Published in final edited form as:

J Immunol. 2016 May 1; 196(9): 3517–3523. doi:10.4049/jimmunol.1600164.

Assembly and expression of shark Ig genes¹

Ellen Hsu^{*}

^{*}Department of Physiology and Pharmacology, State University of New York Health Science Center at Brooklyn, Brooklyn, NY 11203, USA

Abstract

Sharks are modern descendants of the earliest vertebrates possessing Ig superfamily receptor-based adaptive immunity. They respond to immunogen with antibodies which, upon boosting, appear more rapidly and show affinity maturation. Specific antibodies and immunological memory imply that antibody diversification and clonal selection exist in cartilaginous fish. Shark antigen receptors are generated through V(D)J recombination, and since it is a mechanism known to generate autoreactive receptors, this implies that shark lymphocytes undergo selection. In mouse, the ~2.8 Mb IgH and IgL loci require long-range, differential activation of component parts for V(D)J recombination, allelic exclusion, and receptor editing. These processes, including class switching, evolved with and appear inseparable from the complex locus organization. In contrast, shark Ig are encoded by 100–200 autonomously rearranging miniloci. This review describes how the shark primary antibody repertoire is generated in the absence of structural features considered essential in mammalian Ig gene assembly and expression.

Keywords

Gene Rearrangement; Repertoire Development

Introduction

With the structure of rabbit Ig and myeloma proteins established in the 1960s, investigators tried to determine the evolutionary origins of serum components in adaptive immunity. In response to injections of a variety of immunogens, chickens, turtles, bullfrogs, lungfish, carp, sharks and other vertebrates produced specific antibody activity in the form of serum glycoproteins that sedimented at 7S or 14S–19S, molecules that produced equimolar heavy (H) and light (L) chains after reduction and alkylation [1]. A physicochemical distinction in the agglutinins and precipitins could be made between two classes representing the earliest vertebrates, jawless fish (lamprey, hagfish) and jawed (cartilaginous) fish, separated by tens of millions of years in evolution and a multitude of ancestral and transitional extinct species. Ig was reproducibly obtained in sharks and rays, which showed that the ability to mount humoral responses like that in mammals arose in ancestral jawed vertebrates >450 million years ago. Divergence times and vertebrate relationships are depicted in Fig. 1.

¹This work was supported in part by the National Institutes of Health grant GM068095

correspondence: Ellen Hsu, ; Email: hsue@hscbklyn.edu, phone: (718) 270-1154 fax: (718) 270-3103.

Elasmobranchs diversify their antibody repertoire and show immunological memory. This implies that clones of lymphocytes are generated and selected, and the presence of characteristic junctions in their Ig V regions, together with the presence of recombination activating genes (RAG), shows that their antigen receptors are created by V(D)J rearrangement [8, 9]. Since rearrangement is a random diversification mechanism that inevitably generates autoreactive molecules in the process [10], this implies that lymphocytes in sharks and skates must undergo selection. In mammals the primary B cell repertoire results from V(D)J recombination, allelic exclusion, and receptor editing. These are three processes whose operations are inextricably bound in the particular genomic arrangement of Ig H and L chain genes, which is generally similar in all tetrapod animals [11]. The Ig gene organization of cartilaginous fishes, however, is very different.

This review discusses how the generation of antibody diversity and clonal expression of B cell receptors (BCR) are achieved in cartilaginous fish, in the absence of the structural features of the DNA deemed essential to Ig assembly and expression in mammalian systems.

Ig organization in cartilaginous fish

In 1986 Hinds and Litman reported [12] that the Ig μ genes in horned sharks have a radically different organization from the “translocon” arrangement characterized in mammals (Fig. 2A). Instead of hundreds of rearranging gene segments and several constant region genes at one extended site, there were 100–200 miniloci, with a few (2–3) gene segments and one set of C exons (Fig. 2B). L chain genes were also arranged in the “multiple cluster” style [13], in fact, as are all genes encoding the classical serum Ig: the two H chain classes (μ , ω) of IgM and IgW, all four L chain types (kappa and lambda orthologs, sigma, sigma-2) [14]. A third unique Ig molecule, IgNAR, consists of H chain dimers or multimers [15, 16]. Being without L chain, each chain of IgNAR carries a single-domain V region that binds independently (Fig. 2C).

The numbers of clusters can vary considerably among species. For instance, compared to horned shark there are only 9–12 distinct, functional Ig μ and three pseudogenes in nurse shark [17]. Whereas within a cluster, recombining gene segments such as V_H-D1 or V_L-J_L or V_{NAR}-D1 are 250–500 bp apart (Fig. 2B), the clusters themselves are spatially distant: 120 to >200 kb between Ig μ , 82 to >200 kb between Ig ω [18, 19]. Clusters can be chromosomally dispersed, as shown in clearnose skate by fluorescence in situ hybridization [20].

It is generally accepted that RAG is derived from a DNA transposase [21, 22] that arrived through lateral transfer. In an ancestral vertebrate its transposase recognition sites (the RSS) became inserted into an Ig superfamily V domain belonging to a V/C structure. The segmented V-J became potentially recombinable. It would have been from a V-J/C structure that modern Ig and TCR genes evolved. Although it has sometimes been suggested that the minimalist cartilaginous fish cluster represents the primordial segmental-rearranging organization, clusters could also have evolved independently, as had TCR μ genes in marsupial and monotreme mammals [23]. Additionally, cartilaginous fish TCR genes are in

the translocon organization, showing that regulation and expression of genes arranged as in higher vertebrates was established by the time of chondrichthyan divergence [24, 25].

