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

Yuko Ohta, E. Churchill McKinney, Michael F. Criscitiello and Martin F. Flajnik

*J Immunol* 2002; 168:771-781; doi: 10.4049/jimmunol.168.2.771

http://www.jimmunol.org/content/168/2/771

---

**References**

This article cites 55 articles, 14 of which you can access for free at: http://www.jimmunol.org/content/168/2/771.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts

---

*The Journal of Immunology* is published twice each month by The American Association of Immunologists, Inc., 1451 Rockville Pike, Suite 650, Rockville, MD 20852

Copyright © 2002 by The American Association of Immunologists All rights reserved.

Print ISSN: 0022-1767 Online ISSN: 1550-6606.
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

Yuko Ohta,* E. Churchill McKinney,† Michael F. Criscitiello,‡ and Martin F. Flajnik*‡

Cartilaginous fish (e.g., sharks) are derived from the oldest vertebrate ancestor having an adaptive immune system, and thus are key models for examining MHC evolution. Previously, family studies in two shark species showed that classical class I (UAA) and class II genes are genetically linked. In this study, we show that proteasome genes LMP2 and LMP7, shark-specific LMP7-like, and the TAP1/2 genes are linked to class I/II. Functional LMP7 and LMP7-like genes, as well as multiple LMP2 genes or gene fragments, are found only in some sharks, suggesting that different sets of peptides might be generated depending upon inherited MHC haplotypes. Cosmid clones bearing the MHC-linked classical class I genes were isolated and shown to contain proteasome gene fragments. A non-MHC-linked LMP7 gene was also identified on another cosmid, but only two exons of this gene were detected, closely linked to a class I pseudogene (UAA-NC2); this region probably resulted from a recent duplication and translocation from the functional MHC. Tight linkage of proteasome and class I genes, in comparison with gene organizations of other vertebrates, suggests a primordial MHC organization. Another nonclassical class I gene (UAA-NC1) was detected that is linked neither to MHC nor to UAA-NC2; its high level of sequence similarity to UAA suggests that UAA-NC1 also was recently derived from UAA and translocated from MHC. These data further support the principle of a primordial class I region with few class I genes. Finally, multiple paternities in one family were demonstrated, with potential segregation distortions. The Journal of Immunology, 2002, 168: 771–781.

Major histocompatibility complex class I and class II molecules are fundamental components of the adaptive immune system. Antigenic peptides, bound and presented to T cells by class I molecules, are produced by the cytosolic proteasome. In mammalian cells stimulated by IFN-γ, three components of the constitutive catalytic proteasome core, X/PSMB5/β5, Y/PSMB6/β1, and Z/PSMB7/β2, are replaced by LMP7/PSMB8/β5i, LMP2/PSMB9/β1i, and MECL-1/PSMB10/β2i, respectively, to produce immunoproteasomes, which are believed to generate peptides suitable in length and sequence for binding to class I molecules (2–4). Peptides transported into the endoplasmic reticulum via a heterodimer of TAP proteins are loaded onto class I molecules, which then are competent to be expressed at the cell surface (5).

In eukaryotes, one of the first examples of genetic linkage between functionally related, yet structurally unrelated loci was shown for the class I Ag pathway genes, LMP2, LMP7, TAP1, and TAP2, which are closely linked and encoded in the MHC of all vertebrates studied to date (6–17). In mammals, these genes are located in the class II region, while both classical and nonclassical class I genes are encoded in another region of the MHC. The MHC organization of chickens (11) and all teleost fish (9, 13–16) examined is quite different from mouse and human: classical class I genes are closely linked to proteasome and transporter genes (except in chicken, the inducible proteasome elements have been deleted from the MHC and perhaps from the genome entirely). Such close linkage among class I and the functionally related proteasome and transporter genes has been proposed to be a true class I region, which may allow coevolution of these components (9–11). The class I region in teleosts is further highlighted because class II genes are found in another region of the genome entirely, so that there are actually (at least) two MHCs in the bony fish (14, 18). Because class I is closely linked to proteasome/transporter genes in all nonmammalian species studied to date, it is likely that this MHC organization is primitive.

In addition to differences in MHC genomic organization in birds and teleosts, suggesting coevolution of class I pathway genes, there are lineages of TAP alleles that have coevolved to function with sets of class I alleles in the rat (19), hamster (20), and probably the chicken (11, 21). In the frog, *Xenopus* allelic lineages of TAP and LMP7 are linked to particular class I lineages (22–24). Coevolution of the proteases, transporters, and presenting molecules in the class I Ag presentation pathway may be the selecting force to preserve linkage of the genes in many vertebrate species. Consistent with this idea is the relatively low number of class I genes that are linked to the MHC proper in most nonmammalian vertebrates (excluding the cichlid (25, 26), the axolotl (27, 28), and perhaps the cod (29)), i.e., a stable class I region.

Cartilaginous fish (sharks, skates, rays, and chimeras) are derived from the oldest vertebrate ancestor with an adaptive immune system, and as such are important models for study of the MHC.
(7). Previously, we have shown that, in contrast to teleost fish, class I and class II genes are closely linked in sharks (1). In addition, previous work demonstrated the existence of TAP1, TAP2 (30), LMP7 homologues (31), and a proteasome gene unique to the cartilaginous fish, called LMP7-like (31). Based on differences in their catalytic sites, it was suggested that proteasomes containing LMP7 or LMP7-like might provide a diverse repertoire of peptides for class I molecules, consistent with the idea that certain class I alleles could bind peptides preferentially produced by one or the other enzyme. In this study, we begin to focus on whether the class I region found in all teleosts and chickens (and probably amphibians) exists in cartilaginous fish by investigating the linkage between class I and class I-related genes (LMP2, LMP7, LMP7-like, TAP1, and TAP2). We also analyze the number of class I genes and their linkage to each other and to other MHC-encoded genes.

