Primordial Linkage of β2-Microglobulin to the MHC

Yuko Ohta, Takashi Shiina, Rebecca L. Lohr, Kazuyoshi Hosomichi, Toni I. Pollin, Edward J. Heist, Shingo Suzuki, Hidetoshi Inoko and Martin F. Flajnik

*J Immunol* 2011; 186:3563-3571; Prepublished online 14 February 2011;
doi: 10.4049/jimmunol.1003933

http://www.jimmunol.org/content/186/6/3563

**Supplementary Material**

http://www.jimmunol.org/content/suppl/2011/02/14/jimmunol.1003933.3.DC1

**References**

This article cites 47 articles, 14 of which you can access for free at:

http://www.jimmunol.org/content/186/6/3563.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
Primordial Linkage of \( \beta_2 \)-Microglobulin to the MHC

Yuko Ohta,†,1 Takashi Shiina,†,1 Rebecca L. Lohr,∗ Kazuyoshi Hosomichi,†,‡
Toni L. Pollin,§ Edward J. Heist,‡ Shingo Suzuki,† Hidetoshi Inoko,†,†,1
and Martin F. Flajnik*,1

\( \beta_2 \)-Microglobulin (\( \beta_2 \)M) is believed to have arisen in a basal jawed vertebrate (gnathostome) and is the essential L chain that associates with most MHC class I molecules. It contains a distinctive molecular structure called a constant-1 Ig superfamily domain, which is shared with other adaptive immune molecules including MHC class I and class II. Despite its structural similarity to class I and class II and its conserved function, \( \beta_2 \)M is encoded outside the MHC in all examined species from bony fish to mammals, but it is assumed to have translocated from its original location within the MHC early in gnathostome evolution.

We screened a nurse shark bacterial artificial chromosome library and isolated clones containing \( \beta_2 \)M genes. A gene present in the MHC of all other vertebrates (\textit{ring3}) was found in the bacterial artificial chromosome clone, and the close linkage of \textit{ring3} and \( \beta_2 \)M to MHC class I and class II genes was determined by single-strand conformational polymorphism and allele-specific PCR. This study satisfies the long-held conjecture that \( \beta_2 \)M was linked to the primordial MHC (Ur MHC); furthermore, the apparent stability of the shark genome may yield other genes predicted to have had a primordial association with the MHC specifically and with immunity in general. The \textit{Journal of Immunology}, 2011, 186: 3563–3571.
Materials and Methods

Animals
Genomic DNA was isolated from RBCs for mapping analysis from the nurse shark family as previously described (21). The procedure of animal use was reviewed and approved by the Institutional Animal Care and Use Committee at the University of Maryland.

Bacterial artificial chromosome library screening
The 17 bacterial artificial chromosome (BAC) filters with 11-fold genomic coverage (22) were screened with radiolabeled full-length β2M or ring3 probes under high-stringency conditions (23). Membranes were exposed to x-ray film for various lengths of time to obtain positive signals and the desired background. Putative positive clones were then re-spotted on nylon membranes for colony hybridization and tested by Southern blotting to confirm true positives. BAC insert DNA was isolated using the PhasePrep BAC DNA kit (Sigma-Aldrich), and the sequence was determined by shotgun sequencing at the sequencing facility at Tokai University with 7.5× coverage.

Sequence alignment and phylogenetic tree
Amino acid sequences of constant-1 (C1)-IgSF domains were aligned using the ClustalX2 program with minor adjustments. A rooted neighbor-joining (NJ) bootstrapped (1000 runs) phylogenetic tree (24) was constructed, and the consensus tree was then viewed with the TreeView program (25).

Database searches
Genome synteny in various species was retrieved and analyzed from publicly available Web sites as noted. Genes from mouse, chicken, human, opossum, and zebrafish were retrieved from GenBank (http://www.ncbi. nlm.nih.gov), and information on other genomes was retrieved from the following Web sites: elephant shark genome (http://blast.fugu-sg.org/); Anolis genome (http://genome.ucsc.edu/cgi-bin/hgGateway?db=anoCar1); Xenopus genome (http://genome.xji-psf.org/Xenon/Xenotax.home.html); and Fugu genome (http://genome.jgi-psf.org/Takru4/Takru4.home.html).

