

Robert B. Walker. INTRASPECIES SEQUENCE VARIABILITY OF THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS II β CHAIN IN STRIPED BASS, *Morone saxatilis*. (Under the direction of Dr. Thomas J. McConnell) Department of Biology, December 1993.

This study reports the initial isolation and characterization of Major histocompatibility complex (*Mhc*) class II β genes in the teleost, *Morone saxatilis*.

Polymerase chain reaction (PCR) and dideoxy sequencing techniques were employed to amplify and characterize seven subclones (Mosa-C-1, Mosa-C-2, Mosa-C22, Mosa-R41, Mosa-S-1, Mosa-S-2, and Mosa-S-3) encoding the entire mature *Mhc* class II β chain. Striped bass from three different locations (Roanoke River, Chesapeake Bay, and Santee-Cooper Reservoir) were used in this study. Extensive polymorphism was found in the β_1 domain encoding region of the cDNA, which corresponds with the functional peptide binding region (PBR) of known *Mhc* class II molecules (Figure 4). Variable amino acid residues in the β_1 domains from the striped bass clones correspond with polymorphic sites of β_1 domains from other teleosts and higher vertebrates. Other regions of the striped bass *Mhc* class II translated proteins correlated well with those of known vertebrate *Mhc* class II proteins. Seventy-one percent of the variability within the presumed PBR of the striped bass *Mhc* class II β chain gene was found to correspond with that of the PBR of a human *Mhc* class II β chain gene (Brown et al. 1993).

INTRASPECIES SEQUENCE VARIABILITY
OF THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS II β CHAIN
IN STRIPED BASS, *Morone saxatilis*

A Thesis
Presented to
The Faculty of the Department of Biology
East Carolina University

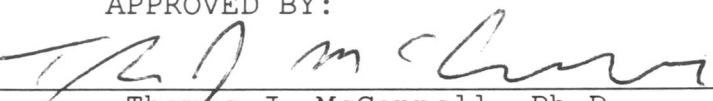
In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Molecular Biology/Biotechnology

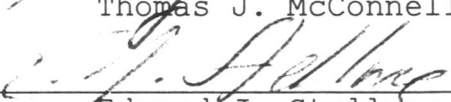
by
Robert B. Walker
December 1993

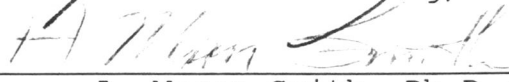
INTRASPECIES SEQUENCE VARIABILITY
OF THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS II β CHAIN
IN STRIPED BASS, *Morone saxatilis*

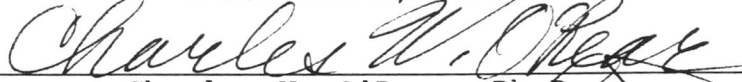
by
Robert B. Walker

APPROVED BY:

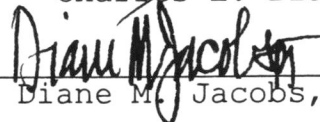
DIRECTOR OF THESIS 
Thomas J. McConnell, Ph.D.

COMMITTEE MEMBER 
Edmund J. Stellwag, Ph.D.

COMMITTEE MEMBER 
A. Mason Smith, Ph.D.

COMMITTEE MEMBER 
Charles W. O'Rear, Ph.D.

CHAIR OF THE DEPARTMENT OF BIOLOGY 
Charles E. Bland, Ph.D.

DEAN OF THE GRADUATE SCHOOL 
Diane M. Jacobs, Ph.D.

DEDICATION

This thesis is dedicated to my wife, Lisa, for her unending patience and understanding; and to my parents, Judy L. Dickerson and James B. Walker for instilling in me the desire for knowledge and the patience to persevere.

ACKNOWLEDGMENTS

This thesis was only made possible through the suggestions and helpful comments of my committee members. I am grateful to: Dr. Charlie O'Rear for his assistance with my fish biology; Dr. Mason Smith for his critical comments in the early stages of this thesis; Dr. Ed Stellwag for his wisdom, verbiage, and attention to detail; Dr. Tom McConnell for his time and patience in directing this study.

The assistance of Michael Turnage for his immense knowledge of dideoxy sequencing and Dave Bova for his assistance in bleeding the many unused striped bass blood samples, is gratefully acknowledged.

I would also like to thank a fellow graduate student, Anthony Pavell, for his encouragement and critical reading of this manuscript.