Cosmid library screening

The cosmid clone 32 was digested with BamHI, SacI, ScaI, XhoI, and all possible combinations of two enzymes, and Clal, SmaI, BamHI, and SacI were used for digestion of cosmid 19 and 65. Gels were stained with ethidium bromide to visualize DNA fragments. Membrane was sequenced with either 32P-labeled double-strand probes (labeled with the random-primed labeling method; Boehringer Mannheim, Indianapolis, IN) or end-labeled exon-specific oligonucleotides to identify the digested fragments. Cosmid DNAs were also partially digested with RE and hybridized with end-labeled T3 or T7 oligomers to construct restriction maps. The LMP7-positive BamHI-XhoI 4.5-kb fragment (cosmid 32) and the Clal-Smal 1.9-kb fragment (cosmid 65) were subcloned into the pBluescript cloning vector and sequenced to identify the exons. For comparison with hybridizing genomic DNA fragments, 10 μg of cosmid DNA was digested with 15 U of RE for 3 h, and digests were checked on a 1% agarose gel and diluted to convenient concentrations; ~50 ng of cosmid DNA was loaded onto the gel compared with 5 μg of genomic DNA. A class I α1 domain-specific primer (5′-TCTCACAGTCTCCCGTATT3′) and a highly conserved α3 domain primer (5′-CCATCGTGTAACACAGACAGA-3′) were used to amplify the class I sequence from the cosmid 32 DNA; PCR was performed for 4 min at 94°C, followed by 35 cycles of 94°C for 1 min, 56°C for 1 min, 72°C for 3 min, and a final extension at 72°C for 15 min. A 1.5-kb fragment was amplified, cloned into pCR2.1 TA cloning vector (Invitrogen, San Diego, CA), and sequenced.

LMP2 isolation

PCR via 5′ RACE with a degenerate primer was performed to isolate the LMP2 gene. The primer 5′-GCTGCTAATCTGCTCCT-3′, in which R is A or G, N is A or G or C or T, W is A or T, D is G or A or T (aa position 135–141 (33), Fig. 1A), and the T3 primer on the Uni-ZAP phage vector (Stratagen, La Jolla, CA) were used with a horn shark (Heterodontus francisci) spleen cDNA library as template (34). PCR was performed for 4 min at 94°C, followed by 35 cycles of 94°C for 30 s, 45°C for 1 min, 72°C for 1 min, and a final extension at 72°C for 15 min. A single 600-bp fragment was amplified, cloned into the pCR2.1 TA cloning vector (Invitrogen, San Diego, CA), and sequenced.

cDNA library screen

The 600-bp LMP2 PCR fragment was used to screen the horn shark spleen cDNA library. This horn shark library was also screened with the full-length nurse shark LMP7 probe (accession D64057). A full-length LMP2 insert was used to probe a nurse shark spleen cDNA library prepared from siblings 34 and 37 (35). This nurse shark spleen cDNA library was also screened with horn shark TAP2 (AF108387, AF108386) and TAP1 (AF108385) (30) to obtain the orthologous nurse shark cDNAs, and with nurse shark nonclassical class I (UAA-NCl-1) using a probe encoding the α1 and α2 domains (AF220360) to obtain full-length cDNAs. All cDNA screenings were done under high stringency conditions (34).

Sequence alignment and phylogenetic tree

The mature LMP2 and LMP7 protein sequences were aligned using Clustal W with minor adjustments, and an unrooted tree was made using the TreeView 1.5 program (36).

Materials and Methods

Animals

A pregnant nurse shark (Ginglymostoma cirratum, order Orectolobiformes) was captured off Little Torch Key, FL, in October 1998, and 39 pups were delivered by Caesarean section (1). Nurse sharks are ovoviviparous, i.e., maintain their young within the uterus, where they are nourished by egg yolk; the size of the yolk sac was an indication of the pups’ ages (32).

Shark Yellow used for preparation of the cosmid library is maintained in an outdoor facility in Miami, Florida, with running seawater.

Southern blotting

A total of 1–3 ml blood was collected from each pup from the caudal vein, and DNA was prepared from nucleated erythrocytes, as described previously (37). Genomic DNA (5 μg) was digested with various restriction enzymes (RE) to obtain useful restriction fragment-length polymorphism (RFLP). Full-length nurse shark cDNA inserts were used as probes, except for TAP2 (it lacked the region encoding the N-terminal 160 aa (30)) and UAA-NCl (the α1 domain exon was used), and all hybridizations were performed under high stringency conditions.

The nurse shark erythrocute genomic DNA cosmid library prepared from shark Yellow was custom-made (SuperCos1 cosmid library; Stratagene) and has average insert sizes of ~40 kb. Membranes from library lifts were prepared per the manufacturer’s recommendation with minor modifications (38), and ~800,000 clones were screened with the full-length nurse shark LMP7 (D64057 (31)) or class Ia probes (UAA01: AF220063 (1)) under high stringency conditions.

