Primitve synteny of vertebrate major histocompatibility complex class I and class II genes

Yuko Ohta*, Kazuhiiko Okamura†, E. Churchill McKinney*, Simona Bartl§, Keiichiro Hashimoto†, and Martin F. Flajnik*‡

*Department of Microbiology and Immunology, University of Maryland, Baltimore, MD 21201; †Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Aichi 470-1192, Japan; ‡Department of Microbiology and Immunology, University of Miami, Miami, FL 33101; and §Moss Landing Marine Laboratories, Moss Landing, CA 95039

Edited by Hugh O. McDevitt, Stanford University School of Medicine, Stanford, CA, and approved February 23, 2000 (received for review December 1, 1999)

Major histocompatibility complex (MHC) class I and class II molecules bind to and display peptidic antigens acquired from pathogens that are recognized by lymphocytes coordinating and executing adaptive immune responses. The two classes of MHC proteins have nearly identical tertiary structures and were derived from a common ancestor that probably existed not long before the emergence of the cartilaginous fish. Class I and class II genes are genetically linked in tetrapods but are not syntenic in teleost fish, a phylogenetic taxon derived from the oldest vertebrate ancestor examined to date. Cartilaginous fish (sharks, skates, and rays) are in the oldest taxon of extant jawed vertebrates; we have carried out segregation analyses in two families of nurse sharks and one family of the banded houndshark that revealed a close linkage of class Iα and β genes both with each other and with the classical class I (class Ia) gene. These results strongly suggest that the primordial duplication giving rise to classical class I and class II occurred in cis, and the close linkage between these two classes of genes has been maintained for at least 460 million years in representatives of most vertebrate taxa.

Materials and Methods

Animals. Nurse sharks (Ginglymostoma cirratum, order Orectolobiformes) were captured off of Little Torch Key, Florida in October, 1996 and October, 1998. Two broods of 17 (family 1) and 39 (family 2) pups were delivered by Caesarian section and were maintained in aerated seawater. Banded houndsharks Triakis scyllium (order Carcharhiniformes) were captured off the coast of Japan, and one family typed for the class Ia gene was found to be the offspring of one father and mother.
(25). One to three milliliters of blood was obtained from each pup. Genomic DNA was prepared from nucleated erythrocytes as described (26).

cDNA Library Screening. The exon encoding the α3 domain of the horned shark \textit{Heterodontus francisci} class Ia cDNA \textit{Hefr}-20 (GenBank accession no. AF028559) was used to select class I clones from a nurse shark spleen cDNA library under low stringency conditions. The probe was PCR amplified from cloned DNA by using the primer set 5′-GAG TGC AAT GGC TAC CAT-3′ (accession no. AF028557) and 5′-GTT GTT GCA TAC GAT CCT-3′ as probe to isolate class Ib (α2 domain exon) and 5′-TCT CAC ATG CCC TGG TAT TTT-3′ (α1 domain exon) and 5′-TGT CAT GTT GAT GAG TGC GTT TTT-3′ as probe to isolate class Ib (α2 domain exon). Conditions were 94°C for 4 min, followed by 35 cycles of 94°C for 1 min, 52°C for 1 min, 72°C for 4 min, and final extension at 72°C for 15 min. DNA fragments were cloned into the pCR2.1 TA-cloning vector (Invitrogen) and were sequenced by the Biopolymer Laboratory of the University of Maryland by using an automated DNA sequencing system (Applied Biosystems).

