACAACGTGGGACGTATGATATTTGGAAAAGGAACAAAACTTATTGTTGATTCAGGT
traj57  GATTTTTGAAATGATGTTTCCTGTTGTG----TGAATGTGGGATTGGAAAAGATCATTTTTGGGAAAGGCACACAATTACTAGTTGAATCAAGT
traj58  AGATATTGCTTAGAATTTAACTATTGTG-------TGACACGTGCTGACAAAATTGTATTTGGTCAAGGCACAAAGCTTAATGTCGATTACAGT
traj59  TGTTTTCGTCATGATTGTTTGTGTTGTG----TGGATGGTTCTGGAAGAAAGATCATCTTTGGTCAAGGGACAGCTTTAACAGTTCAGTCAAGT
traj60  TGTTTTTGTAAAGCCACATAACTCTGTG-TGACTATGACTACAATGGGACGTATGATATTTGGAAAAGGAACAAAACTTACTGTCCATTCAGGT
traj61  TGTTTTTGCTGTGAGGTCAGTTAGTGTG----TGAATGATGGGACCTGGAAGCTTCATTTTGGAAAAGGGATGACTTTATATGTGCAGTCAAGT
```

Figure B-5 continued

traj62 **TGTTTTTGT**CATGAGTGTTTG**TGTTGTG**----TGGATGGTTCTGGGCTGAAAATCATT**TTTGGTCAAGGA**ACAACTTTAATAGTTGAATCAAGT
traj63 **GCTTTTGGA**AGAGGTGCTTGT**TACTGTG**----TGAATGCTGCAACACAAAAGATCGTC**TTTGGATCTGGA**ACAAAAGTATTCATAGAAACAAGT
traj64 **GTTTTTTGT**CATGATTGTTTC**TATTGTG**-TGGATGCTGGTGGAGGGAGAAAGATCATT**TTTGGTCAAGGA**ACAGCCTTAAAAGTTGAATCAAGT
traj65 **ACATTTTGT**CTGAGTGTTTGT**TACTGTG**----TGAATTATGGGAACAACAAGATCACT**TTTGGAAGTGGA**ACAAAACTGACAGTTGAATCTCGT
traj66 **TTTTTGTGA**AGGCTGCTTAGA**TTGTGTG**------ACTTTGAATGCAAACAAGATCATC**TTTGGAGAAGGA**ACAATGCTACATGTAAAAGCAAGT
traj67 **TGTAAAAAT**TGTGGTGTTCAA**CACAGTG**----------TGAATTACAATAAGATCACT**TTCGGTCAGGGA**ACACAGCTTTATGTTTACTCCAGT
traj68 **AGTTTTTGC**TCAGGCTTTTGT**TAGTGTG**----------TGAATTACGACAAGATCTTT**TTTGGAGCTGGA**ACCAAACTAATCATGGAATCTCGT
traj69 **AGTTTTTGA**AGTTAACCAGAT**CGCTGTG**-------TGACAACTGGAGTAAAGATAATA**TTTGGAAGGGGT**ACTCGACTTATTGTGGAGTCAAGT
traj70 **TGTATAAGT**AATTGTAATGTA**GTTTGTG**-TCTGTGTGACTGCTGGTAACAAGATAATC**TTTGGAAAAGGA**ACACAAGTTCATGTTGAAACAAGT
traj71 **AATGTTTTG**GCTGTGAGGTCA**GTTAGTG**--TGTGAATGATGGGACTAGGAAGATTCAT**TTTGGAAAAGGG**ACGACTTTATATGTGCAGTCAAGT
traj72 **TGTTTTTGT**CATGATTGTTTG**TGTTGTG**----TGGATGGTGCTGCAAGAAAAATCTTT**TTCGGTCAAGGA**ACAACTTTAACTGTTCAATCAAGT
traj73 **GGTTTTGGA**AGTGGTGTTTGT**TACTGTG**----TGAATGACGCAACACGAAAGATCATC**TTTGGGTCTGGA**ACAAAACTATTCATAGAATCAAGT
traj74 **AGTTTTTGT**CATGATTGTTTG**TGTTGTG**-------TGGATGCTGCTAGGAAGATCATT**TTTGGTCAAGGA**ACAACTTTAATAGTTGAATCAAGT