Restriction mapping and subcloning

To obtain the transmembrane/cytoplasmic region genomic sequence for the UAA-NCl gene, a UAA/UAA-NCl exon 5 primer, 5′-TGAACAAAAGAC CAC-3′, and a UAA-NCl-exon 6 primer, 5′-TGAGATTGAT CAGGTGTCCTG-3′, were used. PCR was performed at 94°C for 4 min, followed by 35 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 3 min, with final extension of 72°C for 15 min using 2–500 ng “mother” genomic DNA as template. A 1-kb product was cloned into pCR2.1 TA cloning vector (Invitrogen) and sequenced. To obtain the UAA transmembrane/cytoplasmic genomic sequence, PCR was performed using cosmid 65 DNA as template. The same exon 5 primer was used as the forward primer with a UAA-specific reverse primer: 5′-AAGGTTTAAATCGCAGCTGCTG-3′. PCR was performed as above. The 2.5-kb product was cloned into the pCR2.1 TA cloning vector and sequenced. The PCR-amplified fragment for UAA exon 5 and intron 5 of cosmid 19 was also sequenced.

PCR of genomic DNA

To obtain the transmembrane/cytoplasmic region genomic sequence for the UAA-NCl gene, a UAA/UAA-NCl exon 5 primer, 5′-TGAACAAAGAC CAC-3′, and a UAA-NCl-exon 6 primer, 5′-TGAGATTGAT CAGGTGTCCTG-3′, were used. PCR was performed at 94°C for 4 min, followed by 35 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 3 min, with final extension of 72°C for 15 min using 2–500 ng “mother” genomic DNA as template. A 1-kb product was cloned into pCR2.1 TA cloning vector (Invitrogen) and sequenced. To obtain the UAA transmembrane/cytoplasmic genomic sequence, PCR was performed using cosmid 65 DNA as template. The same exon 5 primer was used as the forward primer with a UAA-specific reverse primer: 5′-AAGGTTTAAATCGCAGCTGCTG-3′. PCR was performed as above. The 2.5-kb product was cloned into the pCR2.1 TA cloning vector and sequenced. The PCR-amplified fragment for UAA exon 5 and intron 5 of cosmid 19 was also sequenced.

Abbreviations used in this paper: RE, restriction enzyme; RFLP, restriction fragment-length polymorphism; UTR, untranslated region.

FIGURE 1. A, Comparison of the amino acid sequence of nurse shark (N-shark) (AF357920) and horn shark (H-shark) (AF357921) LMP2 with LMP2 from human (U01025), mouse (D43620), Xenopus (D87687), trout (AF115541), Fugu (AF217723), and teleost fish (AF155579, LMP2B: AF155580; LMP2A: AF155581). B, The TAP1 and TAP2 from Fugu (AF217723), trout (AF115540), and zebrafish (AF155576); and with δ from human (D29012), mouse (U13394), Xenopus (D87689), (AF303292), and lamprey (D87690). Dashes and dots show sequences identical to human LMP2 and gaps, respectively. The mature protein begins at +1, and the primer position used for 5′ RACE is indicated with an arrow above the sequence. The percentage of similarity to nurse shark LMP2 is shown at the branch of each sequence. B, LMP2 phylogenetic tree using the mature protein. The scale shows the genetic distance. Bootstrap values were obtained with a total of 500 runs and are shown at the branches. C, LMP7/XLMP7-like phylogenetic tree using the mature protein. GenBank accession numbers are human (Z14982), mouse (U22033), rat (P28064), Fugu (AF217723), zebrafish (AF302390), N-shark (D64057), salmon (AF184938), Xenopus7B (I51537), Xenopus7A (I51536), H-shark (AF365383) LMP7, N-shark (D64056), H-shark (AF365382) LMP7-like, hagfish (D64054), lamprey (D64055), medaka fish (D89725), tuna (X97729), nurse shark (D64058), human (P28074), mouse (O55234), and chicken (X57210). X. Scale as in B.

Downloaded from http://www.jimmunol.org/ by guest on October 29, 2017
Northern blotting

RNA was isolated from various tissues using the TRIzol reagent (Life Technologies, Rockville, MD). Total RNA (20 μg) was electrophoresed in 1% agarose gels using MOPS as running buffer (Quality Biologicals, Gaithersburg, MD), and then transferred to an Optitran nitrocellulose membrane (Schleicher & Schuell, Keene, NH). The membranes were hybridized with UAA 3′ untranslated region (UTR), UAA-NC1 (α1 domain and intron 2), and nucleoside diphosphate kinase probes (39) on the tissue distribution blot. The 3′ UTR of LMP7 and LMP7-like were used as probes with RNA made from PBL of 8 sibling sharks, shark Yellow, and 11 wild animals. All blotting was done under high stringency conditions (34).

RT-PCR

cDNA was prepared from 5 μg of total RNA from shark Yellow PBL with the Superscript premiocrination system for first-strand cDNA synthesis kit (Life Technologies). PCR was then performed by 3′ RACE using a 5′/3′ domain forward primer, 5′-TCTCACAGTCTCGTATTT-3′, and the RACE reverse primer, 5′-CTCCAGAAGCTGTTCCGATCC-3′ at 94°C for 4 min, followed by 35 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 4 min, followed by 35 cycles of 94°C for 1 min, 72°C for 15 min final extension. A 1.7-kb single band was cloned into pCR2.1 TA cloning vector (Invitrogen) and sequenced.

