

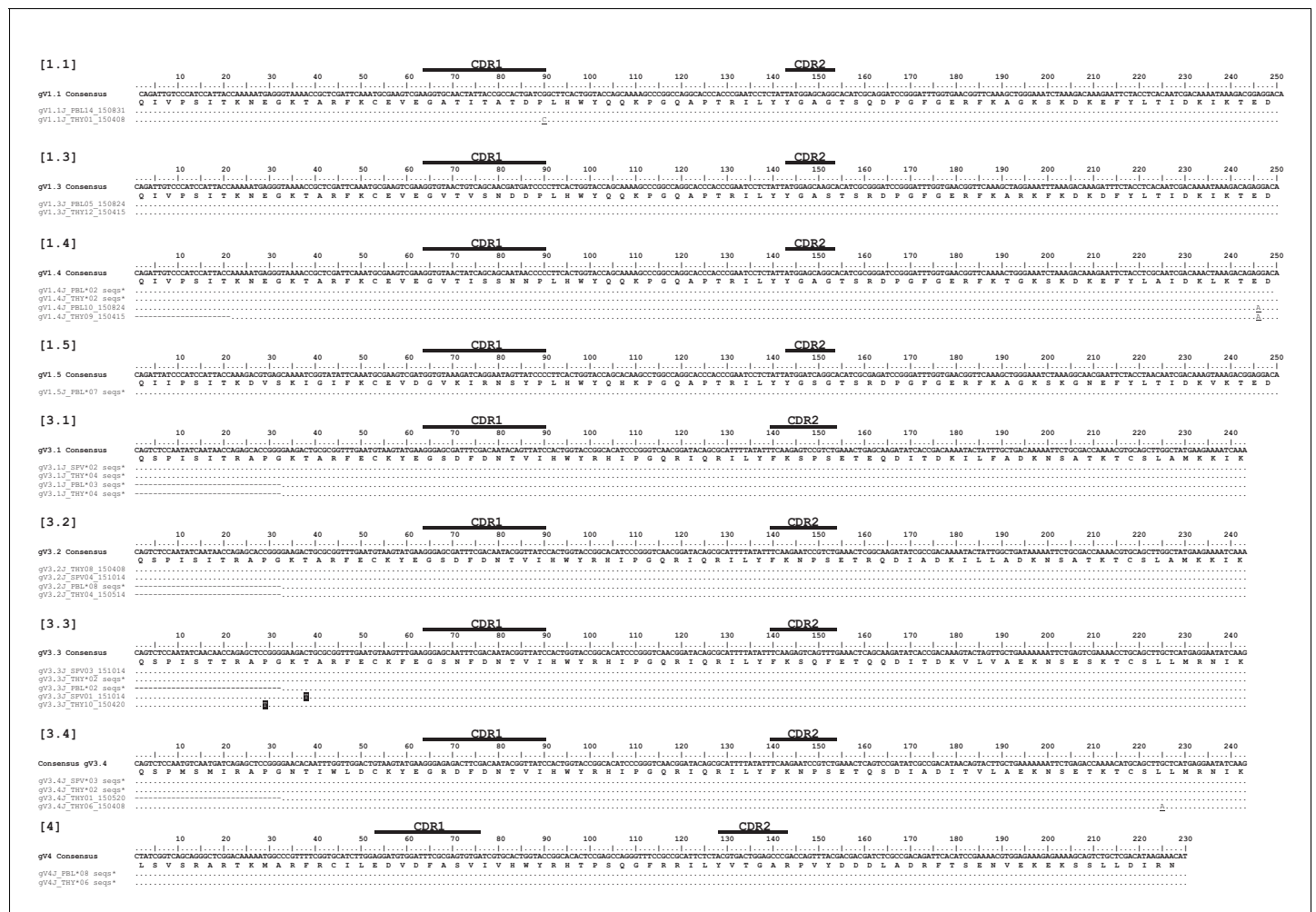


---

## Figures and figure supplements

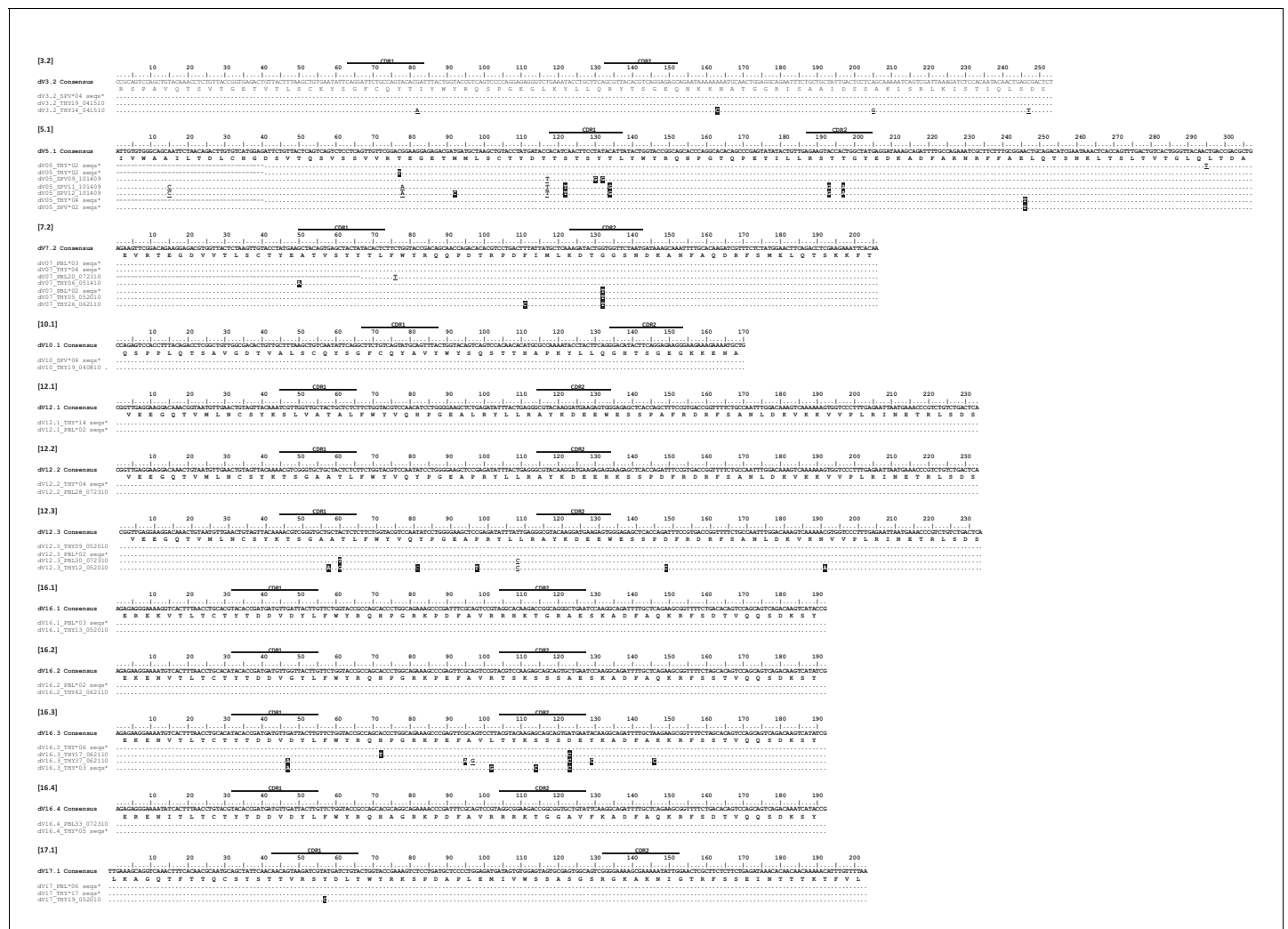
Somatic hypermutation of T cell receptor  $\alpha$  chain contributes to selection in nurse shark thymus