traj75 **AGTTTTTGT**GAAGGCTGCTTA**TATTGTG**----TGACTTTTGGTGCAACCAAGATCATC**TTTGGAGCAGGA**ACGATGTTATACGTAGAAGCAAGT
traj76 **AGTTGTTGT**TCTGGTGTTCAA**CACTGTG**-------TGAATTACGGCAACAAGATCATA**TTTGGTCAGGGA**ACACAGCTTTATGTTTATGCCAGT
traj77 **AGAGTTTTT**GCTCAGGCTTTT**GTTAGTG**--------TGTGAATTACAACAAGATCACT**TTTGGAGCTGGA**ACCAAACTAATTGTGGAATCTCGT
traj78 **GGTTTTTGT**AGTTATCCAGAT**CGTTGTG**-------TGACATCTGGAGTAAAGCTAATA**TTTGGAAGTGGT**ACTAGACTAATTGTGGAGTCAAGT
traj79 **AGTTTTTGC**ATTGTAGTTTGT**GTCTGTG**-------TGACTGCTGGTTACAAGTTAATA**TTTGGAAAAGGA**ACACAAGTTCACGTTGAAGCGAGT
traj80 **AGTTTTTGT**GCTGTGGTTAAT**TGATGTG**-------TGTATGTCGGCAATAAGATTGTG**TTTGGAAGAGGG**ACACAATTATATGTCACTGCACGT
traj81 **GGTATTTGT**GTAGAAGTCATT**AGTTGTG**-------TGAATGACGGAGTTAGAATAATA**TTTGGCAAAGGA**ACAGAACTCAAAATACAACAAAGT
traj82 **ATTTTGCTT**GCACGGTTTTCA**TAGCGTG**---------ACATCAAACTGGAAAGTATTT**TTTGGTTCAGGA**GTTAAACTACAAGTTGAGCAAAGT
traj83 **AGTTTTTGG**TGCATGTCTAAA**GACTGTG**-------TGAATACTGGCTATAAGATTATC**TTTGGCACTGGA**ACCAAACTTATCATTCAAACAAGT
traj84 **TGTTTATGT**TATGGATTCAGC**TTATGTG**-------TGACTAATGCTCAAAAGATAATC**TTTGGGACGGGG**ACAACACTCACTGTACAAACACGT
traj85 **GGTTTTGAG**TTAATGGTTAAT**GACTGTG**-------TGAATACTGGATATAAGATGGTT**TTTGGAAAAGGG**ACCAAACTTGTCGTACAAACGAGT
traj86 **TGTTTTCGT**TATGGATTCAGC**AAATGTG**-------TGACTAACAACCGAAAGATAGTC**TTTGGGACGGGC**ACATCACTCACTGTACACACACGT
traj87 **TGTTTTTGT**AATGGATTCAGC**ATATGTG**-------TGACTAACAACCAAAAGCTAATC**TTTGGGACGGGG**ACAACACTCACTGTACATACCCGT
traj88 **TGAGGTTGA**GATGAAGTCAGT**ATTTGTG**-------TGAATGCTGCCAACAAGATTATA**TTTGGGAGTGGG**ACAAGATTAATAACAGAAACAAGT
traj89 **TGTTTTTGC**ACAGACAAAGTT**CATAGTG**----TGACATCCTCACAATGGAAATTAGTG**TTTGGGGAAGGA**ACTAAATTAACCGTTGAACCAAGT
traj90 **GTAGTTTTT**GGCACATGCTCA**AAGGGTG**--------TGTGAATGTCAATAAGATTACT**TTTGGCAAAGGA**ACCAGACTGATTGTGCGAACAAGT
traj91 **TTAATGTAG**TGTACTGTACTG**CTGTGTG**---------AATAATAACTACAAGATCATA**TTTGGAAGCGGC**ACCAGTCTCTTTGTTTCCTCAAGT
traj92 **TTTTTGGTA**TGTGACCAATCG**CTGTGTG**---------AATGATGCGTATAAGATTTAC**TTTGGTAAAGGG**ACCAAAGTCATAATACAGTCTAGT
traj93 **TAAATGTAC**CATGCTGTACTA**ATGTGTG**------AAAGCTGGATACAAGAAAATCATA**TTTGGACAAGGC**ACTAAACTAACAGTTGATTCACGT
