Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Mol Immunol. Author manuscript; available in PMC 2021 March 31.

Published in final edited form as:

Mol Immunol. 2020 December ; 128: 125–138. doi:10.1016/j.molimm.2020.10.003.

Cartilaginous fish class II genes reveal unprecedented old allelic lineages and confirm the late evolutionary emergence of DM

Tereza Almeida[a,b,c], **Arnaud Gaigher**[a], **Antonio Muñoz-Mérida**[a], **Fabiana Neves**[a], **L. Filipe C. Castro**[b,d], **Martin F. Flajnik**[c], **Yuko Ohta**[c], **Pedro J. Esteves**[a,b], **Ana Veríssimo**[a,*]

[a]CIBIO-InBIO, Centro de Investigacão em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[b]Department of Biology, Faculty of Sciences - University of Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal

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

[d]CIIMAR – Interdisciplinary Centre of Marine and Environmental Research, University of Porto, 4450-208 Matosinhos, Portugal

Abstract

Cartilaginous fish (chimaeras, rays and sharks) are the most basal extant jawed vertebrates with an adaptive immune system based on the Major Histocompatibility Complex (MHC). Despite being a key taxon in the evolution of vertebrate adaptive immunity, no comprehensive characterization of MHC class II genes has been undertaken for the group. We performed extensive bioinformatic searches on a taxonomically diverse dataset of transcriptomes and genomes of cartilaginous fish targeting MHC class II sequences. Class IIα and IIβ sequences were retrieved from all taxa analyzed and showed typical features of classical class II genes. Phylogenetic trees of the immunoglobulin superfamily domain showed two divergent and remarkably ancient lineages of class II genes in Selachians (sharks), originating >350 million years ago. Close linkage of lineage-specific pairs of IIα and IIβ genes was found, confirming previous results, with genes from distinct lineages segregating as alleles. Nonclassical class II DM sequences were not retrieved from these data and classical class II sequences lacked the conserved residues shown to interact

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*Corresponding author at: CIBIO-InBIO, Centro de Investigacão em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal. averissimo@cibio.up.pt (A. Veríssimo).

Author contributions

A.V., M.F.F. and Y.O designed the study and the experiments; T.A., A.G., A.M.-M., F.N. and Y.O. performed experiments; T.A., A.G., F.N., M.F. F., Y.O. and A.V. analyzed the data; and T.A., A.G., F.N., L.F.C., M.F.F., Y.O., P.J.E. and A.V. wrote the manuscript.

CRediT authorship contribution statement

Tereza Almeida: Data curation, Formal analysis, Methodology, Writing - original draft. **Arnaud Gaigher:** Data curation, Formal analysis, Methodology, Writing - original draft. **Antonio Muñoz-Mérida:** Data curation, Formal analysis, Methodology. **Fabiana Neves:** Formal analysis, Methodology, Writing - original draft. **L. Filipe C. Castro:** Writing - original draft. **Martin F. Flajnik:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing - original draft. **Yuko Ohta:** Conceptualization, Formal analysis, Methodology, Writing - original draft. **Pedro J. Esteves:** Writing - original draft. **Ana Veríssimo:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Conflict of Interest: The authors declare no commercial or financial conflict of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2020.10.003.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with DM molecules, supporting claims that the DM system arose only in the lobe-finned fish lineage leading to tetrapods. Based on our search methods, other divergent class II genes are unlikely in cartilaginous fish.

Keywords

Adaptive immunity; MHC; Antigen presentation; Basal jawed vertebrates; Sharks and rays

1. Introduction

The major histocompatibility complex (MHC) is a genomic region encoding genes involved in antigen presentation and processing, which are essential for immune defense mechanisms. One of the key immune gene families is MHC class II, which initiates an adaptive immune response by presenting exogenously acquired antigens in the form of peptides to CD4+ T-cells. Class II molecules are heterodimers of α and β chains (IIα and IIβ) encoded by separate genes. Each chain is composed of two extracellular domains, namely the peptide-binding region (PBR, i.e. α1 and β1) and the immunoglobulin superfamily domain (IgSF, i.e. α2 and β2), in addition to a signal peptide, a connecting piece (CP), a transmembrane domain (TM), and the intracellular cytoplasmic tail (CYT) (Klein and Sato, 2000). A hallmark of most class II genes in jawed vertebrates is the high level of polymorphism and evolutionary positive selection on the codons encoding the peptide binding residues, resulting in high diversity at the PBR and allowing the recognition of a wider array of pathogen-derived peptides conferring broad pathogen resistance (Radwan et al., 2020).

Binding of antigenic peptides by classical MHC class II molecules is preceded by several steps associated with heterodimer assembly and association with the invariant chain (Ii) in the endoplasmic reticulum (ER). The class II-Ii complex transits from the ER to a specific late endosomal compartment where Ii breakdown occurs, leaving the class II–associated invariant chain peptide (CLIP) in the class II PBR. The exchange of CLIP by antigenic peptides is facilitated by the nonclassical class II molecule DM, that interacts most closely with the α1 domain of the CLIP-class II complex (Pos et al., 2012). The DM system also promotes binding of high-affinity peptides that form stable complexes with class II proteins (Lazarski et al., 2006; Schulze and Wucherpfennig, 2012). Class II peptide processing is, in turn, regulated by another nonclassical class II molecule found in mammals, DO, that works in concert with the DM molecules (Denzin et al., 1997; Kropshofer et al., 1998).

Classical MHC class II genes have been described in all jawed vertebrate lineages, including mammals (reviewed in Yeager and Hughes, 1999), birds (reviewed in Burri et al., 2010 and Parker and Kaufman, 2017), amphibians (Flajnik et al., 1991; Kiemnec-Tyburczy et al., 2010), bony fishes (reviewed in Dijkstra et al., 2013), and cartilaginous fishes (Bartl and Weissman, 1994; Kasahara et al., 1992). The *Ii* genes have also been reported for all jawed vertebrate lineages (Criscitiello et al., 2012), while *DM* genes have been found in sarcopterygians (i.e. lobe-finned fishes), from lungfish to mammals (Dijkstra et al., 2018; Fling et al., 1994; Kaufman, 1999; O'hUigin et al., 1998; Ohta et al., 2006), but so far not in basal jawed vertebrates such as cartilaginous and actinopterygians (i.e. ray-finned fishes),

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

including teleosts (Dijkstra et al., 2013). Teleost fish genomes show many divergent features from other jawed vertebrates, with tetrapods and cartilaginous fish genomes showing comparatively more similarities (reviewed in Ravi and Venkatesh, 2018). In fact, teleosts lack the primordial linkage between classical class I and II genes observed in all other jawed vertebrates (Klein and Sato, 1998; Ohta et al., 2000). Given the paucity of immunogenetic studies on Chondrichthyan taxa, it remains to be confirmed if the loss of DM is (another) derived feature of teleosts, or if the absence of the DM is a common feature in basal jawed vertebrates implying its late emergence in the lobe-finned fish leading to tetrapods.

Cartilaginous fish (Class Chondrichthyes) are the oldest living jawed vertebrates (~500 MY old) possessing an adaptive immune system and represent a key evolutionary group to study the evolution of vertebrate adaptive immunity. The group includes Elasmobranchs (sharks and rays) and Holocephalans (chimaeras) comprising a diverse assemblage of species occupying all aquatic ecosystems (marine, estuarine and freshwaters) and exploiting a variety of habitats (e.g. coastal, open-ocean, or deep waters) (Carrier et al., 2010), and whose lineages date back several hundred million years (Heinicke et al., 2009). Thus, exposure to different environmental stressors and pathogenic agents, both currently and through time, may have shaped putatively different immune response mechanisms and genetic makeup in this group. However, our current knowledge of Chondrichthyan immunobiology has relied mostly on a single model species, the nurse shark *Ginglymostoma cirratum* (reviewed in Smith et al., 2014).

Here we provide the first comprehensive assessment of MHC class II gene diversity in cartilaginous fishes, making use of accessible genetic, genomic and transcriptomic resources made available for sharks, rays and chimaeras in the last few years. Chondrichthyan class II sequences were compared with those from classical and nonclassical class II genes of other vertebrates regarding their biochemical features and phylogenetic relationships. By doing so, we also objectively searched for DM-like homologs in cartilaginous fish. Furthermore, we assessed class II gene diversity among sharks, rays and chimaeras to infer the number and diversity of class II lineages in each group. Ultimately, this study serves as a baseline of Chondrichthyan MHC class II immunogenetics and aims to promote additional studies exploring the evolutionary history of MHC genes in the group at the origins of adaptive immunity. These data also set the stage for future studies of the drivers of MHC diversity within and among species exhibiting remarkable variety in biology and ecology, as well as having long evolutionary histories.

2. Material and methods

2.1. Taxonomic breadth and data availability

Sequence Read Archive (SRA), Transcriptome Shotgun Assembly (TSA) and Whole Genome Sequence (WGS) databases publicly accessible on NCBI were screened for MHC class II-like sequences on available Chondrichthyan taxa. A total of 33 species were included in the analysis (Supplemental Table 1), covering the Holocephalans (or chimaeras; 1 species) and the two sister lineages of Elasmobranchs (Selachii, or sharks: 23 species, and Batoidea, or rays: 9 species). All but one order of Batoids (sensu Aschliman, 2011) and five out of the nine orders of Selachii (sensu Naylor et al., 2012) are represented in the dataset.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Some SRA datasets from RNAseq projects on Selachian taxa were excluded when referring to species whose genus had multiple datasets and if including non-immunity-related tissues (e.g. retina, pectoral fin bud or ampullary receptor cells). In turn, Batoid and Holocephalan taxa were underrepresented in public databases and thus all available data were retained for MHC screening.

2.2. Bioinformatic searches and filtering

The different "*omic*" databases available on NCBI (i.e. TSA; SRA; WGS) were screened using different protocols to extract class II-like sequences. Briefly, for TSA data, the available transcripts were downloaded and blasted against a set of five class IIα and five IIβ reference sequences (species name followed by accession number: *Homo sapiens* HLA-DR ARB08440.1 and AAB60387.1, and HLA-DM CAA54169.1 and ARB08367.1; elephant shark *Callorhinchus milii* AFK10583.1 and AFP09377.1, nurse shark *Ginglymostoma cirratum* AAF66123.1 and AAF82681.1, and winter skate *Leucoraja ocellata* GEZH01010955.1 and GEZH01040536.1). We only retained transcripts with a blast e-value under 0.001 and covering more than 50 % of each reference sequence. For SRA data, raw reads were downloaded, cleaned using Trimmomatic, and assembled using SPAdes to obtain the final transcripts. The resulting transcripts were screened as mentioned above for TSA data. Regarding WGS data, available contigs were screened using the same reference sequences mentioned above through tblastn, to locate the positions in the contigs using stringent e-values (10^{-7}). The corresponding hits were visually inspected and filtered to avoid overlaps and redundancy, and the corresponding gene sequence was bioinformatically extracted from target contigs using a flanking region of 200 bp to maximize inclusion of the whole gene sequence.

To refine our search and retain only class IIα and IIβ proteins, all retrieved sequences were translated to amino acids and only those showing the basic features of class II genes were retained, namely (at least) full sequences containing the PBR (α1/β1) and IgSF domains (α2/β2). This Chondrichthyan sequence dataset was complemented with additional class II sequences deposited in the GenBank nt database (see Supplemental Files 1 and 2 for full list of sequences and references). Furthermore, classical and nonclassical class IIα and IIβ sequences from other jawed vertebrate lineages (ray-finned fishes, amphibians, reptiles/birds and mammals) were downloaded from GenBank to compare and contrast MHC diversity of Chondrichthyans to that of other vertebrates.