**Jeannine A Ott et al**



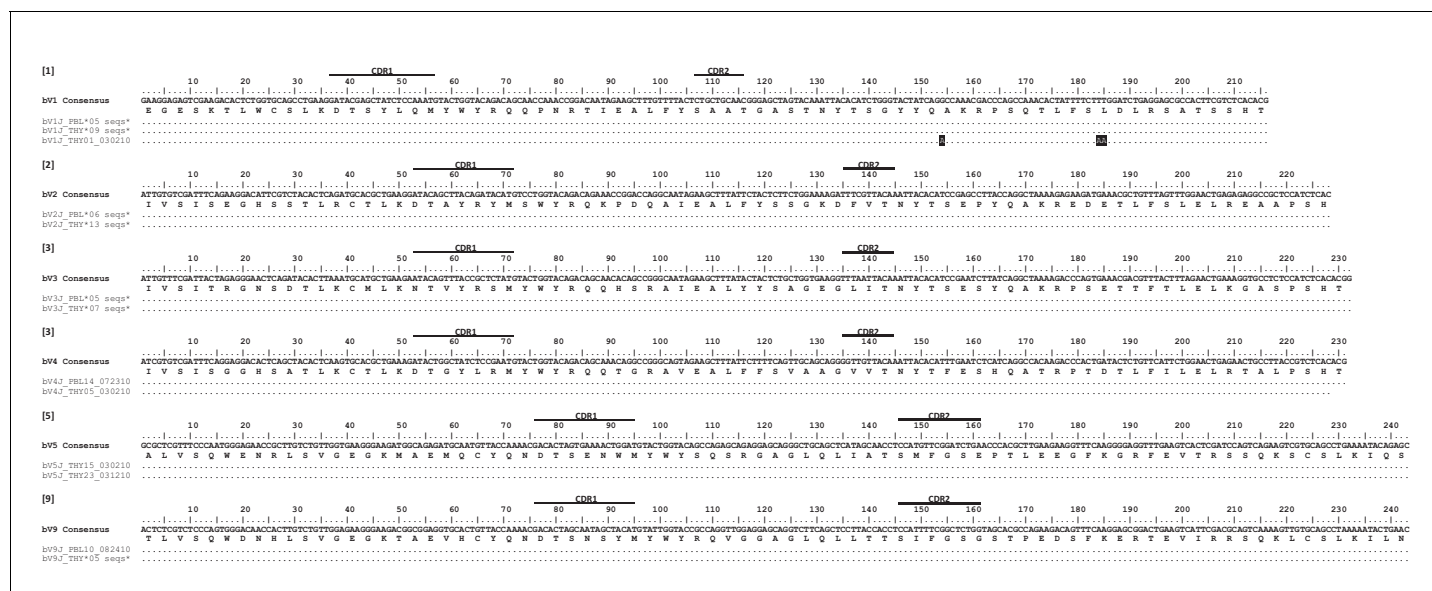
**Figure 1.** Alignment of Gamma V clones suggests minimal somatic hypermutation. Thymocyte clones for nine  $\gamma$ V groups from three different predicted V genes. CDR regions are marked above the scale for each  $\gamma$ V alignment. Amino acids are shown under the nucleotide consensus sequence, and dots represent identity to this sequence. We highlighted nonsynonymous changes in black; synonymous changes are underlined. Gaps are used for alignment purposes and indicate a shortened sequence (at the beginning or end). Sequences are identified by a single clone number or a group of identical clones condensed to a single line (the number of clones are indicated). Clone numbers that contain 'THY' are from thymus, 'PBL' are from peripheral blood leukocytes, and 'SPV' are from spiral valve (intestine). We deposited all 69 sequences into GenBank under accession numbers KY351639 – KY351707.

DOI: <https://doi.org/10.7554/eLife.28477.002>



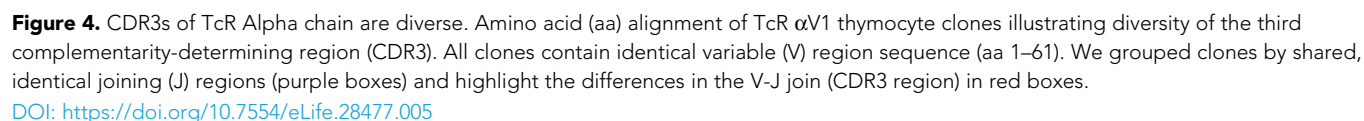
**Figure 2.** Alignment of Delta V clones suggests somatic hypermutation. Thymocyte clones for 12  $\delta V$  groups from seven different predicted V genes. CDR regions are marked above the scale for each  $\delta V$  alignment. Amino acids are shown under the nucleotide consensus sequence, and dots represent identity to this sequence. We highlighted nonsynonymous changes in black; synonymous changes are underlined. Gaps are used for alignment purposes and indicate a shortened sequence (at the beginning or end). Sequences are identified by a single clone number or a group of identical clones condensed to a single line (the number of clones are indicated). Clone numbers that contain 'THY' are from thymus, 'PBL' are from peripheral blood leukocytes, and 'SPV' are from spiral valve (intestine). We deposited all 112 sequences into GenBank under accession numbers KY346705 – KY346816.