4

5

Table of Contents

List of Tables	vi
List of Figures	vii
List of Abbreviations	vii
INTRODUCTION	1
Striped Bass	1
Geographical habitat and range	1
Economic importance	2
Aquaculture	4
Major Histocompatibility Complex	6
Major Histocompatibility Complex in Teleosts	8
Statement of Purpose	11
MATERIALS AND METHODS	14
Striped Bass.	14
Genomic DNA isolation.	14
Total RNA isolation.	16
Primer selection and construction.	16
Polymerase chain reaction (PCR) amplification	18
Extraction of PCR fragments from NuSieve agarose	22
Preparation of PCR fragments for ligation	23
Ligation of PCR fragments into a plasmid vector	24
Cloning and plasmid minipreparations of PCR fragments	26
Sequencing of PCR fragments	27
35S Sanger Method	27
Fluorescence-based dideoxy sequencing	28

Results	30
Isolation of DNA encoding <i>Mhc</i> class II β chain.	30
Isolation of cDNA encoding the 5' and 3' ends of the <i>Mhc</i> class II β chain.	30
Sequence Analysis	32
Comparison of the <i>Mhc</i> class II β chains of <i>Morone</i> <i>saxatilis</i> to other teleosts	39
Conservation of <i>Mhc</i> class II β chain sequences in different species	47
Discussion	52
References	59
Appendix: Complete sequence data	64

List of Tables

Table 1. Oligonucleotide primers and sequence used for polymerase chain reaction amplification.

Table 2. Nucleotide and amino acid similarity matrices of the *Mhc* class II β chains from *Morone saxatilis* Mosa clones.

Table 3a. Amino acid similarity matrix of the *Mhc* class II β chain of Mosa, *Morone saxatilis* to other teleosts.

Table 3b. Amino acid similarity matrix of the *Mhc* class II β_1 domain of Mosa, *Morone saxatilis* to other teleosts.

Table 3c. Amino acid similarity matrix of the *Mhc* class II β_2 domain of Mosa, *Morone saxatilis* to other teleosts.

List of Figures

Figure 1. Nucleotide sequences of *Mhc* class II β chain cDNAs from striped bass, *Morone saxatilis*, as deduced the PCR+1 clones.

Figure 2. Deduced amino acid sequence of the *Mhc* class II β chain cDNAs from striped bass, *Morone saxatilis*.

Figure 3. Amino acid sequence comparison of the *Mhc* class II β chain of *Mosa*, *Morone saxatilis* and other teleosts.

Figure 4. Amino acid sequence comparison of the *Mhc* class II β chain of *Mosa*, *Morone saxatilis*, other teleosts, and higher vertebrates.

Figure 5. Wu-Kabat variability plot of *Mosa* *Mhc* class II β_1 domain amino acid sequences from Figure 2.

List of Abbreviations

bp	base pair(s)
<i>Mhc</i>	major histocompatibility complex
ml	milliliter
PBR	peptide binding region
PCR	Polymerase Chain Reaction ¹
μl	microliter
μg	micrograms
T _h	T _{helper} cell
Mosa	<u>Morone saxatilis</u>

INTRODUCTION

Striped Bass

Geographical habitat and range

Morone saxatilis, the striped bass, is one of six species in the genus *Morone*. Of the six species, four are found in North America and two in Europe. The North American species include *M. saxatilis*, striped bass (Walbaum), *M. mississippiensis*, white perch (Jordan and Eigenmann), *M. americanus*, yellow bass (Gmelin), and *M. chrysops*, white bass (Rafinesque). The European species are *M. labrax*, European bass (Linnaeus), and *M. punctatus* (Bloch). (Setzler et al. 1980)

Striped bass have an extensive North American geographic range that is the result of natural migration and stocking. They are found along the entire Atlantic coast from the St. Lawrence River, Canada, to the St. Johns River, Florida and in the Gulf of Mexico (Hill et al. 1989). During the 1870's, striped bass were introduced into the Sacramento River, California (Scofield and Bryant, 1926) and now are found on the Pacific coast from Barkley Sound, British Columbia to Ensenada, Mexico (Hassler 1988). Within the Carolinas, striped bass are found in mostly riverine or estuarine environments such as the Santee-Cooper Reservoir system and the Roanoke River. With the construction of Pinopolis Dam and the formation of the Santee-Cooper Reservoir system, the Santee-Cooper striped bass population

differs from the others in that it is freshwater and landlocked (Braschler et al. 1988) preventing their anadromous migratory behavior.