2.3. Structural and biochemical features of MHC II genes

Amino acid alignments of full MHC IIα and IIβ proteins were performed separately, using the MUSCLE algorithm (Edgar, 2004) implemented in Geneious Prime v2.1. Sequence alignments were manually edited to highlight typical features of functional class IIα and IIβ genes following Brown et al. (1993); Kaufman et al. (1994); Wang et al. (2001) and Dijkstra et al. (2018, 2013). These structural and biochemical features were checked in all retrieved cartilaginous class II sequences and compared against representative lineages of other vertebrate groups (as listed above) including both classical and nonclassical MHC II genes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2.4. Phylogenetic analyses

To infer the phylogenetic relationships of Chondrichthyan class II genes and those of other vertebrate representatives, including classical and nonclassical genes (e.g. DO and DM genes), we built neighbor-joining trees with Jones-Taylor-Thornton model as implemented in MEGAX v10.1.18 (Kumar et al., 2018; branch support evaluated by 1000 bootstrap replicates). MHC IIα and IIβ trees were performed independently based on amino acid sequence alignments of the PBR and IgSF domain (exons 2 and 3), including representative sequences of each Chondrichthyan genus in the dataset in addition to classical and nonclassical sequences representative of different vertebrate lineages (i.e. Actinopterygii/ray-finned fishes: spotted gar *Lepisosteus oculatus*; stickleback *Gasterosteus aculeatus*; salmon *Salmo salar;* zebrafish *Danio rerio* Sarcopterygii/lobe-finned fishes: coelacanth *Latimeria chalumnae,* lungfish *Lepidosiren paradoxa*, frog *Xenopus laevis,* lizard *Sphenodon punctatus*, chicken *Gallus gallus*, rat *Rattus norvegicus* and human *Homo sapiens*).

As different evolutionary dynamics are expected between class II domains, neighbor-joining trees were built independently for PBR and IgSF domains (i.e. exon 2 vs. exon 3) of IIα and IIβ genes of Chondrichthyan taxa, based on nucleotide sequence alignments and using the Kimura 2-parameter distance implemented in MEGAX (branch support evaluated by 1000 bootstrap replicates). An expanded Chondrichthyan sequence dataset was used for this purpose, i.e. including multiple sequences from each species screened, where available, to allow a better perspective on the sequence diversity and divergence patterns in the group. With this in mind, we retained only sequences with less than 97 % identity at the within-species level.

2.5. Selection analysis

To infer codon-specific footprints of evolutionary positive and negative selection across the Chondrichthyan phylogeny we used four Maximum Likelihood (ML) methods available in the HYPHY package (Pond et al., 2005) implemented in the Datamonkey web server (Weaver et al., 2018), namely Single Likelihood Ancestor Counting (SLAC), Fixed-Effect Likelihood (FEL), Random Effects Likelihood (REL) and Fast Unconstrained Bayesian AppRoximation (FUBAR). All these methods are based on the estimation of nonsynonymous (dN) and synonymous (dS) substitutions rates for each site, with $dN/dS < 1$ being indicative of negative selection while $dN/dS > 1$ of positive selection. The selection inferences were conducted independently for each domain (α1, α2, β1 and β2) by using the tree topology described above. To avoid false positives, codons were considered evolving under selection if detected by at least two different methods.

2.6. Amplification of MHC class II genes and pedigree analyses

Based on the results from the phylogenetic analysis, two lineages were found for class IIα (i.e. DAA and DBA) and IIβ (DAB and DBB) (see more details in the Results section). One family of nurse sharks was previously used in demonstrating linkage between class II αDAA and βDAB genes (family#1; Ohta et al., 2000), with two pups exhibiting paternal alleles for two αDBA genes (the mother was negative for DBA genes, but multiple paternity was found in the litter). Here, the same family was screened for the presence/absence of the MHC class

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

II αDBA and βDBB genes to infer possible linkage between specific pairs of α and β genes. In addition to nurse shark family#1, 15 wild nurse sharks were also analyzed to assess consistency of the association between class IIα and IIβ genes in unrelated individuals. Screening was based on lineage-specific amplification with the polymerase chain reaction (PCR) using primers anchored in the IgSF domains of IIα and IIβ genes, namely DAA forward 5' CCTGAAGTCTCTGTGTATTCTG 3' and reverse 5' TACTGGATCCTGTAGGCTCGA 3'; DBA forward CCTCAGATTGCCATGTATCCTG 3' and reverse 5' CGTTGGATCCTGCAGCCCCTC 3'; DAB forward 5' ATCCGAACAAAAGAATCGAC 3' and reverse 5' ATCCACACTCATTGGTGAGG 3'; DBB forward 5' ATCCGACCTAAAGCCTCTCA 3' and reverse 5' TTCCACACTCCTAGGACTCC 3'. PCR amplification was performed in 15 μL reactions using 7.5 μL Qiagen Multiplex master mix, 7.5 pmol of each primer, 3 μL autoclaved water, and 15 ng of genomic DNA. The temperature profile included an initial denaturation at 95 °C for 15 min, followed by 35 cycles of 95 °C for 1 min, annealing at 58 °C for DAA and DAB, 59 °C for DBB, and 62 °C for DBA for 45 s, 72 °C for 1 min, and a final extension at 72 °C for 5 min. The resulting amplicons were sequenced directly.

3. Results

3.1. Chondrichthyan sequences showed conserved features of vertebrate class II genes

Our bioinformatic searches retrieved MHC class II-like sequences from 33 species of sharks, rays and chimaeras for which data were available. In total, 97 and 83 sequences exhibited the typical domains in class IIα and IIβ chains, and showed high similarity to previously described class II sequences of Chondrichthyan taxa (e.g. nurse shark *G. cirratum,* Bartl and Weissman, 1994 and Kasahara et al., 1992; elephant shark *C. milii,* Venkatesh et al., 2014; Fig. 1). Several structural features typical of the PBR domains (α1 and β1) of class II genes were found across the Chondrichthyan sequences (Fig. 1). Specifically, as expected, there were no intradomain disulfide bonds in α1, while there were the two canonical cysteines in β1 (Cβ15 and Cβ79). The α1 domain had a putative N-linked glycosylation site (NXS/T) conserved at Nα75 across Elasmobranchs (except in the requiem sharks - Order Carcharhiniformes, and stingrays - Order Myliobatiformes), although this was shifted from the mammal-like position (i.e. DR Nα78) as noted by Kasahara et al. (1992). In Carcharhiniformes, the glycosylation site is generally conserved as Nα78, while in Myliobatiformes the α75 was polymorphic (N, Q, K or I) (Fig. 1a, Supplemental Files 3 and 4). In contrast, all sequences retrieved from the Holocephalan *C. milii* lacked the N-linked glycan residues in the α1 domain. In turn, the N-linked glycosylation site reported for the β1 domain of higher vertebrates was conserved across all Chondrichthyans (Nβ19; Fig. 1b). Most class IIα sequences exhibited the highly conserved salt bonds in the distal domain (Hα5/Dα27; Fig. 1a), with few sequences exhibiting the Yα5 substitution originally described for the nurse shark (Kasahara et al., 1992); thus, the α1 salt bond is the rule rather than the exception.

The PBR of classical class II genes are characterized by the presence of eight conserved peptide-binding residues, of which four are particularly important in establishing hydrogen-bonds and are often perfectly conserved in tetrapods: αN62, αN69, Wβ61, and Nβ82

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Kaufman et al., 1994; Painter and Stern, 2012). Three of these four residues were generally conserved in Chondrichthyans, except α62 which was highly variable and under positive selection (Table 1). The remaining peptide-binding residues were either highly variable (e.g. α51, α53), or generally conserved for a residue distinct from those in higher vertebrates resulting in amino acid substitutions with very distinct properties (α76, β81; Table 1).

Expected structural features of IgSF domain of MHC class II molecules were also detected across Chondrichthyan sequences (Fig. 1), namely two Cys residues in α2 (Cα107, Cα163) and β2 (Cβ117, Cβ173). Minisatellite motifs were found across all Chondrichthyan β2 sequences (codons 150–156, Fig. 1b), but also across Elasmobranch α1 sequences (codons 40–44, Fig. 1a). Sites associated with putative CD4-binding residues in higher tetrapods were generally conserved in Chondrichthyans but showed alternative residues to those reported in birds and mammals (Fig. 1; Wang et al., 2011; Zhang et al., 2020), supporting previous claims that cartilaginous fish class II genes do not show an obvious tetrapod-like CD4 binding site (Bartl, 2001; Dijkstra et al., 2013).

Additional features associated with intra- and interdomain contact sites typical of classical class II molecules were also found in Chondrichthyan sequences (Fig. 1). Highly conserved residues characteristic of IgSF molecules at Dα29/Eα30 and Rβ93 were conserved across taxa, while β33 was mostly conserved as an Asp except in the elephant shark *C. milii*, stingrays (Myliobatiformes), and some shark species (*H. zebra* and *C. taurus*), where it showed the mammalian condition as Nβ33.

The CP, TM and CYT domains were generally conserved in size (with 1–2 amino acid variations) across sharks, rays and elephant shark, and also exhibited some typical features of vertebrate class II molecules (Fig. 1). In the beta CP, the residue Kβ198, a C-terminal ectodomain binding between α and β chains, is conserved in Elasmobranchs although the elephant shark *C. milii* exhibits a conservative substitution as Rβ198. In the TM domain, all residues involved in helix packing and binding are highly conserved (e.g. α: C195, G198, G202, G205 and G209; and β: G202, G205, G209 and G216), with minor exceptions. In both IIα and IIβ sequences, the CYT of Chondrichthyans differ in length from those of other vertebrates (α:2–7 aa; β:3–16 aa), as well as the CP but to a less extent (1–2 aa difference only).

3.2. Chondrichthyan class II sequences show signs of positive selection and increased diversity at the PBR

When comparing the diversity of MHC class II domains across Chondrichthyan α and β sequences, the TM (average percentage of pairwise amino acid differences: 10 % and 18 %, respectively) and CYT (20 % and 21 %, respectively) were the most conserved, followed by the CP which was more conserved in the IIα than in the IIβ sequences (26 % vs. 41 %). The PBR and IgSF domains were the most variable, with the former exhibiting higher average amino acid differences (alpha: 36 % vs. 32 %, beta: 42 % vs. 27 %). The PBR domains had a higher proportion of non-synonymous substitutions compared to the IgSF domains (average values in PBR vs. IgSF: alpha: 20 % vs. 16 %, beta: 27 % vs. 16 %).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Both the PBR and IgSF domains of Chondrichthyan class IIα and IIβ proteins showed residues under negative selection, although these were most prevalent in the latter (α1: 20 sites, α2: 39 sites, β1: 13 sites, β2: 39 sites). Positions under negative selection generally coincided with important structural residues such as the cysteines involved in disulfide bridges, or the Hα5 residue essential in establishing a salt bridge (Fig. 1). Sites under positive selection were detected only in the PBR domains (α1: 10 sites, β1: 6 sites) and largely coincide with polymorphic residues that are used for peptide binding in mammalian class II molecules (Fig. 1, Table 1).

3.3. Elasmobranchs show remarkably ancient class II lineages

All Chondrichthyan sequences of MHC IIα and IIβ genes formed well-supported monophyletic clades, with classical and nonclassical class II gene sequences from other jawed vertebrates forming separate clusters (Fig. 2). Within Chondrichthyans, MHC IIα and IIβ trees had congruent topologies with sequences clustering into well-supported hierarchical clades consistent with the currently accepted taxonomic arrangement in the group, namely two sub-classes: Holocephali (chimaeras) and Elasmobranchii (sharks and rays), and two sister lineages of Elasmobranchs: Batoidea (rays) and Selachii (sharks) (Supplemental Fig. 1). Furthermore, taxon-based clustering was also observed towards the inner tree branches, with sequences generally clustering according to taxonomic order (but see below for discordances).