In-house EST collection
We constructed the cDNA library using the Gateway System (Invitrogen) from adult nurse shark pancreas. To eliminate Ig genes, we first hybridized with Ig H and L chain probes under high-stringency conditions. Negative colonies (<8000) were then manually picked and sequenced from the vector end. All draft sequences were blast searched against GenBank databases, and we obtained ~1150 sequences not specific to the pancreatic enzymes (Y. Ohta and M.F. Flajnik, personal observations).

Single-strand conformation polymorphism analysis
Nurse shark ring3 primers were designed based on the sequence obtained from BAC GC 614H19 clone. Multiple primers were tried, and we selected the primer set anchoring exons 4 and 5 for the single-strand conformation polymorphism (ssCP) analysis. The primers were exon 4 forward, 5′-GTAAACACCTGCACCAAAAT-3′; and exon 5 reverse, 5′-ATTTGGGACCCTGAGACACAGT-3′. PCR amplification was performed for 4 min at 94°C, followed by 35 cycles of 94°C for 1 min, 56°C for 1 min, and 72°C for 1 min, and a final extension at 72°C for 10 min. An ~550-bp fragment amplion was cloned into the pCRII TA cloning vector (Invitrogen), and individual clones were sequenced. Nurse shark families 2 and 3 were genotyped using 12 DNA microsatellite markers and assigned sires (E.J. Heist, J.C. Carrier, H.L. Pratt, and T.C. Pratt, submitted for publication).

Statistical analysis of linkage
We used parametric linkage analysis to formally assess the evidence for linkage of β2M to the MHC region in the offspring of deletion-carrying sires. This approach assesses the odds of the likelihood of obtaining the observed data set if the two loci are linked versus if the loci are not linked, showing as a log of the odds (LOD) score. The paternal sibships were determined based on consolidated data from combination of Southern blotting, sequencing of MHC class I alleles, and microsatellite analyses (shown in Table 1).

The LOD score is calculated as follows when parental phase (linkage status) is known: LOD = log10{[(P × (1 − P))/((1 − P) × P)]−NR}, where P is the recombination fraction, NR is the number of nonrecombinant offspring, and R is the number of recombinant offspring.

Because the parental phase was unknown in the current study due to a lack of grandparental genotypes, a phase ambiguous LOD score was first calculated for each family by taking the log of the average odds for the two possible phases (1 and 2 in Table 1), and the resulting LOD scores were then summed over the two families to obtain the LOD score at a given recombination fraction. LOD scores were calculated at recombination fractions between 0 and 0.5 to obtain the recombination fraction where the LOD score was maximized (26). The corresponding p value was calculated using a one-sided χ2 test of LOD × 2 (log10) (27).

Results
Characterization of nurse shark β2M
Cartilaginous fish are the oldest living vertebrates having an adaptive immune system centered upon Ig, TCR, and MHC (1). When it was suggested that class I and class II genes may have evolved in separate linkage groups from studies of teleost fish (28), we demonstrated in family studies that the two MHC classes were closely linked in two shark species, nurse shark and houndshark (21). To gain further insight into the primordial MHC organization, we have isolated many shark genes associated with adaptive immunity, including β2M. The full-length β2M clone was found in an in-house EST collection (GenBank accession number HM625831), as well as from a previously published genomic sequence (GenBank accession number GQ865623) (29), and the deduced amino acid sequence was aligned with β2M from other species (S1). As was noted in previous studies, evolutionarily
conserved residues are either found in all C1-IgSF (or just IgSF) domains (29, 30) or are predicted to be at class Ia-chain interaction sites (31). Some cartilaginous fish β2M have potential N-glycosylation sites that are rare in tetrapods but present in several bony fish species (32). Consistent with previous studies (33, 34), phylogenetic tree analysis revealed that cartilaginous fish β2M clustered with the orthologous proteins and to the IgSF domains of MHC class IIA/DMA, suggesting that they share the most recent common ancestor (Fig. 1A). Also consistent with previous studies (33), the IgSF domains of class IIB and class Ia shared the most recent common ancestor. β2M expression pattern seems to coincide with MHC class I expression (Fig. 1B).

