



Published in final edited form as:

J Immunol. 2020 February 15; 204(4): 892–902. doi:10.4049/jimmunol.1901025.

An ancient, MHC-linked, nonclassical class I lineage in cartilaginous fish

Tereza Almeida^{*†‡}, Pedro J. Esteves^{*†}, Martin F. Flajnik^{‡,1}, Yuko Ohta[‡], Ana Veríssimo^{*§}

^{*}CIBIO-InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Universidade do Porto, Vairão, Porto, Portugal

[†]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal

[‡]Department of Microbiology and Immunology, University of Maryland Baltimore, Baltimore, Maryland, 21201 USA

[§]Virginia Institute of Marine Science, William & Mary, Gloucester Point, Virginia, 23062 USA

Abstract

Cartilaginous fishes, or Chondrichthyans, are the oldest jawed vertebrates having an adaptive immune system based on the major histocompatibility complex (MHC) and immunoglobulin superfamily-based (IgSF) antigen receptors. In this basal group of jawed vertebrates we identified a third nonclassical MHC class I lineage (*UDA*), which is present in all species analyzed within the two major cartilaginous subclasses, Holocephali (chimaeras) and Elasmobranchii (sharks, skates, and rays). The deduced amino acid sequences of *UDA* have eight out of nine typically invariant residues that bind to the N- and C- termini of bound peptide found in most vertebrate classical class I (*UAA*); additionally, the other predicted 28 peptide binding residues are perfectly conserved in all elasmobranch *UDA* sequences. *UDA* is distinct from *UAA* in its differential tissue distribution and its lower expression levels, and is mono- or oligomorphic unlike the highly polymorphic *UAA*. *UDA* is in low copy number in elasmobranchs but is multicopy in the holocephalan spotted ratfish (*Hydrolagus colliei*). Using a nurse shark (*Ginglymostoma cirratum*) family, we found that *UDA* is MHC-linked but separable by recombination from the tightly linked cluster of *UAA*, *TAP* and *LMP* genes, the so-called class I region found in most nonmammalian vertebrates. *UDA* has predicted structural features that are similar to certain nonclassical class I genes in other vertebrates, and, unlike polymorphic classical class I, we anticipate that it may bind to a conserved set of specialized peptides.

Keywords

nonclassical class I; adaptive immunity; jawed vertebrates; Elasmobranchs; Holocephalans; ancient lineages; evolution

¹Corresponding author; MFlajnik@som.umaryland.edu ; Phone +1(410)706-5161; fax +1(410)706-6970.

Introduction

The Chondrichthyes or cartilaginous fish, including all its modern representatives, i.e. elasmobranchs (sharks, skates, and rays) and holocephalans (chimaeras), comprise the oldest class of extant jawed vertebrates possessing an adaptive immune system grounded on immunoglobulins (Ig), T cell receptors (TCR), and the major histocompatibility complex (MHC) (reviewed in (1-3)). Thus, cartilaginous fish are essential for studying the evolution of the adaptive immune system, as well as for identifying ancestral and derived features of adaptive immunity. Previously it was thought that chondrichthyans had a very simple immune system; however, many studies of shark immunology conducted in the last three decades have shown their immune system to be quite complex ((4-7); reviewed in (2, 8, 9)), but still preserving several ancestral features, such as the cluster-type organization of Ig genes (10), close genetic linkage of MHC class I processing and presenting genes (11), and the MHC linkage of beta-2 microglobulin (12).

MHC class I molecules play the central role in antigen presentation to CD8-positive T cells. Classical class I (class Ia) has been best studied for its function of presenting peptide antigens to trigger activation of cytotoxic T cells, whereas nonclassical class I (class Ib) molecules have a myriad of functions (review in (13)), mostly examined in mammals. The classical class I protein from most vertebrates generally shows conservation of nine amino acid residues that bind to the N- and C-termini of bound peptide in the peptide-binding region (PBR). Class Ia genes show high levels of polymorphism, linkage to the MHC, and ubiquitous tissue expression. In contrast, nonclassical class I genes are generally monomorphic or minimally polymorphic, have a limited tissue distribution, may or may not be linked to the MHC, and have been recruited for multiple functions, some not even involved in immunity (3, 14-17)

Classical MHC class I genes (*UAA*) have been identified in several cartilaginous fishes, with similar structural and genetic features as are found in other vertebrates (5, 18). Several nonclassical class I genes and lineages also have been described in cartilaginous fish: *UBA* in nurse shark (*Ginglymostoma cirratum*) and horn shark (*Heterodontus francisci*) (5, 19), and a highly divergent gene, *UCA*, so far found only in spiny dogfish (*Squalus acanthias*) (20). To date, no functional study has been performed on any class I molecules in cartilaginous fish. Note, that we designate these previously described nonclassical class I genes in cartilaginous fish for the first time in this report, following the nomenclature criteria for MHC genes in nonmammalian species (21).

In this study, using predominantly the chondrichthyan genome and transcriptome databases, we have identified a new nonclassical class I lineage in all cartilaginous fish examined, and assigned it as *UDA*. We characterized this new lineage focusing on its expression, level of polymorphism, deduced structural features, presence or absence in cartilaginous fish species, and linkage to the MHC.

Material and Methods

Database searches

Several representative MHC class I sequences from GenBank (<http://www.ncbi.nlm.nih.gov>), namely nurse shark (*G. cirratum*) and horn shark (*H. francisci*) class Ia (Gici *UAA*: AAF66110 and Hefr *UAA*: AAC60349) and class Ib (Gici *UBA*: AAC60347 and Hefr *UBA*: AAC60348), spiny dogfish (*S. acanthias*) class Ib (Sqac *UCA*: AAN78091) and human CD1 (AAX49405), were used as templates to search for MHC class I sequences using a blastp in non-redundant protein databases and tblastn in the transcriptome shotgun assembly (TSA) and short read archive (SRA) databases using default parameters in public databases (<http://www.ncbi.nlm.nih.gov>; <http://skatebase.org>) (Supplemental Table I). After searching the nurse shark transcriptome (SRX219865; SRX219866) we assembled the *UDA* gene using the Geneious software 6.0 (22). A 42bp gap (positions 612 to 653 bp) in this assembly was filled with polymerase chain reaction (PCR) in the same nurse shark individual used for the transcriptome (23). The primers $\alpha 2$ Fw 5'-GGTGCTGCAGTACTGAATCG-3' and $\alpha 3$ Rv 5'-GTATCTCCTTCGGTGCAGG-3' PCR was performed at 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec, with a final extension of 72°C for 10 min using GoTaq Master Mix (Promega). The PCR products were cloned into pGEM T-easy vector and sequenced.

Sequence alignments and phylogenetic tree analyses

Deduced amino acid sequences were aligned using the ClustalX program in the Geneious software 6.0 (22) with manual adjustments. The neighbor-joining (NJ) phylogenetic tree of the class I peptide-binding domains ($\alpha 1$ and $\alpha 2$) was constructed in MEGA 6.06 (24) using p-distances, uniform rates among sites, and pairwise deletions, and 10,000 bootstrap replicates. The U and Z lineages of MHC class I from bony fish were also included, and we chose the human and chicken CD1 molecules as outgroups since they are the most divergent class I lineage in vertebrates (16).

Northern blotting

Lineage-specific tissue expression was assessed and compared between *UDA* and *UAA* using northern blotting. Ten micrograms of total RNA from various nurse shark tissues (brain, epigonal, gill, gonad, liver, muscle, pancreas, spiral valve, spleen, stomach, thymus and white blood cells (WBC) were electrophoresed on denaturing 1% agarose gel electrophoresis, and subsequently transferred onto nitrocellulose membranes as previously described (5). Hybridization was done using ³²P-labeled regions encoding the $\alpha 3$ domains of *UAA* and *UDA*, as well as a loading control using the nucleoside diphosphate kinase probe (NDPK) (4) under high stringency conditions (5).

In situ Hybridization

To detect cell types expressing *UDA* within organs, we performed *in situ* hybridization (ISH) on those tissues with highest expression on the northern blotting (epigonal, spiral valve, gill and spleen). Nurse shark tissues were collected and fixed in 4% paraformaldehyde

in 1x SPB solution (0.06M Phosphate Buffer (Na₂HPO₄/NaH₂PO₄)/3% Sucrose/0.15mM CaCl₂ pH 7.4) for 6 days at 4°C overnight. Tissues were rinsed gradually in SPB containing an increasing amount of sucrose from 10% to 30% and infiltrated overnight at 4°C. The fixed tissues were then embedded in O.C.T. medium (Sakura) and frozen in a liquid nitrogen/2-methylbutane bath. Frozen tissues were sectioned (8 µm in thickness) and mounted onto glass slides. Nurse shark *UAA* and *UDA* riboprobes were generated from linearized plasmid DNA using RNA polymerase (Promega) and DIG RNA Labeling Mix (Roche). Tissue slides were prefixed in 4% paraformaldehyde in shark PBS, quenched endogenous peroxidase activities using 0.3% hydrogen peroxide, treated with proteinase K (20µg/ml; Sigma-Aldrich), and acetylated in 0.25% acetic anhydride. The slides were hybridized with riboprobes (6.5ng per slide) in 1x Hybridization Solution (Sigma-Aldrich) containing 50% formamide and baker's yeast tRNA (Sigma-Aldrich) overnight at 67°C. After hybridization, tissue slides were washed twice in 0.2x SSC (0.003M Sodium citrate/ 0.03M Sodium chloride) at 72°C for 30 min. Signals were amplified using the TSA plus Biotin System (PerkinElmer) following the manufacturer's protocol. For colorimetric signal visualization, slides were incubated with Streptavidin-alkaline phosphatase (SA-AP: PerkinElmer) followed by substrate BCIP/NBT (Roche). For fluorescent signal visualization, slides were incubated with SA-Alexa Fluor647 (Thermo Fisher Scientific) and mounted with ProLong Gold plus DAPI (Invitrogen).