V(D)J Rearrangement

No combinatorial diversity

The gene segments in an Ig cluster are flanked by RSS consisting of heptamer, a 12 bp or 23 bp spacer, and a nonamer (Fig. 2B). Analysis of rearranged V regions show that VDJ is formed from recombining the four segments, V_H -D1-D2- J_H and primarily within the cluster [17, 27]. Moreover, as the V_H tend to belong only to one family [17, 27], as do V_K segments [28], the primary antibody repertoire is patently dependent on junctional diversity. This property, existing in ancient vertebrates like cartilaginous fish, perhaps underlines the evolutionary uniqueness of V(D)J recombination [29]: routine generation of sequence of varying lengths (CDR3), loops that bestow topological diversity to the combining site. Since RAG activity is site-specific, the variability is localized to a structurally tolerating position. The emergence of TdT, a lymphoid tissue-specific enzyme, in jawed vertebrates supports the idea that generating loop length diversity was a primary feature in selection.

One may suggest an analogy in the variable lymphocyte receptor (VLR) system in jawless fish (Fig. 1), which evolved an adaptive immune system also based on a somatically assembled repertoire [30]. VLR are not Ig superfamily-derived nor are they generated by RAG, but part of the VLR assembly process involves stepwise insertion of variable numbers of leucine-rich-repeat (LRR) V modules, each motif 24 residues, into the VLR gene. In both adaptive immune systems the scaffold supports a range of dimensions in combining site -- varying loop sizes (CDR3) or inserts (LRRV). Although evolving by independent pathways, the unique feature in common is that the lymphocyte antigen receptors are generated with extensive conformational diversity.

What is the role of sequence diversity in this context? To study the effect of sequence versus conformational diversity, single-domain scaffolds were engineered with binding surfaces consisting of varied loop sizes with repetitive residues of serine and tyrosine [31]. Despite the restricted sequence content it is the compensating conformational diversity provided by the loops that enabled such surfaces to complex with ligands, at a similar affinity as natural antibodies. When protein sequence diversity is incorporated as part of the system, as when generated by varied LRR templates or D_H gene segments and TdT activity, the sequence-diverse surfaces will interfere with or bind more tightly a particular ligand: this makes for the specificity [32].

Rearrangement and orderliness

Sterile transcripts of non-rearranged H and L chains consist of the cluster germline gene segments with splicing of J to C region. They are most easily detected in neonatal tissues [33] and assumed to herald Ig locus activation [34]. The only promoter appears to be that 5' of the V gene since no transcripts containing C region alone have been observed. Although the upstream TATA sequence and octamer are present at all elasmobranch L chain genes, neither the octamer nor other conventional Ig-related motifs have been identified at IgH

clusters [8, 27]. Enhancers have not been characterized, and it is not clear if every cluster possesses complete cis-regulatory elements enabling its activation and functioning.

Rearrangement at the cartilaginous fish IgH clusters occurs as deletional recombination between adjacent gene segments to generate VDJ (Fig. 2B) [8, 17]. That means the “12/23 bp rule” [21] is followed, but not every combination of RSS pairing is used. The Ig μ cluster in Fig. 2B (highlighted field) is depicted with four gene segments and their flanking RSS, ~400 bp apart. Four kinds of deletions and four inversions should be possible, given the sufficient spacing between compatible RSS (400 or 800 bp). However, *any* inversion produces a signal joint within the cluster, and most subsequent recombinations utilizing one RSS of the signal joint will destroy the second RSS and thus ability to complete to VDDJ. When sought, such disrupted clusters have been observed but are very rare [35].

Studies have shown that the three deletional rearrangements (3R) forming VDDJ take place without a pattern common to all the Ig μ [36]. Intermediates (1R, 2R) are infrequent in B cells, as shown in Fig. 3, but can be amplified from sIgM⁺ cell DNA and compared to those from thymic DNA. The rearrangement-order preferences in one particular Ig μ were the same in the two cell populations but differed from other clusters. For instance, in the G2 cluster, 14/16 cloned 1R were VD-D-J and 74/81 2R were VDD-J [35, 36]. This suggests that for G2, V to D1 tends to be the first rearrangement, generally followed by VD to D2, concluding with VDD to J_H to form the VDDJ. The pattern is characteristic for a cluster: in G5, D2 to J_H is typically the first event whereas in G1, D1 to D2 is the preferred step [36]. Rearrangement can commence with any pair of gene segments as long as they are adjacent, and the loosely preferred sequence of events is cluster-specific.

The absence of a strictly imposed order, the short distances separating gene segments, and the rarity of recombination intermediates suggest that IgH rearrangement in shark B cells is not regulated like in mouse and most likely occurs in one stage, once initiated. Because sometimes the second rearrangement event is not physically adjacent to the first (i.e. V_HD1--D2J_H) the process does not nucleate from one particular RAG-bound site as it does in mouse [37]. Perhaps the whole cluster in itself is a recombination center [38] where all gene segments are bound by RAG, and RSS pairs sort out. How RAG distinguishes among three 23-bp RSS and three 12-bp RSS in a cluster is perplexing: of the eight possible pairings one or two are preferred and five are avoided.

The RSS partner preferences and the rarity of inversion recombination suggest RSS coupling or cleavage hierarchies [39] or highly localized epigenetic modifications or both. The enforcement of tandem recombinations, whatever its basis, generates not only orderliness but also the greatest CDR3 diversity. For example, although direct V_H to D2 joining is possible in any Ig μ cluster, this is hardly ever observed in genomic or cDNA, in-frame or non-productive, rearrangements [35]. These observations suggest that, in mouse and shark lymphocytes, a combination of RSS sequence and RSS-availability features have evolved at antigen receptor genes to direct RAG-targeting for productive assembly of VDJ. This aspect has been most obvious with specific RSS cleavage patterns, as in the “beyond 12/23 restriction” [40].