Topologies of the NJ trees differed when the PBR and IgSF domain were analyzed independently (Fig. 3 and 4, respectively). The PBR trees had short inner branches (weakly supported) and long outer branches, with well-supported clusters generally coincided with Order-level (i.e. taxon-based) clades. In contrast, the IgSF trees exhibited longer inner branches and shorter outer branches with good support at different levels of the tree, but the major clusters were not fully consistent with taxon-based clades. In this case, sequences from the same species may occur in very divergent clusters not including their closest relatives. This was particularly evident among shark taxa where two divergent and well supported lineages were found for IIα and IIβ genes (Fig. 3 and 4, respectively). Both phylogenetic trees imply that the lineages within each gene are remarkably old considering previous molecular-based divergence time estimates (Heinicke et al., 2009; Fig. 5). The lineages of class IIα and IIβ genes were found across several long-diverged shark orders implying a lineage split in the Selachian ancestor, i.e. after the divergence between sharks and rays and prior to the radiation of sharks between 350–393 Mya (Heinicke et al., 2009). Based on the placement of previously reported nurse shark sequences on the phylogenetic trees, we will refer to the lineages of class IIα genes as DAA and DBA following Kasahara et al. (1993), and of class IIβ genes as DAB and DBB (for lineage 1 of Bartl and Weissman, 1994; and lineage 2 of Bartl, 2001, respectively).

3.4. Class II genes form lineage-specific pairs of linked α and β genes

A consistent feature between the class IIα and IIβ datasets is the unbalanced sequence numbers and unequal taxonomic representation between lineages. One lineage in each gene (αDBA and βDBB) included the vast majority of the shark sequences retrieved, and representatives from all Selachian taxonomic orders present in the dataset. The second

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

lineage (αDAA and βDAB) included comparatively fewer sequences and representatives of only three shark orders (Orectolobiformes - *G. cirratum*, Lamniformes - *C. taurus* and Squaliformes - *S. acanthias* and *E. spinax*; Figs. 3 and 4). The association of lineage-specific pairs of α and β chains was confirmed by PCR amplification using lineage-specific primers, consistent with the previously reported linkage between DAA and DAB genes (Ohta et al., 2000) but now also between DBA and DBB genes (Fig. 6A). Indeed, results from family#1 of nurse sharks showed absence of DBB genes in all of the offspring except in the two pups previously shown to have two DBA genes (Fig. 6B; Ohta et al., 2000), with the same results also found in additional wild sharks.

Insights into lineage-specific genetic diversity and among-lineage genetic divergence were obtained considering sequence alignments of the PBR and IgSF domains of the class IIα and IIβ genes of three shark species exhibiting both lineages per gene (Table 2). In general, αDBA and βDBB showed higher genetic diversity compared to αDAA and βDAB. As expected, this diversity was mostly distributed within the PBR domain while the IgSF was more conserved within lineages. Despite its lower polymorphism levels, among-lineage divergence was more pronounced at the IgSF domain (Table 2; Fig. 3 and 4) where several lineage-specific amino acid differences were found (Fig. 7). The between-lineage amino acid differences changed the hydrophobicity profile of the corresponding proteins, particularly of the beta chain (Fig. 7), suggesting conformational changes in the resulting heterodimers of class II α and β chains. Other fixed amino acid differences found between lineages included the absence of the Hα5 in DAA (replaced by Yα5) and its associated salt bonds, and conservation of alternative residues for β150 between DAB (Nβ150) and DBB (Dβ150) at the conserved minisatellite motif on the β2 domain (codons 140–144).

3.5. Chondrichthyan fish lack DM homologues

Nonclassical class II DM molecules exhibit characteristic residues, such as Nα125 (here Nα111) which directly interacts with the classical class II Wα43 residue (Pos et al., 2012), or the presence of an endosomal sorting motif (i.e. YXXΦ, X can be any residue, Φ denotes a hydrophobic residue) in the β CYT required for transport of DM to the cellular compartments where classical class II proteins are loaded with antigen (Dijkstra et al., 2013). Here, none of the Chondrichthyan class II sequences exhibited the required Nα111 residue, which was generally conserved as Gly (Fig. 1, Supplemental File 3). Likewise, none of the β sequences had CYT with the Tyr residue required for the endosomal sorting motif.

In turn, all Chondrichthyan class II alpha sequences lacked the conserved Wα43 residue (Fig. 1a) and instead showed high variability of amino acids at this site across taxa, none of which was a Trp or had hydrogen-bonding abilities. Other classical class IIα1 residues have been proposed as important for the human HLA-DR and HLA-DM interaction and show some conservation across higher vertebrates (highlighted in Fig. 1a; Dijkstra and Yamaguchi, 2019; Pos et al., 2012; Zhang et al., 2020). However, these residues were either not conserved (in fact, some were under positive selection, e.g. α51; Fig. 1a), or were entirely different in Chondrichthyan class II. Among the variable sites, some exhibited conservative amino acid substitutions (e.g. αV34 M, αV42I) but lacked the required hydrogen-bonding ability, which may compromise DM binding.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4. Discussion

MHC class II sequences were obtained from all Chondrichthyan lineages, i.e. sharks, rays and chimaeras, showing features typical of those described for other vertebrates. These sequences also exhibited highly conserved residues binding peptide main-chain atoms although for alternative amino acids and at fewer residues than those previously described for tetrapods (Table 1; Kaufman et al., 1994; Painter and Stern, 2012), similar to previous reports on classical class II genes of teleosts (Dijkstra et al., 2013). The presence of conserved peptide-binding residues in addition to other features found here and in previous studies indicate that Chondrichthyan class II sequences conform to expectations of classical MHC genes, such as ubiquitous expression across tissues, linkage to the MHC and individual polymorphism. Ubiquitous tissue expression is suggested by the large amount of sequences retrieved from transcriptome data from a variety of Chondrichthyan taxa and tissues (including brain, heart, kidney, spleen, pancreas, liver, gonads, skin, eye; Supplemental Table 1). This pattern is consistent with studies showing MHC class II α and β genes expression in several tissues of the nurse shark *G. cirratum* and of the whitespotted bamboo shark *Chiloscyllium plagiosum* (Criscitiello et al., 2012; Ma et al., 2013; Ohta et al., 2004). It is likely, but not proven, that the ubiquitous class II expression in non-lymphoid tissues is due to the infiltration of hematopoietic cells like lymphocytes and antigen-presenting cells. Previous studies showed very high expression of class II on splenic dendritic cells (Rumfelt et al., 2004). Previous work has also shown linkage between MHC class I and II genes in the nurse shark *G. cirratum* (Ohta et al., 2000), and the same linkage was found in the recently assembled genome of the bamboo shark *C. plagiosum* (Zhang et al., 2019). Data on the polymorphism levels of Chondrichthyan MHC II genes is still limited, but we found several sequences for both alpha and beta genes within the same individual (e.g. 3–11 sequences per lineage of alpha or beta genes per individual; Supplemental Files 1 and 2) suggesting the presence of multiple alleles and gene copies. These observations are in line with studies of a few shark species showing copy number variation and allelic diversity within species at MHC class II genes (e.g. IIα Kasahara et al., 1993; IIβ Ma et al., 2013). Moreover, we also found higher genetic diversity and residues under positive selection in the PBR of IIα and IIβ genes, in contrast to the more conserved IgSF domain, as expected in classical class II genes (Radwan et al., 2020).

Chondrichthyan sequences comprised a monophyletic group of class II genes clustering separately from those of bony fishes, amphibians, birds/reptiles, and mammals (both classical and nonclassical). Within Selachians (i.e. sharks), two distinct and well supported lineages of class IIα and IIβ genes were found based on sequence divergence at the IgSF domain. Most notably, the class II gene trees imply a remarkably old age of lineage divergence, with the split of class IIα and class IIβ lineages occurring earlier than 350 MYA (sensu Heinicke et al., 2009), i.e. prior to the major shark radiation (Fig. 5). Such an old age of classical class II lineages is unprecedented among vertebrates and highlights not only the long evolutionary history of cartilaginous fish but likely their slow evolution rates (Martin, 1999; Martin and Palumbi, 1993; Renz et al., 2013). However, dedicated molecular phylogenetic analyses including higher taxonomic coverage are advised to ascertain the evolutionary history of MHC class II genes in Elasmobranchs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Phylogenetic trees based on the PBR did not show similar lineage splits and, instead, exhibited a taxon-based sequence clustering. Despite its higher genetic diversity compared to the IgSF domain, the PBR showed lower between-lineage divergence. Such contrasting topologies between the two functionally distinct extracellular domains are in line with expectations of different selective pressures and recombination levels: the PBR may be under selection by taxon-specific pathogen communities and may evolve to better fit with the species' antigen repertoires, including a higher recombination rate (via gene conversion), while differences at IgSF domain may be more limited given its conserved function. Indeed, the IgSF domain has been shown to provide a better evolutionary signal among MHC class IIβ genes compared to the PBR (e.g. in birds; Burri et al., 2010). Although similar functions and selective pressures may be expected for class I genes, lineage distinction in this case is generally based on the genetic divergence levels and distinct structural features of the PBR, while the IgSF domain shows among-lineage conversion (Almeida et al., 2020; Flajnik et al., 1993; Ohta et al., 2019) perhaps due to the close interaction with the conserved β2-microglobulin.

Previous studies in the nurse shark *G. cirratum* showed linkage between IIα and IIβ genes, as well as linkage to class I genes (Ohta et al., 2000). These results were based on Southern blot analyses of family#1 (used here) using full cDNA probes of DAA and DAB at high stringency, and also showed the presence of paternal DBA "alleles" in two siblings (pups 10 and 13). Here we build on previous data to show that IIα and IIβ genes are found as lineage-specific pairs, namely the presence of DAA genes was exclusively associated with the presence of DAB genes, while the same is true for DBA and DBB genes. In line with previous observations of linked DAA and DAB genes, our results also suggest linkage between DBA and DBB genes. Based on the current results, it is expected that (at least) sharks may exhibit one set of class II genes while others may exhibit an extra set (Fig. 6). Genes in the two lineages segregate as alleles despite belonging to separate loci (Fig. 6B; Kasahara et al., 1993; Ohta et al., 2000). Given the stable nature of Chondrichthyan genomes and evidence of an MHC region linking genes in class I, class II and class III regions (Ohta et al., 2011, 2002, 2000; Terado et al., 2003), coupled with reports of specific sets of genes organized into MHC haplotypes (this study; Ohta et al., 2002), it may be hypothesized that organization into haplotypes may extend to the whole MHC region.

In nurse sharks, DAA and DAB genes were found in all individuals analyzed here and appear to be most common gene set, in line with previous reports (Kasahara et al., 1993; Ohta et al., 2000). However, this may not be true across taxa as our results suggest that the most common gene set refers to DBA and DBB genes whose sequences were predominantly detected for the vast majority of the shark taxa analyzed (Fig. 3 and 4). Additional confirmation on genomic DNA are needed to ascertain if both lineages (and gene sets) are also present across shark taxa. In fact, one important caveat of the present dataset is its reliance on genes being expressed in high enough amounts to be detected by non-targeted RNAseq projects (such as those used here); such an approach may limit detection of non-expressed/underexpressed genes that may provide a distinct perspective of the MHC class II evolution in Elasmobranchs. In fact, preliminary data from gene expression analysis on the bull shark *C. taurus*, for which sequences from all IIα and IIβ lineages were retrieved,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

showed higher relative expression of DBA and DBB genes compared to DAA and DAB genes (A. Veríssimo, unpublished data).