Mapping of β2M to the MHC in family studies

Two families of nurse sharks previously were used to map several genes to the MHC (21, 36, 37). All of these families showed multiple paternity, at least five fathers in family 1 and seven in family 2. Southern blotting analysis using many restriction enzymes demonstrated that β2M is a single-copy gene (five representative digestions are shown in Fig. 1C); unfortunately, no RFLPs were obtained to test the linkage status, and thus we sequenced the gene from animals with different MHC haplotypes, hoping to find polymorphisms. A two-nucleotide deletion was detected in one of the paternal β2M alleles “p3” from groups “i” (p3/m2) and “j,” (p3/m1) from family 2 with 39 members (Fig. 1B).

FIGURE 1. A, Phylogenetic tree analysis of β2M. GenBank accession numbers used for this analysis are as follows. β2M: M17987 (human), X69084 (bovine), NM_009735 (mouse), Y00441 (rat), P01885 (rabbit), P01886 (guinea pig), M84767 (chicken), P21612 (turkey), AAM9836 (opossum), BQ389924 (X. tropicalis), AAF37230 (X. laevis), L05536 (carp), NP_571238 (zebrafish), L63534 (trout), CAI0761 (cod), AAG17535 (salmon), CAB61324 (Siberian sturgeon), AAN40738 (Japanese flounder), CAD44965 (African barb), O42197 (catfish), CA30181 (Fugu), AAN62852 (skate), and CX197532 (dogfish). Class Ia: AA966123 (nurse shark), AAL59857 (nurse shark), NP_001001762 (chicken), NP_001001762 (chicken), and NP_942035 (rat). DMA: NP_006111 (human), NP_006111 (human), NP_942035 (rat). DMG: NP_006111 (human), NP_942035 (rat), DMA: NP_006111 (human), NP_942035 (rat), ACY01474 (chicken), XP_001377359 (opossum). The NJ tree was rooted with the fourth constant IgSF domains of IgM, and bootstrapping analysis was done after 1000 runs. Values are noted at the branch nodes, and the asterisk (*) indicates no significant value. The scale indicates divergence time (genetic distance). Teleost fish that underwent a third round of genome expansion (“3R”) are omitted from this analysis because the sequences were more divergent and skewing the tree topology. DM genes have not been identified in any fish.
FIGURE 2. The shark $\beta 2M$ is linked to the MHC. A. The two-nucleotide (CC) deletion polymorphism was found in intron 2 of $\beta 2m$ sequences in "p3" paternal allele from siblings belonging to the groups "i" and "j." Thus, allele-specific primers were designed based on this polymorphism. All primers are underlined. The ends of coding regions are boxed. The (AG) at the end of intron 2 is underlined. 

B. PCR was carried out with a combination of allele-specific and universal NSB2mEx3Rev reverse primers. Presence or absence of the amplicon using the "p3"-specific primers was used for typing (top gel) the family 2 with 39 offsprings. Maternal primers were used for the positive control (bottom gel). Forward primers are indicated on the left side of the gels, and mother and sibling numbers are indicated above the gel along with MHC groups (36).