Interaction between clusters

In mouse, chromatin conformational changes cause the 2.8 Mb IgH to form compacted domains, a pre-rearrangement structure believed to enable encounters between linearly distant gene segments [41]. There is no information about chromatin folding during V(D)J recombination in shark. As some IgH are linked [18, 19], chromatin folding may bring the distant clusters into proximity. However, intercluster recombination is rarely observed [17, 26], and the reason could be that in any pro-B cell few (1–3) Ig μ genes are accessible to RAG at the same time (see H chain exclusion).

Some intercluster Ig μ recombination has been found in thymocytes, where many (3–7) Ig μ have undergone some rearrangement in single cells, indicating multiple activated genes [35, 36]. Interestingly, most (82%) of recombined thymic μ clusters are 1R and 2R; Fig. 3 contrasts Ig μ configuration in B cells and thymocytes, disparities that preclude the rearrangement events having occurred in common lymphoid precursors. With RAG-accessible IgH in thymocytes, “trans-rearrangements” have been recovered, involving V gene segments of Ig μ or Ig δ and the TCR δ D/J and TCR α J gene segments [42]. The role, if any, of these chimeric transcripts is unknown [24].

These cDNAs indicate that the shark Ig V_H promoter functionally communicates with TCR cis-regulatory elements, once the physical proximity was established. In contrast, despite the many intracluster-recombined Ig μ , no μ transcripts containing frequently observed configurations (e.g. VDD--J) were detected in thymocyte RNA, nor was any significant signal observed with V_H probe by northern blotting [36]. The Ig genes in thymocytes are not under normal B-lineage regulation, and it is tempting to speculate about a relationship between the incompletely rearranged Ig μ genes and the poor transcriptional activity.

Accessibility to RAG is dependent on transcription [34], an indication of locus activation, and transcription is associated with chromatin remodeling and histone modification, both of which stimulate RAG activity [43, 44]. In shark thymocytes, the many partially recombined Ig μ clusters demonstrate that effective “processivity” by RAG to VDJ is under B lineage-regulated chromatin modifications. Thymic IgH show uniquely suboptimal but stable RAG accessibility, where RAG proteins are perhaps not effectually recruited [38].

No stepwise VDJ formation

Since pairs of gene segments at shark Ig μ recombine over 400 bp, the formation of its VDJ differs in two well-known aspects of V(D)J rearrangement at mouse IgH, (a) the ordered two-step process where D recombines with J_H, followed by V_H to the DJ, with (b) the attendant locus compaction providing equal opportunity for the many V_H. Put simplistically -- because of the V_H are located over >2 Mb, the activation of the H chain locus occurs in two stages, in part enabled by an insulator barrier in the V-D intergenic region that enforces regulated activation of the V_H and the step-wise assembly [45, 46]. The complex spatial dynamics in mouse antigen receptor systems, necessary to form a VJ or VDJ, have evolved with and are bound up with the extended translocon configuration and multiple dispersed gene segments.

As discussed previously, rearrangements at *Igu* in developing shark B cells most likely occur in a single stage, in which case activation of the IgH cluster can be correlated with directly forming the VDDJ.

H chain exclusion

Single cell studies showed that few IgH clusters are active per elasmobranch lymphocyte. In the clear-nose skate [47] the IgH genes number in the hundreds but per cell, primarily one in-frame VDJ out of 1–3 transcripts was isolated by RT-PCR. Other experiments conducted in nurse shark [36, 18] inspected 9–10 defined, functional *Igu* loci (18–20 genes) in the individuals examined, and in single sIgM+ cells only 1–3 genomic VDJ were recovered; the other *Igu* were in germline configuration. As in the skate, only one VDJ was in-frame per cell, an observation suggesting that the nonproductive VDJ may have been earlier, failed attempts to generate a H chain polypeptide. When detected, the allele of almost any VDJ, productive or not, was in germline configuration, demonstrating that gene activation did not occur per locus, in allelic pairs. Moreover the *Igu* do not appear to be cis-activated, en bloc, since three linked *Igu* rearranged independently [18]. These results support the idea that Ig clusters are individually activated.

In the mouse multiple mechanisms have been proposed for inducing monoallelic expression of Ig H chain, κ L chain, and TCR β [48, 49]. Without entering into controversial details, it can be said that in all cases asynchronous formation of the V region is observed at the two active alleles, and the disagreements are on the nature of the temporal difference. Since some mouse models are based on the allelic relationship, explanations involving allele-pairing [50] or delayed demethylation retarding activation of one allele [51] are difficult to reconcile with the multiplicity of elasmobranch clusters, linked or chromosomally dispersed.

In both skate [46] and shark studies [36, 18], RAG activity had occurred at many fewer than the number of genomic *Igu* in a B cell. This at first appears odd, considering that numerous clusters could offer additional rounds of rearrangement, reducing cell wastage due to nonproductive VDJ and to activation of defective clusters. But the paucity of VDJ events suggests a limited time period during which successful H chain formation can take place. Such a restriction also clarifies why the total locus number, which ranges from 12–15 in nurse shark to 100–200 in horned shark, is not relevant to H chain exclusion.

The nurse shark *Igu* genes were individually scored, and because there was no discernible pattern in B cells, it was concluded that the *Igu* clusters are stochastically recruited [36, 18]. Gene activation begins with the binding of nuclear factors to cis-regulatory elements, bringing on modifications of the chromatin that allow recruitment of RNA polymerase to the promoter and transcription initiation. This series of events also makes RSS available to RAG [34]. In shark, B lineage-regulated *Igu* cluster activation is followed by three rearrangements forming VDDJ. To obtain asynchronous production of the VDJ, at least one of the processes, cluster activation or V(D)J recombination, would have to be rate-limiting, which implies that the clusters are available but quiescent until activation is initiated.