Results

Isolation of shark LMP2 and horn shark LMP7-like

A 600-bp PCR product was amplified by 5′ RACE from a horn shark spleen cDNA library with a primer specific for shark LMP2 genes (33). The sequence of this product showed high similarity to all known LMP2 genes, and it was used to screen the same library to isolate LMP2 cDNA clones. The longest horn shark clone (1044 bp, with a 654-bp open reading frame, 304-bp 5′ UTR, and 863-bp 3′ UTR, Fig. 1A) was then used to screen a nurse shark spleen cDNA library. One full-length clone was sequenced and consisted of a 119-bp 5′ UTR, 61-bp 3′ UTR, and 654-bp open reading frame encoding 217 aa (Fig. 1A). The horn and nurse shark clones are 81% identical at the nucleotide level, and the deduced mature nurse shark protein sequence (start at +1 in Fig. 1A) is 87% identical to horn shark LMP2, 67–73% to LMP2 of other species, and 54–62% to proteasome δ (δ throughout the text). Nurse and horn shark LMP2 group with other species’ LMP2, whereas LMP2/δ (an intermediate found only in teleosts (14, 15, 40, 41)) and δ form separate clusters (Fig. 1B).

The nurse shark LMP7-like gene was described previously (31). To determine whether this gene is expressed by other elasmobranchs, we screened a horn shark spleen cDNA library and isolated the orthologous LMP7-like. Rather like the teleost LMP2/δ, the sequence of the resulting horn shark clone forms a cluster with nurse shark LMP7-like in phylogenetic trees, which is equidistant from the LMP7 and X clusters (Fig. 1C). This position of the LMP7-like cluster in the tree suggests that the gene arose not long after the divergence of LMP7 and X, and indeed it may be present in all extant cartilaginous fish.

Linkage studies of TAP, LMP, and class I

In the family of 39 nurse sharks, 13 segregating groups (a–m) were identified previously from classical class I (UAA) RFLP typing (1). Now we have extended the linkage analysis for LMP2, LMP7 (31), LMP7-like (31), TAP1, and TAP2 (30). Representative Southern blots of 20 of these siblings are shown in Fig. 2, and summaries of all experiments are shown in Tables I and II. TAP1 provided limited RFLP and formed only three groups (denoted a–c under the TAP1 blot in Fig. 2. a = U/L, b = L, c = U; U and L designate the upper and lower bands, respectively, in Tables I and II), but each group matched perfectly to the class I groups. RFLP for LMP2, LMP7 (EcoRI), and TAP2 were found only for paternal MHC haplotypes within groups d, e, g, and h (marked with dashes at the side of each blot, and see Table II). One weakly hybridizing LMP7 band was shown not to be MHC linked (BamHI digest in Fig. 2).

On the LMP7-like blot, there are two sets of segregating loci (indicated as 1 and 2). The three groups of 1 bands (a = U/L, b = U, c = L, shown on the bottom of the blot) matched the class

![FIGURE 2.](http://www.jimmunol.org/)  
Linkage of LMP7, LMP2, LMP7-like, TAP1, and TAP2 to the nurse shark MHC. Southern blot analyses with 20 siblings from the family with 39 offspring (1). RE were: BamHI for TAP1; BamHI/EcoRI for LMP7-like; PvuII for TAP2; and LMP2, EcoRI, and BamHI for LMP7. Lower case letters at the side of each blot indicate MHC groups, and large arrows indicate the non-MHC-linked bands. Two loci revealed on the LMP7-like blot are indicated as 1 and 2, and groups are shown under the blot. Segregants sorted with previous UAA typing (1) are shown above the TAP1 blot as MHC groups. TAP1 groups are shown under TAP1 blot. M, mother; Sib, siblings (numbers indicate individuals from Ref. 1). Band sizes are as follows: TAP2: e, g, 9 kb; d, h, 6.5 kb. Non-MHC: 5, 3.5 kb; e, d, g, h, 2.5 kb. LMP7: EcoRI, 3 kb; BamHI, 9 kb. TAP1: 10, 9 kb. LMP2: e, d, g, h, 6.6 kb; e, g, 6 kb. LMP7-like: 1, 5.3 and 4.8 kb; 2, 6.3 and 4 kb.
I MHC groups perfectly, and the presence vs absence of the \( \text{set} \) set of precisely matched paternal haplotypes in groups d, e, g, and h; thus, both the \( \text{F} \) and \( \text{G} \) \( \text{LMP7-like} \) genes are MHC linked, but only one set of bands is found in all haplotypes (and see later).

RFLP of all 39 siblings are summarized in Table I, in which results with the maternal \( \text{UAA} \) class I a1 domain are shown to illustrate the inheritance of maternal haplotypes identified definitively from siblings having distinctive paternal bands (1). The predicted maternal haplotype \( \text{m1} \) was in 21 offspring and is \( \text{L} \) (\( \text{UAA} \)), \( \text{U} \) (\( \text{TAP1} \)), and \( \text{U} \) (\( \text{LMP7-like} \); \( \text{L} \)); \( \text{m2} \) was in 21 offspring and is \( \text{U} \) (\( \text{UAA} \)), \( \text{L} \) (\( \text{TAP1} \)), and \( \text{L} \) (\( \text{LMP7-like} \); \( \text{L} \)); neither maternal haplotype contains the \( \text{LMP7-like} \) set of bands. Unfortunately, no \( \text{LMP2} \), \( \text{LMP7} \), or \( \text{TAP2} \) RFLP were detected for the maternal haplotypes; nevertheless, as mentioned above, several paternal haplotypes did have distinctive RFLP mapping to MHC for all of these genes.