The particular genomic arrangement of class II genes in Elasmobranchs and their segregation as alleles (pseudoalleles) may perhaps explain the absence of between-lineage gene conversion of the IgSF domain, in contrast to that found in class I genes (see details above). In turn, given the linkage between specific lineages of IIα and IIβ genes, we think it is likely that each has co-evolved to better interact with each other. Similarly, Burri et al. (2010) report on a gene duplication event of IIβ genes prior to major avian radiations in which positive selection lead to changes in residues associated with IIα and IIβ chain interaction, suggesting co-evolution of the two genes. Given that between-chain interaction is mostly centered on the α2 and β2 domains (Bondinas et al., 2007; Brown et al., 1993), it may be expected that between-lineage differences may be found predominantly in the IgSF domains, as shown here (Table 2).

Our search identified neither the nonclassical class II gene *DM* from the sequences retrieved from the Chondrichthyan "omic" datasets, nor did the classical class II proteins display the typical residues that interact with DM in mammals (*sensu* Dijkstra et al., 2013). The absence of sequences showing classical MHC class II features required for DM binding coupled to the lack of typical DM sequences in the present Chondrichthyan dataset are consistent with the absence of a DM-regulating system in basal jawed vertebrates. These results are reinforced by the fact that, although we specifically searched for MHC class II sequences in the cartilaginous "*omics*" datasets, other divergent sequences were also retrieved such as class I and immunoglobulin sequences. Based on our search strategy, it is likely that no other class II genes are present in cartilaginous fish genomes aside from the classical sequences described above. The evidence gathered so far in this and previous studies indicates that the existence of DM is a derived feature of the lobe-finned fish lineage leading to tetrapods (Dijkstra and Yamaguchi, 2019). Dijkstra et al. (2013) proposed that the critical residues involved in DM function, i.e. Wα43 and Nα125, may have evolved early in the lobe-finned fish lineage as suggested by their simultaneous presence in classical class II sequences of the coelacanth, a basal lobe-fin fish. On the other hand, the Holocephalan *C. milii* differed from Elasmobranchs by exhibiting Cys at α13 and α66, described as being important to form disulfide bonds in the nonclassical DM (van Lith and Benham, 2006). However, these Cys are present not only in DM molecules but also in the DA, DB and DE lineages of teleosts and in the classical class IIα of the coelacanth (Dijkstra et al., 2013; Dijkstra and Yamaguchi, 2019) and may represent an ancestral trait lost in most vertebrates.

It remains to be clarified how peptide loading onto classical MHC class II proteins occurs without a DM molecule, a condition now observed in cartilaginous and bony fish. Several alternative mechanisms assisting peptide-CLIP exchange without DM have been proposed, such as the presence of Ii isoforms with a thyroglobulin domain, reported in basal jawed vertebrates, that may facilitate CLIP dissociation from and peptide binding to MHC class II molecules (Criscitiello et al., 2012); expression of classical class II molecules that bind CLIP with low affinity at low pH or existence of other types of nonclassical class II molecules that may interact with class II proteins in a similar way to the DM (Dijkstra et al., 2013). Along this line, recent studies have shown that small molecules may interact with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

MHC class II in the endosomal compartment by enhancing CLIP dissociation and peptide binding (reviewed in Call, 2011); these may serve as a simpler alternative to a more complex DM-based system. A final possibility is that fish class II molecules themselves might act in a DM-like fashion as ‘dimers of dimers’ that aid each other in peptide exchange. If this scenario is possible, then it would make sense that the catalyst for peptide exchange in tetrapods is a class II duplicate (DM) rather than any other enzyme found in endosomes.

5. Conclusions

We provide the first comprehensive overview of MHC class II diversity in the most basal jawed vertebrate lineage, a key taxon to infer ancestral traits in adaptive immunity. All shark, ray and chimaera species analyzed here showed MHC class II sequences exhibiting features reported as conserved across vertebrates, and consistent with classical class II molecules. Notably, our data showed ancient lineage diversification of IIα and IIβ genes in Elasmobranchs in addition to linkage between lineage-specific pairs of IIα and IIβ genes. Finally, our results also support previous claims on the absence of a DM system in basal jawed vertebrates and instead support its late emergence in the lineage leading to tetrapods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This work was supported by Portuguese funds through Portuguese Foundation for Science and Technology Grant PD/BD/114542/2016 (to T.A.) and Contracts IF/ 00376/2015 (to P.J.E.) and DL57/2016 (to A. V.); Y.O. and M.F.F. were supported by National Institutes of Health GrantsAI140-326-27 and AI02877; and AMM was supported by the project GenomePT (POCI/010145/FEDER/022184). This work is also financed by FEDER Funds through the Operational Competitiveness Factors Program COMPETE and by national funds through the Foundation for Science and Technology within the scope of Project PTDC/ASP-PES/28053/2017, which supported also A.G. and F.N.

Abbreviations:

MHC	Major Histocompatibility Complex
PBR	peptide binding region
IgSF	immunoglobulin superfamily
CLIP	class II associated invariant chain peptide
Ii	invariant chain
CP	connecting piece
TM	transmembrane domain
CYT	cytoplasmic tail
NJ	neighbor-joining

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

References

Almeida T, Esteves PJ, Flajnik MF, Ohta Y, Veríssimo A, 2020. An ancient, MHC-Linked, nonclassical class I lineage in cartilaginous fish. J. Immunol 204, 892–902. 10.4049/jimmunol.1901025. [PubMed: 31932500]

Aschliman NC, 2011. The Batoid Tree Of Life: Recovering The Patterns And Timing Of The Evolution Of Skate, Rays And Allies (Chondrichthyes: Batoidea). Thesis, p. 198.

Bartl S, 2001. New Major Histocompatibility Complex Class IIB Genes From Nurse Shark. Advances in Experimental Medicine and Biology, pp. 1–11. 10.1007/978-1-4615-1291-2_1.

Bartl S, Weissman IL, 1994. Isolation and characterization of major histocompatibility complex class IIB genes from the nurse shark. Proc. Natl. Acad. Sci. U. S. A 91, 262–266. [PubMed: 8278377]

Bondinas GP, Moustakas AK, Papadopoulos GK, 2007. The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. Immunogenetics 59, 539–553. 10.1007/s00251-007-0224-8. [PubMed: 17497145]

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC, 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364, 33–39. 10.1038/364033a0. [PubMed: 8316295]

Burri R, Salamin N, Studer RA, Roulin A, Fumagalli L, 2010. Adaptive divergence of ancient gene duplicates in the avian MHC class II β. Mol. Biol. Evol 27, 2360–2374. 10.1093/molbev/msq120. [PubMed: 20463048]

Call MJ, 2011. Small molecule modulators of MHC class II antigen presentation: mechanistic insights and implications for therapeutic application. Mol. Immunol 10.1016/j.molimm.2011.05.022.

Carrier JC, Musick JA, Heithaus MR (Eds.), 2010. Sharks and Their Relatives II - Biodiversity, Adaptive Physiology, and Conservation, first edit. CRC Press, Boca Raton, FL, U.S.A.

Criscitiello MF, Ohta Y, Graham MD, Eubanks JO, Chen PL, Flajnik MF, 2012. Shark class II invariant chain reveals ancient conserved relationships with cathepsins and MHC class II. Dev. Comp. Immunol 36, 521–533. 10.1016/J.DCI.2011.09.008. [PubMed: 21996610]

Denzin LK, Sant'Angelo DB, Hammond C, Surman MJ, Cresswell P, 1997. Negative regulation by HLA-DO of MHC class II-Restricted antigen processing. Science (80-) 278, 106–109. 10.1126/science.278.5335.106.

Dijkstra JM, Yamaguchi T, 2019. Ancient features of the MHC class II presentation pathway, and a model for the possible origin of MHC molecules. Immunogenetics 71, 233–249. 10.1007/s00251-018-1090-2. [PubMed: 30377750]

Dijkstra JM, Grimholt U, Leong J, Koop BF, 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 Evol. Biol 13, 260. 10.1186/1471-2148-13-260. [PubMed: 24279922]

Dijkstra JM, Yamaguchi T, 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. 10.1007/s00251-017-1050-2. [PubMed: 29270774]

Edgar RC, 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. 10.1093/nar/gkh340. [PubMed: 15034147]

Flajnik MF, Canel C, Kramer J, Kasahara M, 1991. Evolution of the major histocompatibility complex: molecular cloning of major histocompatibility complex class I from the amphibian Xenopus. Proc Nati Acad Sci USA 88, 537–541.

Flajnik MF, Kasahara M, Shum BP, Salter-Cid L, Taylor E, Du Pasquier L, 1993. A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian Xenopus. EMBO J. 12, 4385–4396. [PubMed: 8223448]

Fling SP, Arp B, Pious D, 1994. HLA-DMA and -DMB genes are both required for MHC class II/peptide complex formation in antigen-presenting cells. Nature 368, 554–558. 10.1038/368554a0. [PubMed: 8139690]

Heinicke MP, Naylor GJP, Hedges SB, 2009. Cartilaginous Fishes (Chondrichthyes), the Timetree of Life. Oxford University Press, Oxford, United Kingdom.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Kasahara M, Vazquez M, Sato K, McKinney EC, 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. 10.1073/pnas.89.15.6688. [PubMed: 1495958]

Kaufman J, 1999. Co-evolving genes in MHC haplotypes: the “rule” for nonmammalian vertebrates? Immunogenetics. 10.1007/s002510050597.

Kaufman J, Salomonsen J, Flajnik M, 1994. Evolutionary conservation of MHC class I and class II molecules—different yet the same. Semin. Immunol 6, 411–424. 10.1006/SMIM.1994.1050. [PubMed: 7654997]

Kiemnec-Tyburczy KM, Richmond JQ, Savage AE, Zamudio KR, 2010. Selection, trans-species polymorphism, and locus identification of major histocompatibility complex class IIβ alleles of New World ranid frogs. Immunogenetics 62, 741–751. 10.1007/s00251-010-0476-6. [PubMed: 20844870]

Klein J, Sato A, 1998. Birth of the major histocompatibility complex. Scand. J. Immunol 47, 199–209. 10.1046/j.1365-3083.1998.00292.x. [PubMed: 9519857]

Klein J, Sato A, 2000. The Hla System. N. Engl. J. Med 343, 702–709. 10.1111/j.1755-3768.1984.tb03060.x. [PubMed: 10974135]

Kropshofer Harald, Vogt Anne B., Thery Clotilde, Armandola Elena A., Li B-C, Moldenhauer G, Amigorena S, Nter G, Hä J, Kropshofer H, Vogt AB, Thery C, Armandola EA, 1998. A role for HLA-DO as a co-chaperone of HLA-DM in peptide loading of MHC class II molecules. EMBO J. 17, 2971–2981. 10.1093/emboj/17.11.2971. [PubMed: 9606180]

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. 10.1093/molbev/msy096. [PubMed: 29722887]

Lazarski CA, Chaves FA, Sant AJ, 2006. The impact of DM on MHC class II–restricted antigen presentation can be altered by manipulation of MHC–peptide kinetic stability. J. Exp. Med 203, 1319–1328. 10.1084/jem.20060058. [PubMed: 16682499]

Ma Q, Su YQ, Wang J, Zhuang ZM, Tang QS, 2013. Molecular cloning and expression analysis of major histocompatibility complex class IIB gene of the Whitespotted bambooshark (Chiloscyllium plagiosum). Fish Physiol. Biochem 39, 131–142. 10.1007/s10695-012-9685-2. [PubMed: 22752338]

Martin AP, 1999. Substitution rates of organelle and nuclear genes in sharks: implicating metabolic rate (again). Mol. Biol. Evol 16, 996–1002. 10.1093/oxfordjournals.molbev.a026189. [PubMed: 10406116]

Martin AP, Palumbi SR, 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. U. S. A 90, 4087–4091. 10.1073/pnas.90.9.4087. [PubMed: 8483925]

Naylor GJP, Caira JN, Jensen K, Rosana KAM, White WT, Last PR, 2012. A DNA sequence–Based approach to the identification of shark and ray species and its implications for global elasmobranch diversity and parasitology. Bull Am Museum Nat Hist 367, 1–262. 10.1206/754.1.