Alternatively, all clusters could be suppressed in shark pro-B cells, and it is the individual release that is limiting. In the mouse, regulation of allelic exclusion has also been associated

with interaction of the Ig or TCR β genes with repressive nuclear compartments. The non-tethered allele is thought to become recombinationally active; in TCR β , it preferentially accumulates DNA repair protein 53BP1, reflecting the aftermath of RAG-induced breaks [52]. It is possible that in shark a few clusters dissociate at random and get activated, but here the species-variable IgH number would affect cluster availability if (an equal) probability of cluster dissociation were the sole basis for release from suppression.

Would elasmobranch H chain exclusion have developed independently because of the unique Ig gene organization? Maybe not. The IgH and IgL clusters cannot be regarded as *manqué* or remaining primitive since co-evolving TCR genes have organizations similar to higher vertebrates and their regulated expression is presumably mechanistically similar as well. For example, it is reasonable to speculate that if both mouse Ig as well as TCR β associated with repressive chromatin during lymphocyte differentiation in one regulatory phase determining allelic exclusion, the basis of this shared feature may have preceded their ancient divergence and survives in cartilaginous fish.

AID-induced changes

Somatic hypermutation

The role played by the AID in mammals and probably in all jawed vertebrates it is that of initiating SHM and class switch recombination (CSR) at Ig genes and, in chickens and rabbits, gene conversion [53]. The result of AID activity is post-rearrangement diversification.

SHM in mammals is responsible for affinity maturation of the antibody response, enhancing specificity and memory, two basic features of an adaptive immune system. Criteria set in mammals, i.e. faster response time with higher antibody titers of increased affinity, have been problematic to interpret in amphibians and fish, although they have B cell repertoires generated by V(D)J recombination. Clearly, other factors influence the course of their antibody response. For instance, affinity ceilings for mutating Ig must differ in cold-blooded vertebrates, and this will be reflected in mutant selection in a way that has not been explored.

The situation in sharks was clarified when Ig species were separately tested [54, 55]. IgM in serum is present as 19S pentamer and 7S monomer, probably from separate B cell lineages. Although both are produced upon immunization, it is the 7S that predominates as the antigen-specific antibody. Flajnik and coworkers [55] showed that, although antibody levels after boosting did not exceed the primary response, there was a faster response as well as affinity maturation for 7S IgM and IgNAR. Subsequent affinity studies on clonally-related V regions against hen egg lysozyme, possible only with the single-domain binders of IgNAR, demonstrated mutants being antigen-driven and selected [56].

SHM in shark Ig consists of point mutations and tandem substitutions [29, 57, 58]. Both changes occur at RGYW/WRCY hotspots, and lesions at G:C sites are assumed initiated by AID. The tandem mutations, consisting of 2–5 bp adjacent changes, are unique to cartilaginous fishes and may result from recruitment of repair processes able to extend

mismatches or a greater suspension of proofreading capacity in their B cells. Mutations at the shark Ig μ occur over a similar span as in mammalian genes, ~1.5–2 kb including VDJ and the downstream flank. The functional shark V region is under selection, as indicated by higher replacement frequency in the CDRs. However the adjacent non-coding flank DNA revealed frequent duplications, deletions (indels), and insertions, suggesting that double-strand DNA breaks (DSB) routinely occur and are resolved during shark SHM [58].

CSR-like products in sharks

Reviews on CSR remark that isotype switching arose in the amphibians. This conclusion is reached on the basis of an absence of canonical switch (S) regions in the J_H-C_H intron at shark IgH clusters [59]. Moreover, the nature of the dispersed clusters would not seem to encourage events that in tetrapods are guided, directionally, by the IgH organizational structure with its tandem array of C_H region exons. In fact, antibodies in cartilaginous fish were never inspected for switching; and there is no easy physicochemical distinction for switching among the IgM or IgW classes. For example, ω VDJ spliced to C μ would be indistinguishable from native μ chain.

All the nurse shark IgH clusters are distinct in their V_H, J_H and C_H [17, 19], so that it was possible to distinguish chimeric sequences from adult spleen cDNA libraries that carried VDJ from one IgH gene and C region from another [60, 19]. Combinations were found between Ig μ clusters or Ig ω clusters or between Ig μ and Ig ω clusters. Compared to normal H chain cDNA, more chimeric sequences showed evidence of expanding clones, as indicated from multiple related sequences with a common VDJ that carried common and individual substitutions. Significantly, one telling pair of “switched” and non-switched sequences, both μ , shared VDJ and 10 out of 13 mutations but differed in the C region. These observations suggested that (a) SHM and AID activity took place both before and after the switching event, (b) there can be clonal expansion after switching, and (c) the switched configuration was stable over several cell divisions.

How did an exchange between two IgH clusters take place? Two lymphocyte-specific enzymes capable of inducing DSB are RAG and AID. If the switched cDNA were the result of trans-rearrangement occurring between IgH, let's say in the J-C intron, chimeric sequences would be present throughout life; however, there were none in neonates or pups nor were any observed in neonatal cDNA libraries [60, 61]. On the other hand, there was a correlation of switching with SHM and AID expression, all of which were below detection in neonatal/pup spleens but present in adults. Moreover, mRNA from non-immunized and immunized adults showed an increase of switching with deliberate immunization [60].

Indels occurring downstream of VDJ are evidence of DSB, generated through AID activity [58], and switch-junctions of intercluster recombinants were in fact found in this region. The joints are mostly located <1 kb 3' of the J_H flank and repaired by NHEJ. Clonally related switch-junctions in J-C shared the upstream CDR3 and various multiple mutations, demonstrating that the joint itself was a stable feature. It is not known if and how this process is antigen-driven or if it occurs by chance as a result of the higher frequency of DSBs due to AID activity near VDJ. Yet, the switched combinations tend to be biased and

the clones expanded, observations not explainable by serendipity; certainly selection is playing a role.