traj94 **TGTTGCTAG**TTTTTAGTATGT**GGCTGTG**-------TGACTGATGGACTTAAACTTTTC**TTTGGCACTGGG**ACGAAACTTCTCGTACAAATTAGT
traj95 **TTTATGTGG**TCCAGTAAACCA**CTGTGTG**------ACTCAAGGGGTTAACAGAATTATT**TTTGGAAAAGGC**ACTCAACTCACTGTGCTTTCAAGT
traj96 **TTTTTAGTA**CGTAATGAATGA**CTGTGTG**---------AATGATCGAATTAAACTGCTA**TTTAGCAAAGGA**ACAAAACTCATTGTACAGTCAAGT
traj97 **TAAATGTAC**TGTGCTTTACTG**ATGTGTG**------AATCCTGGAGGAACCAAGCTCATA**TTTGGCTCAGGC**ACCAAACTAACAGTTGTTTCAAGT
traj98 **ATATGAGTA**TGCGATCAACAG**CAGTGTG**------ACTCAAAGTGGATTCAAATTTATA**TTTGGAAAAGGC**ACTAAACTAACCGTGCTTTCCAGT
traj99 **AGTTTTTGG**CATATAGTTAAA**GACTGTG**-------TGAATGCTGGACTTCAACTTTTC**TTTGGAACAGGG**ACAAAACTTGTGGTACAAATCAGT
traj100 **AGTTATTAT**CCCTCTACAAAC**CACTGTG**----TGAATAACAATTTTGGTAAATTAATA**TTTGCGAGTGGG**ACACAACTGTATGTTCAACCAAGT
traj101 **TTTCTGTAA**TGAGATTTATAG**TTGTGTG**------ACTCAAAATGCATTCAAACTGATA**TTTGGAGAGGGG**ACGCAGCTCTTTGTGTTATCCAGT
traj102 **AGTATTTAT**AACACAGACAGT**CACTGTG**----TGAATGACAACAATTACAAAGTGATA**TTTGCGAACGGA**ACTCAGTTGTTTGTACAAACAAGT
traj103 **AGTTTTTGG**TATGTGGCTAAA**GACTGTG**-------TGAATGACGGACGTAGACTTTTC**TTTGGTGCTGGG**ACAAAACTTCTTGTACAAATCAGT
traj104 **AGTTTTTAT**TCCACCACTAAC**CACTGTG**----TGAATGGCAATACTGGTAAAATAATA**TTTACGAATGGC**ACCCGACTATATGTACAAACAAGT
traj105 **TGTTTCTGT**GCTGAAGTGTAT**AACTGTG**----TGAATCAAAACACAAACAAACTAATA**TTTGGACAGGGG**TCGCAGCTTGTTGTGTTATCCAGT
traj106 **AGTATTTAT**AACACAGACAGT**CACTGTG**----TGAATGACAACTACGGCAAAATGATA**TTTGCGAAGGGA**ACTCGGTTGTTTGTACAGACGAGT
traj107 **GGTTTTTGT**GCTGATAGATCA**GCCTGTG**-------GGAATAATAATATCAAAATTTTC**TTTGGAAGAGGA**ACAAAATTAACAGTACAATCAAGT
traj108 **AGTTTTTCG**TATTAGGCTGAA**GACTGTG**-------TGAATGATGGATTTAAGCTTTTC**TTTGGCACGGGG**ACAAAAGTTGTAGTACAAATTAGT
traj109 **AGTTATTAT**CCCTCGACTAAC**CACTGTG**----TGAATGACAATATCGGCAAAATTATA**TTTGCGACTGGG**ACACAAGTGTATGTACAACCAAGT
traj110 **GGTTTTTGT**ACTCGTATATCA**ATCTGTG**-------TGACTACTAGCGTTAAAATTGTT**TTTGGAAGAGGA**ACAAAATTAACAGTGCAATCAAGT
traj111 **TTATTATAT**GTGCATTTATCA**TTGTGTG**------ACGAACACTGCTGACAAAATCATA**TTTGGAAAAGGG**ACACAACTATTTGTGCACTCAAGT

Figure B-6 Sequences of clones Sanger sequenced.