**Sorting out paternal haplotypes**

Twelve paternal haplotypes were detected (p1 through p12), demonstrating that there were at least six fathers of this family (Table II). Besides the MHC-linked genes, we used the non-MHC-linked \( \text{LMP7} \) RFLP (Fig. 2) and other nonlinked genes (not shown), as well as yolk sac and body size, to determine haplotype pairs for each father. The p1 and p2 haplotypes had RFLP for the nonlinked \( \text{LMP7} \), strongly suggesting that they were derived from one father (designated father 1, responsible for 17 of 39 offspring). In addition,

<table>
<thead>
<tr>
<th>Siblings</th>
<th>Mother 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex ( \alpha )</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Yolk sac diameter (cm)</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>1.2</td>
<td>0.6</td>
<td>2.5</td>
<td>2.5</td>
<td>1.8</td>
<td>2.5</td>
<td>0.6</td>
<td>2.5</td>
<td>0.6</td>
<td>1.8</td>
<td>1.2</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body length (cm)</td>
<td>23</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>23</td>
<td>25.5</td>
<td>28</td>
<td>23</td>
<td>31</td>
<td>24</td>
<td>29</td>
<td>27</td>
<td>25.5</td>
<td>27</td>
<td>25.5</td>
<td>25.5</td>
<td>25.5</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>MHC group ( \beta )</td>
<td>a</td>
<td>b</td>
<td>a</td>
<td>c</td>
<td>d</td>
<td>e</td>
<td>f</td>
<td>g</td>
<td>a</td>
<td>c</td>
<td>h</td>
<td>c</td>
<td>d</td>
<td>i</td>
<td>d</td>
<td>h</td>
<td>g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal haplotype</td>
<td>m1/m2</td>
<td>m1</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td>m1</td>
<td>m2</td>
<td></td>
</tr>
<tr>
<td>( \text{UAA} ) a1/1</td>
<td>L/L</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U/L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>U</td>
<td></td>
</tr>
<tr>
<td>( \text{TAP1} )</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td></td>
</tr>
<tr>
<td>( \text{LMP7-like} )</td>
<td>1</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td></td>
</tr>
<tr>
<td>( \text{TAP2} ) non-MHC (3.5 kb)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>( \text{LMP7} ) non-MHC (BanII)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>( \text{UAA-NC1} )</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td>U/L</td>
<td></td>
</tr>
</tbody>
</table>

\( \alpha \) F, female; M, male.  
\( \beta \) Data taken from Ref. 1.

<table>
<thead>
<tr>
<th>MHC Group</th>
<th>( \text{UAA} ) a1/Leader</th>
<th>( \text{LMP7} ) (EcoRI)</th>
<th>( \text{LMP7-like} )</th>
<th>( \text{TAP1} ) (9, 6.5, 2.5 kb)</th>
<th>Paternal-Maternal Haplotype</th>
<th>( \text{LMP7} ) non-MHC Total</th>
<th>Yolk Sac Average (cm)</th>
<th>Body Length Range (cm)</th>
<th>Predicted Father</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>U/U/U</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>p1-m1</td>
<td>2</td>
<td>3</td>
<td>2.3</td>
<td>23–24</td>
</tr>
<tr>
<td>f</td>
<td>U/U/U</td>
<td>-</td>
<td>+</td>
<td>L</td>
<td>p1-m2</td>
<td>1</td>
<td>1</td>
<td>2.5</td>
<td>23</td>
</tr>
<tr>
<td>d</td>
<td>L/L/L</td>
<td>U</td>
<td>+</td>
<td>+</td>
<td>p2-m1</td>
<td>5</td>
<td>8</td>
<td>1.7</td>
<td>23–27</td>
</tr>
<tr>
<td>h</td>
<td>L/L/L</td>
<td>U</td>
<td>+</td>
<td>+</td>
<td>p2-m1</td>
<td>3</td>
<td>5</td>
<td>1.4</td>
<td>23.5–27</td>
</tr>
<tr>
<td>j</td>
<td>L/L</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>p3-m1</td>
<td>0</td>
<td>2</td>
<td>0.3</td>
<td>27–28</td>
</tr>
<tr>
<td>e</td>
<td>L/U</td>
<td>U</td>
<td>+</td>
<td>L</td>
<td>p3-m1</td>
<td>0</td>
<td>5</td>
<td>0.5</td>
<td>25.5–28</td>
</tr>
<tr>
<td>g</td>
<td>L/U</td>
<td>U</td>
<td>+</td>
<td>L</td>
<td>p3-m1</td>
<td>0</td>
<td>2</td>
<td>0.9</td>
<td>25.5–28</td>
</tr>
<tr>
<td>d</td>
<td>L/U</td>
<td>-</td>
<td>+</td>
<td>U</td>
<td>p5-m2</td>
<td>0</td>
<td>3</td>
<td>1.4</td>
<td>24–25.5</td>
</tr>
<tr>
<td>b</td>
<td>L/L</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>p7-m1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>k</td>
<td>L/L</td>
<td>U</td>
<td>+</td>
<td>L</td>
<td>p8-m1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>l</td>
<td>M/L/L</td>
<td>U</td>
<td>+</td>
<td>+</td>
<td>p9-m1</td>
<td>0</td>
<td>1</td>
<td>2.5</td>
<td>25.5</td>
</tr>
<tr>
<td>m</td>
<td>L/L</td>
<td>L</td>
<td>-</td>
<td>-</td>
<td>p10-m1</td>
<td>0</td>
<td>1</td>
<td>0.3</td>
<td>27</td>
</tr>
<tr>
<td>c</td>
<td>L/L</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>p11-m2</td>
<td>0</td>
<td>3</td>
<td>0.2</td>
<td>25.5–29</td>
</tr>
<tr>
<td>g</td>
<td>L/U</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>p12-m2</td>
<td>0</td>
<td>1</td>
<td>0.6</td>
<td>25.5</td>
</tr>
</tbody>
</table>