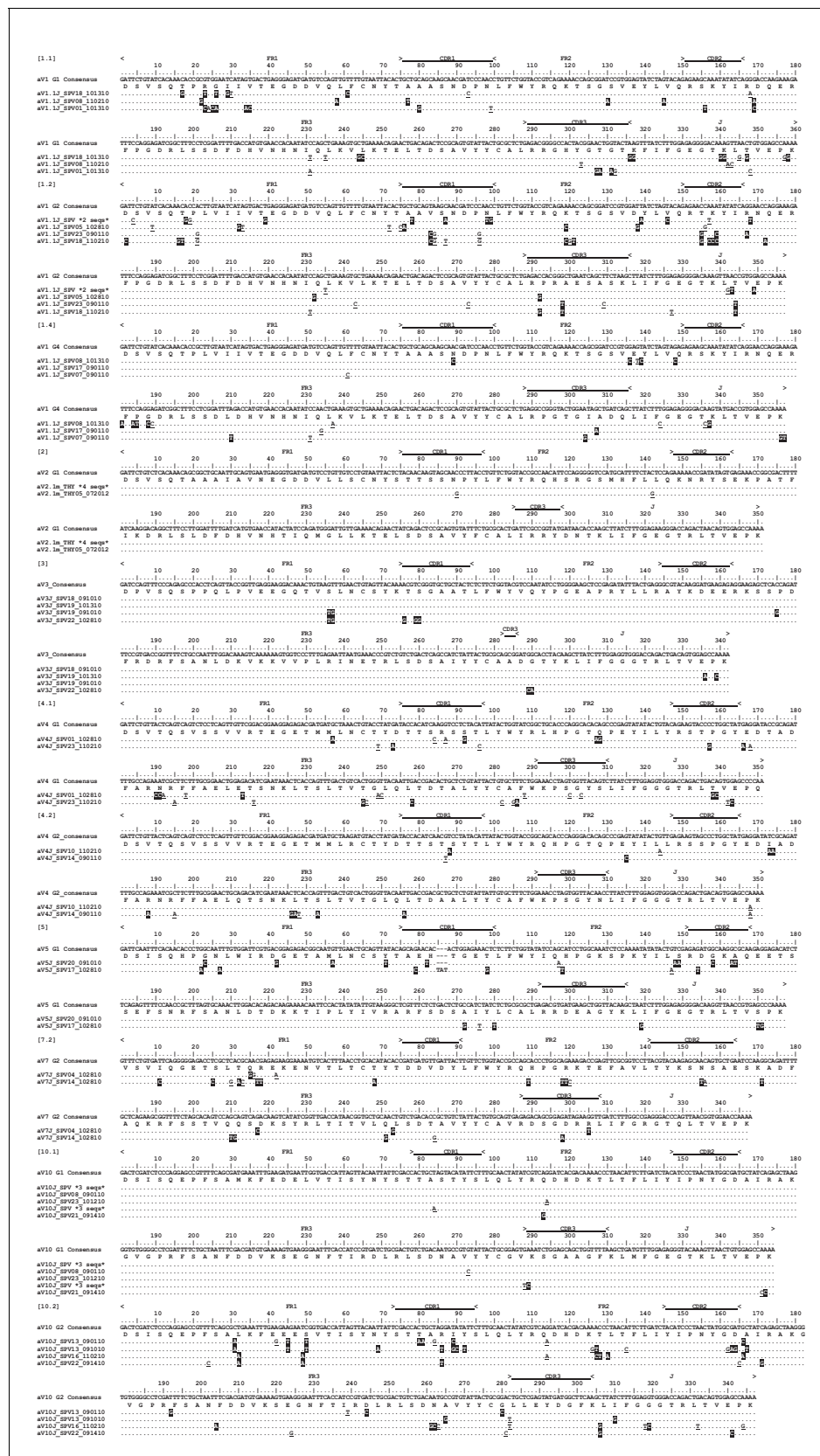
DOI: <https://doi.org/10.7554/eLife.28477.003>



**Figure 3.** Alignment of Beta V clones illustrates a lack of somatic hypermutation. Thymocyte clones for  $\beta V$  groups from six different predicted V genes. CDR regions are marked above the scale for each  $\beta V$  alignment. Amino acids are shown under the nucleotide consensus sequence, and dots represent identity to this sequence. We observed three nonsynonymous changes within a single sequence (highlighted in black). Sequences are identified by a single clone number or a group of identical clones condensed to a single line (the number of clones is indicated). Clone numbers that contain 'PBL' are from peripheral blood leukocytes and 'THY' are from thymus. We deposited all 57 sequences into GenBank under accession numbers KY351708 – KY351764.

DOI: <https://doi.org/10.7554/eLife.28477.004>





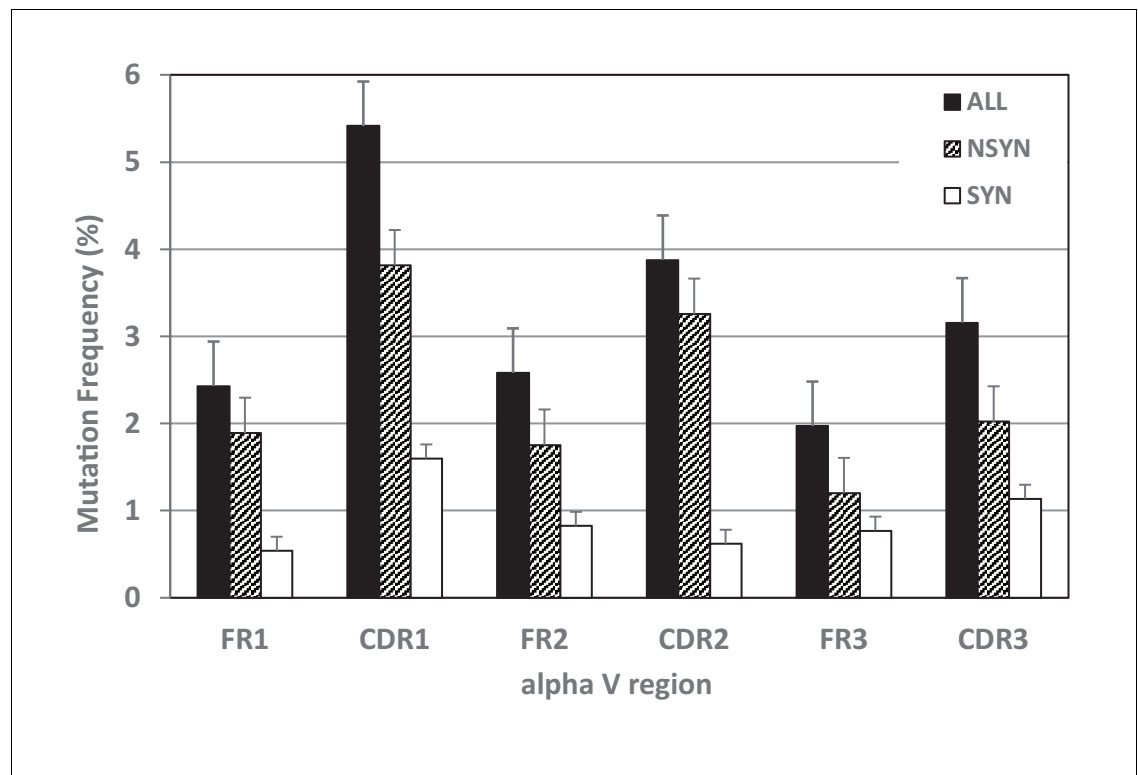
**Figure 5.** Alignment of Alpha V cDNA clones suggest somatic hypermutation at shark TCRα. Thymocyte clones for all 11 αV groups with the same CDR3 from six different predicted V genes. Locations of framework regions (FR), complementarity determining regions (CDR), joining regions (J), and

Figure 5 continued on next page

*Figure 5 continued*

constant (C) regions are marked above the scale for each  $\alpha V$ . In absence of germline sequence information, we used a Geneious-derived nucleotide consensus sequence for analysis of nucleotide changes in thymocyte clones. Amino acids are shown under the consensus sequence, and dots represent identity to this sequence. Nonsynonymous changes are highlighted in black; synonymous changes are underlined. Gaps are used for alignment purposes and indicate either a shortened sequence (at the beginning or end) or insertions or deletions within the sequence. Sequences are identified by a single clone number or a group of identical clones condensed to a single line (the number of clones are indicated). Clone numbers that contain 'THY' are from thymus, 'PBL' are from peripheral blood leukocytes, and 'SPV' are from spiral valve (intestine). We did not use clones from  $\alpha V1.3$  and  $\alpha V7.1$  because they did not contain mutations in FR or CDR regions. We deposited all 42 sequences into GenBank under accession numbers KY189332 – KY189354 or KY366469 – KY366487.