Results

Cloning and Genomic Analysis of the Nurse Shark Class Ia Gene. Our previous studies reported the isolation of class Ia (\textit{Hefr}-20) and class Ib (\textit{Hefr}-19) genes from the horned shark (29), class Ia genes (\textit{Trsc}-UA) from the banded houndshark (25), and a predicted class Ib gene from the nurse shark (\textit{Gici}-11) (29). Because we had not unearthed the expected nurse shark class Ia gene in the first set of experiments, we used the exon encoding the α3 domain of horned shark class Ia as probe to isolate additional nurse shark cDNA clones. A clone differing in sequence from \textit{Gici}-11 was isolated and named \textit{Gici-UA} (Fig. 1); it was the only class Ia-like clone encountered from this library screen, and its gene was shown to have class Ia-like high expression in intestine, spleen, and kidney in Northern blotting experiments (not shown). The deduced protein sequence of \textit{Gici-UA} displays class Ia features, notably canonical evolutionary- conserved peptide-binding residues that anchor the termini
of associated peptides at both ends of the class I PBR (Fig. 1). The class Ia genes from the three divergent shark species (Hefr-20, Trsc-UAA, Gici-UAA) are more similar to each other than they are to class Ib genes (Hefr-19 and Gici-11), especially in the C-terminal halves of the PBR α1 and α2 domains and the cytoplasmic region (Fig. 1), suggesting with these limited data that, in contrast to mammalian class I genes, class Ia and class Ib lineages are exceptionally old in the cartilaginous fish (31).

Southern blots of genomic DNA with probes encoding the Gici-UAA α3 domain (see Fig. 3) and α2 domain (not shown) suggested the presence of at least two class Ia-like genes in the nurse shark genome; indeed, in addition to the two bands seen for class Ia genes in the spleen. A probe encoding the class Ia 1 domain, in contrast to the α2 and α3 domain probes, yielded only one or two bands under high stringency conditions, strongly suggesting that only one gene, Gici-UAA, was detected (see Figs. 3 and 4; two bands indicate heterozygocity in α1 domain hybridizations, revealed in the segregation analyses). We further confirmed that no other class Ia-like gene except Gici-UAA was detected in our experiments: the sequence of the class Ia 1 domain probe (two animals, sibs 9 and 13 shown in Fig. 2A), leaving little doubt that the expressed class Ia gene Gici-UAA is the only one detected in our experiments (hybridizations with the nonclassical gene described above did not detect any of the major hybridizing bands in Fig. 2). Thus, we initiated family studies to examine whether the single expressed nurse shark class Ia gene was linked to nurse shark class II α and β genes.

Family Studies. Two nurse shark families (17 and 39 sharks) were surveyed for RFLP with class IIα, IIβ, and class Ia probes (Figs. 3 and 4). It was fortunate that the mothers of both families were MHC heterozygotes with readily detectable class Ia and class II RFLP. Class Ia probes specific for the α3 domain, leader, and 3′ UT, or especially the class II probes in delineating the MHC type or segregating “group” of each offspring (lowest letters a–j) whereas probes encompassing entire cDNA clones for the class IIα and β genes (class II) revealed three groups (a–c). MU and ML refer to maternal upper and lower bands in the class IIα locus that segregates as an allele of DAA (see text and ref. 32). Restriction enzymes were HindIII (class IIα and -β and class II α1 domain and leader) and BamHI (class II α3 domain and 3′ UT). Approximately molecular sizes of all discriminatory bands are as follows. Class IIα locus: MU, 10 kb, ML, 7 kb; class IIβ: MU, 6 kb, ML, 4 kb; class Iα: M, 21 kb, ML, 17 kb; class I leader: common band, 4 kb, g, 3 kb; class Iα: M+, 12 kb; class I 3′ UT b: e group, 15 kb, nonpolymorphic band, 6.5 kb.

Fig. 2. The class Ia gene appears to be the only nurse shark gene that hybridizes to the Gici-UAA α1 probe on Southern blots. Genomic DNA digested with BamHI results in a band of 225 bp (A) predicted to be in the expressed class Ia sequence, but not in a Gici-UAA-related class I gene isolated by PCR (B).