\( \alpha \) Data taken from Ref. 1.

\( \beta \) The group b a1/leader band migrates slightly faster than groups i and j.
siblings belonging to the four groups, a, f, d, and h, were younger than the others, having larger yolk sacs and smaller body sizes. Unfortunately, no other haplotypes produced RFLP, or yolk sac and body size were not informative enough to discriminate the pairs (the e and g groups had the same paternal RFLP for all genes tested; however, yolk sac and body length in these two groups are distinct, suggesting different fathers, or the same father after a long refractory period). Interestingly, there is an allelic preference for father 1 since 76% (13 of 17) of the offspring have the p2 haplotype. We observed the same phenomenon in another family of 18 nurse shark pups: one haplotype was found in six siblings, and the other haplotype in a maximum of three pups (1).

Cosmid clones bearing class I and proteasome genes

A genomic cosmid library was prepared from shark Yellow to study physical linkage of the shark MHC genes. Three class I-positive clones (cosmids 19, 32, 65) were isolated that hybridized weakly to the full-length LMP7 probe, with cosmids 32 displaying the highest intensity signal. Comparison of BamHI digests of cosmids 32 and shark Yellow genomic DNA resulted in comigration of the strongly hybridizing band (the weakly hybridizing band is due to cross-hybridization to the vector) in the cosmids clone and the weakly hybridizing, non-MHC-linked band in genomic DNA (Figs. 3A and 2. LMP7 BamHI blot). Two LMP7-positive fragments were identified in cosmids 32, and sequence of the XbaI-BamHI fragment revealed LMP7 exon 4 (Fig. 3C, accession AF357928). Exon 5 was identified by Southern blotting and PCR with exon-specific primers and sequencing of the product (data not shown).

As mentioned above, we previously showed that there is only one class Ia gene per MHC haplotype in the nurse shark (1). Two cosmid clones (clones 19 and 65) were mapped (Fig. 3C) and shown to bear class Ia genes with sequences precisely matching the two expressed UAA cDNAs expressed by shark Yellow, the presumed alleles (UAA03 and UAA05, Fig. 4). To identify the weakly hybridizing LMP7 band in each cosmids, the 1.9-kb Clai/Smal DNA fragment was subcloned, and the sequence revealed an LMP7-like exon 1 (accession AF357927). Double digestions of the cosmid clones with BamHI and EcoRI demonstrated that the LMP7-like genes on cosmid clones 19 and 65 are the U and L bands, respectively, detected on the genomic blot in Fig. 2 (Fig. 3B). As expected from mapping of the cosmids (the polymorphic BamHI site downstream of LMP7-like exon 1), shark Yellow, like the mother of the family shown in Fig. 2, has both bands (U/L) (Fig. 3B). The bands present in shark Yellow and in some of the fathers of the family encode at least one complete LMP7-like gene (there are additional bands not shown in Figs. 2 and 3B). As with the LMP7 gene in cosmids 32, LMP7-like is a pseudogene, as only exon 1 was identified from sequencing analysis. In summary, the cosmid clones bearing the nurse shark MHC-linked classical class I genes were isolated, and such genes are closely linked to proteasome gene fragments.

To prove conclusively that some animals lacked a complete LMP7-like gene, we performed Northern blots with PBL RNA from eight siblings, three positive (groups d and h) and five negative for the set of bands (Fig. 2 and Table II), and with shark Yellow, which is also positive (Fig. 3B). Only those animals that were band positive on the Southern blot expressed LMP7-like transcripts, while all nine animals expressed LMP7 (Fig. 5A).

To further investigate expression of the LMP7 and LMP7-like genes, we performed Northern blots of 11 unrelated (to our knowledge) nurse sharks. As expected, the mother of the family lacking the set of LMP7-like bands detected by Southern blotting (Figs. 2 and 5B) did not express this gene (Fig. 5B). However, we were surprised to find that only four of the wild animals expressed LMP7, correlating with the presence of a 2.8-kb RFLP detected on the LMP7 Southern blot (arrow in Fig. 5B). Thus, from the family study and examination of the wild animals, functional LMP7 and LMP7-like behave like alleles (pseudoalleles), but LMP7 and...
LMP7-like gene fragments are nevertheless present in all haplotypes (Figs. 2, 5B, 9, and 10). The presence of two LMP7 alleles with quite different sequences is reminiscent of the situation in Xenopus, in which ancient LMP7 lineages have been detected (23) (see below).