Southern blotting

To estimate the presence/absence and number of *UDA* genes in various chondrichthyan species (including those with no transcriptome or genome sequences), we performed Southern blotting. For the cartilaginous fish blot ("Chondroblot"), we digested 10 µg of genomic DNA (gDNA) extracted from erythrocytes, with BamHI for 48 hours and electrophoresed in a 0.8% agarose gel. The digested gDNA was transferred onto a nitrocellulose membrane via capillary transfer, and a ³²P-labeled α3 domain probe of nurse shark *UDA* was hybridized to the membranes under low-stringency conditions (5). The membrane was exposed to X-ray film for different periods to obtain the optimal signal strength. For the nurse shark family blot, the same α3 probe was used for hybridization, but under high-stringency conditions (5).

Statistical analysis of linkage

We validated the linkage status of *UDA* to the MHC using Parametric linkage analysis. We calculated the odds of the likelihood of whether two loci are linked vs. non-linked using a MHC-typed family of 39 siblings (19, 25). Family-based linkage analysis makes use of information of at least one of the parents (e.g. in this case, the mother) and a large number of descendants to detect co-segregation of markers. We compared the restriction fragment length polymorphism (RFLP) banding pattern of *UDA* to the MHC haplotypes and determined concordance or non-concordance patterns between the *UDA* and MHC haplotypes. The log of the odds (LOD) score was calculated as previously described (12):

$$\text{LOD} = \log_{10} \left\{ \frac{[(\theta)^R * (1-\theta)^{NR}] + [(\theta)^{NR} * (1-\theta)^R]}{2} / (0.5)^{R + NR} \right\}$$

Phase 1 Phase 2

where θ is the recombination frequency, R is the number of recombinant offspring and NR is the number of non-recombinant offspring. Since we do not know whether the parental phase is linked or non-linked, we calculated both phases (i.e. phase 1 and phase 2: switching R and NR) and averaged. Since the recombination frequency is not known, we calculated LOD scores with recombination frequency (θ) ranging between 0-0.5 to obtain recombination frequency at maximum LOD score (26). The p value was further calculated using a one-sided χ^2 test at the maximum LOD score (27)

Determination of degree of polymorphism

To amplify the *UAA* and *UDA* genes from nurse shark genomic DNA we performed a PCR using the GoTaq Master Mix (Promega) and the following primers: *UAA* primers $\alpha 1$ Fw 5'-GGTCTCACAGTCTCCGGT - 3' and $\alpha 2$ Rv 5'-GGTCTCAGTTCCACATTTCC - 3'. PCR for *UAA* was performed at 95°C for 2 min, followed by 35 cycles of 95°C for 45 sec, 50°C for 30 sec, and 72°C for 45 sec, with a final extension of 72°C for 10 min. *UDA* primers were based on the new sequence obtained from the nurse shark (GenBank Accession no. [MN339476](#)), and were as follows: $\alpha 1$ Fw: 5'-CTGAGGTATTACTACACCTC - 3', $\alpha 1$ Rv: 5'-GTTCTCCTTGTTGCTATCTG - 3'; $\alpha 2$ Fw: 5'-CAGGTTTGAACCTACCTGCA - 3', $\alpha 2$ Rv: 5'-CGATTCAGTACTGCAGCACC - 3'. PCR was performed using GoTaq Master Mix (Promega) at 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 58°C for 30 sec, and 72°C for 30 sec, with a final extension of 72°C for 10 min. All PCR products were cloned into pGEMT-easy vector (Takara), and at least three clones per individual (in a total of 4 individuals) were sequenced to eliminate PCR errors. Note that we used $\alpha 1$ -5' and $\alpha 2$ -3' primers for amplification of genomic DNA for *UAA* because of the short intron size between the $\alpha 1$ and $\alpha 2$ -encoding exons, but the same strategy did not work for *UDA* and we had to use intra-exonic primers and amplified the $\alpha 1$ and $\alpha 2$ regions separately. To further confirm the monomorphic feature and eliminate PCR errors of *UDA* alleles, we set up separate PCR reactions, extracted the corresponding bands from the gel, cleaned and sequenced them directly.

Results

Identification of a new cartilaginous fish class I lineage

Our initial approach was to identify all annotated MHC class I genes in the available cartilaginous fish genomes and transcriptomes, using published MHC class Ia and class Ib sequences from different shark species, as well as human CD1 (the most divergent vertebrate class I gene (16)) (Supplemental Table I) as templates for blast searches. One group of sequences (representative sequences in Figure 1, entire list of sequences in Supplemental Table I, Supplemental Figures 1 and 2) was found to be quite different from the three previously published class I lineages, *UAA*, *UBA*, and *UCA* at the nucleotide (51%, 50% and 41%, respectively) (data not shown) and protein level (37%, 36%, and 23%, respectively) (Supplemental table II), and thus we named it *UDA* following the MHC nomenclature for nonmammalian vertebrates (21). The deduced amino acid sequence of *UDA* has the basic features of all class I molecules with the typical peptide-binding domains ($\alpha 1$ and $\alpha 2$) and the IgSF $\alpha 3$ domain. Residues in dark shade in Figure 1 are highly conserved in all class I proteins, primarily for maintaining structural integrity (16, 28). The

$\alpha 2$ domain has the canonical disulfide bridge found in most class I molecules, and there is a typical Asp/Glu-containing connecting piece and a hydrophobic transmembrane (TM) region (Figure 1; Supplemental Figure 1 and 2). The Asn-linked glycosylation site at the C-terminus of the $\alpha 1$ domain, present in all classical and most nonclassical class I molecules, is not found in all UDA sequences (Figure 1, Supplemental Figure 1). Most conspicuously, UDA was found to have much longer cytoplasmic (cyt) tail (Figure 1, Supplemental Figure 2) than most class I lineages; this characteristic is similar to the bony fish “typical” Z class I lineage (29) but unlike the highly conserved length of classical class I in most vertebrates (30, 31). The UDA lineage does not have conserved Tyr (Y320, position based in HLA-A2 sequence) in the cytoplasmic tail found in the classical class I (Figure 1; Supplemental Figure 2). The conserved Ser (S335, position based on the HLA-A2 sequence) in the classical class I cytoplasmic tail may be present in UDA, although this is highly speculative due to the high Ser content in all UDA cytoplasmic sequences and the high divergence between the UDA sequences across cartilaginous fish making the alignment of the cytoplasmic region and ascertainment of the conserved Ser position ambiguous (Figure 1; Supplemental Figure 2). No distinct signaling motifs were identified in the UDA cytoplasmic tail using ScanProsite (<http://prosite.expasy.org/scanprosite>), nor were there long stretches of conservation among UDA in the different cartilaginous fish species.

In contrast to the poor conservation of the cytoplasmic tail between UAA and UDA, UDA shared 8 of the 9 invariant residues that bind to the N- and C-termini of bound peptides in almost all classical class I PBR, which lock peptide into the class I groove (28) (Table I). This feature is also found in the bony fish “typical” nonclassical Z lineage (32). Furthermore, interestingly, the other 28 predicted peptide binding residues (16) of UDA are perfectly conserved in elasmobranch sequences, and among them 18/28 residues are hydrophilic (polar or charged), a feature shared with classical class I molecules but not the lipid-binding CD1 protein (only 3/28 hydrophilic residues). Note that the peptide-binding residues of the “typical” Z lineage in all bony fish sequences are also nearly perfectly conserved, but these Z residues are different from those in UDA. Based on the conservation of these peptide-binding residues (but not the rest of the PBR between UDA in all of the cartilaginous fish species; percent identities between the different domains shown in Supplemental table II), we predict that UDA (and Z) binds to a set of peptides with the same or similar anchor residues, unlike UAA. In contrast, preliminary results suggest that UCA, like CD1, has a strongly hydrophobic binding groove, suggesting that it also may bind to glycolipids.

