



Analysis of shark *NCR3* family genes reveals primordial features of vertebrate NKp30

Allison Kinlein¹ · Morgan E. Janes¹ · Jacob Kincer¹ · Tereza Almeida^{2,3} · Hanover Matz¹ · Jianxin Sui¹ · Michael F. Criscitiello⁴ · Martin F. Flajnik¹ · Yuko Ohta¹

Received: 24 August 2020 / Accepted: 7 February 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract

Natural killer (NK) cells play major roles in innate immunity against viruses and cancer. Natural killer receptors (NKR) expressed by NK cells recognize foreign- or self-ligands on infected and transformed cells as well as healthy cells. NKR genes are the most rapidly evolving loci in vertebrates, and it is generally difficult to detect orthologues in different taxa. The unique exception is NKp30, an activating NKR in mammals that binds to the self-ligand B7H6. The NKp30-encoding gene, *NCR3*, has been found in most vertebrates including sharks, the oldest vertebrates with human-type adaptive immunity. *NCR3* has a special, non-rearranging VJ-type immunoglobulin superfamily (IgSF) domain that predates the emergence of the rearranging antigen receptors. Herein we show that *NCR3* loci are linked to the shark major histocompatibility complex (MHC), proving *NCR3*'s primordial association with the MHC. We identified eight subtypes of differentially expressed highly divergent shark *NCR3* family genes. Using in situ hybridization, we detected one subtype, NS344823, to be expressed by predominantly single cells outside of splenic B cell zones. The expression by non-B cells was also confirmed by PCR in peripheral blood lymphocytes. Surprisingly, high expression of NS344823 was detected in the thymic cortex, demonstrating NS344823 expression in developing T cells. Finally, we show for the first time that shark T cells are found as single cells or in small clusters in the splenic red pulp, also unassociated with the large B cell follicles we previously identified.

Keywords NKp30 · Comparative genomics · Sharks · MHC · T cells · Secondary lymphoid tissue

Introduction

Natural killer (NK) cells are components of the innate immune system that target virally infected cells and tumors, using a combination of activating and inhibitory cell surface receptors. Inhibitory receptors survey the cell surface for normal, healthy markers, including major histocompatibility

complex (MHC) class I, thereby inhibiting NK cell activation. Activating receptors detect specific cell-surface molecules on infected or transformed cells and activate NK cells to eliminate those cells via direct killing or cytokine production. The balance of signals transduced by activating and inhibitory receptors dictates the outcome of NK cell activation during the immune response (Vitale et al. 2019). Most NK receptors (NKR) contain either C2-type immunoglobulin superfamily (IgSF) (e.g., KIR) or C-type lectin (e.g., LY49) domains, and NKR gene numbers for each type vary greatly in different vertebrate species (Yoder and Litman 2011). There are other receptors on NK cells, such as natural cytotoxicity triggering receptors (NCR) 1–3. Although all NKRs play roles in NK cell lysis, the three NCR are different in structure and are derived from a different evolutionary origin, although each of them is encoded in the MHC or an MHC paralogous region. NCR1 has a C2-type IgSF (Pessino et al. 1998) and is encoded near the Leukocyte Receptor Complex (LRC) on human chromosome 19. NCR2 (Cantoni et al. 1999) has a special V-IgSF that is found in TREM and

✉ Yuko Ohta
yota@som.umaryland.edu

¹ Department of Microbiology and Immunology, University of Maryland Baltimore, Baltimore, MD 21201, USA

² Centro de Investigação Em Biodiversidade E Recursos Genéticos, CIBIO-InBIO, Campus Agrário de Vairão, Universidade Do Porto, Vairão, Porto, Portugal

³ Departamento de Biologia, Faculdade de Ciências da Universidade Do Porto, Porto, Portugal

⁴ Department of Veterinary Pathobiology, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843, USA

PIGR family genes and is encoded in the MHC (Allcock et al. 2003). *NCR3* has a unique type of IgSF domain (see below) and is also encoded in the human MHC (Pende et al. 1999). Most NKR evolve rapidly and are thus generally species-specific, including *NCR1* and *NCR2*. Note that outside mammals, only bony fish possess *NCR2* (Stet et al. 2005). However, we have previously identified *NCR3* (NKp30) genes from sharks and amphibians, revealing that *NCR3* is the most evolutionarily conserved NKR preserving some primordial features of NKR and rearranging antigen receptors (Flajnik et al. 2012).

Mammalian *NCR3* is a single-copy gene encoding an activating receptor that contains positively charged amino acid residues (arginine) in the transmembrane domain that make an ionic bond with aspartic acid or glutamic acid residues in the transmembrane domains of adaptor proteins (e.g., CD3 ζ) (Pende et al. 1999). Adaptors have the immunoreceptor tyrosine-based activation motif (ITAM) in their cytoplasmic tails that initiate activation signaling cascades in NK cells resulting in either the production of cytokines or cellular cytotoxicity (Kruse et al. 2014). Unlike most NKR with C-type lectin or C2 IgSF domains, *NCR3* has a single extracellular domain of the so-called VJ-type IgSF (VJ-IgSF) that resembles antigen receptors (i.e., immunoglobulin (Ig) and T cell receptor (TCR)) (Flajnik et al. 2012; Du Pasquier et al. 2004). However, unlike antigen receptors in which the VJ-IgSF domains are generated by recombination activating gene (RAG)-mediated somatic gene rearrangement of several gene segments (Flajnik 2016; Tonegawa 1983), the VJ-IgSF of *NCR3* is non-rearranging and encoded by a single exon. It has been hypothesized that such single-exon VJ-IgSF domains predated the emergence of antigen receptors (Du Pasquier et al. 2004; Sakano et al. 1979), and thus, receptors containing these domains may be related to ancient immune receptors (Du Pasquier et al. 2004; Fu et al. 2019; Ohta et al. 2019).

Mammalian *NCR3* is expressed on the surface of NK cells, recently expanded $\gamma\delta$ T cells, activated CD8 cells, and innate lymphocytes (Hudspeth et al. 2013; Correia et al. 2018). *NCR3* binds to the self-ligand B7H6, which is upregulated on tumor cells in humans. Interaction of *NCR3* with B7H6 leads to activation of NK cells and induces cytotoxicity of target cells resulting in apoptotic cell death (Brandt et al. 2009). We have observed a strong positive correlation between the presence, absence, and number of *NCR3* and *B7H6* genes in all vertebrates examined, suggesting coevolution between *NCR3* and B7H6 throughout vertebrate evolution (Ohta and Flajnik 2015).

Sharks are Chondrichthyans, the oldest jawed vertebrate class possessing a mammalian-type adaptive immune system based on presence of Ig, TCR, RAG-based gene rearrangement, and the MHC (Flajnik 2018). Previous studies demonstrated that the shark genome is stable yet preserves

some primordial features, and therefore, Chondrichthyans are suitable models to study genome evolution (Ohta et al. 2011; Venkatesh et al. 2007, 2014). Since *NCR3* genes are MHC-linked in mammals, we sought to determine whether the *NCR3* linkage to MHC is primordial by examining its linkage to the shark MHC.

In addition to genomic analyses, we examined expression of the shark *NCR3* genes. Over 35 years ago, Pettey and McKinney discovered spontaneously lysing shark killer cells proposed to be NK cells (Pettey and McKinney 1983), but to date there have been no molecular tools to definitively identify such cells. In this study, we identified divergent shark *NCR3* family genes and analyzed *NCR3* expression using the nurse shark. Our studies revealed some unexpected expression patterns of this ancient NKR.

Materials and methods

Database searches

We searched for nurse shark *NCR3* transcripts with tBLASTn using previously reported (Flajnik et al. 2012) *Xenopus* and dogfish shark (ES788778; Fig. 1) *NCR3* sequences in the nurse shark RNAseq databases generated from spleen and thymus (NCBI (ncbi.nlm.nih.gov) GSM1893239–GSM1893252 (Venkatesh et al. 2014)). All hits were further validated using BLASTx against vertebrate “nr” databases to confirm the best matches were *NCR3* from other vertebrate species including mammals. *NCR3* homologs of other vertebrates were also found during the process, and they were further confirmed against mammalian databases with BLASTp. *NCR3* genes were also searched in the cartilaginous fish protein databases in NCBI by BLASTp using translated nurse shark *NCR3* genes. In addition, for all cartilaginous fish, we searched their genomic scaffolds by tBLASTn and identified VJ-IgSF exons. Those exons only present in the genome sequences were further evaluated by BLASTx to confirm homology to *NCR3*. In the species whose genome has been assembled, we obtained location of the *NCR3* genes in the “Gene” page at NCBI and examined other genes in the scaffolds. All accession numbers are taken from the NCBI GenBank site. Sequence information can be retrieved by searching the “Gene” page with given accession numbers.

Bacterial artificial chromosome library screening

We screened 17 BAC library (Luo et al. 2006) filters with radio-labeled variable (V) IgSF domains of NS344823 and NS348460 under high stringency conditions (Bartl et al. 1997). Membranes were exposed to X-ray film