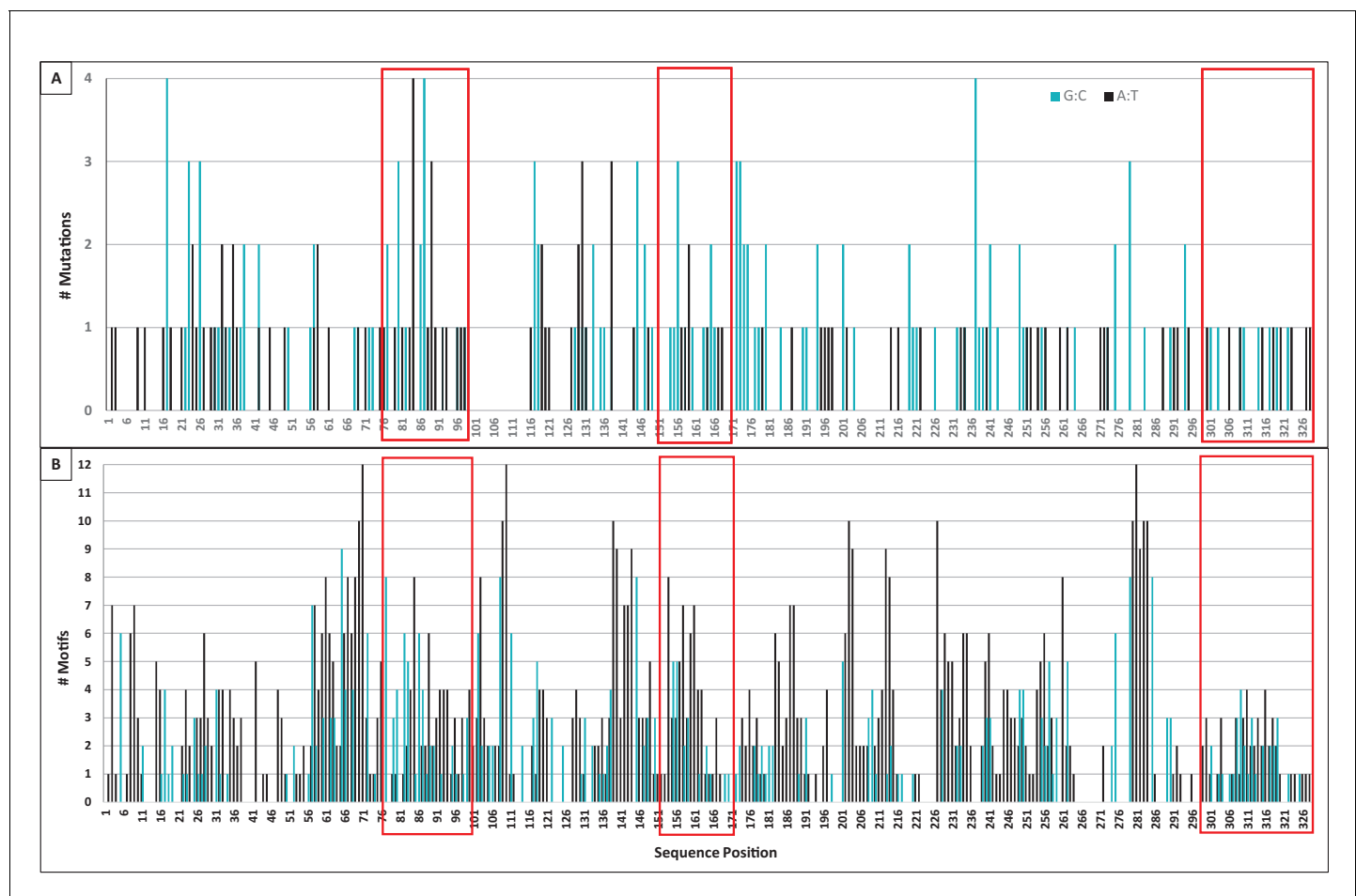
DOI: <https://doi.org/10.7554/eLife.28477.009>



**Figure 6.** Mutation frequencies differed within TcR V regions. Mutability of complementarity determining regions (CDR), especially CDR1, exceeded that of framework regions (FR) for all mutations together (black bars) and for nonsynonymous mutations alone (NSYN, hatched bars). We found no statistical difference in synonymous mutations (SYN, white bars) between CDRs and FRs.

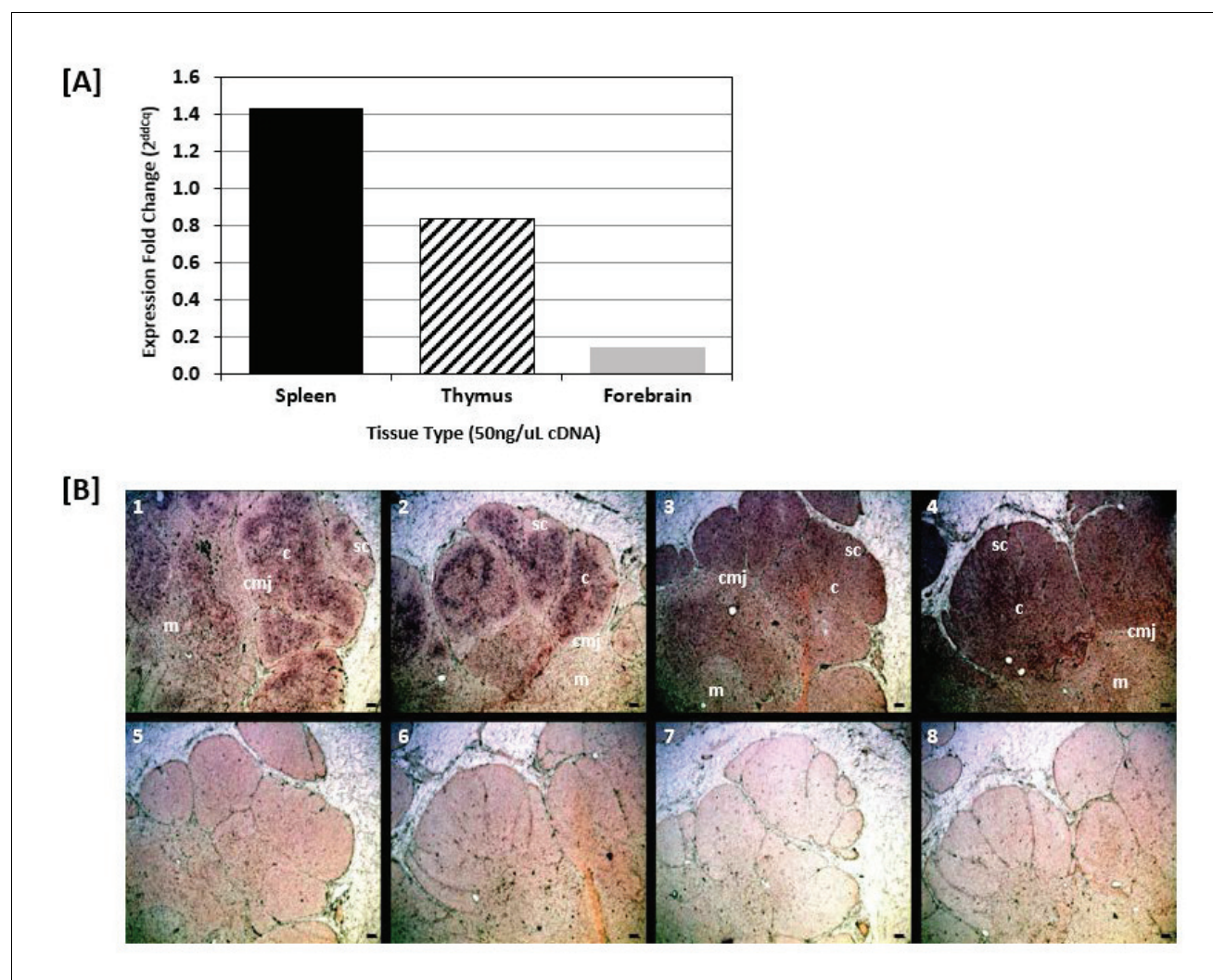
DOI: <https://doi.org/10.7554/eLife.28477.012>