O’hUigin C, Sültmann H, Tichy H, Murray BW, 1998. Isolation of mhc class II DMA and DMB cDNA sequences in a marsupial: the gray short-tailed opossum (Monodelphis domestica). J. Mol. Evol 47, 578–585. 10.1007/PL00006414. [PubMed: 9797408]

Ohta Y, Okamura K, McKinney EC, Bartl S, Hashimoto K, 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. 10.1073/PNAS.97.9.4712. [PubMed: 10781076]

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. [PubMed: 11777971]

Ohta Y, Landis E, Boulay T, Phillips RB, Collet B, Secombes CJ, Flajnik MF, Hansen JD, 2004. Homologs of CD83 from Elasmobranch and teleost fish. J. Immunol 173, 4553–4560. 10.4049/jimmunol.173.7.4553. [PubMed: 15383588]

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. 10.4049/JIMMUNOL.176.6.3674. [PubMed: 16517736]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Ohta Y, Shiina T, Lohr RL, Hosomichi K, Pollin TI, Heist EJ, Suzuki S, Inoko H, Flajnik MF, 2011. Primordial linkage of β2-microglobulin to the MHC. J. Immunol 186, 3563–3571. 10.4049/jimmunol.1003933. [PubMed: 21321107]

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, ji1900597 10.4049/jimmunol.1900597.

Painter CA, Stern LJ, 2012. Conformational variation in structures of classical and non-classical MHCII proteins and functional implications. Immunol. Rev 250, 144–157. 10.1111/imr.12003. [PubMed: 23046127]

Parker A, Kaufman J, 2017. What chickens might tell us about the MHC class II system. Curr. Opin. Immunol 10.1016/j.coi.2017.03.013.

Pond SLK, Frost SDW, Muse SV, 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679. 10.1093/bioinformatics/bti079. [PubMed: 15509596]

Pos W, Sethi DK, Call MJ, Schulze MSED, Anders AK, Pyrdol J, Wucherpfennig KW, 2012. Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. Cell 151, 1557–1568. 10.1016/j.cell.2012.11.025. [PubMed: 23260142]

Radwan J, Babik W, Kaufman J, Lenz TL, Winternitz J, 2020. Advances in the evolutionary understanding of MHC polymorphism. Trends Genet. 10.1016/j.tig.2020.01.008.

Ravi V, Venkatesh B, 2018. The divergent genomes of teleosts. Annu. Rev. Anim. Biosci 6, 47–68. 10.1146/annurev-animal-030117-014821. [PubMed: 29447475]

Renz AJ, Meyer A, Kuraku S, 2013. Revealing less derived nature of cartilaginous fish genomes with their evolutionary time scale inferred with nuclear genes. PLoS One 8, e66400. 10.1371/journal.pone.0066400. [PubMed: 23825540]

Rumfelt LL, Diaz M, Lohr RL, Mochon E, Flajnik MF, 2004. Unprecedented multiplicity of ig transmembrane and secretory mRNA forms in the cartilaginous fish. J. Immunol 173, 1129–1139. 10.4049/jimmunol.173.2.1129. [PubMed: 15240702]

Schulze M-SE, Wucherpfennig KW, 2012. The mechanism of HLA-DM induced peptide exchange in the MHC class II antigen presentation pathway. Curr. Opin. Immunol 24, 105–111. 10.1016/J.COI.2011.11.004. [PubMed: 22138314]

Smith SL, Sim RB, Flajnik MF, 2014. Immunobiology of the Shark, first edit. CRC Press, Boca Raton, FL, U.S.A.

Terado T, Okamura K, Ohta Y, Shin D-H, 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. 10.4049/JIMMUNOL.171.5.2461. [PubMed: 12928394]

van Lith M, Benham AM, 2006. The DMα and DMβ chain cooperate in the oxidation and folding of HLA-DM. J. Immunol 177, 5430–5439. 10.4049/jimmunol.177.8.5430. [PubMed: 17015729]

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 B-H, 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. 10.1038/nature12826. [PubMed: 24402279]

Wang JH, Meijers R, Xiong Y, Liu JH, Sakihama T, Zhang R, Joachimiak A, Reinherz EL, 2001. Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. Proc. Natl. Acad. Sci. U. S. A 98, 10799–10804. 10.1073/pnas.191124098. [PubMed: 11535811]

Wang XX, Li Y, Yin Y, Mo M, Wang Q, Gao W, Wang L, Mariuzza RA, 2011. Affinity maturation of human CD4 by yeast surface display and crystal structure of a CD4-HLA-DR1 complex. Proc. Natl. Acad. Sci. U. S. A 108, 15960–15965. 10.1073/pnas.1109438108. [PubMed: 21900604]

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL, 2018. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. Mol. Biol. Evol 35, 773–777. 10.1093/molbev/msx335. [PubMed: 29301006]

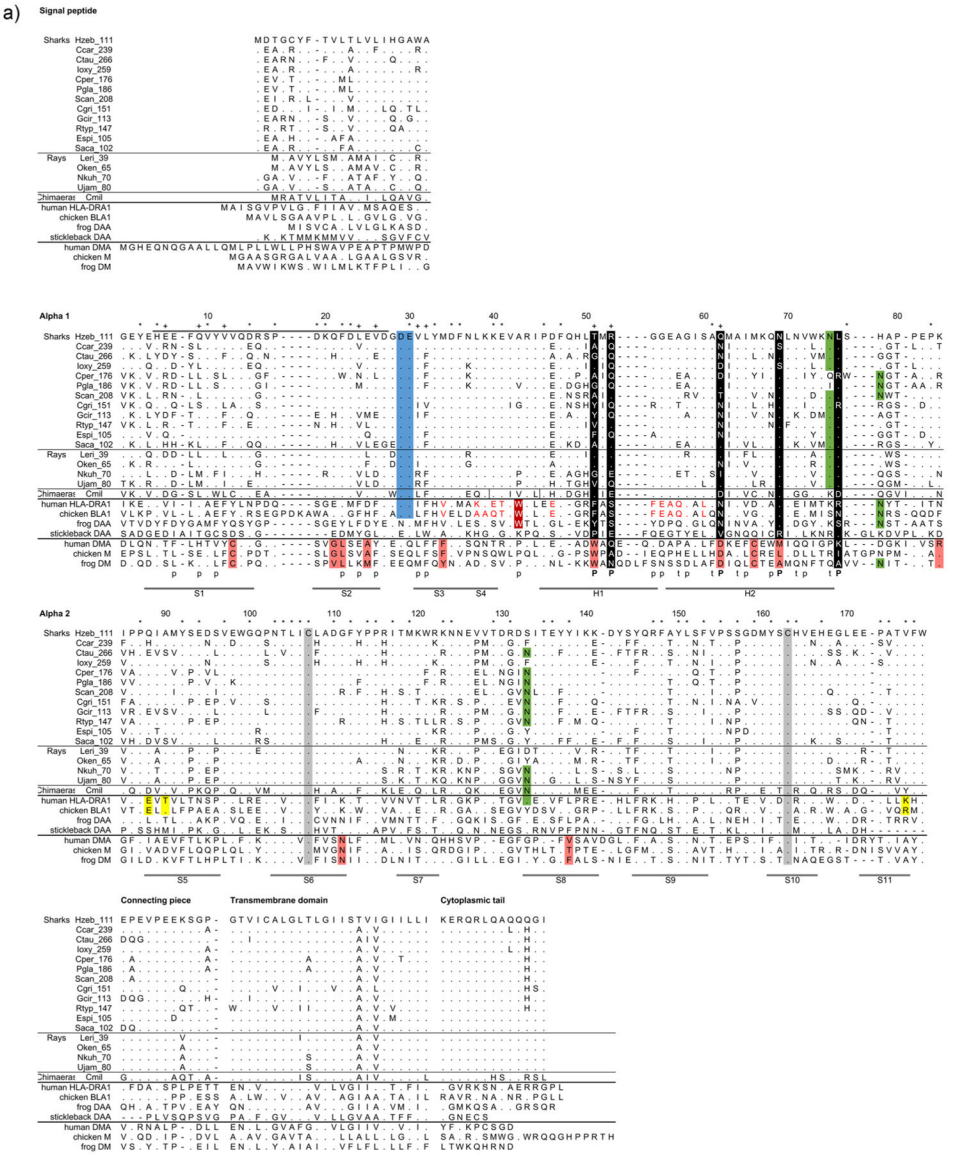
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Yeager M, Hughes AL, 1999. Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. Immunol. Rev 167, 45–58. 10.1111/j.1600-065X.1999.tb01381.x. [PubMed: 10319250]

Zhang Y, Gao H, Li H, Guo J, Wang M, Xu Q, Wang Jiahao, Lv M, Guo X, Liu Q, Wei L, Ren H, Xi Y, Guo Y, Zhao Q, Pan S, Liu C, Sang L, Ding X, Wang C, Xiang H, Song Y, Liu Y, Liu Shanshan, Jiang Y, Shao C, Shi J, Liu Shiping, Sabir JSM, Sabir MJ, Khan M, Hajrah NH, Lee SM-Y, Xu X, Yang H, Wang Jian, Fan G, Yang N, Liu X, 2019. Dynamic chromosome rearrangements of the white-spotted bamboo shark shed light on cartilaginous fish diversification mechanisms. bioRxiv 602136. 10.1101/602136.

Zhang L, Li X, Ma L, Zhang B, Meng G, Xia C, 2020. A newly recognized pairing mechanism of the α- and β-chains of the chicken peptide–MHC class II complex. J. Immunol 204, 1630–1640. 10.4049/jimmunol.1901305. [PubMed: 32034060]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