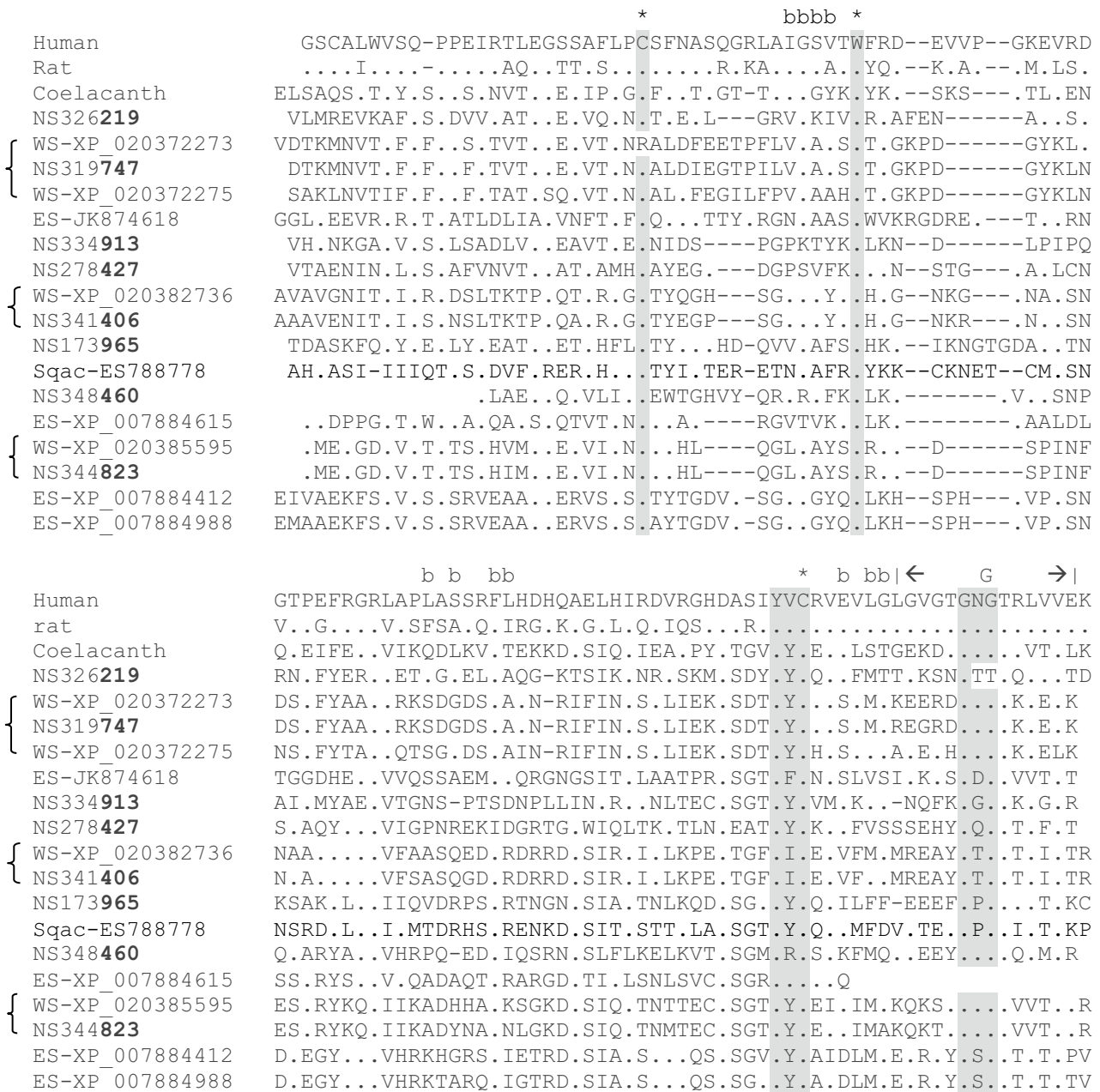


Fig. 1 Alignment of the nurse shark VJ exon and other vertebrate NCR3 protein sequences. The VJ domains of eight nurse shark (NS) sequences identified from RNAseq and NCR3 family genes identified in the whale shark and elephant shark were aligned using clustal omega along with human, rat, and coelacanth NCR3 sequences. The grey-shaded residues are amino acid residues of typical IgSF (also marked with asterisk) and “GXG” in the G-strand. B7H6 interaction sites of human NCR3 are denoted with “b” above alignment based on the co-crystal structure of human NCR3 and B7H6 (Joyce et al.

2011; Li et al. 2011). Dots and dashes indicate the same amino acid residues compared with human or a gap, respectively. Sequences starting with NS, WS, and ES are from nurse shark, whale shark, and elephant shark, respectively. Gene models (i.e., predicted genes) starting with “XP_” are accession numbers from the whale shark (WS: *Rhincodon typus*) and elephant shark (ES: *Callorhynchus milii*) genomes. Bracketed sequences are orthologous gene groups among different sharks based on the phylogenetic tree analyses (Fig. 4)

to identify locations of the positive clones. Positive clones were purchased from Arizona Genome Institute (www.genome.arizona.edu) and grown in media with chloramphenicol (12.5 µg/ml). BAC DNA was extracted

using the alkaline-lysis method and sequenced at the University of Maryland, Genomics Resource Center (<http://www.igs.umaryland.edu/grc>), using PAC Bio platform. Sequences were submitted to GenBank (accession numbers

MT914168–MT914170). Note that the BAC131D21 contains an ambiguous block region, thus account for two accession numbers.

Southern blotting

Genomic DNA (5 µg) from shark erythrocytes was digested with various restriction enzymes (RE) to obtain allele-specific restriction fragment-length polymorphism patterns (RFLP). To obtain the number of each NCR3 subfamily gene, we selected four siblings (sib 3, 17, 35, and 39) from the previously MHC-typed family members (family of mother and 39 siblings (Ohta et al. 2002)). Since this family is the result of multiple paternity, the selected four siblings carry different combination of maternal and paternal MHC haplotypes based on the genotypes of MHC class I, MHC class I processing genes, and class II genes. We also used five unrelated animals for the NS334913 probe. Each V domain of nurse shark *NCR3* genes was used as ³²P-labeled-probes, except for the NS348460 probe, which contains parts of the transmembrane- and leader-encoding regions. All hybridizations were performed under high-stringency conditions (50% Formamide/2× saline sodium citrate (SSC)/5× Denhard's/2% sodium dodecyl sulfate (SDS) at 42 °C), and washes under moderate-stringency washes (medium (0.5× SSC/0.2% SDS at 55 °C) followed by high stringency (0.2× SSC/0.2% SDS at 65 °C) to detect all genes belonging to each subfamily. Hybridizing bands were detected by exposing membranes to X-ray film.

Phylogenetic tree analyses

We determined the exons encoding the VJ-IgSF domains from genomic sequences of most genes at NCBI (<https://www.ncbi.nlm.nih.gov/gene/>). When genes were not annotated, exonic regions were translated and aligned with those of other species to determine the exon-intron boundaries. These exons were aligned using “Muscle” and conducted maximum-likelihood (ML) with JTT-matrix model (Jones et al. 1992) and UPGMA tree analyses using the MEGAX program (Kumar et al. 2018). Bootstrap was set to 500 runs. We included all cartilaginous fish NCR3 family genes as well as other species including coelacanth, *Xenopus*, alligators, turtles, and mammals. As control genes, we included genes containing non-rearranging VJ-IgSF domains, especially those mapping near the MHC region in non-mammalian species, such as PRARP (Fu et al. 2019; Ohta et al. 2019), XMIV (Ohta et al. 2006), and *Xenopus* VJC1₂₅₈ (Ohta et al. 2019). We also included amphioxus IgVJ-C2 (Chen et al. 2018), an invertebrate gene that contains an VJ-IgSF domain. As a distantly related gene, we added CD83, which has a V-IgSF and maps near the MHC. We also included V-IgSF domains from B7H6 and TAPBP,

which share a common ancestor among themselves, but do not share a recent common ancestor with NCR3. To root the trees, we used the Constant-type 1 IgSF (C1-IgSF) domain of β2-microglobulin (B2M). The *B2M* genes map in the shark MHC (Ohta et al. 2011). Accession numbers used for these analyses are provided in Fig. 4 and Online Resource 4.

RT-PCR

First-strand cDNA was synthesized from 500 ng of total RNA isolated from various nurse shark tissues using the SuperScript™ IV first-strand synthesis system (Invitrogen) following the manufacture's recommendations. The cDNA from various tissues was used for PCR using inter-exon primers to eliminate amplification of residual DNA. Primers used for expression patterns are NS344823 (5'-aatctg-caccattacgtgctg-3'; 5'-agcactatctaagcttgattc-3'), NS348460 (5'-caatcagtcctcatcccgtg-3'; 5'-atacttcctgttagtcacc-3'), NS334913 (5'-aaaggagccgtggtgcagtc-3'; 5'-ctggttcgcc-cacaatgga-3'), NS319747 (5'-gatacgaag atgaatgtcac-3'; 5'-ctgccacgattagaacgcac-3'), and NS278427 (5'-gttctgcattct-tagtcttgg-3'; 5'-actgcttgagacaaattccac-3'). PCR was performed using the GoTaq enzyme (Takara Bio) with cycling conditions: 95 °C for 2 min (hot start), then cycled 35 times at 95 °C for 45 s, annealing temperature for 30 s, then 72 °C for 30 s, with final extension at 72 °C for 5 min. The annealing temperatures for NS348460 was 49 °C, NS319747 was 53 °C, NS344823 and NS334913 was 55 °C, and NS278427 was 57 °C. There was no amplification with primers for NS314406, NS173965, and NS326219. This could be either to technical problem in PCR or low expression within the tested tissues. Nucleoside diphosphate kinase (NDPK) (Kasahara et al. 1991) as PCR control: annealing at 50 °C with primers 5'-aacaaggaacgaaccttc-3'; 5'-atcttcaccagacaa-gcaac-3'. Equal amounts of cDNA were used as template, and equal amounts of PCR reactions were loaded onto a 1% agarose gel to examine the amplification products.

In situ hybridization

Freshly harvested nurse shark tissues were immediately fixed in 4% paraformaldehyde/phosphate buffer/sucrose solution for 6 days. After the sucrose equilibrated, tissue blocks were embedded in O.C.T. in a 2-methylbutane/liquid nitrogen bath. Frozen tissues were sectioned at 7 µm using Cryostat (Leica) and mounted onto glass slides. Slides were re-fixed in 4% paraformaldehyde in shark PBS (1.87 mM sodium phosphate monobasic/8.2 mM sodium phosphate dibasic/470 mM sodium chloride), permeabilized in Proteinase K (20 µg/ml) (Sigma), and acetylated in 0.25% acetic anhydride. Hybridization was performed in a 50% formamide

solution in 1× hybridization solution (Sigma) supplemented with yeast tRNA at 67 °C overnight. Slides were then washed in 0.2× SSC (0.003 M sodium citrate/0.03 M sodium chloride) twice at 72 °C for 30 min, blocked with TNB (0.1 M Tris, pH 7.5/0.15 M sodium chloride/0.5% w/v blocking reagent (Perkin Elmer) for 60 min at room temperature, and incubated with anti-DIG-POD (Perkin Elmer) for 30 min at room temperature. Signals were amplified using the TSA-plus Biotin System (Akoya Biosciences) following the manufacturer's protocol. Colorimetric signals were visualized after incubation with Streptavidin-alkaline phosphatase (SA-AP; Perkin Elmer) followed by substrate BCIP/NBT (Roche). Sense and antisense probes were labeled from the dual promoter vector pCRII (Invitrogen) containing part of the gene of interest. After the plasmids were singly digested with restriction enzyme that cleaves at either 5' or 3'-side of the polylinker, RNA was synthesized using RNA polymerases (T7 RNA polymerase or SP6 RNA polymerase depending on the orientation of the insert) with DIG RNA labeling mix (Roche). A full-length coding region of NS344823 was labeled. We used shark "G" (spleen) and "W" (thymus) for these experiments. These sharks were immunized with various antigens and boosted with original antigens approximately 10 days before harvesting tissues.

Magnetic cell separation and RT-PCR

Shark "LB" was originally caught off the Florida Keys and was maintained in captivity with no immune stimulation. Peripheral blood lymphocytes (PBL) (2×10^8) were resuspended into shark PBS supplemented with 10% FCS and 0.00025% DNase. A monoclonal antibody (mAb) specific for shark IgM, GA16 (Rumfelt et al. 2001), was mixed with cells and incubated on ice for 1 h. Cells were washed in shark PBS. Secondary antibody, goat anti-mouse IgG microbeads (Miltenyi Biotec) were then added on ice for 1 h. Cells were washed with MACS buffer (shark PBS supplemented with 0.5% BSA/2 mM EDTA/0.00025% DNase I) three times, resuspended in the recommended volume of MACS buffer, and loaded on the LS column (Miltenyi Biotec) following the manufacturer's recommendations. RNA was extracted from positively and negatively sorted cell populations as well as pre-sorted PBL, using the Trizol (Invitrogen) protocol, and cDNA was synthesized using SuperScript IV™ first-strand synthesis system (Invitrogen) from 500 ng of total RNA from both samples. The same PCR condition was used as described in Fig. 5 for NS344823 and NDPK. Other primers used for this experiments are TCR beta (5'-caccgggttaatttc-ccaa-3'; 5'-cagtcacacacctttgtgtgag-3') and IgM C2-TM (5'-catgggcagctgatcacg-3'; 5'-cacttagccttcaccagcgtg-3') at annealing temperatures 54 °C and 52 °C, respectively.