C. Allele-specific PCR in the families 1 and 3. Only two animals belonging to the MHC groups "h" possessed the "CC-deletion" allele, and two animals belonging to the "g" groups had this allele in family 3. We partially typed family 3 based on the MHC groups by sequencing of the PBR of the class Ia alleles (maternal and paternal alleles are designated as numbers above the gel) and by Southern blotting with a probe containing MHC class Ia leader and $\alpha 1$ domains (small dot, band for maternal haplotype 1; large dot, maternal haplotype 2). The "p2" allele of the "g" group is the only haplotype possessing the "CC-deletion" allele of $\beta 2M$. D. Plot of LOD scores at corresponding recombination fractions. The sums of the two families were used (Supplemental Table I).
formally the evidence for linkage of a paternal intra-MHC recombination event in sib 36. To quantify "i" and "j" groups; thus this father had the MHC haplotypes shown to have been sired by the same father as offspring in the "i" versus no linkage, equivalent to obtained a maximum LOD score of 3.14 [1378:1 odds of linkage included in the analysis. We performed a parametric linkage analy-
microsatellite-characterized, and therefore family 1 was not in-
submitted for publication) (Table I). Family 1 sires have not been
(36), sequences of MHC class I alleles (Fig. 2C), Table I), and microsatellite analysis (E.J. Heist et al., submitted for publication) consistent with a paternal intra-MHC recombination event in sib 36. To quantify formally the evidence for linkage of \( \beta 2M \) to the MHC, we considered all offspring of the two deletion-carrying sires (found within families 2 and 3) as assigned by Southern blotting with class I probes (Fig. 2C) (36), sequences of MHC class I alleles (Fig. 2C, Table I), and microsatellite analysis (E.J. Heist et al., submitted for publication) (Table I). Family 1 sires have not been microsatellite-characterized, and therefore family 1 was not included in the analysis. We performed a parametric linkage analysis (26) to evaluate the evidence for \( \beta 2M \) and MHC synteny and obtained a maximum LOD score of 3.14 [1378:1 odds of linkage versus no linkage, equivalent to \( p = 7 \times 10^{-7} \) (27)] at a \( \theta = 0.056 \) (Supplemental Table I). \( \beta 2M \) is adjacent to MHC-linked Ring3

**Ring3** (or **BRD2**) is a putative nuclear transcriptional regulator and a nuclear kinase required for early development (38–41) with no defined immune functions but nevertheless linked to the MHC of all other gnathostomes and to the "proto-MHC" in lower deuter-ostomes (42). A portion of **ring3** was initially cloned via de-generate PCR from nurse shark spleen cDNA, and this short fragment was used as a probe to isolate a full-length CDNA from a phage library. BLAST searches and phyllogenetic tree analysis confirmed the orthology of nurse shark **ring3** to that of other species (GenBank accession number HM625830) (Fig. 3A). The nurse shark **ring3** is ubiquitously expressed (Fig. 1B). To ensure that the shark **ring3** is linked to the MHC as in all other species examined (8), we performed ssCP analysis using siblings of family 2 (Fig. 3B). Two distinguishing **ring3** bands corresponding with the maternal MHC allele \( m2 \) were found in those siblings possessing this allele (groups "i" and "d" in Fig. 3) with 100% fidelity, demonstrating that **ring3** is closely linked to the MHC and further confirming the \( \beta 2M \) linkage. We identified other BAC clones that were either \( \beta 2M \)- or **ring3**-single-positive; unfortunately, none of them was positive for other MHC genes, again consistent with larger intergenic distances in sharks compared with those of other species (36). Chen et al. (29) drew a premature conclusion of non-MHC linkage; however, determining the linkage status of \( \beta 2M \) (or almost any gene) based on a single BAC sequence is not sufficient for the shark genome, where there are large intragenic and intergenic distances. Several nurse shark BAC clones (22) were isolated with the **ring3** and \( \beta 2M \) probes, and some of them were positive for both genes. As previously reported (29), the \( \beta 2M \) gene contains at least three exons, having a similar genomic organization and size to other species. The shark **ring3** gene spans \( \sim 20 \) kb and contains 12 exons, which is approximately twice as large as mammalian **ring3** genes (e.g., 12.8 kb and 9.7 kb for human and mouse, respectively), consistent with a larger gene size found in most shark MHC genes (36). Sequencing through an entire BAC clone (GC_614H19) confirmed that the \( \beta 2M \) and **ring3** genes were adjacent to each other \( \sim 45 \) kb apart (Fig. 4).