UDA is an ancient class I lineage present in all cartilaginous fish

To understand the evolutionary relationship between all chondrichthyan class I sequences, we constructed class I neighbor-joining phylogenetic trees using the PBR ($\alpha 1$ and $\alpha 2$ domains) and rooted with the nonclassical CD1 protein (Figure 2; alignment in Supplemental Figure 1). As was suggested by Wang et al. (2003) (20), the shark UCA lineage is highly divergent, clustering outside of all vertebrate class I sequences besides CD1 (Figure 2). UDA class I sequences were found in all holocephalans and elasmobranchs tested, and they formed a single clade supported with high bootstrap values (99%). The UDA lineage is more closely related to the groups of UAA, UBA, and U sequences and to

the bony fish “typical” Z lineage, when compared to UCA lineage. We also found that the cartilaginous fish *UBA* sequences (5) were only found in elasmobranch taxa, showing on average ~50% similarity to UAA at the amino acid level between all Cartilaginous fish sequences ((Supplemental table II), and most likely arose via duplication of the *UAA* lineage based on the tree topology (Figure 2).

Tissue-specific expression of UDA

We examined *UDA* expression in different nurse shark tissues and compared it to that of *UAA* by northern blotting (Figure 3). *UDA* was most highly expressed in the epigonal organ and spiral valve, moderately expressed in gill, spleen, stomach, and white blood cells (WBC), and poorly expressed in thymus, gonad, and brain. *UAA* exhibited the typical ubiquitous classical class I expression in many tissues, with highest expression in immune organs such as spleen, thymus, and mucosae. To detect cell types expressing *UDA* within organs, we performed *in situ* hybridization on various tissues with highest expression by northern blotting (Figure 3). In the gill, we detected expression of *UDA* within the filaments (especially in the so-called pillar cells) and weak on the gill epithelium, whereas *UAA* was highly expressed only by the epithelium (Figure 4). We did not detect *UDA* expression in the other tissues with high expression by northern blot (epigonal, spiral valve and spleen; data not shown), suggesting that *UDA* is expressed broadly but at low levels by many different types of cells or regions in these tissues. We should add that we have only studied the baseline expression of *UDA* and many nonclassical class I molecules are induced under a variety of stimulatory conditions (33).

UDA is single or low copy in elasmobranchs

In agreement with our bioinformatic searches, *UDA* genes were present in all chondrichthyan species tested by Southern blotting (“Chondroblot” in Figure 5). Similar to *UAA* (5, 19), a single or low copy number of *UDA* genes was detected in all elasmobranchs. In contrast, several *UDA* genes were found in the Southern blot for the holocephalan spotted ratfish (*Hydrolagus collieti*). However, bioinformatic searches of the genome of another holocephalan, the elephant shark (*Callorhynchus milii*), revealed only one *UDA* gene. The *C. milii* genome assembly is incomplete (and unfortunately, no genomic DNA was available for Southern blotting), so our gene estimate *C. milii* for this species might be an underestimate.

UDA seems to be monomorphic in nurse sharks

Classical class I genes are highly polymorphic in almost all vertebrates, while nonclassical class I are generally mono- or oligomorphic. To test the degree of *UDA* polymorphism, we selected four MHC-disparate nurse shark individuals collected from the wild and sequenced the regions encoding the $\alpha 1$ and $\alpha 2$ domains of *UAA* and *UDA* genes. The *UDA* sequences for all four individuals were identical (Figure 6), further confirming *UDA*'s nonclassical status.

UDA is linked to the MHC

Nonclassical class I genes in all vertebrates are found in various genetic regions, some encoded outside of the MHC. Thus, we examined linkage status of *UDA* using an MHC-

typed nurse shark family of 39 siblings (19, 25). Three segregating *UDA* RFLP bands were obtained with the nurse shark $\alpha 3$ probe by Southern blotting (marked U, M, L in Figure 7). Previous studies showed that this family comprises 13 MHC groups (indicated groups a-m in Figure 7, Table II) due to the multiple paternity of at least seven fathers (19, 25). These groups are the combination of maternal and paternal MHC haplotypes that were previously identified (Table II). We observed a strong correlation between m1 (*UAA* L band) and *UDA* U band, and m2 (*UAA* U band) and *UDA* L band. However, there was no correlation in three of the siblings 2, 8, 24, suggesting that they are recombinants between *UDA* and MHC (indicated by R in Table II). Paternal haplotype p2 also had a perfect correlation with the *UDA* M RFLP band in all 13 siblings (Table II). Based on the correlations observed in m1, m2, and p2, we calculated the corresponding LOD score (Table III). Since we cannot distinguish maternal vs paternal bands when the siblings have the U/L band of *UDA*, we eliminated these siblings from our calculations to be conservative. We used a total of 30 maternal alleles, including three potential recombinants, and 13 paternal alleles for the calculations and obtained the LOD score 7.617 (odds are 4,139,967 to 1 with a p value of 3.169×10^{-9}) (Table III) with a recombination frequency (θ) of ~ 0.07 . We anticipate that *UDA* is located at a relatively great distance from *UAA* (as well as from *TAP* and *LMP*, which are closely linked to *UAA*), perhaps to preclude recombination or other gene exchanges between *UAA* and *UDA*.

Discussion

The amino acid sequence alignments and phylogenetic tree analyses clearly indicate that the *UDA* lineage is a unique MHC class I lineage. *UDA*'s tissue distribution, low polymorphism, and unusual cytoplasmic tail show it to be a nonclassical class I molecule, likely having a specialized function. Based on its sharing of the classical class I canonical peptide-binding residues and the general high conservation of the hydrophilic PBR in all species, we propose that *UDA* binds to a specific set of peptides different from the polymorphic *UAA*, which binds to sets of peptides in an allele-specific fashion. Indeed, the *UDA* system may be similar to HLA-E (and mouse orthologue Qa1) that binds to leader peptides of classical HLA and peptides from some other self and foreign peptides in a “non allele-specific” manner (35).

We showed that *UDA* is linked to *UAA* in the shark MHC, and similar scenarios of linked classical and nonclassical class I genes are found in many other nonmammalian species. For example, there are two class I genes identified as classical in the chicken, BF1 and BF2; BF2 is highly polymorphic with ubiquitous and high expression, whereas BF1 is oligomorphic and poorly expressed, more like a class Ib molecule. Kim et al. (2018) (36) found that BF1 is recognized by natural killer (NK) cells, while BF2 is recognized primarily by cytotoxic T lymphocytes. Their results suggest that BF1 may bind only to specialized peptides and not the diverse array of peptides presented by the highly expressed BF2 proteins. Another comparison to consider is the classical U and nonclassical “typical” Z class I lineages in bony fish. Like *UAA* and *UDA*, both bony fish class I lineages are MHC-linked (but note that in bony fish, classical class I and class II genes are not linked; however, Z is linked to U) (29, 37). Also like *UDA*, class I proteins of the Z lineage are predicted to have long cytoplasmic tails, preserve the classical class I canonical peptide-binding residues, and

appear to be monomorphic within a species. However, our phylogenetic analysis does not support that *UDA* and the “typical” Z lineage are derived from a recent common ancestor.

Unlike the multigene family nonclassical *UBA* (5), but similar to *UAA* (and most vertebrate classical class I genes (17, 30)), *UDA* is in a low copy number for all elasmobranchs examined. Since *UDA* is linked to the MHC and preserved in all cartilaginous fish, it emerged over 400 million years ago. The evolutionary conservation of *UDA* suggests that it plays an important role in the cartilaginous fish immune system or even another essential function. We detected three recombinants between *UAA* and *UDA* in our family study, suggesting that unlike *UAA*, *UDA* is not closely linked to the transporter associated with antigen processing (*TAP*) and immunoproteasome (*LMP*) genes (no recombinants have ever been detected between these genes). Ancient lineages of *TAP/LMP/UAA* are found in many nonmammalian vertebrates including sharks (25, 30, 38-40); *UDA* likely uses a different antigen-processing pathway, perhaps like mouse Qa-1 and human HLA-E binding to peptides generated in the endoplasmic reticulum (35).

The Southern blot suggests that the holocephalan spotted ratfish (*H. colliei*) has many *UDA* gene copies as opposed to the elasmobranch species. However, only one copy was found in the genome sequence of its close relative, elephant shark *C. milii*. Whether the high copy number of spotted ratfish *H. colliei* is specific of this species remains to be confirmed, but it raises the question on whether *UDA* copy number may be tied to the biological and ecological features of the different holocephalan taxa or is rather a general aspect of the holocephalans.

One important aspect to be consider in future studies is that previous work on Chondrichthyan immunity has focused primarily on a model species for immune studies, the nurse shark (*G. cirratum*). Although this cartilaginous fish is an important animal model for many different immunological studies, its subtropical range, benthic habit and perhaps unique exposure to pathogens, might have resulted in specific features in its adaptive immune system. Indeed, lifestyle/habit may be a driver of copy number of MHC class I genes in cartilaginous fish. This feature can be found for other immune genes like immunoglobulins: the nurse shark *G. cirratum* has a relatively low number of Ig genes (~15 IgM cluster) (41, 42) compared to horn shark *H. francisci* (~200 IgM clusters) (43). Cartilaginous fish species differing in spatial range and ecological features may provide important and useful comparisons to assess the influence of habit and habitat on the genetic architecture of immune genes in this ancient vertebrate group. As suggested by Wilson (2017) (32), the lifestyle of different members of a vertebrate class, especially one as diverse as the Chondrichthyans, should be reflected in the structure and function of different physiological systems, including the immune system.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

The authors are indebted to L.F. Castro, A. Machado and A. Muñoz-Mérida for their help in preliminary data collection and bioinformatic analyses, and for useful criticism of the data.