Fig. 3. Linkage of class Ia and class IIα and β genes in the first nurse shark family of 17 offspring. The sibling number is noted above and below the blot. α1 domain, α3 domain, leader, and 3′ UT class Ia probes (CL I) revealed 10 segregating groups (lowercase letters a–j) whereas probes encompassing entire cDNA clones for the class IIα and β genes (CL II) revealed three groups (a–c). MU and ML refer to maternal upper and lower bands in the class IIα locus that segregates as an allele of DAA (see text and ref. 32). Restriction enzymes were HindIII (class IIα and β and class II α1 domain and leader) and BamHI (class II α3 domain and 3′ UT). Approximate molecular sizes of all discriminatory bands are as follows. Class IIα locus: MU, 10 kb, ML, 7 kb; class IIβ: MU, 6 kb, ML, 4 kb; class Iα: M, 21 kb, ML, 17 kb; class I leader: common band, 4 kb, g, 3 kb; class Iα: M+, 12 kb; class I 3′ UT b: e group, 15 kb, nonpolymorphic band, 6.5 kb.
relatively low level of class II polymorphism (RFLP)—only three or four groups were detected in all of our experiments—we sequenced the class II α1 and β1 exons after PCR amplification of the alleles from genomic DNA. The predicted maternal class II allele, either DAA*01 or DAA*03 for α (32) or DAB*02 or DAB*01 for β, was found to match perfectly to the “upper (U)” or “lower (L)” RFLP positions for each offspring (see supplemental Figs. 6 and 7 and Table 3, published on the PNAS web site, www.pnas.org). There was only one discordant case (offspring 3), likely a recombinant between the two class II loci.

The paternal class II α genes of offspring 10 and 13 were encoded by two DBA genes (Table 3) that segregate like alleles of DAA. In the previous report of nurse shark class II polymorphism, it was believed that DBA and DAA were different class II α loci (32); rather, our studies here suggest that these are ancient genes that now segregate as alleles, but they should properly be regarded as pseudoalleles. Further studies are needed to clarify this point.

Discussion

Previous studies of the cartilaginous fish MHC showed that the nurse shark class IIα (32) and houndshark class Iα (25) genes are highly polymorphic and under Darwinian positive selection similar to their mammalian homologues. Here we show that the MHC consisted of class Iα and class II genes ever since the common ancestor of cartilaginous fish and all other vertebrates existed between 460 and 540 million years ago (24). These data strongly suggest that one gene duplicated from the other in cis and the two loci subsequently remained closely linked in most vertebrate taxa.

In contrast to all other vertebrates, every teleost species examined so far carries unlinked class Iα and class II genes (23). It was suggested that this unusual feature might have been the primordial condition, with class I and class II arising on two paralogous chromosomes and being “brought together” in a tetrapod ancestor. Our work makes this proposal untenable and suggests the more likely possibility (also recognized by the teleost workers), that the situation in bony fish is derived characteristic. Even if recent mitochondrial DNA studies suggesting that cartilaginous fish have
haplotypes have the same RFLP with this enzyme but not with others (not shown). The entire coding region of the nurse shark class II genes is more stable (reviewed in ref. 1). In teleosts, the class II genes are more plastic, fluctuating in numbers and function in different species, whereas class I genes (the major transplantation loci) over 30 years ago (43). In mammals, linkage disequilibrium studies would suggest this is indeed the case; perhaps, for coordination of effective immune responses, certain combinations of class Ia and class II alleles are advantageous (44). Will the sharks also have a functional advantage to maintaining class I and class II genes together? This has been the important question since class II genes (Ir genes) were shown to be linked to class I genes (the major transplantation loci) over 30 years ago (43). In mammals, linkage disequilibrium studies would suggest this is not necessarily Gici-UAA itself.

Another comparison to be made with teleost and all other vertebrates is the relative stability of the class I and class II genes. In mammals, class I genes are plastic, fluctuating in numbers and function in different species, whereas class II genes are more stable (reviewed in ref. 1). In teleosts, the situation is reversed, with class II being plastic and found on multiple chromosomes whereas class I appears to be more conserved (20, 39). In addition, class II gene number in the teleostean cichlid fishes differs greatly from the shark. The “nonlinkage” in teleosts may have occurred by differential silencing of MHC genes after genome-wide chromosomal duplication events proposed for a teleost-specific ancestor (38). Is there a functional advantage to maintaining class I and class II genes together? This has been the important question since class II genes (Ir genes) were shown to be linked to class I genes (the major transplantation loci) over 30 years ago (43). In mammals, linkage disequilibrium studies would suggest this is indeed the case; perhaps, for coordination of effective immune responses, certain combinations of class Ia and class II alleles are advantageous (44). Will the sharks also have a defined “class I region,” as is certainly true in zebrafish (20).