Results

Identification of nurse shark *NCR3* loci and deduced proteins

Previously, we identified NCR3 sequences in a dogfish shark EST database (ES788778) and the elephant shark genome (updated accession numbers: XP_007884412; XP_007884615) (Flajnik et al. 2012) (Fig. 1). In this study, we searched for NCR3 sequences by BLAST in the nurse shark RNAseq databases constructed from spleen and thymus RNA (Venkatesh et al. 2014) and found eight distinct sequences (Fig. 1) (GenBank accession numbers MT914154-MT914162). Deduced amino acid sequences revealed that all of the protein sequences contain a single VJ-IgSF type domain with the typical disulfide bond between the two canonical cysteine residues (C) in the B and F strands that are stabilized by tryptophan (W) 13–16 amino acids downstream of the first cysteine (marked with asterisk and gray highlights in Fig. 1). Each sequence also contains a typical YXC motif in the F strand and a conserved diglycine (G) bulge (GXG) (gray highlights), that latter of which is usually found in joining (J) segments (the G strand of IgSF domains) of antigen receptors (Flajnik et al. 2003; Williams and Barclay 1988). All sequences contain transmembrane domains (Online Resource 1a). Only two of the eight nurse shark sequences (NS173965 and NS326219) contain positively charged amino acid residues within the transmembrane domains, like human NCR3. As mentioned above, these positively charged amino acid residues are predicted to interact with adaptor molecules with immunoreceptor tyrosine-based activation motifs (ITAMs) (e.g., DAP10, DAP12, CD3z, FCRg) (Humphrey et al. 2005). Note that shark CD3z has been identified (Pettinello et al. 2017), whereas DAP12 and FCRg seem to be present (e.g., GenBank accession numbers AGQ17903, XP_038656720, XP_032898506), but must be further validated. Three sequences have long cytoplasmic tails with immunoreceptor tyrosine-based inhibitory motifs (ITIM) (Long 2008) (NS319747, NS326219, and NS348460), suggesting that they may be inhibitory receptors. Interestingly, NS326219 contains both positively charged amino acid in transmembrane domain and ITIM in the cytoplasmic tail. Three other sequences (NS344823, NS341406, and NS34913) do not contain either positively charged amino acids or ITIMs. NS278427 is a partial sequence, lacking the 3'-tail. Alignment of the VJ domains revealed that the nurse shark sequences are highly divergent, with only 23–38% identity (Online Resource 1b). The degree of divergence is similar between the nurse shark and dogfish shark ES788778 (~23–30%) or elephant shark sequences with exception of ~47% between

ES788778 and NS173965 (Online Resource 1b). To our surprise, we found that the NS314406, NS319747, and NS344823 are highly similar to whale shark NCR3 genes XP_020385595 (97–100% identity between two species, bracketed sequences in Fig. 1). Since the whale shark and nurse shark diverged over 100 million years ago, the high conservation of the NCR3 sequences was unexpected. This level of conservation was not observed for other shark species within the available databases.

We previously reported that the number of *NCR3* genes and those of its ligand, *B7H6*, correlate well within all vertebrate species examined, and we have proposed coevolution between these two gene families (Flajnik et al. 2012; Ohta and Flajnik 2015). Here, we examined potential B7H6 binding capacities in these nurse shark sequences. Compared with the B7H6-binding residues of human NCR3 (marked with “b” in Fig. 1) (Joyce et al. 2011; Li et al. 2011), some residues are also conserved in sharks; “IGSV” is found as “IGSY” in NS314406, phenylalanine (F) in position 85 is well conserved among many shark sequences, and “ExLG” is perfectly conserved in NS334913. The glutamic acid (E) makes a salt bridge with lysine (K) at position 130 of B7H6, and this E is conserved in four shark sequences. Based on these conserved contact residues, we predict that some shark NCR3 subtypes may bind to particular B7H6. We found multiple (at least five) B7H6 sequences in shark (GenBank accession numbers MT914163–MT914167). Similar to the case of NCR3, some NCR3-interaction residues (Li et al. 2011) are well conserved between human and shark B7H6 (Online Resource 2): seven out of eleven residues are conserved, especially “TPLK,” which is perfectly conserved in NS332265 and “TP” is conserved in four shark B7H6 genes. Glycine (G) at positions 62 and 83 are also conserved. Based on the conservation of interaction residues, there may be binding pairs between shark NCR3 and B7H6.

To obtain the genomic organization of the nurse shark *NCR3* genes, we screened a nurse shark bacterial artificial chromosome (BAC) library with $\sim 11\times$ genome coverage (Luo et al. 2006) with the combined probes for NS344823 and NS348460. We identified ~ 30 positive clones, of which 12 clones were analyzed. All clones were exclusive to either the NS344823 or NS348460 probes. We sequenced the two longest BAC clones of each type, 131D21 and 392N18. Both BAC clones contained multiple *NCR3* genes, but there was no overlapping sequence between the clones (Fig. 2a). BAC 131D21 with a 162 kb insert contains four *NCR3* genes, corresponding to NS344823, NS334913, NS319747, and NS513006. NS513006 is a pseudogene having two frameshifts in both the mRNA (from RNAseq) and genomic BAC sequences, resulting in a premature stop codon (Online Resource 1c). There are two other predicted gene models in BAC 131D21, one of which has no orthologous genes in other species, while the other hit nicotinamide riboside

kinase 1 from Thorny Skate (*Amblyraja radiata*) with an *E*-value of $2e^{-115}$; however, as we did not detect orthologs of this gene in mammals, we classified it as a “hypothetical protein.” In the other BAC with a 139 Kb insert, 392N18, there are five *NCR3* genes, one corresponding to NS278427 and four genes highly similar to the NS348460 sequence. There are no other genes besides the *NCR3* genes in 392N18 BAC. In summary, we found a total of nine *NCR3* genes belonging to five subfamilies and one pseudogene in the two BAC clones. We do not know whether all nurse shark *NCR3* genes map in a single genomic region. Since only five subfamily genes were identified in the two BAC clones and there seems to be duplication of particular subfamily genes (i.e., four copies of NS348460 in BAC 392N18), we performed genomic Southern blotting to assess the number of genes in each subfamily and genotypic differences among individual nurse shark. With a single-exon probe, we detected multiple bands for several *NCR3* genes, revealing multiple copies. With the NS348460 probe, there are about ten visible bands. Since the NS348460 probe contains small parts of other exons, weakly hybridizing bands may be accounted for by leader and transmembrane exons. Based on the stronger hybridizing bands, we predict at least 4–5 genes of each allele if each allele is accountable for each band. There are also multiple bands with the NS314406 and NS278427 probes, and we estimate the copy number to be 3–4 genes for NS314406 and at least 5 genes for NS278427. All other genes seem to be single-copy.

Primordial linkage of shark *NCR3* genes to the MHC

The *NCR3* genes map to the MHC in all species examined so far except in *Xenopus*, although a close *NCR3* homologue, *XMIV*, is MHC-linked (Ohta et al. 2006) (Fig. 3) (note that several vertebrate taxa (mainly birds and bony fish) have lost *NCR3* genes). Therefore, it is reasonable to propose that *NCR3* linkage to the MHC is primordial and we predicted that shark *NCR3* genes would map to the MHC. We first examined the existing databases of the shark genomes for the location of *NCR3* genes. In the whale shark genome, we found four *NCR3* genes on three scaffolds, of which two scaffolds contained other genes that map to the mammalian MHC (Fig. 3): In scaffold NW_018056204, the complement *c4* gene maps adjacent to an *NCR3-like* gene, and in NW_018036233 *aif* maps adjacent to two *NCR3* genes. The *AIF* gene maps to the human and *Xenopus* MHC, and we have shown previously that shark *c4* genes are MHC-linked (Terado et al. 2003). All scaffolds containing the elephant shark *NCR3* genes were small, and no other genes were detected in these scaffolds. In both species, the MHC is poorly assembled. However, in the whitespotted bamboo shark genome (Hara et al. 2018), we identified eight *NCR3* VJ-IgSF exons that

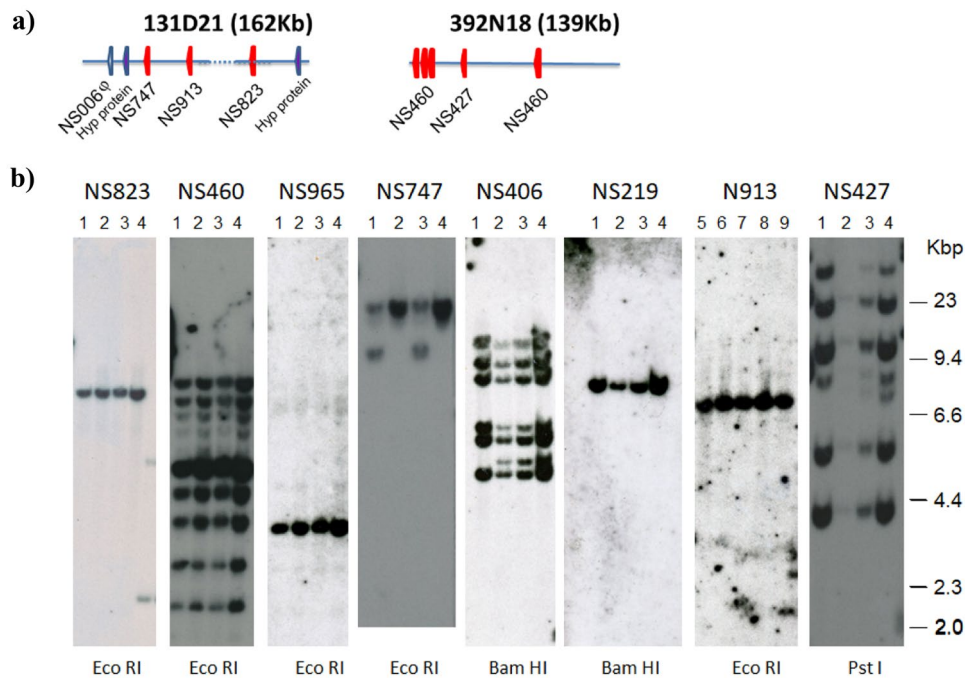


Fig. 2 Number of *NCR3* genes in the nurse shark genome and partial mapping of their loci. **a** Multiple *NCR3* genes were found in two BAC clones. These two BAC clones do not overlap, and thus, the linkage status of these *NCR3* genes is unknown. *NCR3* genes are indicated with red arrows that also show the 5' to 3' direction. Corresponding sequences from RNAseq are shown below (truncated to the last 3-digit numbers), and an expressed pseudogene (NS006) is shown as an open arrow. Two hypothetical proteins that have no similarity to mammalian genes by BLASTp searches are also shown as purple arrows. The locations of genes in the BAC clones are roughly scaled. An ambiguous block region in the middle of BAC 131D21 is marked as a dotted line and not included in the total length. **b** The approximate number of each *NCR3* genes in the nurse shark genome.

Gene names are shown with the last 3-digit numbers from the original identification numbers. Genomic Southern blotting was performed using four nurse shark siblings with different combinations of maternal and paternal MHC genotypes except for NS334913 (shown as NS913), for which we used unrelated sharks. Probes for each gene family were generated from the VJ-IgSF exon, except NS348460 (shown as NS460) which contains short sections of the transmembrane and leader encoding regions. Restriction enzymes used to digest DNA are shown on the bottom of the blots. HindIII digested λ DNA was used as the size standard and shown on the right side of the blots. Lanes/Animals: 1-sibling 3, 2-sibling 17, 3-sibling 35, 4-sibling 39 (from MHC typed family with 39 siblings (Ohta et al. 2002); 5–9-unrelated nurse sharks

mapped around the *c4* genes on chromosome 37, scaffold CM012992 (Fig. 3, Online Resource 3). The region is indeed a part of the shark MHC on chromosome 37 (<https://doi.org/10.1101/602136>), containing MHC class I and class II genes, thus confirming the linkage of *NCR3* genes to the cartilaginous fish MHC.