Other nonclassical class I

In our previous study, we isolated a nonclassical class I gene (UAA-NC1, accession AF202360) by PCR with the primer set used to amplify UAA from genomic DNA (1). UAA-NC1-specific bands segregating in the family were detected (Fig. 6), but the UAA-NC1 groups matched neither the MHC groups nor the non-MHC-linked TAP2 and LMP2 RFLP (Fig. 2 and Table I). To further understand the UAA-NC1 protein structure, the nurse shark spleen cDNA library was screened, resulting in the isolation of a 2072-bp clone. Amino acid identity to Gici-UAA is 81% in leader, 75% in α1, and 89% in α3 domains (Fig. 4). There is a 186-bp short interspersed elements retrotransposon element in the 3′ UTR not detected in UAA, and this element may have been involved in movement of UAA-NC1 out of the MHC (42). The beginning of the transmembrane domain is highly conserved, but there is a 49-bp deletion compared with UAA, resulting in premature termination in UAA-NC1a, suggesting the presence of additional genes or gene fragments. Whether there is a second expressed LMP2 gene or gene fragment (Fig. 9). In groups d, e, g, h, k, and m, the LMP2 L band seems to segregate with the very top band (~15 kb in Fig. 2, e.g., compare siblings 4 and 5 in which this uppermost band is approximately half of the intensity in sib 5), but the LMP2 L band is found only in groups e, g, k, l, and m, suggesting the presence of additional genes or gene fragments. Whether there is a second expressed LMP2 gene in elasmobranchs, and whether this gene is related to the novel teleost LMP7-like gene is not found in all shark MHC. One set of haplotypes in 19 of 39 siblings has a functional, expressed LMP7-like gene and apparently more than one LMP2 gene or gene fragment (Fig. 9). In groups d, e, g, h, k, l, and m, the LMP2 U band seems to segregate with the very top band (~15 kb in Fig. 2, e.g., compare siblings 4 and 5 in which this uppermost band is approximately half of the intensity in sib 5), but the LMP2 L band is found only in groups e, g, h, k, l, and m, suggesting the presence of additional genes or gene fragments. Whether there is a second expressed LMP2 gene in elasmobranchs, and whether this gene is related to the novel teleost LMP7-like gene (14, 15, 40, 41) requires further study.

The LMP7-like gene, previously identified in nurse sharks (31), now has been found in horn sharks as well, and, based on the phylogenetic tree analysis (Fig. 1C), this gene arose early in evolution and may be found in all elasmobranchs. Despite its apparent early emergence, a complete LMP7-like gene is not found in all haplotypes, and the functional gene is likely to be allelic with LMP7 (Figs. 5, 9, and 10). It was already recognized that shark LMP7 and LMP7-like proteins differ in residues near their respective catalytic sites (31), implying that the spectrum of peptides

**FIGURE 4.** Amino acid comparison of all nurse shark class I and human HLA-Cw3. GenBank accession numbers are: UAA01 (AF220063), UAA02 (AF220358), UAA09 (AF220359), UAA05 (AF357925), UAA03 (AF357924), UAA-NC1 (AF220360), UAA-NC2 from cosmid 32 (AF357926), UCA01 (AF028557), and HLA-cw3 (M84172). Dashes and dots show sequences identical to UAA01 and gaps, respectively. Predicted peptide-binding residues and catalytic sites (48) are indicated as underlines and dots above and under the sequences, respectively. boldface, and the regions that differ between the two class Ia alleles are bracketed. Three stop codons in the cosmid 32 class I clone are shown as *.

![Image](http://www.jimmunol.org)
might be altered in animals that express both LMP7 genes compared with those with only one. Fig. 11 displays stereoimages of LMP7, LMP7-like, and the constitutive enzyme LMPX, modeled onto the crystal structure of yeast LMPX (45). Residues in red include the active site Thr-1, the obligatory (proton acceptor?) Lys33, and Met45, which is believed to coordinate chymotryptic-like cleavages (after hydrophobic residues) in this active site (46). Five residues in positions 28 (green), 29 (yellow), 30 (blue), 31 (black), and 32 (purple) (Ser-Gln-Thr-Val-Lys, Ser-Val-Asp-Ala-Asn, and Thr-Gln-Met-Phe-Lys in LMPX, LMP7, and LMP7-like, respectively) are diverse among the three proteins (31), and their positions relative to the active site are displayed. Note especially the proximity of the bulky Phe31 (black) to Met45 in LMP7-like as compared with Ala31 in LMP7 and Val31 in LMPX. We think that it would be surprising if these differences in the homologous catalytic subunits did not alter the spectrum of peptides produced by proteasomes containing them. Amazingly, for species separated from each other by ~500 million years, the two old allelic lineages of Xenopus LMP7 genes also differ from each other primarily in
the vicinity of the catalytic site, one lineage having Phe\textsuperscript{31} and the other Ala\textsuperscript{31} (Ref. 23 and Y. Ohta, S. J. Powis, and M. F. Flajnik, manuscript in preparation). Although we are just beginning our analyses of the inducible proteasome genes in nonmammalian vertebrates, based upon 1) the very old gene/allelic lineages for several MHC genes (22–24, 47); 2) different proteasome genes, in particular MHC haplotypes (23; this study); and 3) the close linkage of proteasome genes to classical class I (6, 11, 13, 15, 17), one can propose that the generation of diverse sets of peptides by immunoproteasomes may be more crucial for adaptive responses in these nonmammalian species than in mice or humans.