Grant support: This project was supported by Portuguese funds through FCT – Portuguese Foundation for Science and Technology, to TA (PD/BD/114542/2016), PJE (IF/00376/2015), AV (DL57/2016), and National Institutes of Health Grants supported YO and MFF (AI140326-26 and AI02877). This work is also financed by the FEDER Funds through the Operational Competitiveness Factors Program - COMPETE and by National Funds through FCT - Foundation for Science and Technology within the scope of the project PTDC/ASP-PES/ 28053/2017.

Abbreviations:

(PBR) peptide-binding region

References

1. Warr GW 1997 The adaptive immune system of fish. *Dev Biol Stand.* 90: 15–21. [PubMed: 9270830]
2. Flajnik MF, and Rumfelt L. 2000 The immune system of cartilaginous fish In: *Origin and evolution of the vertebrate immune system.* Springer Berlin, Heidelberg 249–270.
3. Flajnik MF 2018 A cold-blooded view of adaptive immunity. *Nature Reviews Immunology.* 18: 438–453.
4. Kasahara M, Vazquez M, Sato K, McKinney EC, and Flajnik MF. 1992 Evolution of the major histocompatibility complex: isolation of class II A cDNA clones from the cartilaginous fish. *Proc Natl Acad Sci U S A.* 89: 6688–6692. [PubMed: 1495958]
5. Bartl S, Baish MA, Flajnik MF, and 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. [PubMed: 9550410]
6. Dooley H, and Flajnik MF. 2005 Shark immunity bites back: affinity maturation and memory response in the nurse shark, *Ginglymostoma cirratum.* *Eur J Immunol.* 35: 936–945. [PubMed: 15688348]
7. Criscitiello MF, Ohta Y, Saltis M, McKinney EC, and 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. [PubMed: 20488795]
8. Hsu E 2016 Assembly and Expression of Shark Ig Genes. *J Immunol.* 196: 3517–3523. [PubMed: 27183649]
9. Smith SL, Sim RB, and Flajnik MF. 2014 *Immunobiology of the Shark.* CRC Press
10. Hinds K, and Litman G. 1986 Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature.* 320: 546–549. [PubMed: 3083268]
11. Ohta Y, Powis SJ, Lohr RL, Nonaka M, Pasquier LD, and Flajnik MF. 2003 Two highly divergent ancient allelic lineages of the transporter associated with antigen processing (TAP) gene in *Xenopus*: further evidence for co-evolution among MHC class I region genes. *Eur J Immunol.* 33: 3017–3027. [PubMed: 14579270]
12. Ohta Y, Shiina T, Lohr RL, Hosomichi K, Pollin TI, Heist EJ, Suzuki S, Inoko H, and Flajnik MF. 2011 Primordial linkage of beta2-microglobulin to the MHC. *J Immunol.* 186: 3563–3571. [PubMed: 21321107]
13. Adams EJ, and Luoma AM. 2013 The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules. *Annu Rev Immunol.* 31: 529–561. [PubMed: 23298204]
14. Hashimoto K, Okamura K, Yamaguchi H, Ototake M, Nakanishi T, and Kurosaiva Y. 1999 Conservation and diversification of MHC class I and its related molecules in vertebrates. *Immunol Rev.* 167: 81–100. [PubMed: 10319253]

15. Dijkstra JM, Grimholt U, Leong J, Koopand BF, and Hashimoto K. 2013 Comprehensive analysis of MHC class II genes in teleost fish genomes reveals dispensability of the peptide-loading DM system in a large part of vertebrates. *BMC evolutionary biology*. 13 (1): 260. [PubMed: 24279922]
16. Dijkstra JM, Yamaguchiand T, and Grimholt U. 2018 Conservation of sequence motifs suggests that the nonclassical MHC class I lineages CD1/PROCR and UT were established before the emergence of tetrapod species. *Immunogenetics*. 70: 459–476. [PubMed: 29270774]
17. Kaufman J 2018 Unfinished business: evolution of the MHC and the adaptive immune system of jawed vertebrates. *Annu Rev Immunol*. 36: 383–409. [PubMed: 29677478]
18. Shen T, Lei M, Wang J, He X, Li X, and Li J. 2014 Molecular cloning, organization, expression and 3D structural analysis of the MHC class Ia gene in the whitespotted bamboo shark (*Chiloscyllium plagiosum*). *Vet Immunol Immunopathol*. 157: 111–118. [PubMed: 24315118]
19. Ohta Y, Okamura K, McKinney EC, Bartl S, Hashimoto K, and Flajnik MF. 2000 Primitive synteny of vertebrate major histocompatibility complex class I and class II genes. *Proc Natl Acad Sci U S A*. 97: 4712–4717. [PubMed: 10781076]
20. Wang C, Perera TV, Fordand HL, and Dascher CC. 2003 Characterization of a divergent non-classical MHC class I gene in sharks. *Immunogenetics*. 55: 57–61. [PubMed: 12715247]
21. Ballingall KT, Bontrop RE, Ellis SA, Grimholt U, Hammond JA, Ho CS, Kaufman J, Kennedy LJ, Maccari G, Millerand D, and Robinson J. 2018 Comparative MHC nomenclature: report from the ISAG/IUIS-VIC committee 2018. *Immunogenetics*. 70: 625–632. [PubMed: 30039257]
22. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duranand C, and Thierer T. 2012 Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28: 1647–1649. [PubMed: 22543367]
23. Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, and Hoon S. 2014 Elephant shark genome provides unique insights into gnathostome evolution. *Nature*. 505(7482): 174. [PubMed: 24402279]
24. Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S. 2013 MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 30: 2725–2729. [PubMed: 24132122]
25. Ohta Y, McKinney EC, Criscitiello MF, and 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. [PubMed: 11777971]
26. Ott J 1999 Analysis of human genetic linkage. JHU Press.
27. Lander E, and Kruglyak L. 1995 Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*. 11(3): 241. [PubMed: 7581446]
28. Kaufman J, Salomonsen J, and Flajnik MF. 1994 Evolutionary conservation of MHC class I and class II molecules—different yet the same Seminars in immunology, Elsevier 6: 411–424. [PubMed: 7654997]
29. Grimholt U, Tsukamoto K, Azuma T, Leong J, Koop BF, and Dijkstra JM. 2015 A comprehensive analysis of teleost MHC class I sequences. *BMC evolutionary biology*. 15(1): 32. [PubMed: 25888517]
30. Flajnik MF, Ohta Y, Namikawa-Yomada C, and Nonaka M. 1999 Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunol Rev*. 167: 59–67. [PubMed: 10319251]
31. Lizée G, Basha G, Tiong J, Julien J, Tian M, Biron KE, and Jefferies WA. 2003 Control of dendritic cell cross-presentation by the major histocompatibility complex class I cytoplasmic domain. *Nat Immunol*. 4(11): 1065. [PubMed: 14566337]
32. Wilson AB 2017 MHC and adaptive immunity in teleost fishes. *Immunogenetics*. 69: 521–528. [PubMed: 28695284]
33. D’Souza MP, Adams E, Altman JD, Birnbaum ME, Boggiano C, Casorati G, Chien YH, Conley A, Eckle SBG, Früh K, and Gondré-Lewis T. 2019 Casting a wider net: Immunosurveillance by nonclassical MHC molecules. *PLoS pathogens*. 15(2): e1007567. [PubMed: 30789961]
34. Schnell AH and Witte JS. 2016 Family-based study designs. In: *Molecular Epidemiology*. CRC Press 33–42.

35. Anderson CK, and Brossay L. 2016 The role of MHC class Ib-restricted T cells during infection. *Immunogenetics*. 68: 677–691. [PubMed: 27368413]
36. Kim T, Hunt HD, Parcells MS, van Santen S, and Ewald SJ. 2018 Two class I genes of the chicken MHC have different functions: BF1 is recognized by NK cells while BF2 is recognized by CTLs. *Immunogenetics*. 70: 599–611. [PubMed: 29947944]
37. Grimholt U 2016 MHC and Evolution in Teleosts. *Biology*. 5(1):6.
38. Nonaka M, Namikawa C, Kato Y, Sasaki M, Salter-Cid L, and Flajnik MF. 1997 Major histocompatibility complex gene mapping in the amphibian *Xenopus* implies a primordial organization. *Proc Natl Acad Sci U S A*. 94: 5789–5791. [PubMed: 9159152]
39. Kaufman J, Milne S, Göbel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, and Beck S. 1999 The chicken B locus is a minimal essential major histocompatibility complex. *Nature*. 401(6756): 923. [PubMed: 10553909]
40. McConnell SC, Hernandez KM, Wcisel DJ, Kettleborough RN, Stemple DL, Yoder JA, Andrade J, and Jong JL. 2016 Alternative haplotypes of antigen processing genes in zebrafish diverged early in vertebrate evolution. *Proc Natl Acad Sci U S A*. 113: E5014–E5023. [PubMed: 27493218]
41. Rumpf LL, Lohr RL, Dooley H, and Flajnik MF. 2004 Diversity and repertoire of IgW and IgM VH families in the newborn nurse shark. *BMC immunology*. 5(1): 8. [PubMed: 15132758]
42. Malecek K, Brandman J, Brodsky JE, Ohta Y, Flajnik MF, and Hsu E E. 2005 Somatic hypermutation and junctional diversification at Ig heavy chain loci in the nurse shark. *J Immunol*. 175: 8105–8115. [PubMed: 16339548]
43. Litman GW, Berger L, Murphy K, Litman R, Hinds K, and Erickson BW. 1985 Immunoglobulin VH gene structure and diversity in *Heterodontus*, a phylogenetically primitive shark. *Proc Natl Acad Sci U S A*. 82: 2082–2086. [PubMed: 3920659]

Key Points

- A new, MHC-linked class I gene (*UDA*) was found in cartilaginous fishes.
- *UDA* is monorphic and in single/low copy number in elasmobranchs.
- Chondrichthyans have at least 4 class I lineages, one classical and 3 nonclassical.