In recent years much has been elucidated as to variables influencing the frequency of exchange between two separate DSB in mammalian lymphocytes that, in lymphoid malignancies, are found as recurrent translocations. Factors include: the spatial proximity, linkage, location in active chromatin compartments, DSB frequency, persistence of the DSB [62]. It is of interest to understand how spatially distant shark IgH clusters are brought into proximity to undergo what in mammalian lymphocytes is the type of proceeding that is normally to be suppressed -- specifically, how conditions defined for aberrant exchanges may apply to events that are represented at 3% of μ cDNA in a non-immunized shark and 17% of ω -containing cDNA in an immunized animal [60].

At first sight, it is not clear why switching occurs *within* the IgM (or IgW) class. Although the H chains are highly conserved in the carboxy termini, distinct cluster differences are localized in C μ 1 and C μ 2, the latter sharing 66–70% identity among subfamilies. Although all Ig domains of the Ig μ genes evolved at comparable rates, both V_H and C μ 2 appear to have been under strong positive selection for increased amino acid diversity [17]. The function of V_H is known, but perhaps C μ 2 specifies effector function differences among the Ig μ .

Given the existence of shark immune cells (neutrophils, eosinophils, monocytes, macrophages, dendritic cells) [63, 64] that in mammals mediate antibody-dependent cellular cytotoxicity, opsonization, and cell degranulation, it is anticipated that shark antibodies likewise recruit effector cells during its immune response. The wide array of pathogens recognized by antibodies require different disposal pathways, and this need for flexible processing may be met by acquiring the ability to vary C-region function. All vertebrates with Ig systems can diversify V and C regions through changes mediated by AID, and its early, spatially close coupling of SHM with CSR is demonstrable in the VDJ and its proximal 3' flanking sequence in the shark.

Although there is, strictly speaking, no CSR event as defined in tetrapods – recombination through downstream deletion via DSB in regions of repetitive motifs – CSR products do exist in the shark. It is the classical S regions, not isotype switching itself, that arose in the amphibian ancestor.

Cluster evolution

The two subclasses of cartilaginous fish, elasmobranchs and holocephalans, diverged 420 million years ago (Fig. 1), so that the shared Ig cluster organization is an ancient one. All elasmobranchs express IgM, IgW, and IgNAR [11, 27], but the genomes of holocephalans (ratfish, elephantshark) carry clusters encoding μ and a μ variant [65, 25]. In the variant clusters, duplicated C μ 2 exons have replaced C μ 1 and may enable H chain dimers. If so, this is an independent evolution of IgNAR-like antibodies, evidence of a drive for generating paratopes with properties different from the classical H-L combination. It has been argued [66] that since single-domain binding emerged only in camelids among all tetrapods but

arose perhaps twice in cartilaginous fish, it is because substantial structural changes are more apt to succeed with multiple expendable gene copies.

IgNAR-expressing cells raise another point of interest. Since IgNAR carry no L chains, what happens to developing IgNAR B cells that express autoreactive receptors? Clonal deletion is a default solution, but perhaps in spite of the cluster organization secondary rearrangements are possible. Let's speculate that lymphocyte development is arrested, and RAG is up-regulated and rearranges a replacement IgNAR gene. Continued RAG activity at the original "offending" cluster cannot involve gene segments, which were terminally recombined in the VDJ, but sets of RSS-like motifs may exist downstream that will delete or invert part of the cluster, like the κ deleting element (κ DE) in mammalian systems [67]. κ DE demonstrates that RSS motifs can arise outside of the gene segment cluster and acquire regulated access by RAG. A similar scenario can thus be conjured for inactivating clusters that encode unwanted L chains, for receptor editing in IgM or IgW-expressing cells. The multiple Ig genes operate independently, but a salvage pathway leading to B cell self-tolerance may be possible in shark if in IgNAR and IgL clusters there evolved the seeds for their potential somatic demise.

Conclusions

Cartilaginous fish are phylogenetically the closest living relatives of the ancestral vertebrates in which V(D)J recombination evolved. The Ig gene cluster may or may not represent the primordial unit where segmented units can recombine, but because of their overall unique organization, the 100–200 miniloci encoding H and L chains assemble V(D)J differently than at the three mammalian Ig loci. Shark IgH gene rearrangement to VDJ occurs in one stage, and the clusters are activated stochastically, without respect to allelic or linkage relationships. Generation of VDJ is, like in other vertebrates, asynchronous among the Igu genes, bringing about H chain exclusion. The shark primary B cell repertoire is almost entirely dependent upon junctional diversity generated by V(D)J recombination; this possibly reflects the ancient selecting feature for RAG. However, V(D)J recombination also generates anti-self reactivity. If each IgNAR or IgL cluster possessed signals both for initiating its somatic function (RSS) and signals for its self-inactivation (κ DE) when somatically unwanted, this could enable B cell self-tolerance in cartilaginous fish. During an immune response, AID diversifies antibody through SHM in the V region and CSR in the C_H region; C region diversification is so strongly selected for that there arose a means of isotype switching in sharks.

Acknowledgments

I thank Louis Du Pasquier for reading the manuscript.