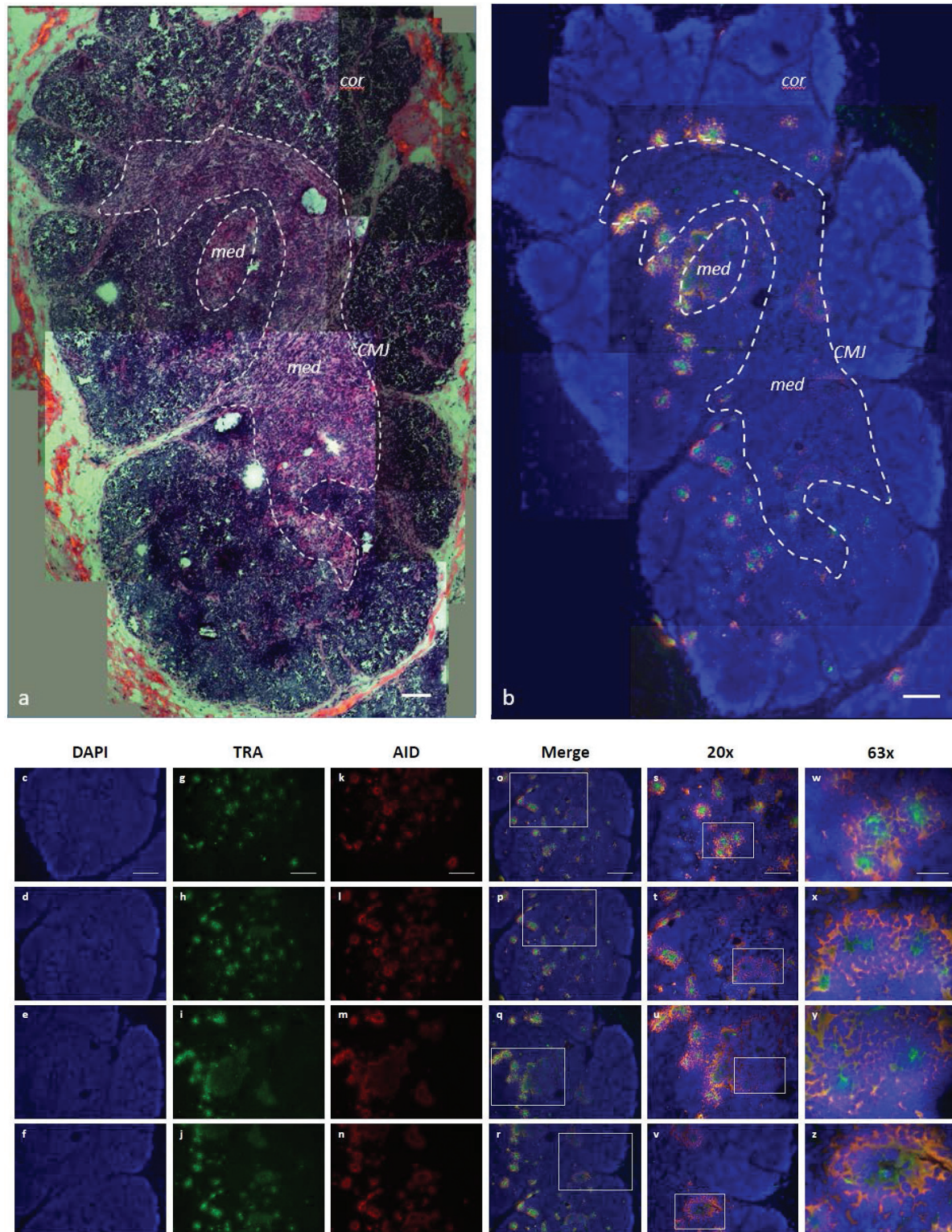
**Figure 7.** Mutation and motif locations within individual domains of TcR αV sequences. (A) Number of mutations (both single and tandem, 283 total) to A:T nucleotides (black bars, 122 mutations) or G:C nucleotides (blue bars, 161 mutations) observed at each position along the sequence length of TcR αV sequences. We counted mutations to Geneious-derived consensus sequences within framework and complementarity-determining region (CDR) domains (41 sequences within 11 αV groups). (B) Number of WA/IV motifs (black bars, indicating possible polymerase η action) or DG/YW/WR/CH motifs (blue bars, indicating hotspots for AID activity, and thus possible mutation) at each position along the sequence length. Position indicates the forward (3' to 5') location of the mutable base of each motif within a Geneious-derived consensus sequence for each TcR αV group (eleven groups). Red boxes indicate the location of CDR1 (positions 76–99), CDR2 (positions 151–171), and CDR3 (positions 300–328) within each panel. [WA/IV: A:T is the mutable position; DG/YW/WR/CH: G:C is the mutable position; W = A/T, D = A/G/T, Y = C/T, R = A/G, and H = T/C/A].

DOI: <https://doi.org/10.7554/eLife.28477.014>



**Figure 8.** Shark thymus expresses AID. (A) Expression of AID in shark spleen (black bar), thymus (hashed bar), and forebrain (gray bar) using real-time quantitative PCR. Expression was measured using the delta delta Cq method and are normalized against shark muscle tissue using  $\beta$ 2M as a reference gene. Data represent expression fold-change differences at four cDNA concentrations. (B) In situ hybridization of mRNA using ribo-probes on adult shark thymus sections. 1,2: Two different fields probing for TcR $\beta$  antisense. 3,4: AID antisense. 5,6: TcR $\beta$  sense. 7,8: AID sense. All micrographs at 10X magnification; black scale bar in lower right of each panel is 100  $\mu$ M. Anatomical structures are designated on the top panels [c: cortex; m: medulla; sc: subcapsular region; cmj: corticomedullary junction].