**Genetic descent of \( \beta 2M \)**
The chromosomal location of the \( \beta 2M \) gene varies greatly among vertebrate species (Fig. 5). Genomic synteny is well conserved in

---

**Table 1. List of sibs used for statistical analysis**

<table>
<thead>
<tr>
<th>Sib No.</th>
<th>Old Group</th>
<th>New Group</th>
<th>MHC Class Ia</th>
<th>( \beta 2m )</th>
<th>Sire</th>
<th>m-Satellite</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>i</td>
<td>i</td>
<td>m2/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>30</td>
<td>i</td>
<td>i</td>
<td>m2/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>21</td>
<td>j</td>
<td>j</td>
<td>m1/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>25</td>
<td>j</td>
<td>j</td>
<td>m1/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>31</td>
<td>j</td>
<td>j</td>
<td>m1/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>33</td>
<td>j</td>
<td>j</td>
<td>m1/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>39</td>
<td>j</td>
<td>j</td>
<td>m1/p3</td>
<td>del</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>20</td>
<td>e</td>
<td>e</td>
<td>m1/p6</td>
<td>ins</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>32</td>
<td>e</td>
<td>e</td>
<td>m1/p6</td>
<td>ins</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>36</td>
<td>e</td>
<td>e</td>
<td>m1/p6</td>
<td>del</td>
<td>4</td>
<td>R</td>
<td>NR</td>
</tr>
<tr>
<td>28</td>
<td>g</td>
<td>g</td>
<td>m2/p6(^b)</td>
<td>ins</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>13</td>
<td>c</td>
<td>g</td>
<td>m2/p6(^b)</td>
<td>ins</td>
<td>4</td>
<td>NR</td>
<td>R</td>
</tr>
</tbody>
</table>

**Family 3**

<table>
<thead>
<tr>
<th>Sib No.</th>
<th>Old Group</th>
<th>New Group</th>
<th>MHC Class Ia</th>
<th>( \beta 2m )</th>
<th>Sire</th>
<th>m-Satellite</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>g</td>
<td>m2/p2</td>
<td>del</td>
<td>2</td>
<td>NR</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>g</td>
<td>m2/p2</td>
<td>del</td>
<td>2</td>
<td>NR</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>d</td>
<td>m1/p4</td>
<td>ins</td>
<td>2</td>
<td>NR</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>d</td>
<td>m1/p4</td>
<td>ins</td>
<td>2</td>
<td>NR</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>d</td>
<td>m1/p4</td>
<td>ins</td>
<td>2</td>
<td>NR</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>d</td>
<td>m1/p4</td>
<td>ins</td>
<td>2</td>
<td>NR</td>
<td>R</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Old group is taken from Ref. 28, and new groups are assigned in this study.

\(^b\)MHC class Ia sequences revealed that sib 13 is further categorized with group g‘ in this study.

del, CC-deletion haplotype; NR, nonrecombinant; R, recombinant.
FIGURE 3. A, Phylogenetic tree analysis of Ring3 and homologues. GenBank accession numbers used in this analysis are as follows. Ring3 (BRD2): CAM25760 (human), AAY34703 (bovine), CAI11405 (dog). CAA15819 (mouse), CAEB3937 (rat), XP_001369391 (opossum), CAN12825 (pig), CAA65449 (chicken), BAC82511 (quail), AAI68574 (X. tropicalis), AAJ0180 (X. laevis), CAG04960 (zebrafish-1), CAD54663 (zebrafish-2), ABQ59964 (salmon), BAD82238 (medaka). Additional accession numbers for Ring3 homologues used for this analysis are the following: BRD3: AAI29055 (X. laevis), NP_031397 (human), NP_075825 (mouse), XP_001365890 (opossum), XP_425330 (Chicken). BRDT: NP_473395 (mouse), NP_997072 (human), NP_537079 (dog). BRD4: NP_490597 (human), NP_065254 (mouse), NP_001104751 (zebrafish), AAH76786 (X. laevis). BRD1: NP_001157300 (horse), XP_698063 (zebrafish), NP_0011085846 (X. laevis), CAG80294 (human). Gene names are noted after species name. BRD1 does not map to an MHC paralogous region, whereas BRDT, BRD3, and BRD4 are found in the MHC paralogous regions. The tree was constructed using the NJ method, rooted with BRD1, and bootstrapping analysis was done with 1000 runs. Values are noted at the branch nodes, and an asterisk (*) indicates no significant value. The scale indicates the divergence time. B, The shark ring3 maps to the MHC. Primers from exons 4 and 5 were used for PCR amplification and ssCP analysis. The ~1440-bp amplicon from the siblings along with mother shark genomic DNA were loaded on a 0.5% MDE gel. Under these conditions, "m2" was identified as two distinctive bands indicated as arrows. Mother and sibling numbers are indicated above the gel along with MHC groups and haplotype combinations from previous work (36).