```
   <M  E  K  Q  L  M  L  I  L  I  L  T  P  G  V  M  T  A  D  Q
C1 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C2 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C3 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C4 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C5 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C6 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C7 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG
C8 ATGGAGAAACAACTGATGCTCATTTTAATTCTGACTCCAGGTGTGATGACTGCAGACCAG

    I  S  P  N  K  E  A  L  T  V  K  E  E  E  T  V  T  F  S  C
C1 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C2 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C3 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C4 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C5 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C6 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C7 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT
C8 ATTAGCCAAATAAAGAAGCTCTTACTGTAAAGGAAAGGAGACAGTGACCTTCAGTTGCCT

                    V REGION  TCRD V23.2.2
    S  Y  D  T  S  S  S  Y  V  R  L  Y  W  Y  R  Q  Y  L  N  G
C1 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C2 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C3 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C4 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C5 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C6 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C7 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG
C8 CATATGATACAAGCAGCAGTTATGTTAGGCTTTACTGGTACAGACAATATCTTAATGGAG

    E  P  Q  Y  L  L  F  K  A  A  R S  S  S  G  G  G  R  P  D
C1 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C2 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C3 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C4 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C5 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C6 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C7 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA
C8 AACCTCAGTATTTATTATTCAAAGCTGCACGATCAAGTAGTGGAGGTGGGAGACCCGATA

   N  P  R  F  K  S  T  T  S  D  S  S  T  E  L  T  I  S  G  V
C1 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
C2 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
C3 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
C4 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
```

```
C5 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
C6 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
C7 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA
C8 ATCCTCGTTTTAAGTCGACTACATCAGACTCATCCACTGAACTCACTATTAGCGGTGTAA

                                               >  <D4>< D6 ><
   T  L  S  D  S  A  L  Y  Y  C  A  L  R  V  G  E  Y  D  Y  A
C1 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C2 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C3 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C4 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C5 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C6 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C7 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA
C8 CTCTGTCAGATTCAGCTCTCTATTATTGTGCTCTAAGAGTAGGAGAGTACGACTACGCTA

                         J2
   T  D  P  L  T  F  G  K  P  I  T  L  T  V  I  P  K  E  T  V
C1 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C2 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C3 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C4 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C5 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C6 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C7 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA
C8 CTGATCCTTTAACATTCGGCAAACCGATCACCCTCACGGTAATACCAAAAGAGACAGTGA

                             ><  C REGION
    N  S  P  P  A  F  L  S  V  L  S  P  I  K  G  H   G  S  D  I
C1 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCA-TGGA-TCTGATAT
C2 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCA-TGGA-TCTGATAT
C3 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCA-TGGA-TCTGATAT
C4 ATTCTCCTCCGGCATTTTTGTCCGTCTTGTCCCCTATAAAGGGCCAATGGAGTCTGATAT
C5 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCAATGGA-TCTGATAT
C6 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCAATGGA-TCTGATAT
C7 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCA-TGGA-TCTGATAT
C8 ATTCTCCTCCGGCATTTTTGTCTGTCTTGTCCCCTATAAAGGGCCA-TGGA-TCTGATAT