DOI: <https://doi.org/10.7554/eLife.28477.018>



**Figure 9.** AID expression localized to inner cortex and cortico-medullary junction. (a) H and E staining of fixed shark thymus tissue illustrating thymic architecture (10x). The densely packed cells at the margins of the image comprise the cortex (cor), while the less densely packed cells in the center

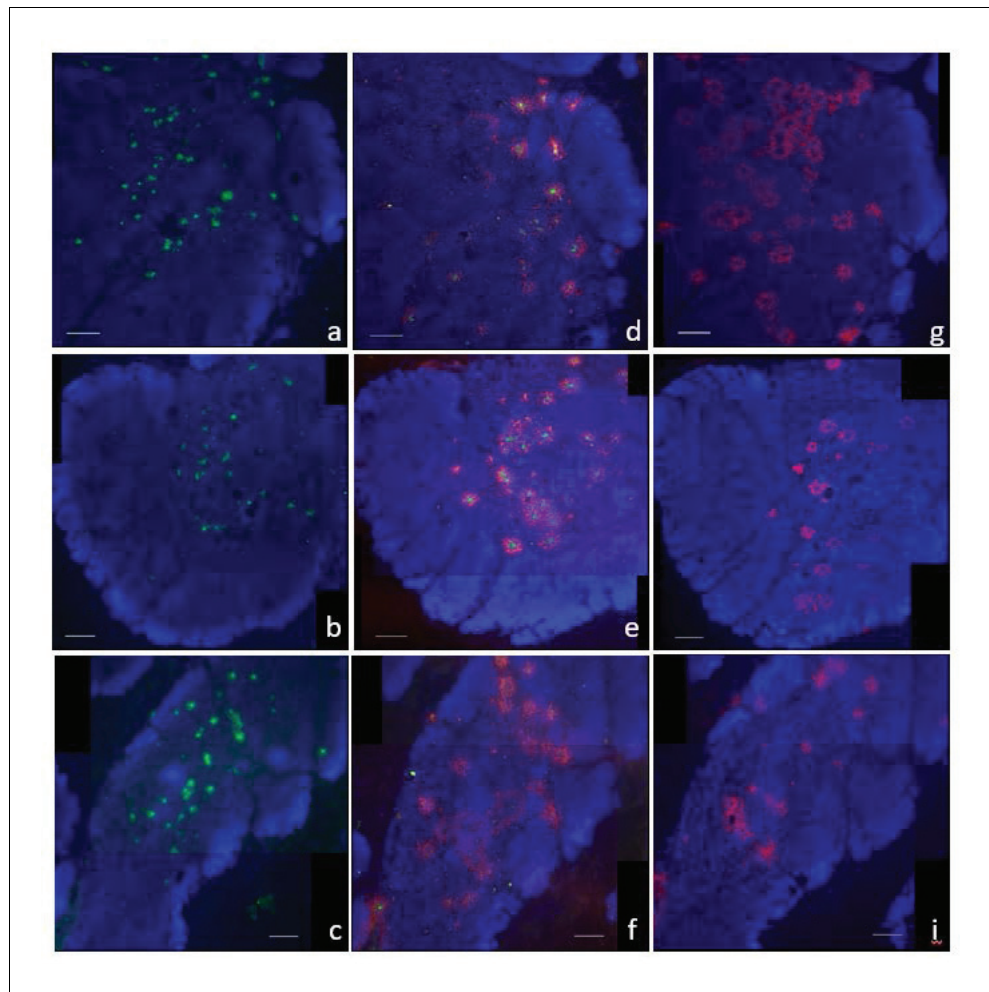
Figure 9 continued on next page



## Figure 9 continued

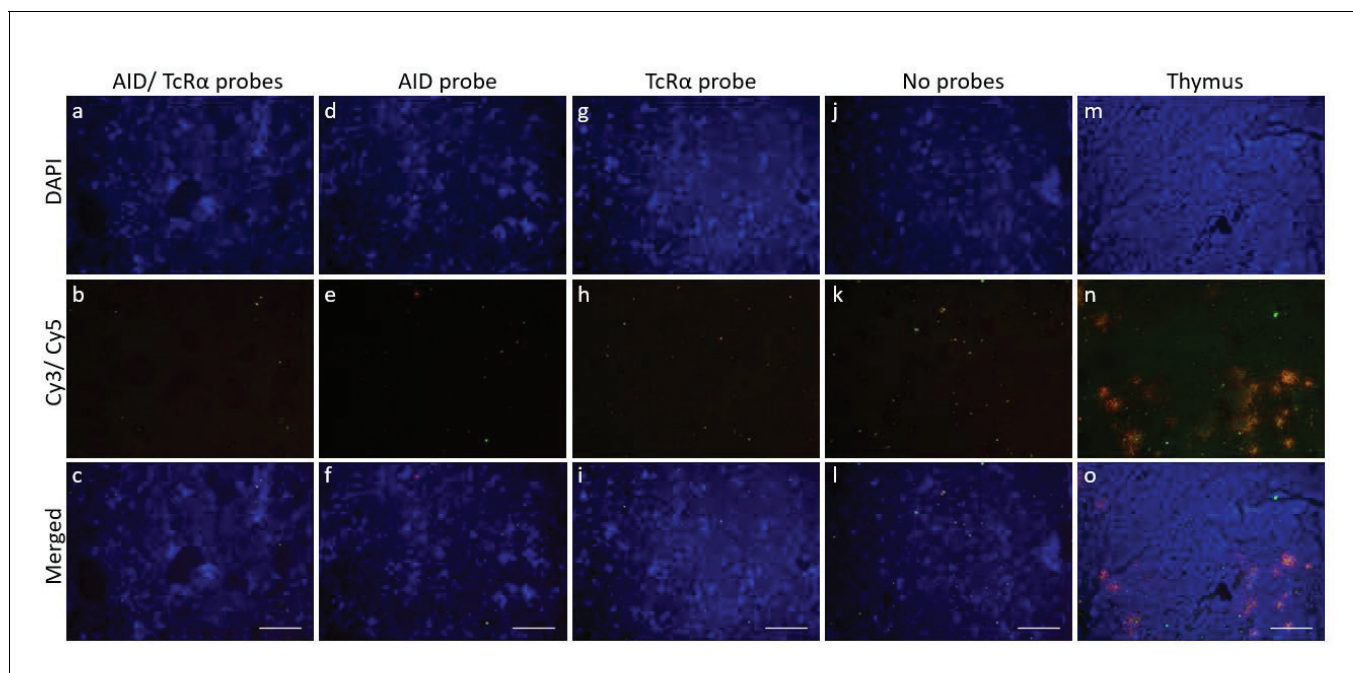
constitute the medulla (med). The region at the junction between cortex and medulla incorporates the corticomedullary junction (CMJ), delineated generally by a hashed white circle. [b-z] Single molecule RNA fluorescence in situ hybridization (FISH) probing fixed thymus sections simultaneously for AID (probes labeled with Quasar 670; pseudo colored *red*) and TcR $\alpha$  (probes labeled with CalFluor Red 610; pseudo colored *green*) and counterstained with DAPI (*blue*). (b) Composite of seven Z-stacked images (10x) depicting overall thymic architecture and the localization of AID expression to the inner cortex and cortico-medullary junction regions of shark thymus. We superimposed (and minimally adjusted) the outlined CMJ boundaries from (a) onto (b) to elucidate the junction between cortex and medulla. [c-z] We obtained images of each fluorophore using 10x, 20x, and 63x magnification and merged Z-stacked images together. Individual (10x) fluorophore images of DAPI [c-f], TcR $\alpha$  [g-j], and AID [k-n] and Z-stacked merged images [o-r] illustrate AID and TcR $\alpha$  expression in four locations of shark thymus. White boxes indicate the magnified regions of the 10x and 20x images shown in the 20x [s-v] and 63x [w-z] images, respectively. Scale bars [a,b,c,g,k,o] 150  $\mu$ m, [s] 75  $\mu$ m, and [w] 30  $\mu$ m. [cor: cortex; med: medulla; CMJ: corticomedullary junction].