Discussion

Compared with other vertebrate models (e.g., chicken or teleost fish), the shark genome seems to be stable, first demonstrated with the linkage of MHC class I and II genes (21), which was lost in bony fish (28), and later with linkage conservation of genes found in the mammalian MHC class III region (37). These MHC linkage data are consistent with global genomic studies in the elephant shark suggesting that cartilaginous fish have greater preservation of synteny than is found in any teleost model (45, 46). The β2M linkage to the shark MHC demonstrated here is likely the primordial condition, thus further supporting the conservation of the cartilaginous fish genome. Furthermore, the close proximity of class I, class II, and β2M is consistent with the theory that they were derived from a common ancestor by tandem (cis) duplication. The close linkage of β2M and class I may have regulated their original coordinated expression and upregulation. Class I and β2M expression is nearly identical in the nurse shark (Fig. 1B), but in other vertebrates β2M is made in excess (47). Furthermore, the number of β2M loci is expanded in rainbow trout (48) and polyplloid Xenopus species (18).

Unlike class II genes, class I genes are extraordinarily plastic. Besides the MHC-linked classical class Ia genes, there are also many nonclassical class Ib genes with varied functions, some encoded in the MHC and others not. The majority of class Ib proteins associates with β2M as well, and it has been speculated that there was an advantage of translocation of β2M out of the MHC so that it would not be subject to duplications and deletions (19), like class I genes in many vertebrates. Consistent with the idea of maintaining genomic stability, but in contrast to class I and class II genes, both β2M and ring3 genes are in a very stable part of the shark MHC, with very few polymorphisms and transposable elements (Fig. 4); there was no polymorphism detected by using restriction enzymes/Southern blotting with either the ring3 or β2M probe. Although there are a few bony fish species in which...
the number of $\beta 2M$ loci has been expanded (49), and there are two loci in the tetraploid *Xenopus laevis* (18), generally these species are exceptions. There seems to be only one $\beta 2M$ locus in the nurse shark genome, because genomic Southern blotting with many restriction enzymes yielded a single band with an exon-specific probe (Fig. 1C).

The primordial linkage of $\beta 2M$ to the MHC does not contribute to the debate on which gene came first, class I or class II. Among the various IgSF domains, the C1-type is a rare form, found primarily in molecules associated with adaptive immunity (50). Therefore, it is reasonable to propose that C1-type IgSF-encoding genes like $\beta 2M$ were present in the “proto-MHC,” which then acquired the PBR from another gene family. Furthermore, it has been speculated that all molecules containing C1-type IgSF domains arose from a common ancestor, and thus an Ig/TCR precursor may have originated from the “proto-MHC” (20). Consistent with previous studies dating back almost 30 y (3, 5, 33, 34), our phylogenetic analysis demonstrated a common origin for the class IIA/DMB/$\beta 2M$ and the class Ia/DMA/class IIB lineages, and all of these genes share an ancestral C1 domain-encoding exon that emerged after the split between Ag receptors and MHC genes (Fig. 1B). Whereas class IIA, $\beta 2M$, class IIB, and class Ia share an immediate common ancestor that arose by tandem duplication from the ancestral molecule, each DM gene was apparently generated by tandem duplications of class IIA and class IIB, perhaps early after the emergence of tetrapods, as no DM genes have been found in the teleost or cartilaginous fish; the maximum likelihood and Bayesian inference trees favor this scenario (S2). The NJ tree (Fig. 1B), however, suggests that shark class IIA and IIB genes cluster with class II genes from other species rather than at the basal position of class II/DM, suggesting that sharks may indeed possess DM.