Economic importance

Historically, the striped bass, *Morone saxatilis*, has been one of the most important fishes both commercially and recreationally in North America, particularly along the Atlantic coast and in North Carolina. Despite its importance, a significant decline in the overall striped bass catch began during the 1970's. For instance, 14.7 million pounds was the record high for commercial catch in 1973 which declined to 2.4 million pounds by 1982 and to 0.4 million pounds by 1987. Within North Carolina, the commercial catch declined from 614,000 pounds in 1979 to 189,000 pounds in 1986. The decline in commercial catch from 1982-1987 has been due in part to restrictions on commercial fishing; however, the decline began long before the introduction of harvest restrictions (Rago et al. 1989; Deuel et al. 1989). The striped bass population has been slow to recover even with increased restrictions and management of the striped bass stocks (Deuel et al. 1989). As a result of the decrease in the overall catch, in 1979 the United States Congress instituted an amendment to the Anadromous Fish Conservation Act, referred to as the Emergency Striped Bass Study. The purpose of the Emergency Striped Bass Study was to institute a monitoring program to

track the status of striped bass stocks, research probable causes for the decline in the striped bass population, and determine the economic consequences of these events (Deuel *et al.* 1989).

Several proposed causes for the decline were cited. These included over-fishing, habitat deterioration, toxic contaminants, effects of industrial development, and quite possibly a series of natural events that inhibit the development of dominant year classes (Deuel *et al.* 1989). Inhibition of the dominant year classes has great impact on future fish generations and catch size, which results in less mature females spawning in any given year (Rago *et al.* 1989).

To address the problem of stock decline, the United States Congress and many of the Atlantic coastal states (North Carolina, Delaware, Maryland, Maine, New Hampshire, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Pennsylvania, and Virginia) have responded by implementing management plans. In 1981, the Atlantic States Marine Fisheries Commission (ASMFC) instituted the Interstate Fisheries Management Plan for the Striped Bass to provide uniform coastwide management regulations. The main focus of the plan was to reduce the harvest by imposing size limitations and spawning period closures. In addition, stocking of hatchery raised striped bass to their native regions was implemented in 1985 (Deuel *et al.* 1989).

Aquaculture

Aquaculture in the most basic sense is the farming of an aquatic organism, such as oyster, crab, seaweed, and several species of finfish, specifically, catfish, carp, trout, salmon, and bass. There is evidence of fish culture occurring in ancient Egypt, Asia, and central Europe as early as 100-500 B.C. (Pillay, 1990). The earliest species used for fish farming was the common carp (*Cyprinus carpio*), which is a native of China, and is now used all over the world for fish culture. As for North America, early fish culture used trout, salmon, and to some extent black bass. In the 18th century, the government developed trout hatcheries to raise and release fry into open water. Slowly the process of rearing trout (or other fish) in hatcheries moved to the private sector and evolved into production of fish for consumption as a commercial venture (Pillay, 1990). The Atlantic coastal regions began stocking programs in 1985-86 to help restore the declining striped bass populations. This spawned aquaculture in North Carolina coastal regions.

According to the North Carolina Aquaculture Development Report (1989) to the North Carolina Biotechnology Center, North Carolina with its diverse climate, geographic variation, abundant natural resources, multifarious agricultural economy, and established commercial fishing is well-suited for the development of aquaculture facilities.

The state has developed a comprehensive plan for production facilities of catfish, crawfish, soft-shell blue crab, rainbow trout, and hybrid striped bass (Helfrich, 1989).

Perhaps the most significant, but least characterized problem in aquaculture is disease susceptibility. In 1983, infectious pancreatic necrosis virus (IPN or IPNV), was isolated from 10-20 day-old striped bass in a Chesapeake Bay hatchery. IPN was originally classified as a possible cause of the stock declines. Wechsler, however, determined that IPN apparently did not contribute to the decline in striped bass stocks (Rago et. al. 1989). Within the hatchery environment, the close living conditions amplify the potential for widespread bacterial or viral infections resulting in large fish kills in hatchery populations. The crowded conditions can cause stress, facilitate contact between fish, and promote concentrated high waste levels which favor the spread of disease. Considering the rapid growth of aquaculture and the importance of defense against infectious disease, more must be learned about the immune system of important aquaculture fish species such as striped bass. This research has focused on the *Mhc* class II β chain of striped bass, *Morone saxatilis*, in an effort to elucidate a better understanding of the immune system in teleosts.

Major Histocompatibility Complex

The major histocompatibility complex (*Mhc*) is a cluster of highly polymorphic genes that encode cell surface

proteins first identified and used as determinants of graft rejection (Abbas et al. 1991). The concept of foreign tissue rejection came from the work of R. A. Gorer and G. D. Snell in the 1930's. From his work with inbred mice, Gorer identified four groups of genes that encoded blood-cell antigens. Combined work by Snell and Gorer of the group II antigens established their relationship to the rejection of foreign tumor tissue and other transplanted tissues (Kuby, 1992).