DOI: <https://doi.org/10.7554/eLife.28477.019>



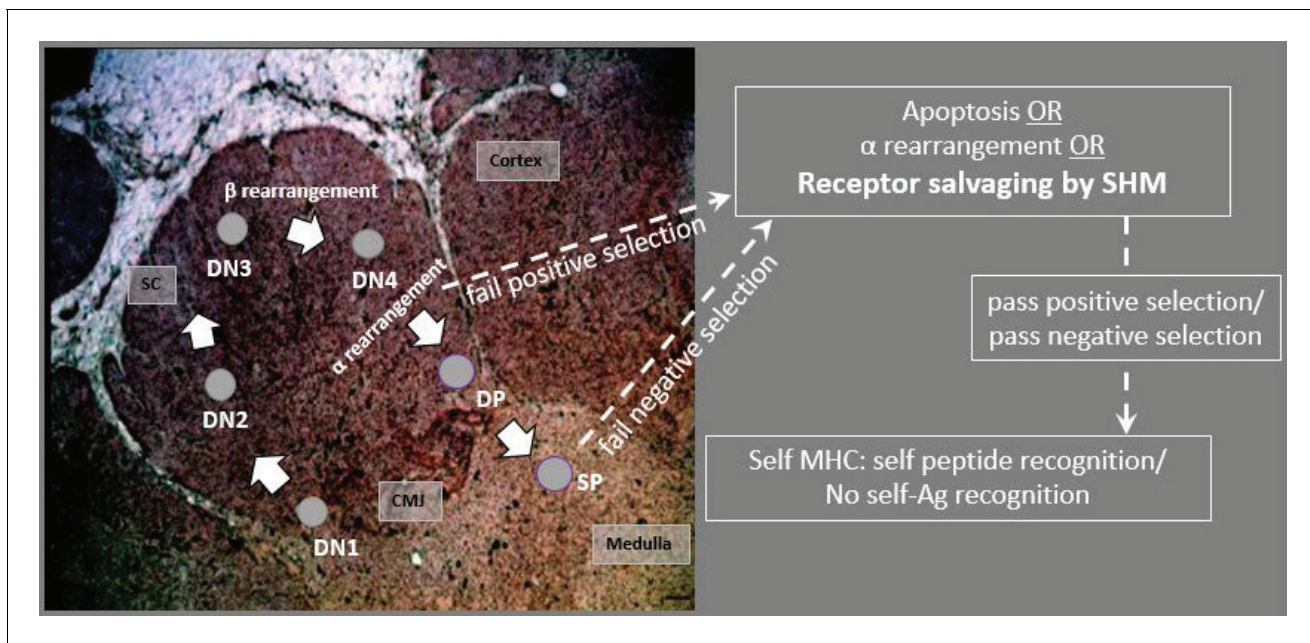
**Figure 9—figure supplement 1.** Localization of AID and TcR $\alpha$  probes is independent. Single molecule RNA fluorescence in situ hybridization (FISH) showing three separate regions of fixed thymus sections demonstrating the independent localization of our probes. We probed individually for TcR $\alpha$  (a-c) (probes labeled with CalFluor Red 610; pseudo colored green), simultaneously for both TcR $\alpha$  and AID (d-f), or individually for AID (probes labeled with Quasar 670; pseudo colored red) and counterstained with DAPI (blue). There were two tissue sections in between imaged regions of TcR $\alpha$  section (a-c) and combined AID/TcR $\alpha$  section (d-f). The combined AID/TcR $\alpha$  section (d-f) was consecutive with the AID only section (g-h). We obtained images of each fluorophore using 10x magnification and merged Z-stacked images together. Scale bar 100  $\mu$ m.

DOI: <https://doi.org/10.7554/eLife.28477.020>



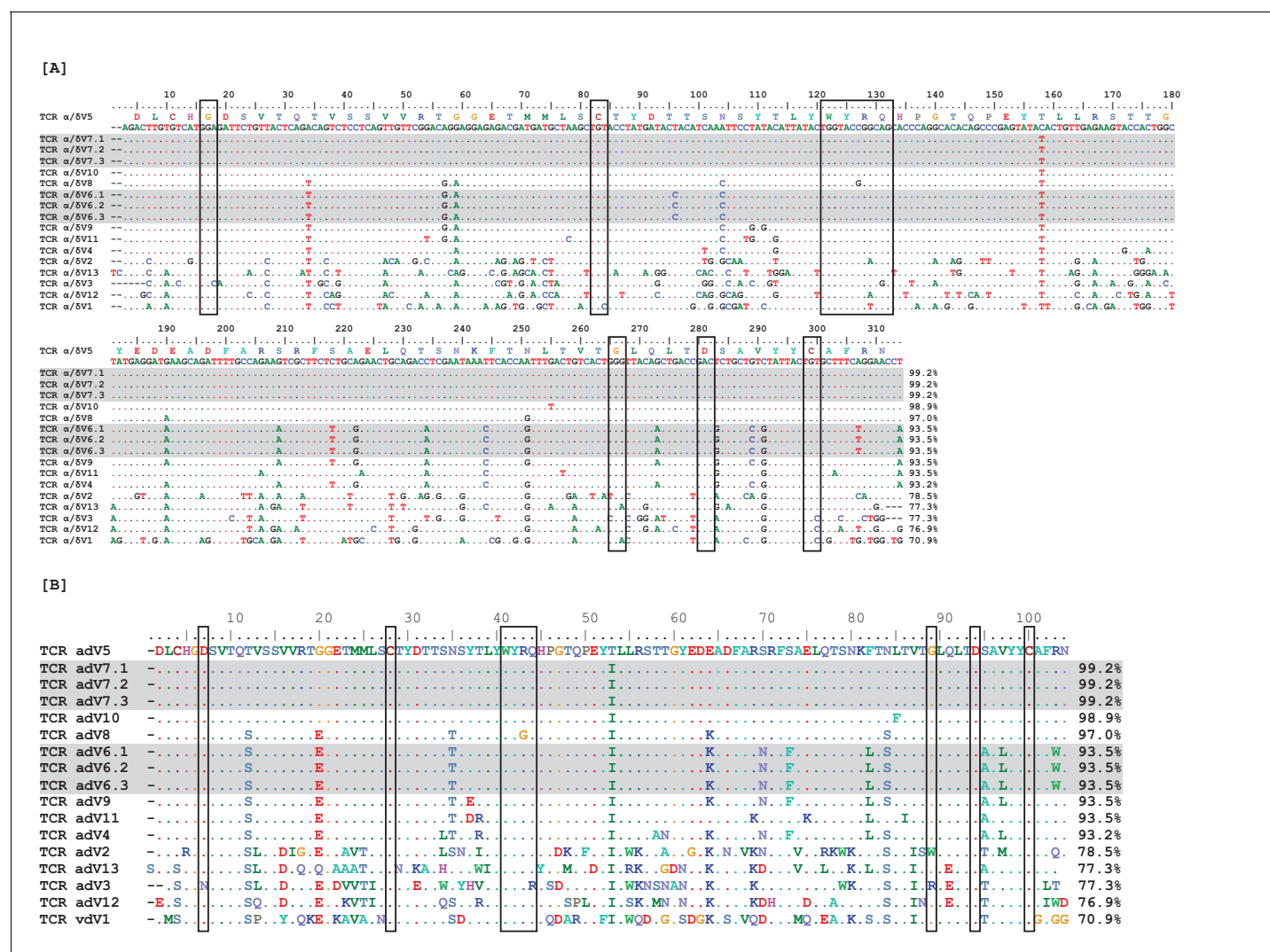
**Figure 9—figure supplement 2.** Lack of AID and TcR $\alpha$  probe hybridization in shark brain. Single molecule RNA fluorescence in situ hybridization (FISH) showing fixed brain sections on four separate slides demonstrating the lack of probe localization in non-immune tissue. We probed simultaneously for both TcR $\alpha$  and AID (a-c), individually for AID (probes labeled with Quasar 670 and imaged with Cy5 filter; pseudo colored red) (d-f), individually for TcR $\alpha$  (probes labeled with CalFluor Red 610 and imaged with Cy3 filter; pseudo colored green) (g-i), or for neither probe (negative control) (j-l). We show probe staining in shark thymus tissue as a positive control [m-o]. We counterstained all slides with DAPI (blue). We imaged fluorescence detected using each filter (DAPI, Cy3/Cy5) regardless of the probe used to illustrate that background fluorescence is not dependent on probe application. We obtained images of each fluorophore using 10x magnification and merged Z-stacked images together (DAPI fluorescence [a,d,g,j,m]; Cy3 (green) and Cy5 (red) fluorescence [b,e,h,k,n]; merged [c,f,i,l,o]). Scale bar 150  $\mu$ m.