     C  V  A  A  G  F  F  P
C1 TTGTGTGGCCGCCGGATTCTTTCC
C2 TTGTGTGGCCGCCGGATTCTTTCC
C3 TTGTGTGGCCGCCGGATTCTTTCC
C4 TTGTGTGGCCGCCGGATTCTTTCC
C5 TTGTGTGGCCGCCGGATTCTTTCC
C6 TTGTGTGGCCGCCGGATTCTTTCC
C7 TTGTGTGGCCGCCGGATTCTTTCC
C8 TTGTGTGGCCGCCGGATTCTTTCC
```

Figure B-7 Alignment of 440 sequences obtained from PacBio sequencing, showing all 440 sequences have identical CDR3s. Sequences presented have been trimmed to show just the terminal portion of the V segment and the initial portion of the C segment to highlight the CDR3

```
        V sequence 23.2.2    >  <D4><  D6 ><                 J2                                 ><Cδ
       Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L   T  F  G  K   P  I  T  L  T  V   I  P  K   E
C1    TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

1     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGC-AAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
2     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGCCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
3     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
4     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACG-TAA-TACCAAAA-GA
5     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
6     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
7     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
8     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
9     TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
10    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
11    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
12    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
13    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
14    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
15    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
16    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
17    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
18    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
19    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
20    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
21    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
22    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
23    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
24    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
25    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
26    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
27    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
28    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
29    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
30    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
31    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
32    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
33    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
34    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
35    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
36    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
37    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
38    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
39    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
40    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
41    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
42    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
43    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
44    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
45    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
46    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
47    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
48    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
49    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
50    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAA-GA
51    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TAC--AAA-GA
52    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
53    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
54    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
55    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
56    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
57    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
58    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
59    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
60    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
61    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
62    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
63    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
64    TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
```

```
     V sequence 23.2.2    >   <D4><  D6 ><                 J2                                 ><Cδ
     Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L   T  F  G  K   P  I  T  L  T  V   I  P  K   E
C1  TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

65  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
66  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
67  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
68  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
69  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
70  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
71  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
72  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
73  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
74  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
75  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
76  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
77  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
78  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
79  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
80  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
81  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
82  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
83  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
84  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
85  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
86  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
87  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
88  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
89  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
90  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
91  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
92  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
93  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
94  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
95  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
96  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
97  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
98  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
99  TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
100 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TAC-AAAA-GA
101 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
102 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
103 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
104 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
105 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
106 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
107 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
108 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
109 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
110 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
111 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
112 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
113 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
114 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
115 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
116 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
117 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
118 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
119 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
120 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
121 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
122 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
123 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
124 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
125 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
126 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
127 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
128 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
129 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
130 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
131 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
132 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
133 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
134 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
135 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
136 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
```

```
      V sequence 23.2.2    >  <D4><  D6 ><                 J2                                 ><Cδ
     Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L   T  F  G  K   P  I  T  L  T  V   I  P  K   E
C1  TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

137 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
138 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
139 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
140 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
141 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
142 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
143 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
144 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
145 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
146 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
147 TATTATTGTGCTCTAAGAGTAGGAGAGTA-G-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
148 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
149 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
150 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
151 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
152 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
153 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
154 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
155 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
156 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
157 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
158 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
159 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
160 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
161 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
162 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
163 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
164 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
165 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
166 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
167 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
168 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
169 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
170 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
171 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
172 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
173 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
174 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
175 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
176 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
177 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
178 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
179 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
180 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
181 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
182 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
183 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
184 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
185 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
186 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
187 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
188 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
189 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
190 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
191 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
192 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
193 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
194 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
195 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATC-TTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
196 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
197 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
198 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
199 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
200 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
201 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
202 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
203 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
204 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
205 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
206 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
207 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
208 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
```

```
      V sequence 23.2.2    >  <D4><  D6 ><                 J2                                 ><Cδ
     Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L   T  F  G  K   P  I  T  L  T  V   I  P  K   E
C1  TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