In addition to having more genes or gene fragments, RFLP for \textit{TAP2}, \textit{LMP2}, \textit{LMP7}, and \textit{LMP7-like} were more apparent in the group d, e, g, h, k, and l paternal haplotypes. One possible explanation for these findings is that there are only two major nurse shark MHC allelic lineages and only one lineage generates RFLP with the restriction enzymes that we used. To date, our analysis has shown two distinct class Ia molecules in all nurse sharks tested that are very different at the C terminus of the α1 helix, one having 77DILSK and the other 77KTALE (bracketed in Fig. 4). This polymorphic region itself could be recognized by different sets of NK cell receptors (e.g., see Ref. 48 for NK receptor interactions with a human killer inhibitory receptor), or perhaps different arrays of peptides generated, depending upon the constellation of inherited proteasome genes, may alter the orientation of this class I helical

**FIGURE 8.** UAA and UAA-NC1 expression in various nurse shark tissues. The ubiquitously expressed nucleoside diphosphate kinase was used as the positive control (39). Size markers shown on the right.

MHC haplotypes:

\[
p_4, p_5 \\
p_8, p_9 \\
p_2 \\
p_1, p_3, p_6, p_10, p_11, p_{12}, m_1, m_2
\]

**FIGURE 9.** Hypothetical organization of nurse shark MHC haplotypes. All genes identified in this report and previous reports are shown in boxes, and \textit{LMP7-like} and \textit{LMP7} pseudogenes are shown as dark-shaded boxes and \(\psi\), respectively. The second \textit{LMP2} gene (fragment) is shown as a lightly shaded box. The order of genes and intergenic distances are not to scale (and are unknown). \textit{TAP2}, \textit{LMP7}, and \textit{LMP2} were definitively shown to map to MHC only in haplotypes from the P2, P4, and P5 groups.

**FIGURE 10.** Postulated evolutionary progression of the pseudoallelic nature of \textit{LMP7} and \textit{LMP7-like}. This model is based upon the expression analyses (Fig. 5) and the fact that \textit{LMP7} and \textit{LMP7-like} gene fragments (i.e., incomplete genes) are detected in MHC haplotypes from which the functional, homologous gene is not present.

**FIGURE 11.** Stereoimage models of shark \textit{LMPX}, \textit{LMP7}, and \textit{LMP7-like} proteins, based upon the crystal structure of the yeast orthologue of \textit{LMPX} (45). PDB files were created in the program Homology Modeling (http://www.embib.kun.nl:1100/WIWWWI/). The models were generated in the three-dimensional imaging program MolScript (http://www.atom.se/molscript/). See text for details of designations for the particular amino acid residues.
p2 haplotype in the d and h groups had RFLP for LMP2, LMP7, and TAP2, and contains a functional LMP7-like gene. Perhaps such a haplotype provided an advantage to offspring through expression of the LMP7-like gene, the particular class I allele, or other genes in the haplotype, so that survival was enhanced, but there are many other explanations for this finding that must be explored.

To date, four class I genes have been isolated from the nurse shark, one classical gene, and three nonclassical class I genes. The nonclassical UCA (formerly Gic-11 (34)) was the first gene to be cloned, and it is expressed in a similar fashion to UAA; unfortunately, we have not been able to map this class I gene because of its low level of polymorphism. UAA-NC1, isolated by a PCR “accident” with primers designed for UAA, was shown in this study not to map to MHC, despite its high sequence similarity to UAA. Some UAA alleles are more similar to UAA-NC1 than others (seen as cross-hybridization in groups d and h in Fig. 6), suggesting that UAA-NC1 arose from one particular UAA lineage. The third nonclassical class I UAA-NC2 was found in a cosmid clone bearing a truncated LMP7 gene. This pseudogene complex was probably the result of recent duplication event, which indicates that MHC class Ia was closely linked to LMP7 in the MHC. In fact, all of the class I genes found on the cosmid clones isolated to date are linked to an LMP7 (or 7-like) gene fragment. In the teleosts Fugu (13, 40) and zebrafish (15, 41), class Ia genes are also physically linked to LMP7 genes (or gene fragments), suggesting that this close physical linkage is ancient.

Based upon Southern blotting analyses with a variety of probes under low stringency conditions, UAA may be the only class I gene linked to MHC. The same feature is found for the Xenopus MHC, in which only one class I gene of the classical type is MHC linked (49), while a large set of linked nonclassical class I genes is found in another gene complex half a chromosome away (50, 51). In chickens, two class I genes (one highly expressed) sandwich the TAP1 and TAP2 genes, and Kaufman and colleagues (11, 21, 52) have proposed that the tight linkage has allowed the particular class I and TAP alleles to coevolve; furthermore, like in frogs, several nonclassical chicken class I genes are found a great distance away from the true MHC on the same chromosome (53). In the banded houndshark Triakis scyllium, there are one or two classical class I genes depending on the haplotype examined (54), and the majority of teleost species examined rigorously (except cod (29) and cichlid (25, 26)) have relatively few class I genes in the class I region. Previously, others and we have suggested that close linkage of this gene, or other proteasome genes, to shark MHC implies a primordial organization.


References