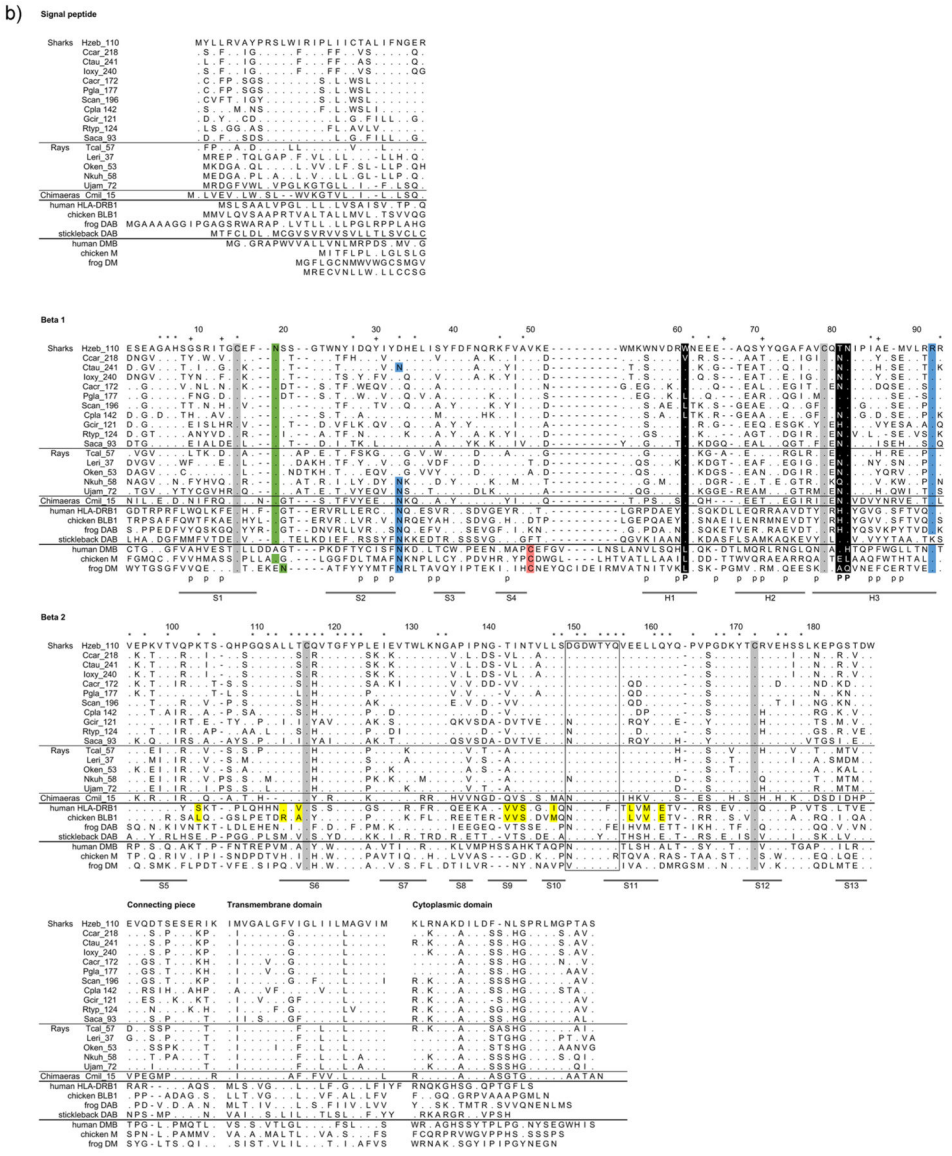


Fig. 1.

Amino acid alignment of full protein of MHC class II genes. Alignments for a) alpha and b) beta genes, for representative Chondrichthyan taxa (one representative sequence per genus) including sharks (Selachii), rays (Batoidea) and chimaeras (Holocephali), as well as other vertebrate taxa including classical and nonclassical (DM) class II sequences. Taxon names for Chondrichthyan representatives are shown as the first letter for genus name (in capital letter) followed by the first three letters of the species name, and a number for cross referencing with the full nucleotide sequences in Supplemental Files 1–2 (alpha sequences) and 3–4 (beta sequences), except for Cmil alpha sequence (GenBank Accession JW875535.1). Abbreviation and full species name are indicated in Supplemental Table 2. GenBank Accessions for non-Chondrichthyan sequences are as follows: human HLA-DRA1 AAA59785.1, HLA-DRB1 BAO73158.1, DMA CAA54169.1, DMB ARB08367.1; chicken BLA1 AAY40298.1, BLB1 BAF62996.1, M alpha CAA18966, M beta BAG69311; frog

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DAA AAL58430.1, DAB BAA02842.1, DM alpha AAH61681, DM beta XP 002942059; stickleback DAA AAU01917.1, DAB AAU01918.1. The salt bonds in class IIα sequences are highlighted by double bars. Residues highlighted in color represent putative N-glycosylation sites (green); disulfide bonds (grey); Ig-like domains (blue); putative residues involved in CD4-binding sites (yellow; Wang et al., 2011; Zhang et al., 2020); conserved peptide binding residues (black); DM specific residues (red; Dijkstra et al., 2019); residues in red font are important for classical class II α chain interaction with DM (Zhang et al., 2020). Predicted locations of β-strands (S1–11 in alpha, and S1–13 in beta) and α-helices (H-2 in alpha, and H1–3 in beta) are shown as double lines below alignments. Numbers above the alignment refer to positions in the HLA-DRA1 and HLA-DRB1, respectively. Residues in Chondrichthyan sequences under positive (+) or negative (*) selection are shown below alignment positions. Residues assumed to interact with peptides (p or P), or with the TCR (t) are indicated below the alignment (Kasahara et al., 1992).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

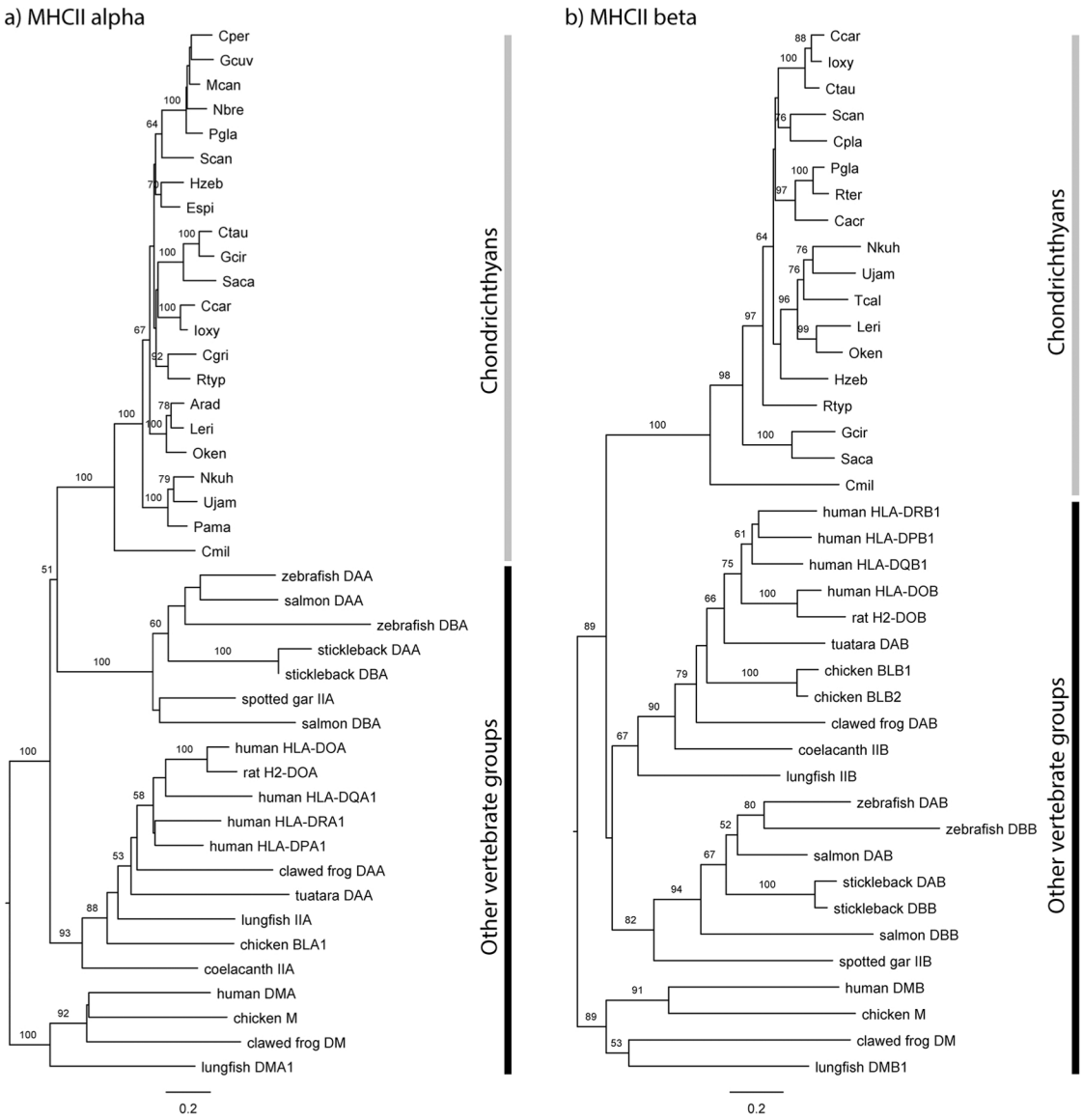


Fig. 2.

Phylogenetic trees of MHC class II genes. Trees for a) alpha and b) beta genes are based on amino acid sequences from the PBR and IgSF domains (exons 2 and 3, respectively) retrieved for Chondrichthyan taxa (names listed on Fig. 1 legend, amino acid sequences provided in Supplemental File 2 and 4 for alpha and beta, respectively), and including other vertebrate representatives. GenBank Accessions for non-Chondrichthyan sequences are as follows: human HLA-DRA1 AAA59785.1, HLA-DRB1 BAO73158.1, HLA-DQA1 AAA59760.1, HLA-DQB1 AAA59770.1, HLA-DPA1 AAH09956.1, HLA-DPB1 AAA59837.1; DOA P06340, DOB P13765, DMA CAA54169.1, DMB ARB08367.1; rat H2-DOA NP 898874, H2-DOB NP 001008846; chicken BLA1 AAY40298.1, BLB1 BAF62996.1, BLB2 BAF62998.1, M alpha CAA18966, M beta BAG69311; tuatara (lizard) DAA AKG62148.1, DAB AAZ77712.1; clawed frog DAA AAL58430.1, DAB BAA02842.1, DM alpha AAH61681, DM beta XP 002942059; coelacanth IIA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

XP006014228.1, IIB XP006010591.1; lungfish IIA GEHZ01055957.1, IIB GEHZ01012825.1; DMA1 GEHZ01056980.1, DMB1 GEHZ01000527.1; stickleback DAA AAU01917.1, DBA AAU01919.1, DAB AAU01918.1, DBB AAU01920.1; spotted gar IIA 501A2 JH591501, IIB 501B1 JH591501; salmon DAA AAL40122.1, DBA ABX44764.1, DAB CAA49725.1, DBB ABX44766.1; zebrafish DAA AAA72019, DBA D8.45A2, DAB AAA50043, DBB AAA87894. Bootstrap support values (10000 replicates) are shown only if above 50 %.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