209 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
210 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
211 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
212 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
213 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
214 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
215 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
216 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
217 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
218 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
219 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
220 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
221 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
222 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
223 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
224 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
225 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
226 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
227 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
228 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
229 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
230 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
231 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
232 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
233 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
234 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
235 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
236 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
237 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
238 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
239 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
240 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
241 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
242 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
243 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
244 TATTATTGTGCTCTAATAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
245 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
246 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
247 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
248 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
249 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
250 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
251 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
252 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
253 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
254 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
255 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
256 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
257 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
258 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
259 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
260 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
261 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
262 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
263 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
264 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
265 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
266 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
267 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
268 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
269 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
270 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
271 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
272 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
273 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
274 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
275 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
276 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
277 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
278 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
279 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
280 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
```

```
      V sequence 23.2.2    >  <D4><  D6 ><                 J2                                  ><Cδ
     Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L    T  F  G  K   P  I  T  L  T  V   I  P  K  E
C1  TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

281 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
282 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
283 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
284 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
285 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
286 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
287 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
288 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
289 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
290 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
291 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
292 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
294 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
295 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
296 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
297 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
298 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
299 TTTTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
300 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
301 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
302 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
303 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
304 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
305 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
306 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
307 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
308 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
309 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
310 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
311 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
312 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
313 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
314 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
315 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
316 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
317 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
318 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
319 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
320 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
321 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
322 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
323 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
324 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
325 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
326 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
327 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
328 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
329 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
330 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
331 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACC-TCACGGTAA-TACCAAAA-GA
332 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACC-TCACGGTAA-TACCAAAA-GA
333 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACC-TCACGGTAA-TACCAAAA-GA
334 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
335 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
336 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
337 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
338 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
339 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
340 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
341 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
342 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
343 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
344 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
345 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
346 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
347 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
348 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
349 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
350 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-C-GATCACCCTCACGGTAA-TACCAAAA-GA
351 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
352 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
353 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
```

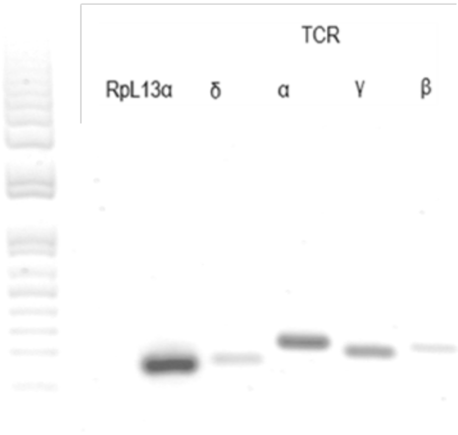
```
      V sequence 23.2.2    >  <D4><  D6 ><                 J2                                 ><Cδ
     Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L   T  F  G  K   P  I  T  L  T  V   I  P  K   E
C1  TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

354 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
355 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
356 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
357 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
358 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
359 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
360 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
361 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
362 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
363 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
364 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
365 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
366 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
367 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
368 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACG-TAA-TACCAAAA-GA
369 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
370 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
371 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
372 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
373 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
374 TATTAT-GTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
375 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
376 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
377 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
378 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
379 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
380 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
381 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
382 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAA--CCGATCACCCTCACGGTAA-TACCAAAA-GA
383 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAAATACCAAAA-GA
384 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
385 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
386 TATTATTGTGCTCTAAGAGTGGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
387 ---------------------------TACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
388 TATTATTGTGCTCTAAGAGTAGGATAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
389 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
390 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
391 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CTGATCACCCTCACGGTAA-TACCAAAA-GA
392 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-C-GATCACCCTCACGGTAA-TACCAAAA-GA
393 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-C-GATCACCCTCACGGTAA-TACCAAAA-GA
394 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAAACATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
395 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
396 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
397 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
398 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
399 TATTAT-GTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACG-TAA-TACCAAAA-GA
400 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
401 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
402 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
403 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
404 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAA--GA
405 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACG-TAA-TACCAAA--GA
406 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAAACCGATCACCCTCACGGTAA-TACCAAAAAGA
407 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
408 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
409 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAAACCGATCACCCTCACGGTAA-TACCAAAA-GA
410 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
411 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
412 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
413 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CACTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
414 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
415 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
416 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
417 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAA-GA
418 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
419 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
420 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
421 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
422 TATTATTGTGCTCTAAGAGTAGGAGAGTACGGACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCCAAAAGA
423 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
424 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
425 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
```