The genetic polymorphism of the *Mhc* is defined by multiple alleles at a specific locus with each allele having a frequency of occurrence greater than 0.01. Nucleotide differences between alleles lead to different encoded proteins with correspondingly distinct peptide binding affinities. Recent data from D. Klein et al. (1993) suggests that these differences are accumulated over time and are passed from one ancestral generation to the next. The *Mhc* is well characterized in humans and mice with respect to structure and function of the genes and gene products (Klein, 1986). These genes have been organized into two classes, I and II, which differ in structure, function, and tissue distribution (Klein, 1986).

The *Mhc* class I molecules are heterodimers that contain a membrane bound glycoprotein referred to as the heavy or α chain and a small non-covalently associated glycoprotein, the β_2 microglobulin. These molecules are distributed on all

nucleated cells throughout the body and present antigen to cytotoxic T cells. Before a foreign protein antigen can be recognized by any T cell, it must first be broken down into smaller peptides. It is these small peptides that associate with the *Mhc* class I molecule expressed on the surface of, for example, virus infected cells. *Mhc* class I molecules are mainly associated with viral proteins synthesized within viral infected host cells (Klein 1986).

The *Mhc* class II molecules are heterodimers that consist of two glycoproteins, the α and β chains, which noncovalently associate with one another on the plasma membrane. In the α chain, only the $\alpha 2$ domain is stabilized with a disulfide bond, whereas both domains in the β chain contain disulfide bonds. These *Mhc* class II molecules are expressed on the surface of antigen-presenting cells (APCs) and function in the presentation of processed foreign antigen, in the form of smaller peptides, to a subset of T-lymphocytes called T_{helper} cells (T_h) (Klein 1986).

Cytotoxic T cells and T_{helper} cells are critical for cell mediated immunity. The T_{helper} cells secrete growth factors (lymphokines) that activate other cells, including macrophages and cytotoxic T cells, which increases their ability to degrade and kill foreign invaders of the host organism with greater efficiency. The role of the activated cytotoxic T cell in cell mediated immunity is to kill altered host cells, such as tumor cells and virus infected

cells. The lymphokines secreted by the T_{helper} cells are crucial to this cycle. The T_{helper} cells only recognize foreign antigen presented with *Mhc* class II molecules on the surface of APCs. Generally, *Mhc* class II molecules are associated with the recognition of non-viral proteins, such as bacteria, extracellular parasites, and fungi (Kuby 1992). The only direct evidence for T-cells in teleost fish has been the isolation of three populations of cells, B cells, macrophages and functionally implied T_{helper} cells from catfish (Miller et al. 1985). Teleost antigen processing and presentation mechanisms have similarities to those in higher vertebrates (Vallejo et al. 1992).

Major Histocompatibility Complex in Teleosts

The structure of the *Mhc* class II β chain gene in teleost fish has been reported to be very similar to the mammalian *Mhc* class II β chain gene (Hashimoto et al. 1990). In general, the class II *Mhc* gene consists of an exon encoding the leader peptide, the two external domains (β_1 and β_2), the transmembrane region and the cytoplasmic region (Hashimoto et al. 1990, Ono et al. 1992). The antigen presenting function of the *Mhc* is performed by a specific peptide binding region (PBR) formed by the amino acid residues from the α_1 and β_1 external domains (Brown et al. 1988; Brown et al. 1993). The PBR regions vary greatly between individuals of the same species (Ono et al. 1992). The framework regions of the *Mhc* are highly conserved

through diverse species. These regions consist of short stretches of amino acids that are clustered together rather than occurring at random. PCR amplification can be used to exploit these conserved regions as primers and amplify the adjoining sequences of the *Mhc* genes (Bourlet et al. 1988, Hashimoto et al. 1990).

Hashimoto et al. (1990) used a PCR amplification approach, with degenerate primers and genomic DNA, to isolate the first putative *Mhc* genes in carp (*Cyprinus carpio*). The initial PCR DNA fragments (KI and KII) were used as probes to hybridize against a carp genomic library which resulted in putative *Mhc* class I α and *Mhc* class II β clones (Hashimoto et al. 1990).