α 1 domain

nurse shark	Gici UAA	ESHSLRYFFTWSTAGSGIPEFVAVGVYDQQFVQYDSDRKEMIPRORWVKESE-----GPEYWERETQTLRGWEPWGKANI	DILSKRTN	OT-G	Asn
horn shark	Hefr UAA	A..WY...Y.SVSGV.DF.Q.....L..KELNY...WK...E..H.VMELAADADEFSDS.SYRTW..R.L..QI..V..QTALS...S-A			
nurse shark	Gici UBA	AT..V...H.GMSSVP.L...I...GV...M...VL...EQ.MVDT.-----AQF.K.KLSLAQEM.DHF.GYVP.FMS...-L			
horn shark	Hefr UBASM.PNP.F...G...V...M...L...EQ.MV.....A.A.KKFLAQER.MTF..R.P..MS.A..S-A			
nurse shark	Gici UDA	E.....YY.SMLNSGDF...IH..VL..SV..ITY...II.KD...P.MRT.Q-----DDQ...Q...R.IDR.KLSL..VQ.ATR...SS-E			
blue shark	Prgl UDA	D.....YY.SMLNSGDL.Q.IH..VL..GVPTITY...NG.TD.A..P.MNT.V-----KEG..AE...R.MDR.KLSM..VR.AT...GSSDK			
spiny dogfish	Sqac UCA	A...SVLH.ELAQCADF...H..VML...TVIMN...N..TSTVHLKHE.IAKHL-----TADQK..L.DLSA.EAHNIM.TG.N.NTM.L..KS-V			
		p P p	p p	p	P
		ssssssssss	ssssssss	ssssss	ss hhh h
					hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh

 α 2 domain

nurse shark	Gici UAA	GHHTYQLMCCGCELRDDGSSNTGFVCHAWDSTDFISLIDKDKMVVWVTPVTWGEITKNKWRDMAFNOGTGKYLEGICIEWLQKYLKNG-NVELRP			
horn shark	Hefr UAA	...V..T.Y.....-TR..Y.DG..GK.Y.NF..EN.....S.....Q.K.Y...W...T...K...L.-ER...			
nurse shark	Gici UBA	...M.VV..D..N..-AVE..N.FG..QO.L.F..H.....MP.A.SI..R..NOHV.NK.EW.YF..EEVQ..R...DY.-QR...V			
horn shark	Hefr UBA	..R.I.M.T..D.....-TS...N.YG..GR..L.F..H.....P..QNIQHN..NQ.RVG.L.W.N..EE...R...LY.-QR...V			
nurse shark	Gici UDA	LNLYL.YTS...LE..-TV..KRHY..FNGREL..F.LEHST..AVSPYA.S.QR...NSNAGD.EYKRY.TKK...N...RE.AA.LN.K			
blue shark	Prgl UDA	ALNFL.YTS...VGN..-TVS..IR..Y..FNGRE...F.LEYAT..AASRFAQS.Q.T..NKNK.D.TYK.Q.TKE...NT...Y.NGALSGK			
spiny dogfish	Sqac UCA	E..LF..W.KKV.VN..V-LVKWD.LFRF..GK...R...QQT.N...VVRSEAQSLERQFTKEPDWRNTRWEN..HTVVVG.NDL.QK.QQFVE..K			
		p p p	p p p	Pp	P PP p pp Pp p p P
		ssssssssss	ssssssssss	ssssss	sss hhhhhhhhhhhhh hhhhhhhhhhh hhhhhhhhhhh hhhh

 α 3 domain

nurse shark	Gici UAA	VKPSVTFTSVRGN---KQLSCVATGFYPHSIEVNLFRDSAKIDETESTGVRPNHEDGSYQLHRSTEFDFNSQAKYSCVVDHIDGLGQQLVVVFY-			
horn shark	Hefr UAA	.A...S..RRGDS---NR...T...QA...T.R..GVSL..SHFD..IL...T..T.KW.DI..AD...F..E...S..KEKM....-			
nurse shark	Gici UBA	.A.T.S..HLGDT---N...LV...QA...T.W..GVV...L.S..IL...RT...RKWI...ED..E...YME..I..EAK--I..-			
horn shark	Hefr UBA	.A.V.S..RPGDS---NR...V...QA...T.W..GVM...IL.S..IL...RT...RKWI...ED..E...R..E..S..TEMM..I..A			
nurse shark	Gici UDA	AV.EANVYTN.H---NVN..H.LV...Q..N.TW..A.GKVPDP.H...IL...T...RVVFQIERDDGRN..V..RIE..SS.PAPKEILW-			
blue shark	Prgl UDA	.V.E.YVYHTKSSDQNLNVH..HV...H..RAN..SWY..GEPL.SPO...IL...T...H..RIHFQTKAGDGRS..V..H..Q..SS..PEI.ETPW-			
spiny dogfish	Sqac UCA	SR..V.YVFDKRKPLDSSVAV..LV...RF...VWQ.NEDFVSGV..P.LL...T...KITLTLAAEDE.E.W.H.T..SS.SETQTEPW-			
		ssssssss	ssssssssss	ssssss	sss ss ss sssssssss sssssss sss

Connecting /transmembrane domain

	Conn	Tm	
nurse shark	Gici UAA	EPKTSSMLP--	VIIGVVVVIVLVIIAVIAGVVVY
horn shark	Hefr UAA	.R.SD.L.--	I.V..L..A..L...V..I.
nurse shark	Gici UBA	D..SH.QV.VT	LG.VFG.IGIIA.AVAVGA.II.
horn shark	Hefr UBA	GL.SD.QI.VT	.G.VL.ILGLIIVLVIV.VA.-.
nurse shark	Gici UDA	-.DGGVV.--	.V..CSILLL.L.FCIG..YYW
blue shark	Prgl UDA	-VQSDYTA---	M.V...LVI.LAGC.G.F.FYW
spiny dogfish	Sqac UCA	V..ER.PTVR--	GVAIAAGILGI.AVVCVTRICT

Cytoplasmic tail

		Tyr	Ser
nurse shark	Gici UAA	KKKAGQKTGYNPAKT--	SDKGDSSSSSLATA
horn shark	Hefr UAA	...E.....	.R.E...N.S.N.
nurse shark	Gici UBA	..GQV.SN.S.TD.--	.ETEF.D.PAVS
horn shark	Hefr UBA	.R.VTGT.S..T.--	...E...S.N.
nurse shark	Gici UDA	.H.DEVNRVITICMGHG.MASN..T..TDGSQTLNGSNSDSSSESGKCLISGPPSSPSNKQTVQTPLIQNPDGPLPHPA	
blue shark	Prgl UDA	.Q.M.SCCDFGGSV--	.VSTTG..NSSQISSEKIEVEVDRNLLSPGTHSSQNRMGTVQEPFLSTPD-PGPCRE
spiny dogfish	Sqac UCA	R..TA.DGQTQQPALLI.VSSRFMEQPKETGGPSGQLLPQPIGV	

Figure 1.