References

1. Kubo, RT.; Zimmerman, B.; Grey, HM. Phylogeny of immunoglobulins. In: Sela, M., editor. The Antigens. Vol. I. Academic Press; New York, NY: 1973. p. 417-477.
2. Benton MJ, Donoghue PC. Paleontological evidence to date the tree of life. Mol Biol Evol. 2007; 24:26–53. [PubMed: 17047029]

3. Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*. 2011; 334:1091–1097. [PubMed: 22116879]
4. Near TJ. Conflict and resolution between phylogenies inferred from molecular and phenotypic data sets for hagfish, lampreys, and gnathostomes. *J Exp Zool Part B Mol Dev Evol*. 2009; 312:749–761.
5. Kuraku S. Impact of asymmetric gene repertoire between cyclostomes and gnathostomes. *Semin Cell Dev Biol*. 2013; 24:119–127. [PubMed: 23291292]
6. Inoue JG, Miya M, Lam K, Tay KH, Danks JA, Bell J, Walker TI, Venkatesh B. Evolutionary origin and phylogeny of the modern holocephalans (Chondrichthyes: Chimaeriformes): a mitogenomic perspective. *Mol Biol Evol*. 2010; 27:2576–2586. [PubMed: 20551041]
7. Renz AJ, Meyer A, Kuraku S. Revealing less derived nature of cartilaginous fish genomes with their evolutionary time scale inferred with nuclear genes. *PLoS One*. 2013; 8:e66400. [PubMed: 23825540]
8. Kokubu F, Litman R, Shambloot MJ, Hinds K, Litman GW. Diverse organization of immunoglobulin VH gene loci in a primitive vertebrate. *EMBO J*. 1988; 7:3413–3422. [PubMed: 3145194]
9. Schluter SF, Marchalonis JJ. Cloning of shark RAG2 and characterization of the RAG1/RAG2 gene locus. *FASEB J*. 2003; 17:470–472. [PubMed: 12551847]
10. Nemazee D. Antigen receptor ‘capacity’ and the sensitivity of self-tolerance. *Immunol Today*. 1996; 17:25–29. [PubMed: 8652047]
11. Flajnik MF. Comparative analyses of immunoglobulin genes: surprises and portents. *Nat Rev Immunol*. 2002; 2:688–698. [PubMed: 12209137]
12. Hinds KR, Litman GW. Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature*. 1986; 320:546–549. [PubMed: 3083268]
13. Shambloot MJ, Litman GW. Genomic organization and sequences of immunoglobulin light chain genes in a primitive vertebrate suggest coevolution of immunoglobulin gene organization. *EMBO J*. 1989; 8:3733–3739. [PubMed: 2511000]
14. Criscitiello MF, Flajnik MF. Four primordial immunoglobulin light chain isotypes, including λ and κ , identified in the most primitive living jawed vertebrates. *Eur J Immunol*. 2007; 37:2683–2694. [PubMed: 17899545]
15. Greenberg AS, Avila D, Hughes M, Hughes A, McKinney EC, Flajnik MF. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature*. 1995; 374:168–173. [PubMed: 7877689]
16. Smith LE, Crouch K, Cao W, Müller MR, Wu L, Steven J, Lee M, Liang M, Flajnik MF, Shih HH, Barelle CJ, Paulsen J, Gill DS, Dooley H. Characterization of the immunoglobulin repertoire of the spiny dogfish (*Squalus acanthias*). *Dev Comp Immunol*. 2012; 36:665–679. [PubMed: 22040740]
17. Lee V, Huang JL, Lui MF, Malecek K, Ohta Y, Mooers A, Hsu E. The evolution of multiple isotypic IgM heavy chains in the shark. *J Immunol*. 2008; 180:7461–7470. [PubMed: 18490746]
18. Zhu C, Feng W, Weedon J, Hua P, Stepanov D, Ohta Y, Flajnik MF, Hsu E. The multiple shark immunoglobulin heavy chain genes rearrange and hypermutate autonomously. *J Immunol*. 2011; 187:2492–2501. [PubMed: 21804022]
19. Zhang C, Du Pasquier L, Hsu E. Shark IgW C region diversification through RNA processing and isotype switching. *J Immunol*. 2013; 191:3410–3418. [PubMed: 23935192]
20. Anderson M, Amemiya C, Luer C, Litman R, Rast J, Nilmura Y, Litman G. Complete genomic sequence and patterns of transcription of a member of an unusual family of closely related, chromosomally dispersed Ig gene clusters in *Raja*. *Int Immunol*. 1999; 6:1661–1670. [PubMed: 7865459]
21. Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. *Annu Rev Genet*. 2011; 45:167–202. [PubMed: 21854230]
22. Kim MS, Lapkouski M, Yang W, Gellert M. Crystal structure of the V(D)J recombinase RAG1-RAG2. *Nature*. 2015; 518:507–511. [PubMed: 25707801]
23. Parra ZE, Baker ML, Hathaway J, Lopez AM, Trujillo J, Sharp A, Miller RD. Comparative genomic analysis and evolution of the T cell receptor loci in the opossum *Monodelphis domestica*. *BMC Genomics*. 2008; 9:111. [PubMed: 18312668]