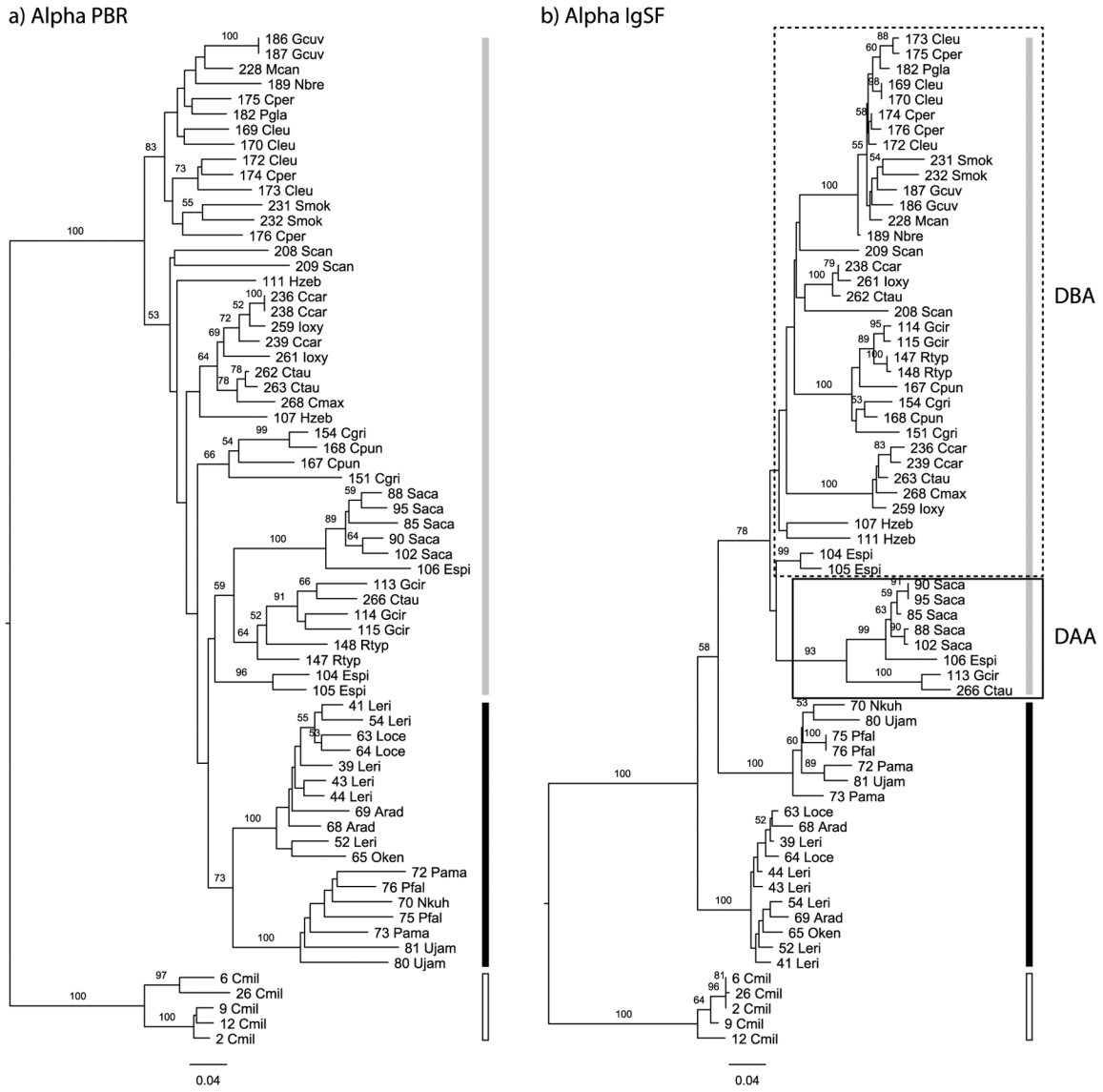


Fig. 3.

Phylogenetic trees of MHC class IIα genes. Trees are based on a) the PBR domain (α1, exon 2) and b) the IgSF domain (α2, exon 3), retrieved for Chondrichthyans. Lineages based on IgSF domain differences are framed by straight lines (*DAA*) and dotted lines (*DBA*). Major Chondrichthyan taxonomic groups are highlighted by colored bars: white – Holocephalans (chimaeras), grey – Selachians (sharks), black – Batoids (rays). Taxon names follow description in Fig. 1 legend (full names listed in Supplemental Table 2A), preceded by a reference number to allow direct comparison between the PBR and IgSF trees. Chondrichthyan nucleotide sequences are provided in Supplemental File 1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

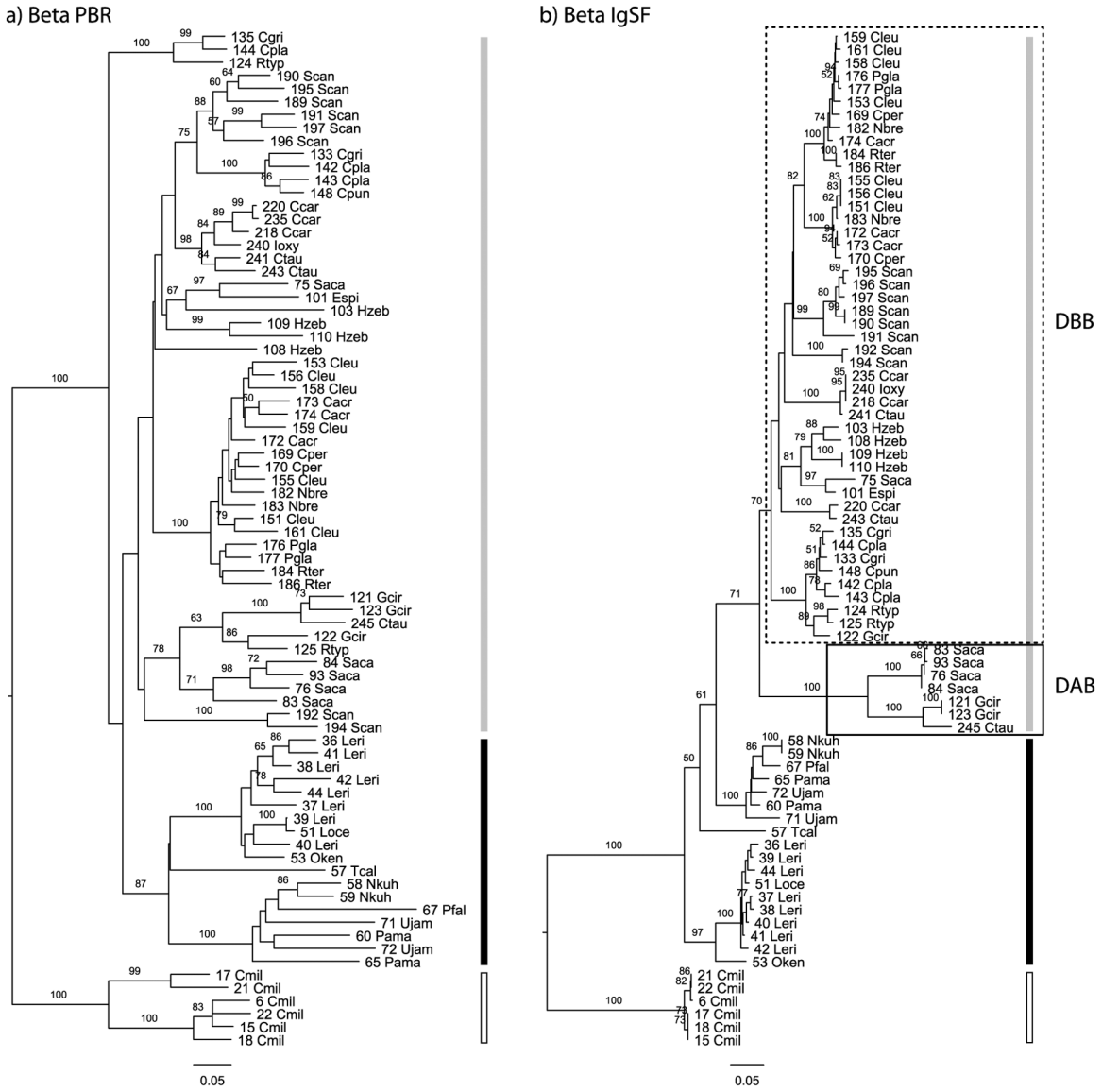


Fig. 4.

Phylogenetic trees of MHC class IIβ genes. Trees are based on a) the PBR domain (β1, exon 2) and b) the IgSF domain (β2, exon 3), retrieved for Chondrichthyans. Lineages based on IgSF domain differences are framed by straight lines (DAB) and dotted lines (DBB). Major Chondrichthyan taxonomic groups are highlighted by colored bars: white – Holocephalans (chimaeras), grey – Selachians (sharks), black – Batoids (rays). Taxon names follow description in Fig. 1 legend (full names listed in Supplemental Table 2B), preceded by a reference number to allow direct comparison between the PBR and IgSF trees. Chondrichthyan nucleotide sequences are provided in Supplemental File 3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

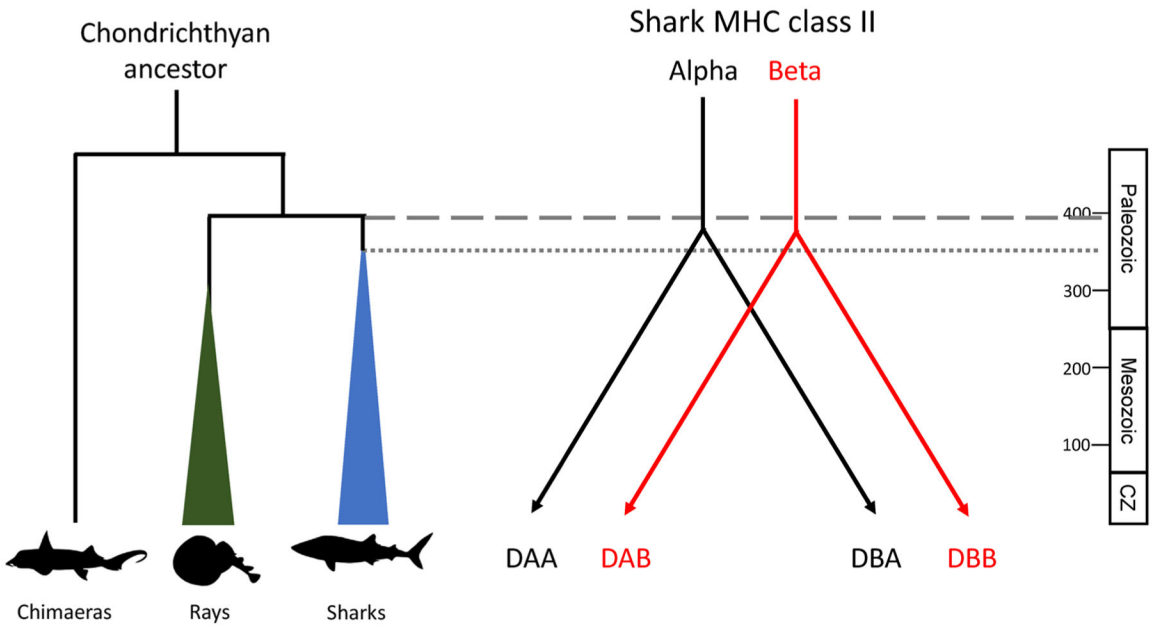


Fig. 5.

Schematic representation of MHC class II lineage split in Chondrichthyans and the hypothetical timings of gene divergence. Dotted grey line indicates divergence of class IIα and IIβ lineages predating the radiation within sharks (represented as a blue triangle), and dashed grey line indicates the split between sharks and rays (radiation within rays represented as a green triangle). Time tree of Chondrichthyan evolution adapted from Heinicke et al. (2009).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

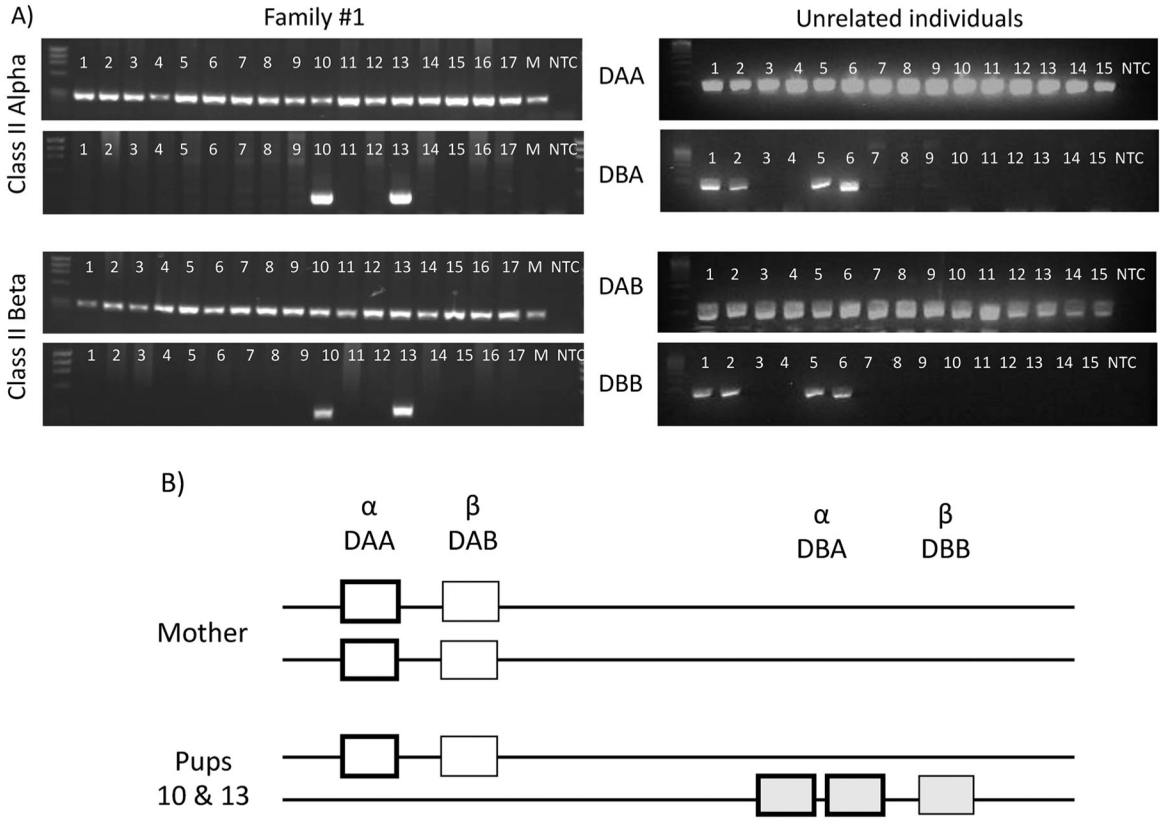


Fig. 6.