Multiple amino acid alignment of the four distinctive MHC class I lineages in cartilaginous fish, UAA, UBA, UCA and the new lineage UDA. GenBank accession numbers are listed in Supplemental Table I. Dots indicate amino acids identical to Gici UAA, and dashes indicate gaps, respectively. Highly conserved residues in all class I proteins are shaded black (16, 28), s and h indicate the β -strands and α -helices, and the line connecting the two Cys in the α 2 domain indicates the class I canonical disulfide bridge. **P** marks the invariant residues that bind to the N- and C-termini of the bound peptide in the classical class I molecules and p indicates the other 28 peptide binding residues. The Asn marks the Asparagine-linked glycosylation site, Asp/Glu indicates the typical Aspartic acid and Glutamic acid residues found in the connecting piece (Conn, light shade), and Tyr and Ser mark the conserved

positions of Tyr and Ser in the cytoplasmic tail (also in light shade) of classical class I molecules.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

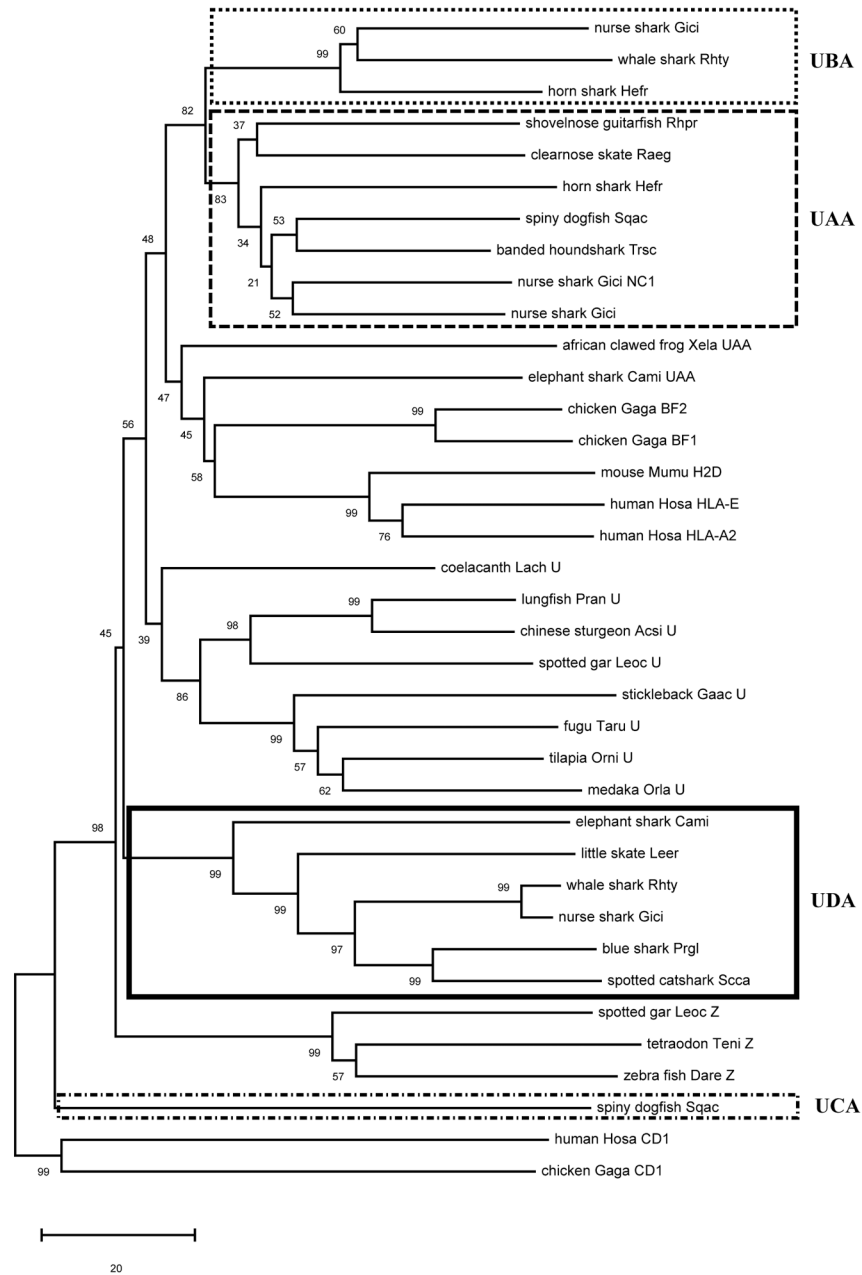


Figure 2.

Phylogenetic tree of cartilaginous fish and other selected vertebrate class I molecules. The tree was constructed using the neighbor-joining (NJ) method with the $\alpha 1$ and $\alpha 2$ domains (PBR), and rooted with CD1. Bootstrap support values are shown as percentages on the branches. The boxes indicate the four distinctive MHC class I lineages in cartilaginous fish. GeneBank accession numbers for all the sequences are listed in Supplemental Table I, and the corresponding alignment is found in Supplemental Figure 1. Common names are shown followed by abbreviation of the scientific name (two letters of the genus and two letters of species). The scale bar indicates the number of amino acid differences per sequence.

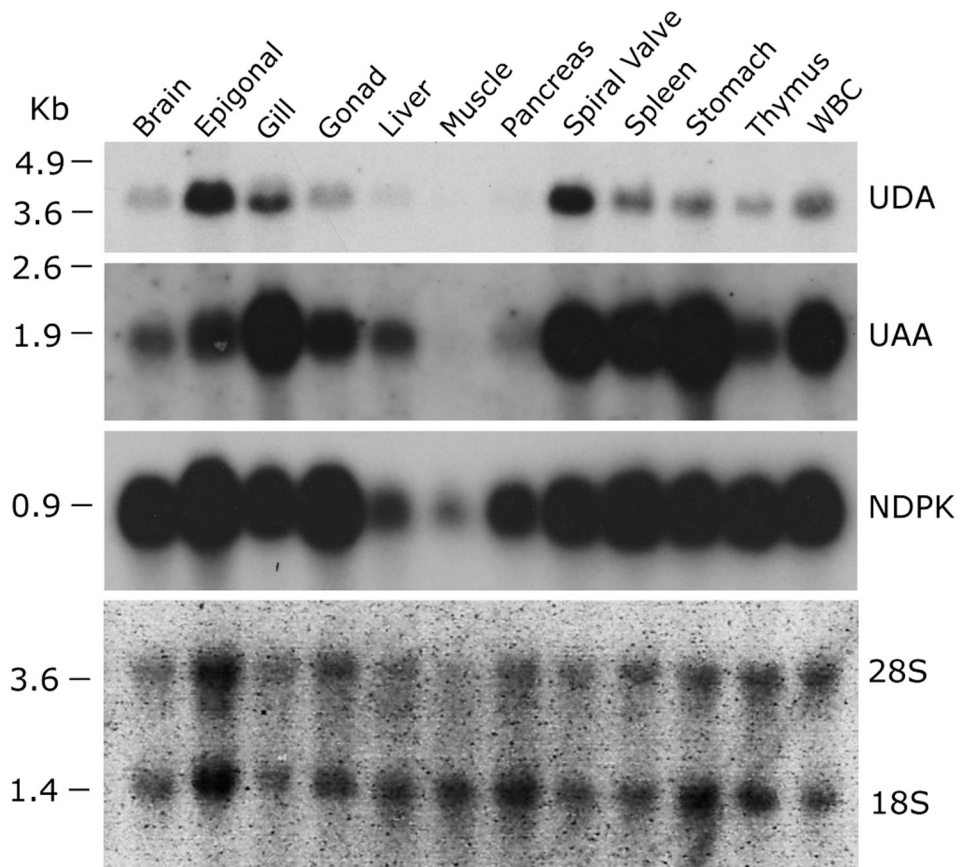


Figure 3. Unique tissue distribution of *UDA* lineage in nurse shark compared to *UAA* via northern blotting. Nucleoside-diphosphate kinase (*NDPK*) was used as a loading control (ubiquitous expression) although we found it to have lower expression in muscle and pancreas in the presence of similar amounts of 28S and 18S rRNA across samples. RNA marker size (kb) is shown on the left of the blot.

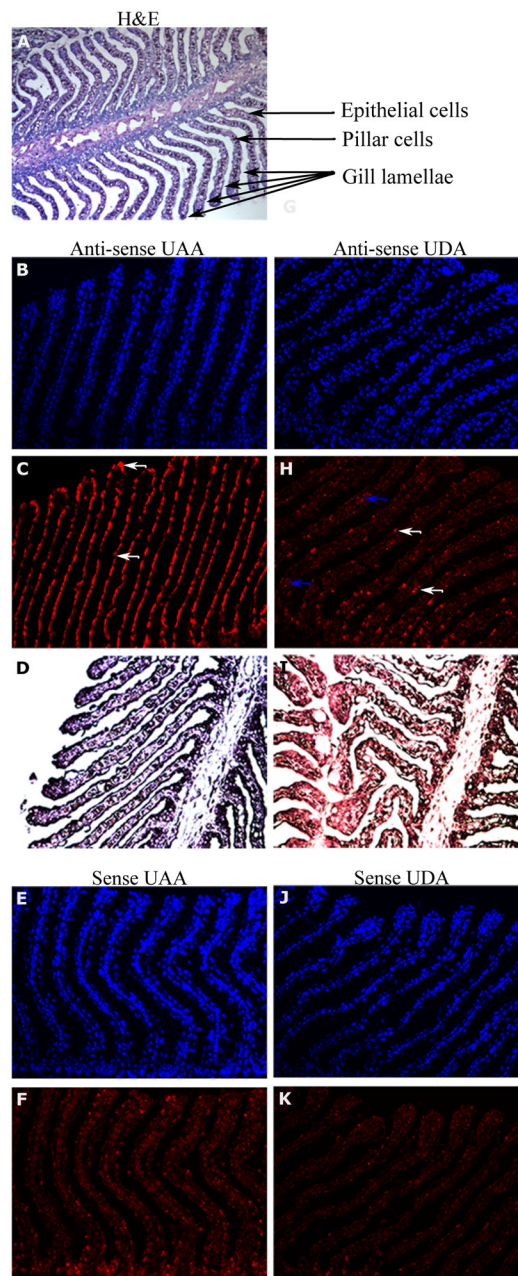


Figure 4:

Differential expression in nurse shark gill section between the classical *UAA* and *UDA* lineages, using *in situ* hybridization of full-length *UAA* and *UDA* region nurse shark riboprobe. A) H&E stained tissue section with the gill filament structures highlighted in the remainder images; B), C) and D) *UAA* antisense riboprobe; E) and F) *UAA* sense riboprobe; G), H) and I) *UDA* antisense riboprobe; J) and K) *UDA* sense riboprobe; B) E) G) and J) DAPI contain staining; C) H) F) and J) riboprobe with SA-AF647 signal detection and D) and I) riboprobe with signal detection into SA-AP and NBT/BCIP substrate. Note the highest expression of *UAA* in epithelial cells (white arrows), and of *UDA* in some pillar

(blue arrows) and epithelial cells (white arrows). All images were taken at 20X magnification except the H&E that was taken at 10X magnification.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

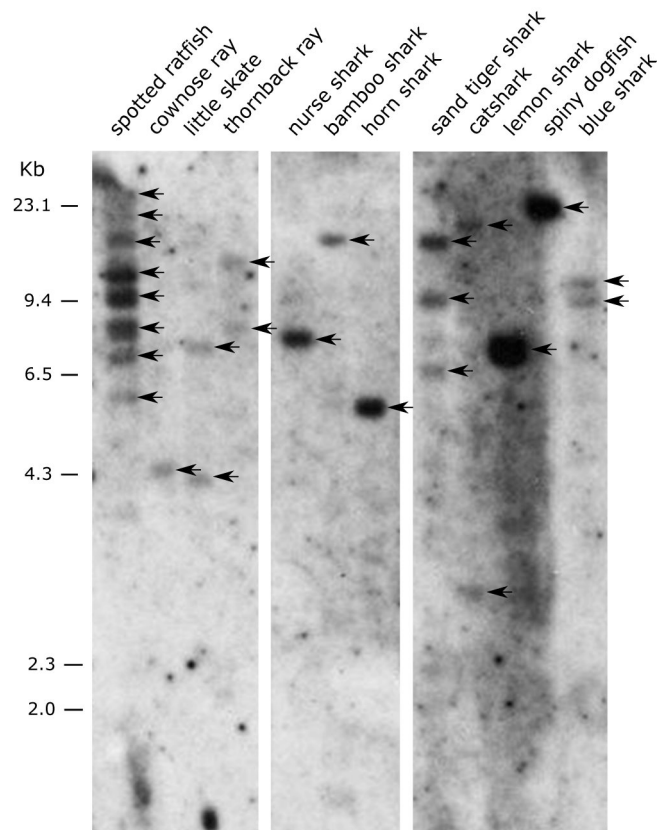


Figure 5. Presence of the *UDA* lineage in Chondrichthyans via Southern blotting. BamHI-digested DNA from different species was hybridized with the nurse shark MHC class I *UDA* $\alpha 3$ domain probe under low stringency conditions. The Chondrichthyan species used were one chimaera (spotted ratfish *Hydrolagus colliei*); 3 rays (cownose ray *Rhinoptera bonasus*, little skate *Leucoraja erinacea*, thornback ray *Raja clavata*) and 8 sharks (nurse shark *Ginglymostoma cirratum*, bamboo shark *Chiloscyllium punctatum*, horn shark *Heterodontus francisci*, sand-tiger shark *Carcharias taurus*, catshark *Scyliorhinus canicula*, lemon shark *Negaprion brevirostris*, spiny dogfish *Squalus acanthias*, blue shark *Prionace glauca*). The marker size (kb) is shown on the left. Bands in the gel are marked with arrows.



Figure 6. *UDA* is monomorphic in wild nurse shark individuals. A) Alignment of nucleotide sequences obtained from cloning of PCR products of $\alpha 1$ and $\alpha 2$ *UAA* and *UDA* from four wild nurse shark individuals (1, 2, 3 and 4). B) Trace files of nucleotide sequences obtained by direct sequencing of PCR products of $\alpha 1$ *UAA* (one individual) and of $\alpha 1$ and $\alpha 2$ *UDA* for the same four wild nurse shark individuals (1, 2, 3 and 4). Dots indicate nucleotides identical to nurse shark #1 (highlight).

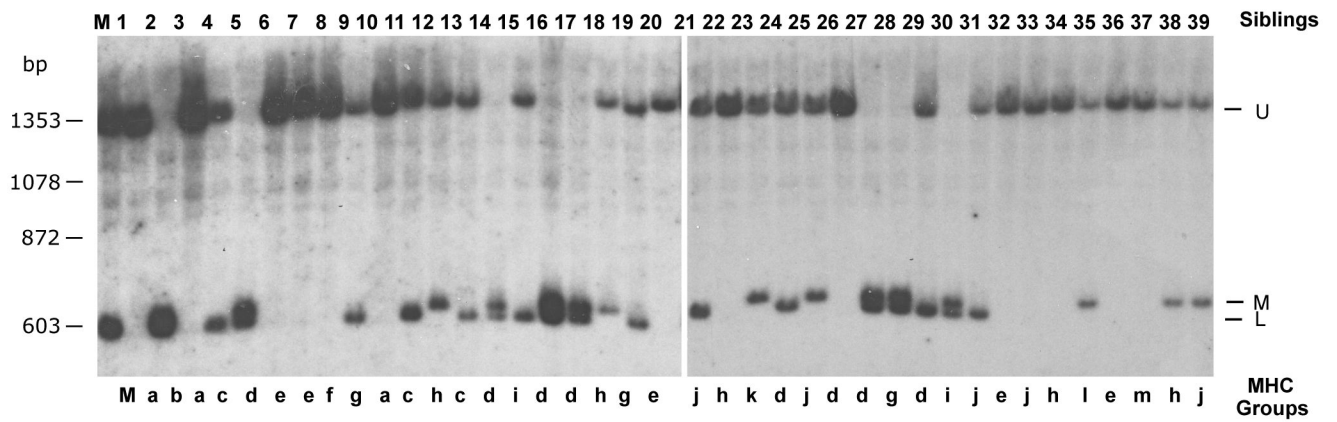


Figure 7.

UDA is linked to the nurse shark MHC. The Southern blot was performed with genomic DNA from 39 siblings in a nurse shark family (animal number shown above the blot) and their mother using BamHI restriction fragments hybridized with nurse shark $\alpha 3$ domain of UDA probe. The previously identified 13 MHC groups a-m identified with a *UAA* probe are indicated below (Ohta et al., 2000, Ohta et al., 2002). Marker size (kb) is shown on the left.

Table 1:

Peptide binding residues in MHC class I molecules from different lineages

Sample	Position of Peptide Binding Residues*																																				
	α1 domain							α2 domain																													
	5	7	9	22	24	34	45	59	63	66	67	70	73	74	77	80	81	84	95	97	99	114	116	118	123	124	143	146	147	152	155	156	159	160	163	167	171
se rk Gici AA	L	Y	F	F	A	V	M	Y	E	T	L	W	W	G	N	I	L	R	Y	L	C	V	H	W	F	I	T	K	W	F	G	T	Y	L	I	W	Y
se rk Gici AA-NCI	L	.	F	T	D	D	.	Q	I	A	W	V	G	N	T	A	.	I	L	V	Y	Y	W	.	I	F	Q	T	.	L	I	W	.
uded inshark c UAA	L	.	H	F	I	V	T	.	N	K	A	W	W	G	N	I	L	.	L	R	S	F	F	W	L	I	.	.	.	L	Q	W	.	L	I	W	.
ny gfish ac UAA	L	.	Y	F	A	V	M	.	E	K	Q	W	I	G	D	T	L	.	L	V	Y	F	Y	W	L	I	.	.	.	L	Q	R	.	L	E	W	.
inshark fr UAA	Y	.	Y	F	A	N	M	S	W	R	L	W	I	G	N	T	A	.	V	T	Y	Y	D	W	Y	I	.	.	.	Y	G	W	.	L	T	W	.
arnose te Raeg A	L	.	Y	F	A	V	L	.	E	K	L	W	G	F	N	T	L	.	I	R	F	S	Y	W	L	I	.	.	.	I	Q	N	.	L	T	W	.
ovelnose tarfish pr UAA	L	.	Y	F	T	T	D	.	Q	I	L	H	T	F	N	I	A	.	F	R	Y	D	F	W	.	I	.	.	W	G	N	.	L	I	W	.	
phant rk rk mi A	L	.	Y	F	I	V	M	.	Q	T	A	W	T	Y	N	T	V	.	V	Q	Y	I	Y	Y	.	I	.	.	L	Q	E	.	L	I	W	.	
inshark fr UBA	L	.	F	F	G	V	L	A	K	L	A	R	T	F	R	I	L	.	I	M	T	N	Y	W	.	L	I	N	.	G	G	W	.	L	E	W	.
ale rk Rhty A	L	.	L	F	I	I	L	A	K	L	A	Q	D	F	F	R	I	.	Y	M	T	N	Y	W	V	L	I	R	.	G	Q	W	.	L	E	W	.
se rk Gici A	V	.	H	F	I	V	L	F	K	L	A	M	H	F	Y	I	F	.	Y	M	V	N	F	W	L	I	I	R	.	N	E	W	F	L	E	W	.
ican wed g Xela A	L	.	Y	F	T	E	D	.	E	K	S	I	T	F	N	V	A	.	V	W	Y	E	H	Y	.	F	.	.	A	R	N	.	L	I	G	.	

