論文内容の要旨

論文題目 EVOLUTION OF THE MAJOR HISTOCOMPATIBILITY COMPLEXCLASS I REGION IN THE TELEOST GENUS *ORYZIAS*

(メダカ属魚類におけるMHCクラスIゲノム領域の進化)

氏名 MEHTA RATNESH BHAI
メタ ラトネイシュ バイ

Major Histocompatibility Complex (MHC) is a genomic region found in jawed vertebrates and harbors various immunologically important genes. All jawed vertebrates from sharks to mammals have the evolutionarily conserved basic gene organization of the MHC. Teleosts are exceptional in that their orthologs of the mammalian MHC-encoded genes are dispersed on several chromosomes. Even in teleost, however, the *MHC class IA* genes and several other genes (transporter associated with antigen presentation 2 (*TAP2*), proteasome subunit beta type (*PSMB8*, *PSMB9*, *PSMB9L*, and *PSMB10*), TAP binding protein (*TAPBP*)) encoding for the molecules directly involved in class I antigen presentation are linked, defining the teleost MHC class I region. The uniquely derived genomic organization of the teleost MHC among the jawed vertebrates makes it a curious target for the structural analysis to understand the evolution of the jawed vertebrate MHC.

Oryzias species commonly known as medaka (ricefish), is a small egg-laying organism distributed throughout the east Asia and more than 20 *Oryzias* and closely related *Xenopoecilus* species have been described thus far which are classified into three species groups, the *latipes*, *javanicus* and *celebensis* groups, based on the karyotypes, allozymes, and mitochondrial and nuclear DNA sequences. Previous comparative analysis of the MHC class I region at the nucleotide sequence level between two inbred strains of *Oryzias latipes* (*latipes* group), Hd-rR and HNI, revealed that an approximately 100 kb block harboring the *PSMB10*, *PSMB8* and two *class IA* (*UAA*, *UBA*) genes were so diverged between these two

strains that almost no sequence similarity was noticed in the intronic and intergenic regions. To understand the evolution of the MHC class I region in genus *Oryzias* at the nucleotide sequence level, I determined the complete nucleotide sequences of the BAC clones covering the MHC class I regions of *O. dancena* (*javanicus* group) and *O. luzonensis* (*latipes* group).

Nucleotide sequences of the MHC class I regions were determined using the BAC clones isolated from the three genomic libraries designated as IMBX and IMBY for *O. dancena* and LMB for *O. luzonensis.* Screening of these libraries to identify the BAC clones containing the MHC class I region were carried out using the primer sets specific to the *PSMB9*, *PSMB8*, *class IA* (*UAA*) and *TAPBP* genes. The sequencing of the isolated BAC clones were carried out using the shotgun approach. In brief, the purified BAC DNA was partially digested with Sau3AI and fragments ranging from 3-5 kb and 6-8 kb were gel purified, cloned into the pGEM3 zf(+) vector and transformed into the DH5α cells. About 1600-1800 sequence reads were collected for each BAC clone, and were assembled into contigs and scaffolds using Phred/Phrap/Consed software. The remaining gaps between the contigs were filled by primer-walking using the BAC DNA as the template. Identification of the genes from the assembled region was carried out by homology search (BLASTn, TBLASTX) and further supported by gene prediction softwares (Genscan, Fgenesh). RT-PCR analysis was performed to confirm the predicted intron-exon boundaries. In addition, RT-PCR analysis was also performed to detect polymorphism of the *class IA* genes.

Twelve and eight clones were identified from the *O.dancena* IMBY and IMBX libraries, respectively. Based on preliminary mapping of these BAC clones, IMBY 58G24, IMBY 68M2 and IMBX 79J15 were chosen for sequence analysis. Insert size of IMBY 58G24 and IMBY 68M2 were 157,124 bp and 201,149 bp, respectively, and these sequences showed 100% identity in the 17,510 bp overlapping region. The combined 340,763 bp sequence covered the region from exon 8 of the *RXRB* gene to further downstream of the *TCF19* gene. The IMBX 79J15 BAC clone had 138,052 bp insert and covered the region from exon 9 of the *RING3* gene to the intergenic region between the MHC class I (*O. latipes UBA*) and *TAPBP* genes. Approximately 4 kb sequence up to exon 6 of the *TAPBP* gene was determined by PCR amplification using *O. dancena* genomic DNA as template followed by sequencing. The 141,664 bp sequence was composed from these sequences.

The sequence analysis of the MHC class I region of *O. luzonensis* was carried out using the LMB 52F15 and LMB 66F20 BAC clones. These two clones showed 100% nucleotide sequence identity at the approximately 15 kb of overlapping region. The193,474 bp continuous sequence was composed of the entire sequence of LMB 66F20 and the extra region of LMB 52F15, and covered the region from exon 2 of the *UDA* gene to exon 3 of the *TAPBP* gene.