Despite the varying numbers of *NCR3* genes found in each vertebrate species so far examined (Fig. 3), their synteny is well conserved from shark to human. In *Xenopus*, *NCR3* genes map outside MHC on a different chromosome, while structurally similar genes (*XMIV*) are MHC-linked and at the expected position (Ohta et al. 2006); this split arrangement is likely the result of differential (or incomplete) silencing of *NCR3* genes during genome evolution, leaving *XMIV* and *NCR3* on two MHC paralogous regions (Ohta et al. 2019). Alternatively, there may have been a translocation of *NCR3* genes out of the MHC in an amphibian ancestor, as the *Xenopus* *NCR3* genes are located at the telomere of chromosome 4. In summary, *NCR3* genes are

linked to the MHC in sharks, demonstrating that their association with the MHC is primordial in gnathostomes.

Phylogenetic analysis of shark *NCR3* family genes

Both the BLAST “hits” and linkage to MHC are evidence that the shark *NCR3* genes are closely related to mammalian *NCR3* genes. However, shark *NCR3* sequences are highly divergent (Online Resource 1b) and some lack both positively charged amino acid residues in the transmembrane region or ITIMs. Thus, we performed phylogenetic tree analyses to examine relationship of shark *NCR3* to those of other vertebrates. For these analyses, we data-mined shark *NCR3* sequences from all available databases. As control genes (Online Resource 4), we used other MHC-linked or MHC-associated (on MHC paralogous regions) “germline-joined” VJ-IgSF-containing genes such as non-mammalian vertebrate PRARP (Fu et al. 2019; Ohta et al. 2019), and *Xenopus*-specific VJC1₂₅₈ (Ohta et al. 2019) and *XMIV*

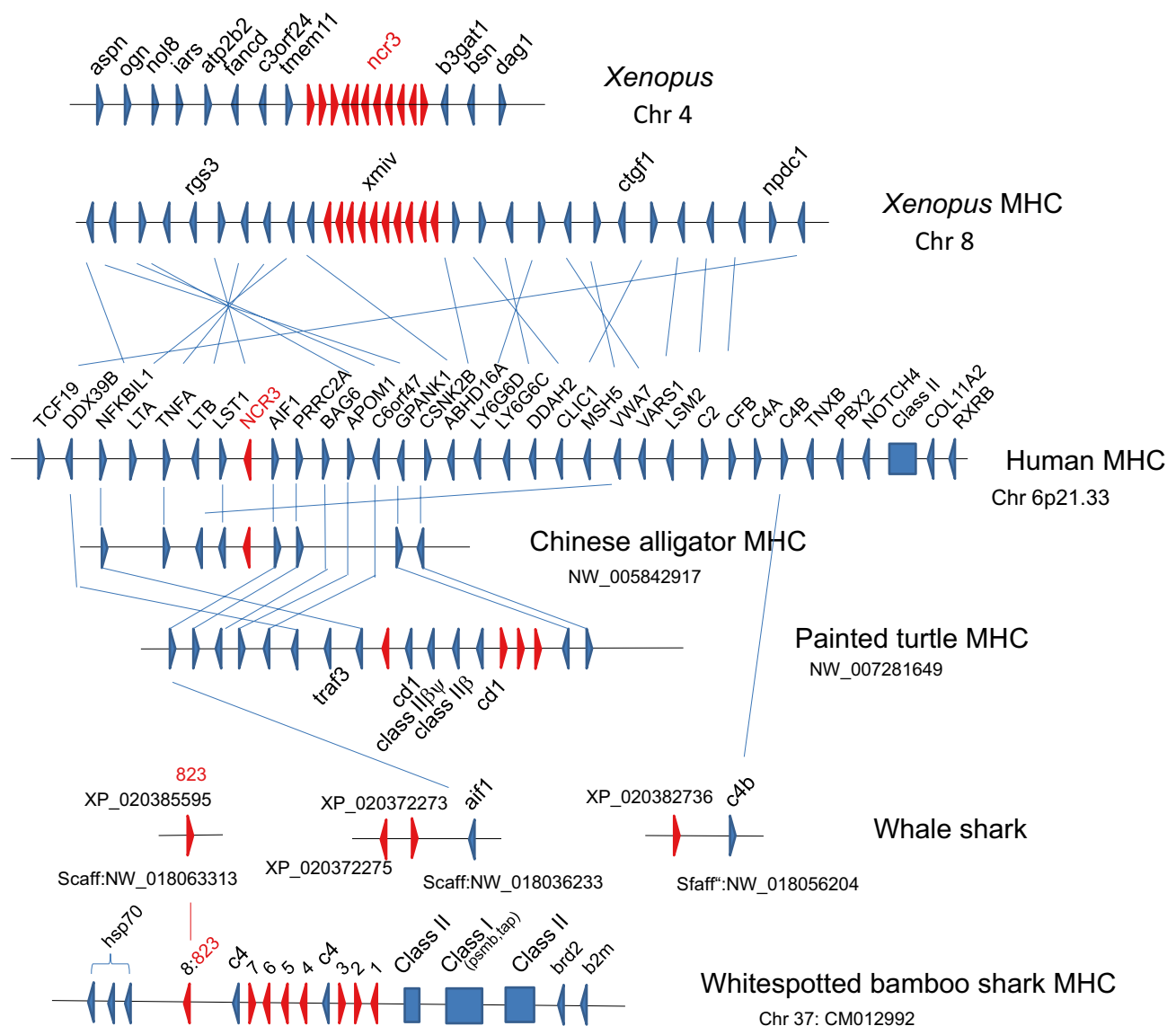


Fig. 3 Comparative synteny of *NCR3* genes in vertebrate genome. *NCR3* genes are shown with red arrows, while neighboring genes are in blue; arrows indicate the 5' to 3' direction. Blocks represent a cluster of genes. Vertical lines indicate conservation of position among the different species. Gene names are shown above the arrows. Only a portion of the genome is shown in human and *Xenopus* (Session et al. 2016), while genes in entire scaffolds are shown in whale shark, Chinese alligator, and painted turtle. GenBank accession numbers for

whale shark scaffolds and *NCR3* genes are shown above. *NCR3* genes in whitespotted bamboo shark are based on the presence of VJ-IgSF domains, and thus, some may not represent functional genes. Note that the gene most proximal to hsp70 is the NS344823 ortholog. Gene locations are not to scale. Note that the *Xenopus* *NCR3* genes map to chromosome 4 instead of chromosome 8 (MHC). However, the structurally related genes *XMIV* map in corresponding location on chromosome 8 (see text)

(Ohta et al. 2006, 2019). We also included amphioxus IgVJ-C2 genes (Chen et al. 2018) since they contain a VJ-IgSF domain and are syntenic with kirrel, which maps to MHC paralogous regions in vertebrate genomes. As an outgroup for V-IgSF, we used CD83, which also maps in the vicinity of MHC. The V-IgSF of CD83 is split into two exons in all species examined and is predicted to have a different common ancestor compared with *NCR3*. Furthermore, we included the V-IgSF of B7H6, the *NCR3* ligand, as another

outgroup. Inclusion of B7H6 serves two purposes: (1) an outgroup and (2) an assessment of orthology of the five nurse shark *B7H6* genes. Since we included B7H6, we also included the V-IgSF of TAPBP (the second IgSF domain), which shares a common ancestor with V-IgSF of B7 family members including B7H6 (Flajnik et al. 2012; Greenwald et al. 2005). Finally, we used B2M as another, distant outgroup, since B2M contains a C1-type IgSF domain. V-, VJ-IgSF domains and C1-IgSF domains are clearly in

different IgSF categories (Du Pasquier et al. 2004; Williams and Barclay 1988).

In both maximum likelihood (ML) and UPGMA trees (Fig. 4), shark NCR3 genes were separated into two groups. One group clustered with the NCR3 of other vertebrate species, suggesting that these are mammalian NCR3 orthologues. The other group seems to be more divergent, only present in sharks, suggesting these may be an ancient NCR3 lineage lost in other vertebrates or a cartilaginous fish-specific NCR3 cluster. Four of eight nurse shark NCR3 genes belong to each cluster: NS341406, NS173965, NS278427, and NS348460 cluster with mammalian NCR3, while NS344823, NS334913, NS326219, and NS319747 cluster only with other cartilaginous fish sequences. Interestingly, three genes in this group, NS334913, NS319747, and NS344823 are found in BAC 131D21. Whether the location of these genes is contributing to the conservation of sequences must be examined in the future. In the whitespotted bamboo shark genome, the NS344823 ortholog (Chpl-8: mapping between *c4* and *hsp70*) is somewhat distant from the other NCR3 genes (Fig. 3). As we showed high conservation among NS344823 and whale shark XP_020372275 (Fig. 1), NS344823 clustered with whale shark XP_020385595 in both trees and also with whitespotted bamboo shark, Chpl-8 (see alignment in Online Resource 3), confirming that these genes are NS344823 orthologues.

Previously, we found *XMIV* genes in the *Xenopus* MHC where mammalian *NCR3* genes map in other vertebrates (Ohta et al. 2006), while *Xenopus NCR3* genes map to an MHC paralogous region. We hypothesized that *XMIV* and *NCR3* arose from a common ancestor and were differentially silenced in the *Xenopus* genome, but an ancient translocation of the *Xenopus NCR3* genes out of the MHC is also a possibility (see above) (Ohta et al. 2019). In these trees, *XMIV* and cartilaginous fish-specific *NCR3* formed a poorly supported sister group, suggesting that *XMIV* and *NCR3* were both present and MHC-linked in an ancient gnathostome ancestor. Thus, it is possible that the *XMIV* and shark-specific *NCR3* make up one *NCR3* lineage and the second *NCR3* lineage is found from sharks to mammals. Alternatively, *XMIV* genes may be under a different selection pressure than *NCR3* and evolved rapidly only in the *Xenopus* (and perhaps other amphibian) group. Amphioxus IgVJ-C2 (Chen et al. 2018) clustered consistently as a sister group with the NCR3/*XMIV* group. Other genes with VJ-IgSF, such as PRARP (Fu et al. 2019) and CD83, clustered outside of the NCR3 group.

As B7H6 and TAPBP do not contain VJ-IgSF, we did not expect them to share a recent common ancestor with NCR3; they were both well separated from NCR3 clusters. The elephant shark B7H6 was previously used to construct a B7 family tree including all known B7 family members (Flajnik et al. 2012); therefore, nurse shark sequences clustering with

other species including elephant shark confirm that the four nurse B7H6 genes are B7H6 orthologues. In the elephant shark scaffold NW_006890061.1, multiple *B7H6* genes map and their synteny is well conserved compared with the human *B7H6* locus in chromosome 11p15.1 (data not shown). One exception is NS331224. NS331224 hit B7H6 by BLAST, but NS331224 is more divergent than other nurse shark B7H6 (Online Resource 2) and contains less conserved NCR3-binding residues. Interestingly, NS331224 has a type II transmembrane composition with NLRC3-like cytoplasmic tail. Based on the position in the phylogenetic trees, NS331224 may not be a true B7H6 homologue.