```
      V sequence 23.2.2    >  <D4><  D6 ><                 J2                                 ><Cδ
     Y  Y  C  A  L  R  V  G  E  Y  D  Y   A  T  D  P  L   T  F  G  K   P  I  T  L  T  V   I  P  K   E
C1  TATTATTGTGCTCTAAGAGTAGGAGAGTACG ACTACGCTACTGATCCTTTAA CATTCGGCAAA CCGATCACCCTCACGGTAA TACCAAAA-GA

426 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
427 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
428 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
429 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
430 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
431 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TAC--AAAAGA
432 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TAC--AAAAGA
433 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATC-TTTAA-CATTCGGCAAA-CCGATCACCCTCACG-TAA-TAC--AAAAGA
434 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAAAGA
435 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
436 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACC-AAAAGA
437 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
438 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
439 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATCCTTTAA-CATTCGGCAAA-CCGATCACCCTCACGGTTAATACCAAAAAGA
440 TATTATTGTGCTCTAAGAGTAGGAGAGTACG-ACTACGCTACTGATACTTTAA-CATTCG-CAAA-CCGATCACCCTCACGGTAA-TACCAAAA-GA
```

Figure B-8 PCR with primers for TCR C region of alpha, beta, gamma, and delta. Showing that just a single band is produced by this primers in preparation for qPCR.



APPENDIX C

SUPPLEMENTAL TABLES

Table 1 Primers used in PCR

Primer Name	For/Rev	Domain	Sequence	Priming Site
MFC525	R	TCR δ C region	5’-GGAAAGAATCCGGCGGCCACAC-3’	VAAGFF
MFC526	R	TCR δ C region	5’-TGCTTGACAAGGACAGGACTGCA-3’	AVLSLSS
MFC527	R	TCR δ C region	5’- TGCCTTTGCTGTTTTTCCATCCA-3’	DGKTAKA
RACE 5’	F		5’-CGACTGGAGCACGAGGACACTGA-3’	
RACE 5’ NESTED	F		5’-GGACACTGACATGGACTGAAGGAGTA -3’	
MFC535	F	TCR δ V region	5’-TACTGGTACCGACAG-3’	YWYRQ
MFC536	F	TCR δ V region	5’-TAYGGTAYMGNCAR -3’	YWYRQ
MFC537	F	TCR δ V region	5’-CTCTAYTGGTAYMGNCARTAT-3’	LYWYRQY
MFC561	F	TCR α C region	5’-CTCATGCCTGGCAACTGACTTCAC-3’	SCLATDFT
MFC562	R	TCR α C region	5’-TCAGCCAGAAGATGCCCAGTGACA -3’	SLGIFWL
MFC565	F	TCR β C region	5’-CCACATAGCCATACAGGACAAGAC -3’	PHSHTGQD
MFC566	R	TCR β C region	5’-CAGGATGTAGCCAAAGCCAACCAGC -3’	QNAKAVDQ