Association of MHC class IIα and IIβ lineages into haplotypes in the nurse shark *G. cirratum*. A) Lineage-specific PCRs in family #1 from Ohta et al. (2000) (numbers refer to pups; M: mother), and in 15 unrelated individuals except no. 1 and 2 which are siblings from family#4 (mother was positive for all lineages; Y. Ohta, unpublished data). Lineage-specific PCR amplification was performed on genomic B) Schematic representation of class IIα and IIβ genomic organization in nurse sharks, with linkage between specific sets of IIα (boxes with thick black lines) and IIβ genes (boxes with light black lines). Two DBA genes were previously found in nurse sharks (Kasahara et al., 1993; Ohta et al., 2000) but copy number may vary in other species. Presence of DAA and DAB genes (shown for the Mother of family#1) was found across all individuals analyzed, while presence of DBA and DBB (shown for pups 10 & 13) was less common. NTC - no template control.

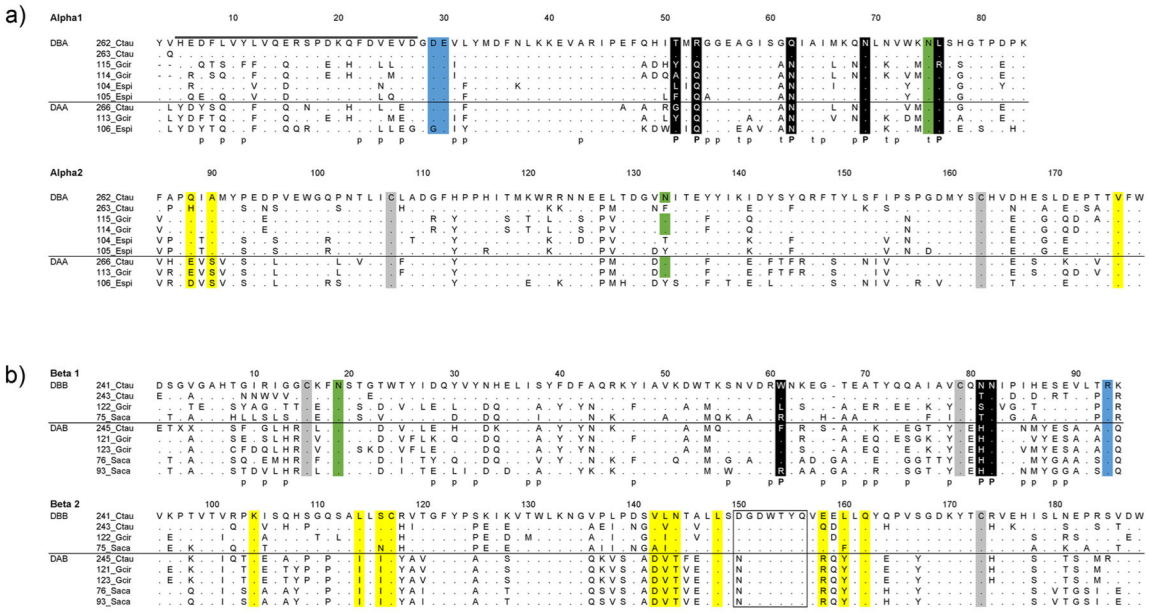


Fig. 7.

Amino acid alignments of the PBR (exon 2) and IgSF (exon 3) domains highlighting the differences between lineages. Alignments are shown for MHC a) IIα and b) IIβ genes for a subset of Elasmobranch taxa. Taxon names follow description in Fig. 1 legend (full names listed in Supplemental Table 2A and 2B for alpha and beta, respectively), preceded by a reference number to allow direct comparison between the PBR and IgSF domain alignments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Mol Immunol. Author manuscript; available in PMC 2021 March 31.

Table 1

Peptide binding residues for MHC class II molecules, with conserved residues responsible for binding to mainchain atoms highlighted in grey shade (letter in parentheses shows conserved residue found in tetrapods, following Kaufman et al., 1994). Residues under positive selection are marked with + on top of their corresponding positions in the alignment. Taxon names for Chondrichthyan representatives are shown as the first letter for genus name (in capital letter) followed by the first three letters of the species name. Abbreviation and full species name indicated in Supplemental Table 2.

											α1									
Taxon		+			+	+	+	+		+	+				+					
	7	**9**	**11**	**22**	**24**	**26**	**31**	**32**	**43**	**51(F)**	**53(S)**	**54**	**55**	**58**	**62(N)**	**65**	**68**	**69(N)**	**72**	**76(R)**
Sharks																				
Hzeb	E	Q	Y	F	L	V	V	L	A	T	R	G	G	G	Q	Q	Q	N	V	L
Ccar	N	L	Y	F	V	V	V	L	A	A	Q	G	G	G	N	Q	Q	S	V	L
Ctau	Y	Q	Y	F	L	E	V	F	A	G	Q	G	G	G	N	Q	N	N	V	L
Ioxy	D	L	Y	F	V	V	V	F	A	T	Q	G	G	G	D	Q	Q	S	V	L
Cper	D	L	S	W	N	L	V	F	A	A	Q	G	G	A	D	Q	Q	N	I	R
Pgla	D	L	S	F	M	V	V	F	V	G	Q	G	G	A	Q	Q	Q	N	V	L
Scan	N	L	Y	F	L	V	V	L	A	A	R	G	G	R	T	Q	N	N	V	L
Cgri	Q	S	Y	F	L	V	I	V	G	Y	Q	R	G	G	N	Q	H	N	I	R
Gcir	F	Q	Y	F	V	E	I	F	A	Y	Q	G	G	A	N	Q	N	N	V	L
Rtyp	E	Q	Y	F	V	V	V	L	A	V	R	G	G	G	N	Q	N	N	V	L
Espi	E	Q	Y	F	L	V	V	F	A	F	Q	A	G	G	N	Q	Q	N	V	L
Saca	H	L	Y	F	V	E	V	F	A	A	R	G	G	A	Q	V	Q	N	V	L
Rays																				
Leri	D	L	Y	F	L	V	V	L	A	T	R	G	G	R	Q	F	Q	N	V	L
Oken	E	Q	Y	F	L	V	I	L	A	T	R	G	G	.	N	V	Q	N	V	L
Nkuh	D	M	F	F	V	D	R	F	P	G	E	G	G	T	I	Q	Q	N	V	L
Ujam	D	M	Y	F	V	D	R	F	P	V	Q	G	G	G	Q	Q	Q	N	V	L
Chimaeras																				
Cmil	G	L	W	F	V	V	L	F	V	T	E	G	G	G	D	V	N	N	G	D

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

												β1												
Taxon	+		+		+											+							+	
	9	**11**	**13**	**28**	**30**	**32**	**37**	**38**	**47**	**56**	**60**	**61(W)**	**65**	**68**	**70**	**71**	**74**	**78**	**81(H)**	**82(N)**	**85**	**86**	**88**	**89**
Sharks																								
Hzeb	G	R	T	D	Y	Y	I	S	F	W	R	W	E	A	S	Y	G	V	T	N	I	A	-	M
Ccar	Y	W	V	D	Y	Y	I	V	F	S	R	V	S	A	A	T	E	I	N	N	I	Y	S	E
Ctau	G	R	G	D	Y	Y	I	S	Y	S	R	W	G	T	A	T	Q	V	N	N	I	H	S	E
Ioxy	Y	R	F	Y	F	Y	I	V	Y	S	R	W	S	A	A	T	E	V	N	N	I	Y	S	E
Cacr	G	L	N	W	Q	Y	I	A	F	G	R	W	Q	A	A	L	E	V	N	N	I	D	S	E
Pgla	F	G	D	W	Y	Y	I	A	F	G	R	L	E	A	A	R	Q	M	T	N	I	A	S	E
Scan	T	N	H	D	T	Y	I	A	Y	S	R	L	S	G	A	E	Q	F	T	N	I	A	S	E
Cpla	H	R	V	A	Y	Y	I	M	F	S	R	L	R	G	A	A	E	F	N	N	I	D	S	E
Gcir	E	S	H	K	Q	Y	I	A	F	S	R	W	G	A	E	Q	S	Y	H	N	I	V	E	S
Rtyp	A	Y	D	N	Y	Y	I	A	Y	G	R	W	E	A	G	T	D	V	N	N	L	A	S	E
Saca	E	A	D	K	L	Y	I	A	F	G	T	W	G	A	A	L	Q	F	N	N	I	I	E	H
Hzeb	G	R	T	D	Y	Y	I	S	F	W	R	W	E	A	S	Y	G	V	T	N	I	A	-	M
Rays																								
Teal	L	K	D	S	G	Y	I	W	A	G	R	W	-	A	S	Y	R	V	T	N	I	H	R	E
Leri	F	R	E	Y	Y	Y	V	D	F	S	K	W	G	A	A	F	E	V	N	N	N	Y	S	N
Oken	G	R	T	D	Q	Y	V	Y	F	S	R	W	G	A	A	N	D	V	N	N	I	Y	R	V
Nkuh	F	H	Q	Y	D	Y	I	D	F	G	R	W	G	A	S	T	G	V	Q	N	I	V	K	W
Ujam	Y	G	H	Y	Q	Y	T	S	F	G	R	W	G	R	A	M	G	M	N	N	I	H	Q	W
Chimaeras																								
Cmil	N	F	Q	Y	E	Y	I	A	Y	P	S	W	-	A	E	T	E	I	N	N	V	Y	R	V

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Estimates of average evolutionary divergence over sequence pairs within and between MHC class IIα and IIβ lineages, estimated in MEGAX v10.1.18 assuming uniform rates and pairwise deletions for indels. Values within lineages correspond to average pairwise amino acid differences; values between lineages correspond to net average amino acid differences between sequences from the two lineages per gene. Values are given for the PBR and IgSF domains separately (exon 2 vs 3, respectively). N – number of sequences used in calculations, including three shark species and corresponding to those shown in Fig. 5.

MHC class II		N	PBR	IgSF
Alpha	DAA	3	22%	16 %
	DBA	6	19 %	19 %
	between lineages		4%	12 %
Beta	DAB	5	27 %	9%
	DBB	4	34%	20 %
	between lineages		13 %	21 %