J Immunol. Author manuscript; available in PMC 2021 February 15.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Position of Peptide Binding Residues*

sample	α1 domain										α2 domain																										
	5	7	9	22	24	34	45	59	63	66	67	70	73	74	77	80	81	84	95	97	99	114	116	118	123	124	143	146	147	152	155	156	159	160	163	167	171
ken	L	.	R	F	D	M	A	.	E	I	V	S	I	N	N	I	L	.	V	W	S	H	A	Y	.	V	.	.	.	Y	G	L	.	L	T	W	.
gaBF2	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
cken	.	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
gaBF1	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
use	.	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
imu	.	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
D	.	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
man	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
SA	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
/A-E	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
man	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
man	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
SA	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
/A-A2	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
SA	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
lacanth	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
chU	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
g fish	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
in U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
nese	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
geon	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
st U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
st U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
itted	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
m Leoc	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
back	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
acU	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
Taru	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
pia	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
daka	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
a U	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
phant	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
rk	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
mi	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
WA	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
le skate	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
er	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.
UDA	L	.	H	Y	D	V	Y	.	Q	I	G	N	S	V	S	T	L	.	V	W	F	R	V	Y	.	I	.	.	.	V	G	W	.	L	T	W	.

J Immunol. Author manuscript; available in PMC 2021 February 15.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Position of Peptide Binding Residues*

sample	α1 domain															α2 domain																					
	5	7	9	22	24	34	45	59	63	66	67	70	73	74	77	80	81	84	95	97	99	114	116	118	123	124	143	146	147	152	155	156	159	160	163	167	171
ale	L	.	Y	F	H	T	D	.	E	R	L	R	L	C	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
ark Rhty	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
ark Gici	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
e shark	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	T	.	D	Y	K	.	T	I	W	.
UDA	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
itted	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
shark	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
a UDA	L	.	Y	F	H	T	D	.	E	R	L	R	L	S	N	I	A	.	L	Y	S	R	Y	F	L	I	.	.	.	D	Y	K	.	T	I	W	.
ted gar	L	.	Y	I	E	R	K	.	G	S	R	K	W	F	N	I	L	.	L	W	H	D	Y	Y	.	L	.	.	L	Y	T	.	L	E	W	F	.
cc Z	L	.	Y	I	E	R	K	.	G	S	R	K	W	F	N	I	L	.	L	W	H	D	Y	Y	.	L	.	.	L	Y	T	.	L	E	W	F	.
adon	L	.	Y	I	E	H	K	.	G	S	R	K	W	F	N	I	L	.	L	W	H	D	Y	Y	.	L	.	.	L	Y	T	.	L	E	W	F	.
ii Z	L	.	Y	I	E	H	K	.	G	S	R	K	W	F	N	I	L	.	L	W	H	D	Y	Y	.	L	.	.	L	Y	T	.	L	E	W	F	.
ara fish	L	.	Y	I	E	R	K	.	G	S	R	K	W	F	N	I	L	.	L	W	H	D	Y	Y	.	L	.	.	L	Y	T	.	L	E	W	F	.
re Z	L	.	Y	I	E	R	K	.	G	S	R	K	W	F	N	I	L	.	L	W	H	D	Y	Y	.	L	.	.	L	Y	T	.	L	E	W	F	.
ny	L	V	H	F	A	M	V	Q	E	S	A	A	I	M	G	N	T	L	F	W	K	V	F	F	.	I	L	Q	F	W	T	W	N	L	T	G	L
fish	L	V	H	F	A	M	V	Q	E	S	A	A	I	M	G	N	T	L	F	W	K	V	F	F	.	I	L	Q	F	W	T	W	N	L	T	G	L
a UCA	L	V	H	F	A	M	V	Q	E	S	A	A	I	M	G	N	T	L	F	W	K	V	F	F	.	I	L	Q	F	W	T	W	N	L	T	G	L
nan	L	V	H	F	A	M	V	Q	E	S	A	A	I	M	G	N	T	L	F	W	K	V	F	F	.	I	L	Q	F	W	T	W	N	L	T	G	L
sa CDI	F	V	Q	T	G	H	A	E	L	I	F	Y	G	F	E	D	F	D	I	G	A	L	G	L	.	L	F	L	I	I	T	V	L	L	T	Y	V
cken	L	L	H	V	G	G	I	D	I	S	I	Y	D	F	L	M	Y	V	F	S	I	F	I	Y	.	L	E	M	A	L	V	I	L	L	T	I	F
ga CDI	L	L	H	V	G	G	I	D	I	S	I	Y	D	F	L	M	Y	V	F	S	I	F	I	Y	.	L	E	M	A	L	V	I	L	L	T	I	F

J Immunol. Author manuscript; available in PMC 2021 February 15.

19 invariant residues that bind to the N- (light gray) and C-termini (dark gray) of bound peptides found in the classical class I and 28 other predicted peptide binding residues are shown. The dots in the 9 invariant residues indicate the same amino acids present in the reference sequence, nurse shark Gici UAA. The UDA peptide-binding hydrophilic or charged residues are indicated in bold.

* position of peptide binding residues is based on the Human HL-A-A2 sequence. Sequence information listed in Supplemental Table 1.

Table II.

Summary of MHC haplotypes and UDA RFLP in a nurse shark family

maternal MHC haplotype*	m1																					
Siblings	1	3	10	12	18	22	34	38	21	25	31	33	39	6	7	20	32	36	2	35	37	
MHC group*	a	a	a	h	h	h	h	h	j	j	j	j	j	e	e	e'	e'	e'	b	l	m	
paternal MHC haplotype*	p1	p1	p1	p2	p2	p2	p2	p2	p3	p3	p3	p3	p3	p4	p4	p6	p6	p6	p7	p9	p10	
UAA aI*	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	
UDA	U	U	U	<u>U/M</u>	<u>U/M</u>	<u>U/M</u>	<u>U/M</u>	<u>U/M</u>	U	U	U	U	U	U	U	U	U	<u>L/M</u>	U	<u>U/M</u>		
potential recombinant																			R			

maternal MHC haplotype*	m2																		m1/2		
Siblings	8	5	14	16	17	24	26	27	29	15	30	9	19	23	4	11	13	28	Mom		
MHC group*	f	d	d	d	d	d	d	d	d	l	i	g	g	k	c	c	c	g'			
paternal MHC haplotype*	p1	p2	p2	p2	p2	p2	p2	p2	p2	p3	p3	p5	p5	p8	p11	p11	p11	p12			
UAA aI*	U	U	U	U	U	U	U	U	U	U	U	<u>U/L</u>	<u>U/L</u>	U	U/L	U/L	U/L	U/L	L/U		
UDA	<u>U</u>	<u>L/M</u>	<u>L/M</u>	<u>L/M</u>	<u>L/M</u>	<u>U/M</u>	<u>L/M</u>	<u>L/M</u>	<u>L/M</u>	<u>L/U</u>	<u>L/U</u>	<u>L/U</u>	<u>L/U</u>	<u>L/U</u>	L/U	L/U	L/U	L/U	U/L		
potential recombinant	R						R														

*MHC haplotypes and UAA RFLP were obtained from (19) and (25)

Note that potential recombinant RFLP bands are underlined. In all cases, recombination was detected in the maternal chromosomes.

Table III.

LOD score calculation from the nurse shark family

θ	Maternal (m1 & m2)			Paternal (p2)			Combined
	"R"	"NR"	LOD score	"R"	"NR"	LOD score	LOD score
0	3	27	∞	0	13	∞	∞
0.01	3	27	2.61	0	13	3.56	6.17
0.02	3	27	3.4	0	13	3.5	6.9
0.03	3	27	3.8	0	13	3.44	7.24
0.04	3	27	4.06	0	13	3.38	7.44
0.05	3	27	4.22	0	13	3.32	7.54
0.06	3	27	4.34	0	13	3.26	7.6
0.07	3	27	4.41	0	13	3.2	7.61
0.08	3	27	4.46	0	13	3.14	7.6
0.09	3	27	4.49	0	13	3.08	7.57
0.1	3	27	4.49	0	13	3.02	7.51
0.15	3	27	4.35	0	13	2.69	7.04
0.2	3	27	4.02	0	13	2.35	6.37
0.25	3	27	3.55	0	13	1.99	5.54
0.3	3	27	2.98	0	13	1.6	4.58
0.35	3	27	2.31	0	13	1.18	3.49
0.4	3	27	1.55	0	13	0.73	2.28
0.45	3	27	0.68	0	13	0.27	0.95
0.49	3	27	0.05	0	13	0.01	0.06
0.5	3	27	0	0	13	0	0