Differential expression patterns of each *NCR3* genes

Unlike human *NCR3*, we found eight subfamilies of *NCR3* transcripts in the nurse shark RNAseq data (Fig. 1), some of which have multiple copies (NS348460, NS278427, NS314406) that were likely generated by *cis*-duplications based on the sequences in BAC clones (Fig. 2a). Thus, we examined the tissue distribution of each gene. Using RT-PCR (Fig. 5), we detected differential expression of the five types of *NCR3* genes. NS344823, NS334913, and NS278427 were expressed in many tissues, while NS348460 and NS319747 expression was limited to certain tissues. Highest expression was observed for each gene in gill, gonad, spleen, and thymus for NS344823; brain, liver, and spleen for NS334913; gill, pancreas, spiral valve (small intestine in sharks), spleen, and thymus for NS348460; gill, spiral valve, and thymus for NS278427; and gill and spleen for NS319747. Such differential expression implies that there is a division of labor for each *NCR3* gene. Note that for multigene families, we cannot rule out that the PCR only amplified one of the highest expressing genes or a combined expression of multiple genes. We did not detect amplification of three genes, NS314406, NS326219, and NS173965, suggesting their expression may be low. Similarly, some B7H6 genes (NS314508 and NS333295) showed tissue-specific expression patterns (Online Resource 5).

Identification of *NCR3*-expressing cells

To identify cell types expressing nurse shark *NCR3* genes, we performed in situ hybridization (ISH) using the DIG-labeled riboprobes. For these analyses, we initially chose NS344823 and NS348460. NS344823 is a single copy gene, belongs to the cartilaginous fish-specific *NCR3* lineage, contains a short cytoplasmic tail with no ITIM (Online Resource 6), and demonstrated robust amplification by RT-PCR (personal observation). NS344823 has orthologs in other shark species (Fig. 4) such as whale shark and whitespotted bamboo shark (Hara et al. 2018) (the NS344823 orthologue is a gene closest to *hsp70* in whitespotted bamboo shark in

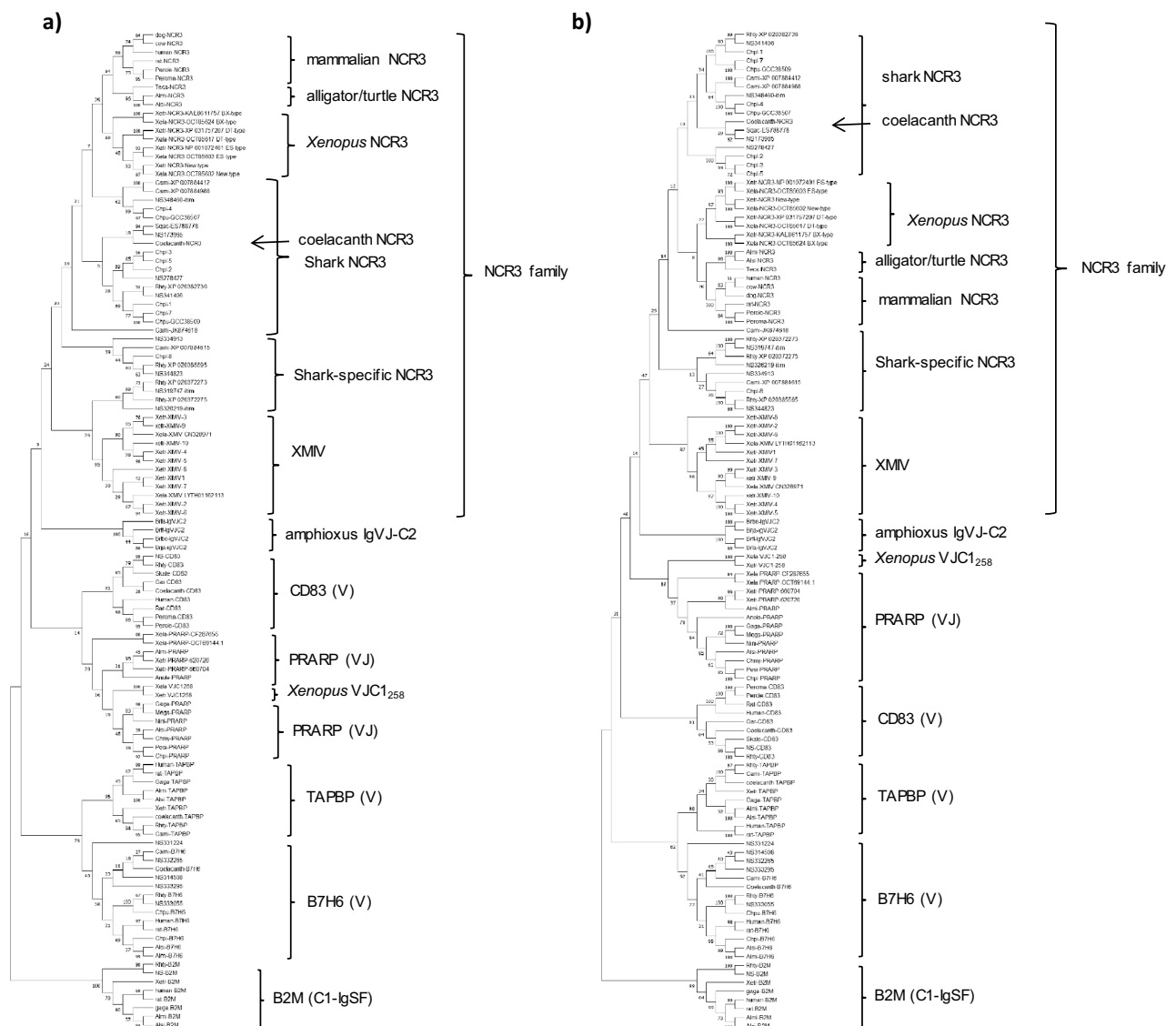


Fig. 4 Phylogenetic analysis of shark NCR3 genes. The VJ-IgSF domains of the cartilaginous fish NCR3 deduced proteins and other genes with VJ-IgSF domains were used for the bootstrapping phylogenetic tree analyses of 500 runs using Maximum-Likelihood (ML) with JTT-matrix based model (Jones et al. 1992) (a) and UPGMA (b) distance with the Poisson correction methods. The shown trees are consensus trees, and the percentage of replicated trees at the given cluster is shown at the nodes. Both trees were rooted with the C1-IgSF domain of vertebrate B2M. GenBank accession numbers used for these analyses are as follows: NCR3: human—NP_001138939, dog—XP_848928, cow—NP_001035614, rat—NP_861543, and deer mice—XP_028743796 (Perole: *Peromyscus leucopus*) and XP_006995306 (Peroma: *Peromyscus maniculatus bairdii*); American alligator—KYO39304 (Almi: *Alligator mississippiensis*) and Chinese alligator—XP_014381422 (Alsi: *Alligator sinensis*); three-toed box turtle—XP_026502865 (Teca: *Terrapene carolina triunguis*); coelacanth—XP_006012356 (*Latimeria chalumnae*); *Xenopus* NCR3: African clawed frog (Xela: *Xenopus laevis*)—OCT85617, OCT85603, OCT85602, and OCT85624 (corresponding to previously reported (Flajnik et al. 2012)) four types (DT, ES, New, BX-types); western clawed frog (Xetr: *Xenopus tropicalis*)—XP_031757207, NP_001072401, KAE8611757, retrieved

from genomic sequence JABAHN010000004 (position 123564763...123564428) which is ortholog of Xela-OCT85602; Shark NCR3: elephant shark (Cami: *Callorhynchus milii*)—XP_007884412, XP_007884988, XP_007884615, JK874618 (Frame+1 translated); whale shark (Rhty: *Rhincodon typus*)—XP_020382736, XP_020372273, XP_020372275, and XP_020385595.1; spiny dogfish (Sqac: *Squalus acanthias*)—ES788778 (Frame+1 translated); brownband bamboo shark (Chpu: *Chiloscyllium punctatum*)—GCC38507 and GCC38509; and *Xenopus* XMIV: Xela—CN328971, LYTH01162113; Xetr—VJ exon sequences of ten XMIV genes were retrieved from genome. All sequences were highly similar to either XP_031748205 or KAE8585365, B7H6: human—CAD97811, rat—XP_017443726, painted turtle—XP_005308752 (Chpi: *Chrysemys picta bellii*), Alsi—XP_014375990, Almi—XP_019332669, coelacanth—XP_014340774, Chpu—GCC17701, Rhty—XP_020389246, Cami—XP_007886037, nurse shark (MT914163-914167). Other genes containing VJ-IgSF domain are amphioxus IgVJ-C2 (marked as IgVJC2) (Chen et al. 2018), PRARP (Fu et al. 2019), and *Xenopus* VJC1₂₅₈ as well as V-IgSF domains of CD83 and TAPBP were used. The accession numbers except for NCR3 family genes or B7H6 genes are listed in the Online Resource 4. Also, in Online Resource 4, we listed species abbreviations

Fig. 5 Differential expression of *NCR3* genes in nurse shark tissues. Expression of each *NCR3* gene was examined by RT-PCR using gene-specific inter-exon primers. For direct comparison between each gene, equal amounts of templates were used for each PCR reaction and equal volumes were electrophoresed. The upper band of NS344823 is a partially spliced mRNA: exons 1 and 2 are spliced, but it contains intron 2 between exons 2 and 3). The expected sized products for NS348460 and NS278427 are shown with arrows. Size standards are shown on the left side of the gels. Nucleoside diphosphate kinase (NDPK) is used as internal control for RT-PCR. Later we found out that *NDPK* is weakly expressed in three tissues, kidney liver, and muscle and a somewhat decreased amount in pancreas

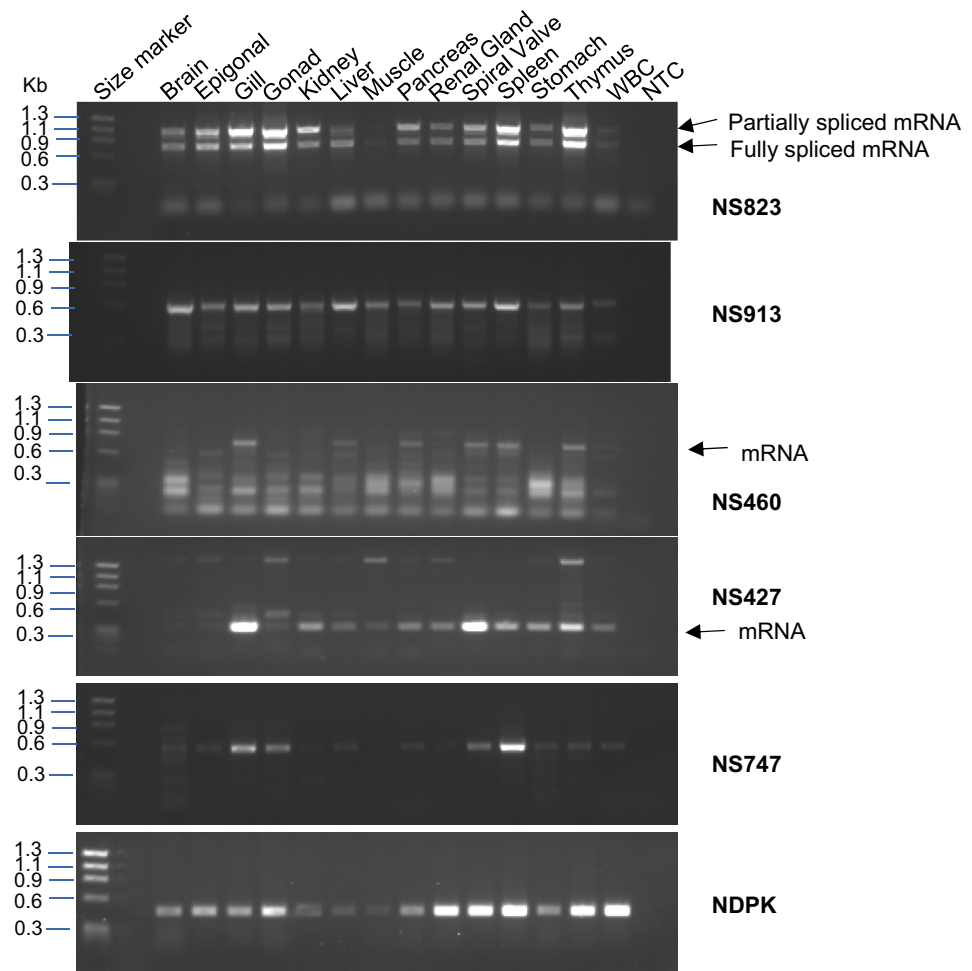


Fig. 3). We also chose to study NS348460; however, there were no obvious signals in ISH.