An orthologous gene related to the ancestor of ring3 is present in the urochordate (e.g., amphioxus) “proto-MHC” (42), and thus the MHC-linkage of ring3 in sharks is not surprising. To determine the linkage status in other cartilaginous fish species, we examined the elephant shark genome. Current analyses of the elephant shark genome (46) has yielded only short (1 kbp) scaffolds (AAVX01540028.1) in which we only identified the $\beta 2M$ C1 domain. Three scaffolds were found to contain some exons of the elephant shark ring3 gene [AAVX01538535 (754 bp), AAVX01069837 (5232 bp), AAVX01012433 (4324 bp)]; however, the assembly is still in its early stages. Further progress in this genome project will reveal the synteny around $\beta 2M$ and all of the other MHC genes and likely provide insight into the natural history of the adaptive immune system by revealing other genes that have been translocated out of the MHC during vertebrate evolution. For example, there is good evidence from various vertebrates that both IgSF- and C-type lectin-containing NK cell receptor genes (in humans, they are encoded in leukocyte receptor complex and NK complex, respectively) and the

![FIGURE 4. Map of BAC clone GC_614H19. Gene orientation is indicated as arrows and exons are shown in boxes. Only one exon for ZFP112-like gene was identified based on the similarity to other species. The positions of repetitive elements are shown above the map classified into four different categories. The total interspersed repeats are found in ~5.35% of the sequences, consisting of ~4.74% of LINEs and ~0.63% of simple repeats. Each exon is indicated as a box, and transcriptional orientations are shown with an arrow in the 5' to 3' direction. The sequence has been deposited in the DNA Data Bank of Japan under accession number AB571627.](http://www.jimmunol.org/)

![FIGURE 5. Inconsistent synteny of $\beta 2M$ among vertebrate species. Genomic synteny of $\beta 2M$ is not consistent in bony fish and *Xenopus*, suggesting that multiple translocations of $\beta 2M$ occurred over evolutionary time. An asterisk (*) indicates the location of the $\beta 2M$ gene, and brackets indicate the genomic regions corresponding with the particular human chromosome. The detailed gene assignments can be obtained in Supplemental Table II. IgH and TCR$\alpha$ loci are marked in opossum chromosome 1.](http://www.jimmunol.org/)
MHC paralogous region by pericentric inversion of 19p13.1. Whether receptor complex (19q13.4). This region had been suggested to be an MHC paralogous region by pericentric inversion of 19p13.1. Whether receptor complex (19q13.3), a nonclassical class Ib molecule, and the leukocyte receptor complex class I and class II genes. We have found a fragment of a zinc finger protein (ZFP), ZFP112-like, in BAC clone GC_614H19, adjacent to β2M (Fig. 5). ZFP112 is found on human chromosome 19q13.2 near FcRn (19q13.3), a nonclassical class Ib molecule, and the leukocyte receptor complex (19q13.4). This region had been suggested to be an MHC paralogous region by pericentric inversion of 19p13.1. Whether the nurse shark ZNF12 is a pseudogene or divergent from human/rodent ZFP112 genes, the linkage of ZFP112 suggests that the linkage of NK receptor(s) and MHC could be preserved in the shark genome. Furthermore, we found β2M on the same chromosome as TCRα/δ in horse (chromosome 1), cow (chromosome 10), and both regions in mammals (20). Such evidence is consistent with our hypothesis that Ag receptors (TCR, Ig), NK receptors, and other genes involved in Ag processing and generally in immune function might have been linked in a “pre-adaptive immune complex” in the ancestral configuration.

Acknowledgments
We thank Dr. Mike Criscitiello and Caitlin Doremus for critical reading.

Disclosures
The authors have no financial conflicts of interest.

References