24. Criscitiello, MF. Shark T cell receptors. In: Smith, SL.; Sim, R.; Flajnik, MF., editors. Immunobiology of the Shark. Vol. Chapter 12. CRC Press; Boca Raton, FL: 2015.
25. Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, et al. Elephant shark genome provides unique insights into gnathostome evolution. *Nature*. 2014; 505:174–179. [PubMed: 24402279]
26. Hinds-Frey KR, Nishikata H, Litman RT, Litman GW. Somatic variation precedes extensive diversification of germline sequences and combinatorial joining in the evolution of immunoglobulin heavy chain diversity. *J Exp Med*. 1993; 178:825–834. [PubMed: 8350056]
27. Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. *Annu Rev Immunol*. 1999; 7:109–147. [PubMed: 10358755]
28. Iacoangeli A, Lui A, Naik U, Ota Y, Flajnik MF, Hsu E. Biased immunoglobulin light chain gene expression in the nurse shark. *J Immunol*. 2015; 195:3992–4000. [PubMed: 26342033]
29. Lee SS, Tranchina D, Ohta Y, Flajnik MF, Hsu E. Hypermutation in shark immunoglobulin light chain genes results in contiguous substitutions. *Immunity*. 2002; 16:571–582. [PubMed: 11970880]
30. Hirano H, Das S, Guo P, Cooper MD. The evolution of adaptive immunity in vertebrates. *Adv Immunol*. 2011; 109:125–157. [PubMed: 21569914]
31. Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S. High-affinity single-domain binding proteins with a binary-code interface. *Proc Natl Acad Sci U S A*. 2007; 104:6632–6637. [PubMed: 17420456]
32. Fellouse FA, Esaki K, Birtalan S, Raptis D, Cancasci VJ, Koide A, Jhurani P, Vasser M, Wiesmann C, Kossiakoff AA, Koide S, Sidhu SS. High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries. *J Mol Biol*. 2007; 373:924–940. [PubMed: 17825836]
33. Miracle AL, Anderson MK, Litman RT, Walsh CJ, Luer CA, Rothenberg EV, Litman GW. Complex expression patterns of lymphocyte-specific genes during the development of cartilaginous fish implicate unique lymphoid tissues in generating an immune repertoire. *Int Immunol*. 2001; 13:567–580. [PubMed: 11282996]
34. Abarrategui I, Krangel MS. Germline transcription: a key regulator of accessibility and recombination. *Adv Exp Med Biol*. 2009; 650:93–102. [PubMed: 19731804]
35. Hsu, EH. Considering V(D)J recombination in the shark. In: Smith, SL.; Sim, R.; Flajnik, MF., editors. Immunobiology of the Shark. Vol. Chapter 10. CRC Press; Boca Raton, FL: 2015.
36. Malecek K, Lee V, Feng W, Huang JL, Flajnik MF, Ohta Y, Hsu E. Immunoglobulin heavy chain exclusion in the shark. *PLoS Biol*. 2008; 6:e157. [PubMed: 18578572]
37. Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell*. 2010; 141:419–431. [PubMed: 20398922]
38. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol*. 2011; 11:251–263. [PubMed: 21394103]
39. Banerjee JK, Schatz DG. Synapsis alters RAG-mediated nicking at Tcrb recombination signal sequences: implications for the “beyond 12/23” rule. *Mol Cell Biol*. 2014; 34:2566–2580. [PubMed: 24797073]
40. Tillman RE, Wooley AL, Hughes MM, Khor B, Sleckman BP. Regulation of T-cell receptor β -chain gene assembly by recombination signals: the beyond 12/23 restriction. *Immunol Rev*. 2004; 200:36–43. [PubMed: 15242394]
41. Lucas JS, Zhang Y, Dudko OK, Murre C. 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. *Cell*. 2014; 158:339–352. [PubMed: 24998931]
42. Criscitiello MF, Ohta Y, Saltis M, McKinney EC, Flajnik MF. Evolutionarily conserved TCR binding sites, identification of T cells in primary lymphoid tissues, and surprising trans-rearrangements in nurse shark. *J Immunol*. 2010; 184:6950–6960. [PubMed: 20488795]
43. Du H, Ishii H, Pazin MJ, Sen R. Activation of 12/23-RSS-dependent RAG cleavage by hSWI/SNF complex in the absence of transcription. *Mol Cell*. 2008; 31:641–649. [PubMed: 18775324]

44. Shimazaki N, Tsai AG, Lieber MR. H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in trans in addition to tethering in cis: implications for translocations. *Mol Cell*. 2009; 34:535–544. [PubMed: 19524534]
45. Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. *J Biol Chem*. 2010; 285:9327–9338. [PubMed: 20100833]
46. Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng HL, Hansen E, Despo O, Bossen C, Vettermann C, Bates JG, Richards N, Myers D, Patel H, Gallagher M, Schlissel MS, Murre C, Busslinger M, Giallourakis CC, Alt FW. CTCF-binding elements mediate control of V(D)J recombination. *Nature*. 2011; 477:424–430. [PubMed: 21909113]
47. Eason DD, Litman RT, Luer CA, Kerr W, Litman GW. Expression of individual immunoglobulin genes occurs in an unusual system consisting of multiple independent loci. *Eur J Immunol*. 2004; 34:2551–2558. [PubMed: 15307187]
48. Brady BL, Steinel NC, Bassing CH. Antigen receptor allelic exclusion: an update and reappraisal. *J Immunol*. 2010; 185:3801–3808. [PubMed: 20858891]
49. Outters P, Jaeger S, Zaarour N, Ferrier P. Long-Range Control of V(D)J Recombination & Allelic Exclusion: Modeling Views. *Adv Immunol*. 2015; 128:363–413. [PubMed: 26477371]
50. Brandt VL, Hewitt SL, Skok JA. It takes two: communication between homologous alleles preserves genomic stability during V(D)J recombination. *Nucleus*. 2010; 1:23–29. [PubMed: 21327101]
51. Mostoslavsky R, Singh N, Tenzen T, Goldmit M, Gabay C, Elizur S, Qi P, Reubinoff BE, Chess A, Cedar H, Bergman Y. Asynchronous replication and allelic exclusion in the immune system. *Nature*. 2001; 414:221–225. [PubMed: 11700561]
52. Chan EA, Teng G, Corbett E, Choudhury KR, Bassing CH, Schatz DG, Krangel MS. Peripheral subnuclear positioning suppresses Tcr β recombination and segregates Tcr β alleles from RAG2. *Proc Natl Acad Sci U S A*. 2013; 110:4628–4637. [PubMed: 23487760]
53. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem*. 2007; 76:1–22. [PubMed: 17328676]
54. Voss EW, Sigel MM. Distribution of 19S and 7S IgM antibodies during the immune response in the nurse shark. *J Immunol*. 1971; 106:1323–1329. [PubMed: 5574406]
55. Dooley H, Flajnik MF. Shark immunity bites back: affinity maturation and memory response in the nurse shark. *Ginglymostoma cirratum*. *Eur J Immunol*. 2005; 35:936–945. [PubMed: 15688348]
56. Dooley H, Stanfield RL, Brady RA, Flajnik MF. First molecular and biochemical analysis of in vivo affinity maturation in an ectothermic vertebrate. *Proc Natl Acad Sci U S A*. 2006; 103:1846–1851. [PubMed: 16446445]
57. Diaz M, Velez J, Singh M, Cerny J, Flajnik MF. Mutational pattern of the nurse shark antigen receptor gene (NAR) is similar to that of mammalian Ig genes and to spontaneous mutations in evolution: the translesion synthesis model of somatic hypermutation. *Int Immunol*. 1999; 11:825–833. [PubMed: 10330287]
58. Zhu C, Hsu E. Error-prone DNA repair activity during somatic hypermutation in shark B lymphocytes. *J Immunol*. 2010; 185:5336–5347. [PubMed: 20921520]
59. Kokubu F, Hinds K, Litman R, Shambloot MJ, Litman GW. Complete structure and organization of immunoglobulin heavy chain constant region genes in a phylogenetically primitive vertebrate. *EMBO J*. 1988; 7:1979–1988. [PubMed: 3138109]
60. Zhu C, Lee V, Finn A, Senger K, Zarrin AA, Du Pasquier L, Hsu E. Origin of immunoglobulin isotype switching. *Curr Biol*. 2012; 22:872–880. [PubMed: 22542103]
61. Rumfelt LL, Lohr RL, Dooley H, Flajnik MF. Diversity and repertoire of IgW and IgM VH families in the newborn nurse shark. *BMC Immunol*. 2004; 5:8. [PubMed: 15132758]
62. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell*. 2013; 152:417–429. [PubMed: 23374339]
63. Rumfelt LL, McKinney EC, Taylor E, Flajnik MF. The development of primary and secondary lymphoid tissues in the nurse shark *Ginglymostoma cirratum*, B-cell zones precede dendritic cell immigration and T-cell zone formation during ontogeny of the spleen. *Scand J Immunol*. 2002; 56:130–148. [PubMed: 12121433]