In nurse shark spleen (Fig. 6a), we detected NS344823 expression as single cells or sporadic cell clusters in the splenic red pulp, clearly outside of splenic white pulp which is exclusively a B cell zone ((Rumfelt et al. 2002) (Castro et al. 2013) (*PAX5*-positive cells within a white pulp “circle,” and marked with B in Fig. 6a). ISH with *TCRβ* and *TCRδ* probes showed, for the first time, small clusters of T cells outside of B cell zone (marked with arrows in Fig. 6a), some of which may be co-expressing NS344823; however, more work is needed to examine co-expression. Since we did not detect NS344823 in clusters, these data suggest that NS344823 is expressed by cells outside of the large B cell follicles and small “T cell zones.” To our surprise, many NS344823-expressing cells were detected in the thymus by in situ hybridization (Fig. 6b), especially in the cortex but not the medulla (Criscitiello et al. 2010), suggesting that *NCR3*-344823 is expressed by immature T cells.

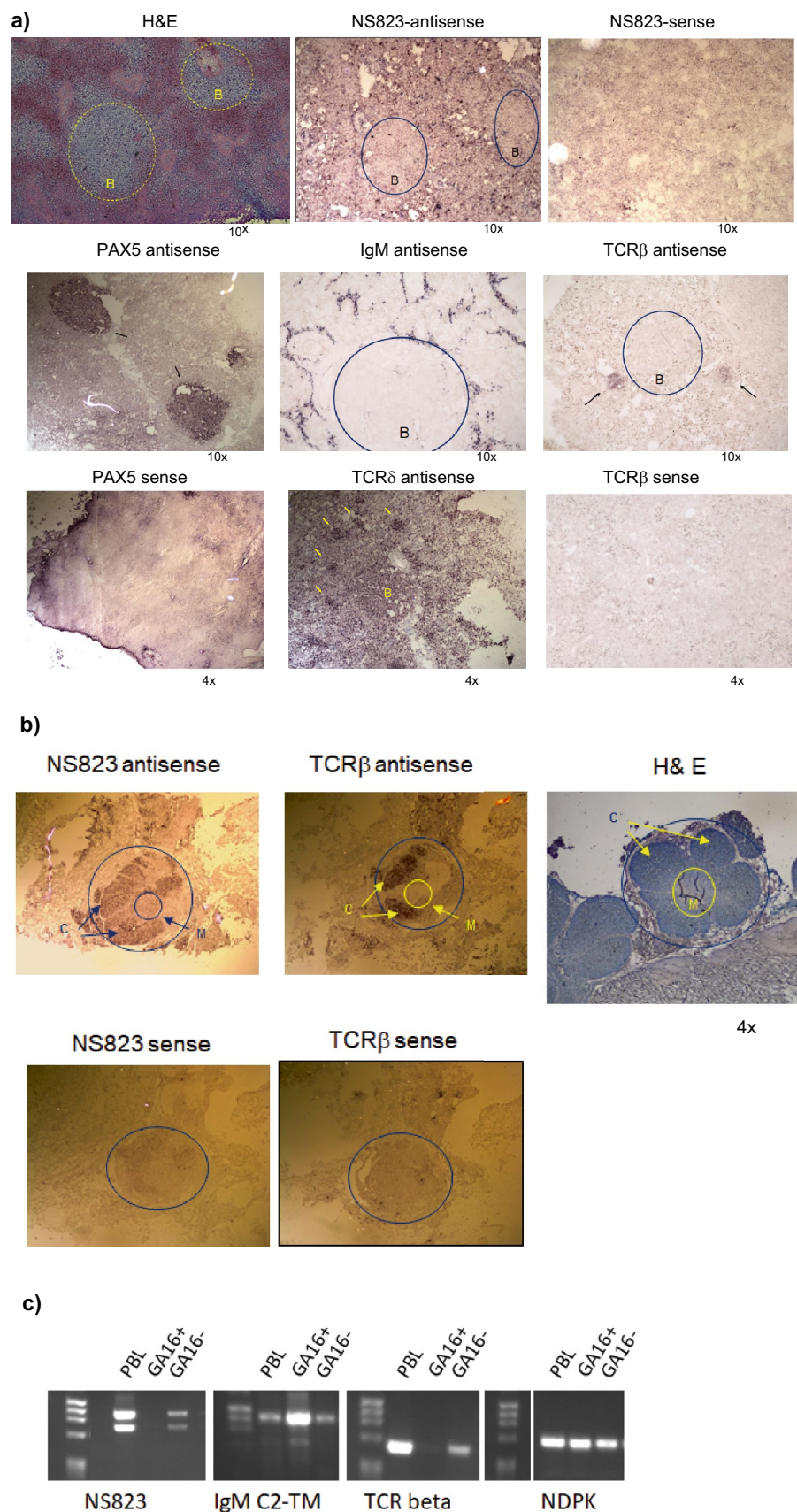
We also MACS sorted nurse shark PBL using GA16, a mAb specific for nurse shark B cells (Rumfelt et al. 2001). Consistent with the data from the spleen ISH, NS344823 was exclusively expressed in the GA16-negative cell population separated from the PBL, as were *TCRβ*-expressing cells (Fig. 6c).

Discussion

NCR3 is a non-rearranging antigen receptor with an ancient association with the MHC

Generally, NKRs evolve rapidly and it is difficult to identify orthologous genes across different vertebrate classes (Yoder and Litman 2011). However, *NCR3* is unique among NKRs since it is evolutionarily conserved (Flajnik et al. 2012) (Ohta and Flajnik 2015), which prompted us to investigate its evolutionary history and expression. Here, we showed

Fig. 6 Shark *NCR3* is expressed in the non-B cell zone and thymic cortex by in situ hybridization. DIG-labeled RNA probes were hybridized to frozen spleen (**a**) and thymus (**b**) sections. Signals were amplified using the TSA system and detected with NBT-BCIP; thus, positive signals are in purple. Sense probes were used as control. IgM-positive cells are plasma cells mostly found outside of the white pulp (B cell zones) that are marked with “B.” The B cell zones containing naïve B cells were identified with the PAX5 anti-sense probe. Image magnification is marked in each figure for spleen, but 4× for thymus. Tissues were sectioned to 7 µm thickness. **b**) Shark thymi are embedded in the connecting tissues. Thymus is roughly outlined with a larger circle. **c**) NS344823 expression by non-B cells isolated from PBL. GA16-positive B cells were MACS-sorted from PBL. RNA was extracted from GA16-positive or negative population as well as pre-sorted PBL. The expression of NS344823 was examined by RT-PCR with the same primers and condition used for Fig. 5. The upper band of NS344823 is a partially spliced mRNA. PBL: peripheral blood lymphocytes (i.e., pre-sorted population). HaeIII digested ϕ x174 DNA was used as the size standard and shown on the left side of the gel. IgM C2-TM primer amplifies the transmembrane-form of IgM and thus selects naïve B cells. Note that the GA16-negative population still contains a small population of naïve B cells



the linkage of *NCR3* to the shark MHC, demonstrating its primordial association with the MHC. Thus far, we have not detected *NCR3* or potential homologues in jawless fish or lower deuterostomes; however, we have detected homologues on MHC paralogous chromosomes (Janes et al., manuscript in preparation), strongly suggesting that the *NCR3* precursor was present before emergence of jawed vertebrates.

NCR3 has a special variable (V)-type IgSF domain, similar to that found in antigen receptors (immunoglobulins and T cell receptors). All *NCR3* have a VJ-IgSF domain with the canonical GXG motif of an antigen receptor J segment in the G-strand of the Ig fold. However, unlike antigen receptors whose VJ-IgSF domains are generated via somatic gene rearrangement, the VJ-IgSF of *NCR3* is a non-rearranging type, where the V and J segments are encoded within a single exon. It has been speculated (Du Pasquier et al. 2004) (Sakano et al. 1979; Thompson 1995) that the rearranging antigen receptors were evolutionarily derived from a non-rearranging VJ-IgSF domain early in vertebrate evolution; thus, extant VJ-IgSF may be related to the precursor of antigen receptors (Chen et al. 2018; Fu et al. 2019; Ohta et al. 2019). Furthermore, such VJ-IgSF domains may have been expressed by ancient “killer” cells and used to detect pathogens or stress-induced self-ligands like B7H6.

Our data suggest that there are two lineages of cartilaginous fish *NCR3* genes, one lineage orthologous to the *NCR3* of other vertebrates and the other specific of cartilaginous fish. Both lineages that are MHC-linked yet are highly divergent. One possibility is that the cartilaginous fish-specific *NCR3* is a sister lineage of *Xenopus XMIV*, which was lost in all other vertebrate groups. Studies of other amphibian and reptile *NCR3* may help in resolving this question.

Shark *NCR3* genes may be the major NKR in cartilaginous fish and other ectotherms

All eight nurse shark *NCR3* subfamily genes encode VJ-IgSF and transmembrane domains with distinctive cytoplasmic C-terminal tails. Among them, we identified different types based on the length and presence or absence of ITIM (Online resource 6). Mammalian *NCR3* is an activating receptor which contains a short cytoplasmic tail and a positively charged amino acid residue (Arginine) in the transmembrane region that interacts with adaptor molecules containing ITAMs (e.g., CD3z and FcRg) (Pessino et al. 1998) (Lanier et al. 1998; Yoder and Litman 2011). There are two nurse shark *NCR3* genes with a positively charged amino acid residue in the transmembrane domain. NS173965 fits into this category but requires future investigation as to whether it will interact with potential shark adaptor molecules. NS326219 also contains a positively charged amino acid; however it also contains long cytoplasmic tail with

ITIM. The potential dual activating and inhibitory function within the same receptor has been reported for an isoform of *NCR2* (Pazina et al. 2017). Whether NS326219 functions in a similar fashion requires further studies. Unfortunately, we were unable to PCR amplify either NS173965 or NS326219 to further investigate their expression. Other nurse shark *NCR3* with long cytoplasmic tails tend to contain ITIMs; thus, they are likely to be inhibitory receptors, but NS334913 does not contain an obvious ITIM. Such variabilities in a particular gene family with activating or inhibitory types (if they function similar to those found in mammals) is a universal mechanism as found in various mammalian immunoreceptors like NKR (e.g., KIR, Ly49) (Yoder and Litman 2011). A similar situation is also found in bony fish-specific receptors such as NITR (Stafford et al. 2006; Yoder et al. 2001) and IpLITR (Stafford et al. 2006), in which expanded receptors contain either activating or inhibitory potentials. Here, we found that *NCR3* is conserved among jawed vertebrates. Therefore, shark *NCR3* genes likely represent the more primordial NKR family universally used in many lower vertebrates. We speculate that all inhibitory types of *NCR3* were lost in mammals and only a single activating *NCR3* has remained. Several shark *NCR3* contain neither positively charged amino acid in the transmembrane nor ITIM in the cytoplasmic domains. These *NCR3* may associate with adaptor molecules differently, use different signaling pathways, or have a totally different function.