Table 2. Segregation of maternal haplotypes reveals class Ia/II linkage

<table>
<thead>
<tr>
<th>Animal</th>
<th>Ila</th>
<th>llb</th>
<th>1-α1/α3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>L(01)</td>
<td>U(02)</td>
<td>U/+</td>
</tr>
<tr>
<td>U(03)</td>
<td>L(01)</td>
<td>L/-</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>U(03)</td>
<td>L(01)</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>U(03)</td>
<td>L(01)</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>L(01)</td>
<td>U(02)</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>U(03)</td>
<td>L(01)</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>L(01)</td>
<td>U(02)</td>
<td>14</td>
</tr>
<tr>
<td>16</td>
<td>U(03)</td>
<td>L(01)</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>U(03)</td>
<td>L(01)</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Animal</th>
<th>Ila</th>
<th>llb</th>
<th>1-α1/α3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>U</td>
<td>L</td>
<td>U</td>
</tr>
<tr>
<td>4</td>
<td>U</td>
<td>L</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>U</td>
<td>L</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>U</td>
<td>L</td>
<td>27</td>
</tr>
<tr>
<td>11</td>
<td>U</td>
<td>L</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>U</td>
<td>L</td>
<td>31</td>
</tr>
<tr>
<td>15</td>
<td>U</td>
<td>L</td>
<td>33</td>
</tr>
<tr>
<td>17</td>
<td>U</td>
<td>L</td>
<td>36</td>
</tr>
</tbody>
</table>

The offspring shown here had only a single band for class Ila and β and clearly identifiable maternal class Ia bands. L, lower band; U, upper band. Numbers in the first family refer to class Ila (DAA*01 or 03) and class Iβ (DAB*01 or 02) alleles. †Because there are at least two genes that hybridize to the Gici-UAA a3 domain probe, the polymorphic band (+/-) detected in Fig. 3 may be class II-linked but not necessarily Gici-UAA itself.

Although we are certain that the α1 domain probe hybridizes to the expressed class I gene, the leader-hybridizing fragment in Fig. 4 is only deduced to be part of Gici-UAA.

Fig. 5. Linkage of class Ila to class Ia genes in the banded houndshark. Sibling number is noted at the top of the blot, and the maternal and paternal alleles are designated on the right side. The segregating diploid groups are noted at the bottom of the figure and match the previous class Ia typings (25). Note that the haplotype of the top band (a, c, d) was deduced from the relative intensities of the signals in the segregants and is somewhat ambiguous; the a and d haplotypes have the same RFLP with this enzyme but not with others (not shown). The entire coding region of the nurse shark class Ila gene was used in this experiment, and EcoRV was used for digestions. Approximate molecular sizes of bands: a, c, and d, 23 kb; c, 16 kb; top b, 9 kb; bottom b, 7 kb.
and chicken (41) and most likely in Xenopus (42)? Will there be a cluster of shark class Ib genes that is not closely linked to the functional MHC, as is found in chickens (45) and frogs (46)? Detailed studies of shark MHC, in comparison to all other vertebrates, should help define the primordial condition of this complex.

We thank Penny Conner for technical support, Mike Criscitiello for help in the marathon fishing expedition, and Pierre Pontarotti for critical reading of the manuscript. M.F.F. and Y.O. are supported by National Institutes of Health grant AI27877. K.H. and K.O. are supported by PROBRAIN and Fujita Health University; K.H. is also supported by the Japan Society for the Promotion of Science and Ministry of Education, Science, Sports, and Culture.