The previous study in *O. latipes* reported the presence of the highly diverged dichotomous haplotypic lineages of the *PSMB8* and *PSMB10* genes termed as d- and N-, originally reported in Hd-rR and HNI, respectively. Based on the typing of the *PSMB8* gene, the IMBY sequence was judged to represent the N- haplotypic lineage of *O. dancena*, whereas the IMBX sequence represented the d- lineage of *O. dancena*. Similarly, the LMB sequence represented the d- lineage of *O. luzonensis*. The order and orientation of the predicted genes in the assembled sequences of d- and N- haplotypes of *O. dancena,* and d-haplotype of *O. luzonensis* were identical to those of d- and N- haplotypes of *O. latipes,* except for the number of the MHC *class IA* genes which showed species-specific variation. Whereas one *UAA* and one *UBA* genes were present in the HNI and Hd-rR strains (*O. latipes*), IMBX sequence contained four copies of the *UAA* gene, *Orda-UAA1*0201*, *Orda-UAA2*0201N*, *Orda-UAA3*0201*, and *Orda-UAA4*0201*. The IMBY sequence also contained four copies of the UAA gene, *Orda-UAA1*0101S*, *Orda-UAA2*0101N*, *Orda-UAA3*0101N* and *Orda-UAA4*0101*. No *UBA* like *class I* gene was detected in *O. dancena*. The LMB sequence contained three *UAA* genes, *Orlu-UAA1*0101*, *Orlu-UAA2*0101*, and *Orlu-UAA3*0101*, and a single *UBA* gene, *Orlu-UBA*0101N*.

Dotplot comparison among these haplotypes detected the diagonal line on both sides of the central variable block which showed no clear diagonal line, indicating that the most of the MHC class I region of *Oryzias* shows usual evolutionary pattern. The central variable block extended from *PSMB10* to the beginning of the *TAPBP* and included the genes directly involved in the MHC class I antigen presentation, *PSMB10*, *PSMB8* and *class IA* genes. The *class IA* genes in the variable block showed species-specific variation not only in copy number but also in its α3 domain sequence. On the other hand, phylogenetic analysis of the deduced amino acid sequence of the mature *PSMB10*, and *PSMB8* showed d- and N- lineage specific clustering, indicating that these genes show trans-species dimorphism between *O. latipes* and *O. dancena*.

In the current study, I explored the evolution of *Oryzias*'s MHC class I region using

the sequence information from the *d-* and *N-* lineages of *O. latipes, d- lineage* of *O. luzonensis* and *d-* and *N-* lineages of *O. dancena.* This study indicated that the *Oryzias* MHC class I region can be divided into conserved block and variable block. The variable block can be further subdivided into two sub-blocks, the *class I* and the *PSMB10-PSMB8.* The class I sub-block showed the most striking inter-species variation. Thus, copy number and type of the *class IA* genes varied from species to species; 1 *UAA* and 1 *UBA* for *O. latipes*, 3 *UAA* and 1 *UBA* for *O. luzonensis*, and 4 *UAA* and no UBA for *O. dancena.* The simplest explanation for this situation may be that the last common ancestor of these three species had only one *class IA* gene that duplicated or multiplied independently in each lineage. However, analysis of other *Oryzias* species is required to clarify the number and type of original *class IA* genes. The "*class I*" sub-block started from little before the 3` end of the first class I gene proximal to the *PSMB8* gene side and extended to the intergenic region of the last class I gene and the *TAPBP* gene. The "*PSMB10-PSMB8*" sub-block spanned from the 6$^{th}$ intron of the *PSMB10* gene to the 3` end of the *PSMB8* gene, and was previously found to show dichotomous haplotypic lineage in *O. latipes.* In addition, these dichotomous haplotypic lineages showed trans-species evolution in the *latipes* and *celebensis* species groups. Here I showed that these dichotomous haplotypic lineages are also found in *O. dancena*, belonging to the third *Oryzias* species group, the *javanicus* group. Therefore it is likely that these dichotomous haplotypic lineages were present in the common ancestor of all *Oryzias* species, and were transmitted from an ancestral species to a descendent species, providing a typical example of balancing selection.

Thus, the MHC class I regions of *Oryzias* species provide a unique opportunity to perform phylogenetic analysis of the three tightly linked sub-blocks evolving independently. Especially, the co-existence of the trans-species dimorphic sub-block and the species-specific variable sub-block is of interest. These gene are involved in the MHC class I antigen presentation and suggested that the two different kinds of selective pressure are working on these two sub-blocks. Although, the merit of the close linkage between these two sub-blocks could be to keep co-segregation of the functionary compatible *PSMB* and *class IA* alleles, independent evolution of the *PSMB* and *class IA* genes makes this explanation unlikely. Another explanation is that the close linkage helps rapid fixation of advantageous *class IA* new alleles. Demonstration of the actual difference in cleaving specificity of two *PSMB8* lineages as well as elucidation of peptide binding spectrum of *class IA* alleles of wild population are required to clarify the evolutionary merit of this linkage.