NCR3 genes are single-copy in mammals but they are multigenic in non-mammalian species, except for alligators, with varying numbers in each species (Fig. 3). Interestingly, the VJ-IgSF domains between different *NCR3* genes within species are highly diverse, showing only ~ 30% similarity at the amino acid level, suggesting that different ligands could be recognized by each *NCR3*. We also found orthologous *NCR3* genes in different shark species (e.g., NS334823 in three shark species; Fig. 4) are highly conserved, suggesting that they are strongly selected to bind common ligands. In mammals, *NCR3* binds to a self-ligand encoded by a single-copy gene, B7H6 (Brandt et al. 2009). Like other lower vertebrate *NCR3* genes, *B7H6* in lower vertebrates also belongs to a multigene family (Ohta and Flajnik 2015). We identified at least five *B7H6* genes (Online Resource 2) from nurse shark, some having expression patterns limited to certain tissues (Online Resource 5). Thus, it is possible that there is preferential pairing between certain shark B7H6 and *NCR3*. In human, B7H6 is induced on tumor and other cells (Brandt et al. 2009). Thus, B7H6 with limited expression (e.g., NS314508; Online Resource 5) may be a ligand for an activating *NCR3* receptor, while ubiquitously expressed B7H6 may be engaging with inhibitory *NCR3* to block NK activation against (presumably) healthy cells.

In order to find interacting partners of *NCR3* and B7H6, we examined interaction sites based on the crystal structure

of human NCR3 and B7H6 (Joyce et al. 2011; Li et al. 2011) (Online Resource 7). Some interaction sites are well conserved between human and sharks. Particularly, positions G51 and F65 in NCR3 both interact with T127 and P128 in B7H6 (orange and blue circles) are well conserved in many shark genes. As we mentioned earlier, NS331224 did not cluster with other shark B7H6 in the phylogenetic tree analyses (Fig. 4), and NCR3 interaction sites are not conserved (Online Resource 7). Thus, NS331224 may not be B7H6 or divergent B7H6. Among shark NCR3, interaction sites are more conserved in NS341406 and NS326219; however, shark may use different interaction sites and/or conservation of several sites may be sufficient. It must be mentioned that human NCR3 also binds directly to other self-ligands like BAG6 (Binici et al. 2013) as well as foreign ligands such as pp65 (CMV) (Arnon et al. 2005) and HA (poxviruses) (Jarahian et al. 2011), as well as β -1,3-glucans (fungi) (Li et al. 2018). Some of the divergent NCR3 we have detected in sharks may also act as pattern-recognition receptors.

NCR3 is expressed in developing T cells

Mammalian NCR3 are expressed by NK cells, but also by innate lymphoid cells (ILC), $\gamma\delta$ T cells (Hudspeth et al. 2013), and distinct CD8+ T cell (Correia et al. 2018). In our study, we detected NCR3 expression by cortical T cells in shark thymus (Fig. 6b), suggesting that most (or all) developing immature T cells express NCR3. We are currently developing a monoclonal antibody against NS344823 which will allow us to further study the ontogeny of T cells in sharks. Preliminary analyses suggest expression of NCR3 genes and proteins in mature T cell subpopulations (data not shown). T and NK cells share developmental pathways in mammals. The expression of NCR3 in developing shark T cells might reveal commonalities among the NK cells and T cells in this primitive animal model. Future studies will determine whether, as expected, there are separate populations of mature NK cells and T cells and whether NCR3 expression can be detected in any mature T cell subset or after T cell activation.

Origins of defined secondary lymphoid tissue

Lastly, for the first time, we have demonstrated the presence of “T cell zones” in the shark spleen (Fig. 6a), or at least that T cells are outside of the well-developed B cell zones. Since sharks lack lymph nodes and mucosal lymphoid tissues, the spleen is the only secondary lymphoid organ expected to play a major role in lymphocyte activation (Flajnik 2018; Neely and Flajnik 2016). While white pulps with clear margins where naïve B cells reside were previously identified (Castro et al. 2013; Rumfelt et al. 2002), the location of T cell zones was not identified (note that we

misidentified adult shark B cell zones as T cell zones in a previous publication (Rumfelt et al. 2002)). Unlike mammals where splenic B cell zones are adjacent to T cell zones and surrounded by a distinct marginal zone (Mebius and Kraal 2005; Neely and Flajnik 2016), small aggregates of T cells are found within the red pulp mostly independent of the B cell follicles. This pattern is similar to the situation in *Xenopus*, where T cells are scattered within the splenic red pulp in naïve animals and only associate with follicles at high levels in immunized animals (Neely et al. 2018). The data in sharks and frogs demonstrate that the primordial splenic organization has clearly defined B cell zones and “scattered” T cells, consistent with the proposal that “B cells need a house, while T cells are happy in a cave” proposed earlier by Du Pasquier and colleagues (Hofmann et al. 2010). Whether any NK cells are found in these same areas is not known, and will require new reagents that can definitively discriminate between shark NK cells and T cells. Our data suggest, however, that the NS344823-expressing cells are not organized into any discrete structure within the red pulp.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s00251-021-01209-6>.

Author contribution Three authors, Allison Kinlein, Morgan E. James, and Jacob Kincer, share the first authorship. The study conception and design were conducted by Martin Flajnik and Yuko Ohta. Material preparation, data collection, and analysis were performed by all authors. The first draft of the manuscript was written by Morgan Janes and further modified by Yuko Ohta and Martin Flajnik. All authors commented on the manuscript preparation and approved the final edits.

Funding MEJ and JK were supported by the University of Maryland Scholars Program, an initiative of the University of Maryland: MPowering the State. MFF and YO are supported by National Institutes of Health Grants AI140-326-27 and AI02877. MFC was supported by National Science Foundation Grants IOS 1257829 and IOS-1656870.

Data availability All sequences had been deposited to GenBank (accession numbers MT914154-MT914170). Reagents generated from these studies are available to the public upon request.

Declarations

Ethics approval All procedures performed in studies involving animals were approved by Institutional Animal Care and Use Committee (IACUC) at University of Maryland.

Conflict of interest The authors declare that they have no conflict of interest.

References

Allcock RJ, Barrow AD, Forbes S, Beck S, Trowsdale J (2003) The human TREM gene cluster at 6p21.1 encodes both activating and

- inhibitory single IgV domain receptors and includes NKp44. *Eur J Immunol* 33:567–577
- Arnon TI, Achdout H, Levi O, Markel G, Saleh N, Katz G, Gazit R, Gonen-Gross T, Hanna J, Nahari E, Porgador A, Honigman A, Plachter B, Mevorach D, Wolf DG, Mandelboim O (2005) Inhibition of the NKp30 activating receptor by pp65 of human cytomegalovirus. *Nat Immunol* 6:515–523
- Bartl S, Baish MA, Flajnik MF, Ohta Y (1997) Identification of class I genes in cartilaginous fish, the most ancient group of vertebrates displaying an adaptive immune response. *J Immunol* 159:6097–6104
- Binici J, Hartmann J, Herrmann J, Schreiber C, Beyer S, Guler G, Vogel V, Tumulka F, Abele R, Mantele W, Koch J (2013) A soluble fragment of the tumor antigen BCL2-associated athanogene 6 (BAG-6) is essential and sufficient for inhibition of NKp30 receptor-dependent cytotoxicity of natural killer cells. *J Biol Chem* 288:34295–34303
- Brandt CS, Baratin M, Yi EC, Kennedy J, Gao Z, Fox B, Haldeman B, Ostrander CD, Kaifu T, Chabannon C, Moretta A, West R, Xu W, Vivier E, Levin SD (2009) The B7 family member B7–H6 is a tumor cell ligand for the activating natural killer cell receptor NKp30 in humans. *J Exp Med* 206:1495–1503
- Cantoni C, Bottino C, Vitale M, Pessino A, Augugliaro R, Malaspina A, Parolini S, Moretta L, Moretta A, Biassoni R (1999) NKp44, a triggering receptor involved in tumor cell lysis by activated human natural killer cells, is a novel member of the immunoglobulin superfamily. *J Exp Med* 189:787–796
- Castro CD, Ohta Y, Dooley H, Flajnik MF (2013) Non-coordinate expression of J-chain and Blimp-1 define nurse shark plasma cell populations during ontogeny. *Eur J Immunol*
- Chen R, Zhang L, Qi J, Zhang N, Zhang L, Yao S, Wu Y, Jiang B, Wang Z, Yuan H, Zhang Q, Xia C (2018) Discovery and analysis of invertebrate IgVJ-C2 structure from amphioxus provides insight into the evolution of the Ig superfamily. *J Immunol* 200:2869–2881
- Correia MP, Stojanovic A, Bauer K, Juraeva D, Tykocinski LO, Lorenz HM, Brors B, Cerwenka A (2018) Distinct human circulating NKp30(+)FcepsilonRIgamma(+)CD8(+) T cell population exhibiting high natural killer-like antitumor potential. *Proc Natl Acad Sci U S A* 115:E5980–E5989
- Criscitiello MF, Ohta Y, Saltis M, McKinney EC, Flajnik MF (2010) Evolutionarily conserved TCR binding sites, identification of T cells in primary lymphoid tissues, and surprising trans-rearrangements in nurse shark. *J Immunol* 184:6950–6960
- Du Pasquier L, Zucchetti I, De Santis R (2004) Immunoglobulin superfamily receptors in protochordates: before RAG time. *Immunol Rev* 198:233–248
- Flajnik MF (2016) Evidence of G.O.D.’s miracle: unearthing a RAG transposon. *Cell* 166:11–12
- Flajnik MF (2018) A cold-blooded view of adaptive immunity. *Nat Rev Immunol* 18:438–453
- Flajnik MF, Miller K, Du Pasquier L (2003) Evolution of the immune system. In Paul WE (ed.) *Fundamental immunology*, Fifth edn. Lippencott Williams & Wilkins, Philadelphia
- Flajnik MF, Tlapakova T, Criscitiello MF, Krylov V, Ohta Y (2012) Evolution of the B7 family: co-evolution of B7H6 and NKp30, identification of a new B7 family member, B7H7, and of B7’s historical relationship with the MHC. *Immunogenetics* 64:571–590
- Fu Y, Yang Z, Huang J, Cheng X, Wang X, Yang S, Ren L, Lian Z, Han H, Zhao Y (2019) Identification of two nonrearranging IgSF genes in chicken reveals a novel family of putative remnants of an antigen receptor precursor. *J Immunol* 202:1992–2004
- Greenwald RJ, Freeman GJ, Sharpe AH (2005) The B7 family revisited. *Annu Rev Immunol* 23:515–548
- Hara Y, Yamaguchi K, Onimaru K, Kadota M, Koyanagi M, Keeley SD, Tatsumi K, Tanaka K, Motone F, Kageyama Y, Nozu R, Adachi N, Nishimura O, Nakagawa R, Tanegashima C, Kiyatake I, Matsumoto R, Murakumo K, Nishida K, Terakita A, Kuratani S, Sato K, Hyodo S, Kuraku S (2018) Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nat Ecol Evol* 2:1761–1771
- Hofmann J, Greter M, Du Pasquier L, Becher B (2010) B-cells need a proper house, whereas T-cells are happy in a cave: the dependence of lymphocytes on secondary lymphoid tissues during evolution. *Trends Immunol* 31:144–153
- Hudspeth K, Silva-Santos B, Mavilio D (2013) Natural cytotoxicity receptors: broader expression patterns and functions in innate and adaptive immune cells. *Front Immunol* 4:69
- Humphrey MB, Lanier LL, Nakamura MC (2005) Role of ITAM-containing adapter proteins and their receptors in the immune system and bone. *Immunol Rev* 208:50–65
- Jarahian H, Fiedler M, Cohnen A, Djandji D, Hammerling GJ, Gati C, Cerwenka A, Turner PC, Moyer RW, Watzl C, Hengel H, Momburg F (2011) Modulation of NKp30- and NKp46-mediated natural killer cell responses by poxviral hemagglutinin. *PLoS Pathog* 7:e1002195
- Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275–282
- Joyce MG, Tran P, Zhuravleva MA, Jaw J, Colonna M, Sun PD (2011) Crystal structure of human natural cytotoxicity receptor NKp30 and identification of its ligand binding site. *Proc Natl Acad Sci U S A* 108:6223–6228
- Kasahara M, Canel C, McKinney EC, Flajnik MF (1991) Molecular cloning of nurse shark cDNAs with high sequence similarity to nucleoside diphosphate kinase genes. In J K (ed.) *Evolution of the major histocompatibility complex*. Springer-Verlag, New York, NY
- Kruse PH, Matta J, Ugolini S, Vivier E (2014) Natural cytotoxicity receptors and their ligands. *Immunol Cell Biol* 92:221–229
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549
- Lanier LL, Corliss BC, Wu J, Leong C, Phillips JH (1998) Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. *Nature* 391:703–707
- Li SS, Ogbomo H, Mansour MK, Xiang RF, Szabo L, Munro F, Mukherjee P, Mariuzza RA, Amrein M, Vyas JM, Robbins SM, Mody CH (2018) Identification of the fungal ligand triggering cytotoxic PRR-mediated NK cell killing of *Cryptococcus* and *Candida*. *Nat Commun* 9:751
- Li Y, Wang Q, Mariuzza RA (2011) Structure of the human activating natural cytotoxicity receptor NKp30 bound to its tumor cell ligand B7–H6. *J Exp Med* 208:703–714
- Long EO (2008) Negative signaling by inhibitory receptors: the NK cell paradigm. *Immunol Rev* 224:70–84
- Luo M, Kim H, Kudrna D, Sisneros NB, Lee SJ, Mueller C, Collura K, Zuccolo A, Buckingham EB, Grim SM, Yanagiya K, Inoko H, Shiina T, Flajnik MF, Wing RA, Ohta Y (2006) Construction of a nurse shark (*Ginglymostoma cirratum*) bacterial artificial chromosome (BAC) library and a preliminary genome survey. *BMC Genomics* 7:106
- Mebius RE, Kraal G (2005) Structure and function of the spleen. *Nat Rev Immunol* 5:606–616
- Neely HR, Flajnik MF (2016) Emergence and evolution of secondary lymphoid organs. *Annu Rev Cell Dev Biol* 32:693–711
- Neely HR, Guo J, Flowers EM, Criscitiello MF, Flajnik MF (2018) “Double-duty” conventional dendritic cells in the amphibian *Xenopus* as the prototype for antigen presentation to B cells. *Eur J Immunol* 48:430–440
- Ohta Y, Flajnik MF (2015) Coevolution of MHC genes (LMP/TAP/class Ia, NKT-class Ib, NKp30-B7H6): lessons from cold-blooded vertebrates. *Immunol Rev* 267:6–15
- Ohta Y, Goetz W, Hossain MZ, Nonaka M, Flajnik MF (2006) Ancestral organization of the MHC revealed in the amphibian *Xenopus*. *J Immunol* 176:3674–3685