64. Luer, CA.; Walsh, CJ.; Bodine, AB. The immune system of sharks, skates, and rays. In: Carrier, JC.; Musick, JA.; Heithaus, MR., editors. Biology of sharks and their relatives. CRC Press; Boca Raton, Florida: 2004. p. 369-395.
65. Rast JP, Amemiya CT, Litman RT, Strong SJ, Litman GW. Distinct patterns of IgH structure and organization in a divergent lineage of chondrichthyan fishes. Immunogenetics. 1998; 47:234–45. [PubMed: 9435342]
66. Flajnik MF, Deschacht N, Muyldermans S. A case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? PLoS Biol. 2011; 9:e1001120. [PubMed: 21829328]
67. Vela JL, Ait-Azzouzene D, Duong BH, Ota T, Nemazee D. Rearrangement of mouse immunoglobulin kappa deleting element recombining sequence promotes immune tolerance and lambda B cell production. Immunity. 2008; 28:161–170. [PubMed: 18261939]

Figure 1.

Phylogenetic relationships among chordate animals. All vertebrates possess adaptive immune systems, but those of the jawed vertebrates (tetrapods, bony fish, cartilaginous fish) are based on V(D)J rearrangement. Cartilaginous fish taxa are indicated in black. The scale shows divergence times according to molecular clock estimations [2–4]. The cyclostome (jawless fish) divergence is problematic [5], shown here by a balance of dating from fossil evidence and by gene sequence (www.timetree.org). Gray bars indicate estimated intervals during which the cartilaginous fish taxa emerged [6, 7]: Osteichthyes (bony fish) and tetrapods from Chondrichthyes (cartilaginous fish), the chondrichthyan subclasses Elasmobranchii from Holocephali (ratfish, elephantshark), and the elasmobranch orders Selachii (sharks) from Batoidea (skates, rays).

Figure 2.

Organization of mouse IgH locus and shark Ig clusters. A. Mouse IgH locus. The mouse IgH consists of multiple V_H (blue boxes), D_H and J_H gene segments (narrow boxes, as labeled), spread over >2 Mb. The C exons (orange boxes) of C_μ and C_δ are shown. B. Shark Ig clusters. The IgM H chain clusters extend 20 kb, V_H gene segments and C exons (TM, transmembrane) are as labeled. The enlarged depiction of V_H gene segments is in highlighted field, with RSS as white (23 bp spacer RSS) and black (12 bp spacer RSS) triangles. Rearrangement occurs in one stage to form VDDJ. Examples of IgL, Ig_ω and IgNAR clusters are shown; the variant forms that exist as “germline-joined” are not included [8]. C. Shark Igs. Top, 19S and 7S forms of IgM. Middle. Diverse IgW generated through splicing of H chain transcripts. Bottom. IgNAR H chain dimer.

Assayed on	Thymocyte	sIgM-positive cells
Single cell DNA	3-7 IgH recombined av. 5 per cell 56 IgH /12 cells	1-3 IgH recombined av. 2 per cell 48 IgH /29 cells
VD-D-J, V-DD-J, V-D-DJ	13 1R	0 1R
VDD-J, V-DDJ, VD-DJ	33 2R	2 2R
VDDJ	10 3R	46 3R
	Av. 9 recombination events per cell	Av. 6 recombination events per cell
Genomic DNA	IgH rearrangements No IgL rearrangements	IgH rearrangements IgL rearrangements
RNA	No H chain detected No L chain detected	H chain detected L chain detected
RT-PCR	No VDD-J detected	VDD-J mutants cloned

Figure 3.

Ig gene configurations scored in shark thymocytes and B cells. Single cell DNA. Single shark lymphocytes were scored for rearrangement configuration at the 9–10 functional shark IgH. The number of genes per cell and the extent of recombination is shown (1R, 2R, 3R), with the average number of individual events. The presence of IgH rearrangements in thymocytes but lack of transcripts is contrasted with sIgM-positive cells. Data summarized from refs. [18, 36].