MFC559	F	TCR δ C region	5’-CAGTCCTGTCCTTGTCAAGCA-3’	AVLSLSS
MFC560	R	TCR δ C region	5’-GTGTGACATTCAAGTGTAGCCG-3’	LLAKCVCV
MFC563	F	TCR γ C region	5’-CCTGGGAAGGACAGTGTTGTGAc-3’	PGKDSVVT
MFC564	R	TCR γ C region	5’-GTGGCTGTGCTCCACAATCTGG-3’	QIVEHSH
MFC534		Barcoded RACE 5’ NESTED	5’-TCAGACGATGCGTCATGGACACTGACATGGACTGAAGGAGTA-3’	
MFC579		Barcoded TCR δ C region	5’- GCATATAGTAGAGATCGGAAAGAATCCGGCGGCCACAC-3’	Barcode/CVAAGFFP

Table 2 GenBank Accession numbers used for Figure 7: Phylogenetic analysis of selected V sequences

Sequence	GenBank Accession Number
Dr IgH V	AF273889.1
Gg TCR D V	AAA51274
Hs IgH V2	AB202787.1
Hs IgH V3	AB202583.1
Hs IgH V	L25544.1
Hs TCRD V1	A31326
Hs TCRD V2	S73843.1
Ip TCRA V1	U58506
Ip TCRA V1.10	JF273623
Ip TCRA V1.15	JF273628
Ip TCRA V1.20	AF178013
Ip TCRA V3.1	AF038158
Ip TCRA V4	FD208564
Ip TCRB 5	AF038161
Ip TCRB V4	U39193
Ip TCRD V 1.26	JN097583
Ip TCRD V 1.8	HQ913597
Ip TCRD V1.1	HQ913590
Ip TCRD V1.16	HQ913586
Ip TCRD V2.1	HQ913580
Ip TCRD V2.7	JN097584
Ip TCRG V1.2	HQ13594
Ip TCRG V1.3	HQ913591
Mm IgH V	M26988.1
MmTCRD V	AAA51274
Po TCRD V	AB076071.1
Po TCRD V(2)	AB0756072.1
Ss TCRAD V1.1	Yazawa et al. 2007
Ss TR V 14.7S1	Yazawa et al. 2007
Ss TR V 44.4	Yazawa et al. 2007
Ss TR V 6.1	Yazawa et al. 2007
Ss TRAD V 26.4	Yazawa et al. 2007
Ss TRAD V 27.1	Yazawa et al. 2007
Ss TRAD V 34.1	Yazawa et al. 2007
Ss TRAD V 52.2	Yazawa et al. 2007

Ss TRAD V23.1	Yazawa et al. 2007
Ss TRAV 10.1	Yazawa et al. 2007
Ss TRAV 12.1	Yazawa et al. 2007
Ss TRAV 14.1	Yazawa et al. 2007
Ss TRAV 16.1	Yazawa et al. 2007
Ss TRAV 17.2	Yazawa et al. 2007
Ss TRAV 2.3	Yazawa et al. 2007
Ss TRAV 21.1	Yazawa et al. 2007
Ss TRAV 25.3	Yazawa et al. 2007
Ss TRAV 28.1	Yazawa et al. 2007
Ss TRAV 31	Yazawa et al. 2007
Ss TRAV 36.1	Yazawa et al. 2007
Ss TRAV 39.1	Yazawa et al. 2007
Ss TRAV 4.1	Yazawa et al. 2007
Ss TRAV 45	Yazawa et al. 2007
Ss TRAV 51.1	Yazawa et al. 2007
Ss TRAV 8.3	Yazawa et al. 2007
Ss TRAV18.1	Yazawa et al. 2007
Ss TRAV3.1	Yazawa et al. 2007
Ss TRAV9.2	Yazawa et al. 2007
Tn TCRAD V1	AL596140.1
Tn TCRAD V10	AL596140.2
Tn TCRAD V11	AL596140.3
Tn TCRAD V12	AL596140.4
Tn TCRAD V13	AL596140.5
Tn TCRAD V2	AL596140.6
Tn TCRAD V3	AL596140.7
Tn TCRAD V4	AL596140.8
Tn TCRAD V5	AL596140.9
Tn TCRAD V6	AL596140.10
Tn TCRAD V7	AL596140.11
Tn TCRAD V8	AL596140.12
Tn TCRAD V9	AL596140.13
Tp TCRD V2.3	AL596140.14