- Ohta Y, Kasahara M, O'Connor TD, Flajnik MF (2019) Inferring the “primordial immune complex”: origins of MHC Class I and antigen receptors revealed by comparative genomics. *J Immunol* 203:1882–1896
- Ohta Y, McKinney EC, Criscitiello MF, Flajnik MF (2002) Proteasome, transporter associated with antigen processing, and class I genes in the nurse shark *Ginglymostoma cirratum*: evidence for a stable class I region and MHC haplotype lineages. *J Immunol* 168:771–781
- Ohta Y, Shiina T, Lohr RL, Hosomichi K, Pollin TI, Heist EJ, Suzuki S, Inoko H, Flajnik MF (2011) Primordial linkage of $\beta 2$ -microglobulin to the MHC. *J Immunol* 186:3563–3571
- Pazina T, Shemesh A, Brusilovsky M, Porgador A, Campbell KS (2017) Regulation of the functions of natural cytotoxicity receptors by interactions with diverse ligands and alterations in splice variant expression. *Front Immunol* 8:369
- Pende D, Parolini S, Pessino A, Sivori S, Augugliaro R, Morelli L, Marcenaro E, Accame L, Malaspina A, Biassoni R, Bottino C, Moretta L, Moretta A (1999) Identification and molecular characterization of NKp30, a novel triggering receptor involved in natural cytotoxicity mediated by human natural killer cells. *J Exp Med* 190:1505–1516
- Pessino A, Sivori S, Bottino C, Malaspina A, Morelli L, Moretta L, Biassoni R, Moretta A (1998) Molecular cloning of NKp46: a novel member of the immunoglobulin superfamily involved in triggering of natural cytotoxicity. *J Exp Med* 188:953–960
- Petty CL, McKinney EC (1983) Temperature and cellular regulation of spontaneous cytotoxicity in the shark. *Eur J Immunol* 13:133–138
- Pettinello R, Redmond AK, Secombes CJ, Macqueen DJ, Dooley H (2017) Evolutionary history of the T cell receptor complex as revealed by small-spotted catshark (*Scyliorhinus canicula*). *Dev Comp Immunol* 74:125–135
- Rumfelt LL, Avila D, Diaz M, Bartl S, McKinney EC, Flajnik MF (2001) A shark antibody heavy chain encoded by a nonsomatically rearranged VDJ is preferentially expressed in early development and is convergent with mammalian IgG. *Proc Natl Acad Sci U S A* 98:1775–1780
- Rumfelt LL, McKinney EC, Taylor E, Flajnik MF (2002) The development of primary and secondary lymphoid tissues in the nurse shark *Ginglymostoma cirratum*: B-cell zones precede dendritic cell immigration and T-cell zone formation during ontogeny of the spleen. *Scand J Immunol* 56:130–148
- Sakano H, Huppi K, Heinrich G, Tonegawa S (1979) Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* 280:288–294
- Session AM, Uno Y, Kwon T, Chapman JA, Toyoda A, Takahashi S, Fukui A, Hikosaka A, Suzuki A, Kondo M, van Heeringen SJ, Quigley I, Heinz S, Ogino H, Ochi H, Hellsten U, Lyons JB, Simakov O, Putnam N, Stites J, Kuroki Y, Tanaka T, Michiue T, Watanabe M, Bogdanovic O, Lister R, Georgiou G, Paranjpe SS, van K, I, Shu S, Carlson J, Kinoshita T, Ohta Y, Mawaribuchi S, Jenkins J, Grimwood J, Schmutz J, Mitros T, Mozaffari SV, Suzuki Y, Haramoto Y, Yamamoto TS, Takagi C, Heald R, Miller K, Haudenschild C, Kitzman J, Nakayama T, Izutsu Y, Robert J, Fortriede J, Burns K, Lotay V, Karimi K, Yasuoka Y, Dichmann DS, Flajnik MF, Houston DW, Shendure J, DuPasquier L, Vize PD, Zorn AM, Ito M, Marcotte EM, Wallingford JB, Ito Y, Asashima M, Ueno N, Matsuda Y, Veenstra GJ, Fujiyama A, Harland RM, Taira M, Rokhsar DS (2016) Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* 538:336–343
- Stafford JL, Bengten E, Du PL, McIntosh RD, Quiniou SM, Clem LW, Miller NW, Wilson M (2006) A novel family of diversified immunoregulatory receptors in teleosts is homologous to both mammalian Fc receptors and molecules encoded within the leukocyte receptor complex. *Immunogenetics* 58:758–773
- Stet RJ, Hermesen T, Westphal AH, Jukes J, Engelsma M, BM LV-vK, Dortmans J, Aveiro J, Savelkoul HF (2005) Novel immunoglobulin-like transcripts in teleost fish encode polymorphic receptors with cytoplasmic ITAM or ITIM and a new structural Ig domain similar to the natural cytotoxicity receptor NKp44. *Immunogenetics* 57:77–89
- Terado T, Okamura K, Ohta Y, Shin DH, Smith SL, Hashimoto K, Takemoto T, Nonaka MI, Kimura H, Flajnik MF, Nonaka M (2003) Molecular cloning of C4 gene and identification of the class III complement region in the shark MHC. *J Immunol* 171:2461–2466
- Thompson CB (1995) New insights into V(D)J recombination and its role in the evolution of the immune system. *Immunity* 3:531–539
- Tonegawa S (1983) Somatic generation of antibody diversity. *Nature* 302:575–581
- Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, Johnson J, Dandona N, Viswanathan LD, Tay A, Venter JC, Strausberg RL, Brenner S (2007) Survey sequencing and comparative analysis of the elephant shark (*Callorhynchus milii*) genome. *PLoS Biol* 5:e101
- Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, Hoon S, Gangu V, Roy SW, Irimia M, Korzh V, Kondrychyn I, Lim ZW, Tay BH, Tohari S, Kong KW, Ho S, Lorente-Galdos B, Quilez J, Marques-Bonet T, Raney BJ, Ingham PW, Tay A, Hillier LW, Minx P, Boehm T, Wilson RK, Brenner S, Warren WC (2014) Elephant shark genome provides unique insights into gnathostome evolution. *Nature* 505:174–179
- Vitale M, Cantoni C, Della Chiesa M, Ferlazzo G, Carlomagno S, Pende D, Falco M, Pessino A, Muccio L, De Maria A, Marcenaro E, Moretta L, Sivori S (2019) An historical overview: the discovery of how NK cells can kill enemies, recruit defense troops, and more. *Front Immunol* 10:1415
- Williams AF, Barclay AN (1988) The immunoglobulin superfamily—domains for cell surface recognition. *Annu Rev Immunol* 6:381–405
- Yoder JA, Litman GW (2011) The phylogenetic origins of natural killer receptors and recognition: relationships, possibilities, and realities. *Immunogenetics* 63:123–141
- Yoder JA, Mueller MG, Wei S, Corliss BC, Prather DM, Willis T, Litman RT, Djeu JY, Litman GW (2001) Immune-type receptor genes in zebrafish share genetic and functional properties with genes encoded by the mammalian leukocyte receptor cluster. *Proc Natl Acad Sci U S A* 98:6771–6776

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.