DOI: <https://doi.org/10.7554/eLife.28477.021>



**Figure 10.** Model predicting how AID acts on T cells in the thymus. CD4/CD8 double negative (DN) thymocytes in the subcapsular region (SC) and cortex rearrange the  $\beta$  chain, using a surrogate pT $\alpha$  receptor to test for expression signaling. Cells with productive  $\beta$  arrangements then proliferate and express both CD4 and CD8, becoming double positive thymocytes (DPs). As DPs move toward the inner cortex and cortico-medullary junction (CMJ) where  $\alpha$  chain rearranges, cells may begin to express AID. Non-productive rearrangements can be rescued from apoptosis by receptor editing or by receptor salvaging, in which AID catalyzes SHM to produce cells with improved affinity to MHC:Ag complexes (to pass positive selection). Salvaged thymocytes then proliferate and express either CD4 or CD8 on their surface as single-positive (SP) cells. AID-mediated receptor salvaging may also reduce recognition of self-peptide, rescuing self-reactive thymocytes from apoptosis (to pass negative selection). [sc: subcapsular region; cmj: corticomedullary junction; green shading indicates region of AID expression].

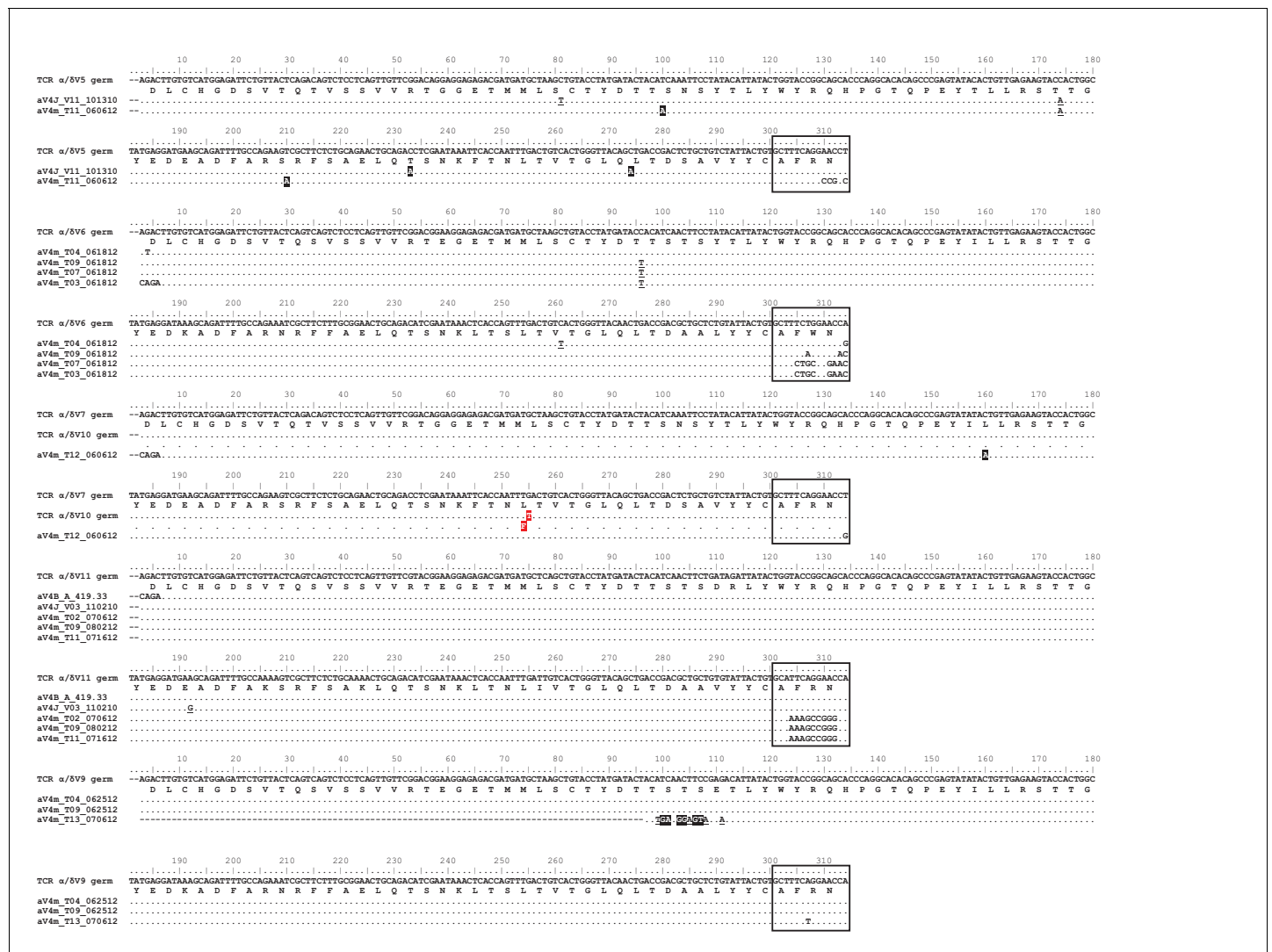
DOI: <https://doi.org/10.7554/eLife.28477.022>



**Figure 11.** Observed TcR Alpha/Delta germline Vs exhibit high sequence identity. Nucleotide (A) and amino acid (B) alignments of 17 germline variable (V) region gene segments. Two V groups contained three identical germline gene segments each (highlighted in gray), leaving only 13 unique V gene segments. Boxes surround conserved amino acids. Numbers at the ends of sequences indicate percent identity to the first germline sequence ( $\alpha/\delta$  V5) within the alignment.

DOI: <https://doi.org/10.7554/eLife.28477.024>





**Figure 12.** Observed germline sequences align only to TcR  $\alpha$ V4 clones. Nucleotide alignments of TcR  $\alpha$ V4 thymocyte clones to known germline V segments. Highlighted and underlined bases denote nonsynonymous and synonymous differences (respectively) to the germline V segment. Boxed regions represent nucleotides of the third complementarity-determining region (CDR3) according to IMGT guidelines, accounting for differences between clone and germline sequences. We highlight the single nucleotide/amino acid change between  $\alpha$ / $\delta$  V7 and  $\alpha$ / $\delta$  V10 germline segments in red.

DOI: <https://doi.org/10.7554/eLife.28477.025>