Subsequently, the methods used by Hashimoto et al. (1990) have been adapted and applied to other teleost species. Juul-Madsen et al. (1992) used a similar approach, but probed against a splenic cDNA library in rainbow trout (*Oncorhynchus mykiss*). After sequencing their one positive clone (O.M.#55), it was concluded to be most of the β_2 domain, the transmembrane and cytoplasmic regions, and a very long 3' untranslated region. Using O.M.#55 as a probe in restriction fragment length polymorphism (RFLP) analysis against the DNA of a rainbow trout family of full siblings, two strong patterns were apparent. As a result, the two patterns were assumed to be genotypes AA and AB (Juul-Madsen et al. 1992). Additionally, Hordvik et al. (1993), D. Klein

et al. (1993) and Ono et al. (1993a) used the same or similar PCR amplification and probing against a library technique to isolate *Mhc* class II β clones in Atlantic salmon (*Salmo salar*), African cichlids (several species), and carp (*Cyprinus carpio*), respectively.

The *Mhc* class II β chain genes of zebrafish (*Brachydanio rerio*) have been characterized extensively by Ono et al. 1992. They sequenced twenty cDNA clones from the β -chain encoding region of the class II genes. After analysis, the clones segregated into three groups representing either three or possibly four loci (the DNA was from a different source from the original twenty). Their research showed extensive polymorphism in the *Mhc* class II β gene of zebrafish with most of the variability confined to the β_1 region within the putative peptide-binding region (PBR). Based on these results Ono et al. (1992) also suggested trans-specific evolution of *Mhc* polymorphism in teleosts inferring that the *Mhc* in fish evolved from a different ancestor than other vertebrates. However, they did state that the overall organization of the *Mhc* in fish seems to be highly conserved when compared with other vertebrates (Ono et al. 1992).

The most extensive characterization to date of the *Mhc* class II β in teleosts was conducted by Ono et al. (1993b) in *Aulonocara hansbaenschii* from the family Cichlidae. They describe the exon-intron organization of fish *Mhc* class II β

in which each gene consists of six exons (E1-E6), "encoding the leader peptide (E1), β 1 region (E1+E2), β 2 domain (E3+E4), connecting peptide (E5), transmembrane region (E5), cytoplasmic domain (E5+E6), and the 3' untranslated region (E6)" (Ono et al. 1993b). Based on these results there is evidence to suggest more than one *Mhc* class II β loci in cichlid fish that arose by duplication from a common ancestral gene. Their observations concluded at least two *Mhc* class II β loci exist that are functionally equivalent to the analogous genes in higher vertebrates (Ono et al. 1993b).

Using a PCR amplification approach, research in this study focused on the class II *Mhc* in striped bass, *Morone saxatilis*. These molecules are derived from the *Mhc* genes and are expressed on the surface of certain immunocompetent cells of the immune system (Klein 1986).

Statement of Purpose

The primary question to be addressed was: what is the variability of the cDNA sequence between different populations of the same species of striped bass?

The *Mhc* class II β genes have been shown to be useful in distinguishing between different populations of the same and/or diverse species (McConnell et al. 1988; Gibbs et al. 1990). The identification of polymorphism in this region will help in understanding the immune system of fish with some possible insight concerning those *Mhc* alleles providing

defense against striped bass pathogens and the application of this understanding to immune defense in an aquaculture setting.

Based on previous research, short, highly conserved sequences of *Mhc* class II β DNA were identified and used to synthesize complementary oligonucleotide primers. These primers were used in specific PCR amplifications of the class II *Mhc* genes from the DNA and RNA of striped bass. After cloning the PCR products into plasmid vectors, the clones were sequenced using the dideoxy method (Sanger, 1977). The sequences were compared with published class II *Mhc* DNA sequences to determine if they are prospective *Mhc* regions.

The entire project can be divided into several interrelated steps:

1. Isolate DNA and RNA from striped bass blood, and spleen.
2. Using several PCR amplification methods, including 3' RACE, 5' RACE, and PCR+1 (Borriello et al. 1990), the isolated DNA and RNA were amplified using the constructed oligonucleotide primers of the conserved regions of the *Mhc* class II β chain genes.
3. Recombinant cloning into a plasmid vector and double stranded sequencing (dideoxy method) of the *Mhc* class II PCR products for similarity comparisons to other known *Mhc* gene sequences.

4. Translate the DNA sequence into prospective amino acid sequence for comparison to other *Mhc* protein sequences.

MATERIALS AND METHODS

Striped Bass.

The striped bass (*Morone saxatilis*) were obtained from two sources: Edenton Fish Hatchery in Edenton, North Carolina, and the North Carolina Research Facility #8 in Aurora, North Carolina, courtesy of Dr. Craig Sullivan and Ron Hodson. The striped bass (Santee-Cooper and Chesapeake Bay) were collected and transported live to the laboratory where fresh blood and spleen tissue samples were collected. The Roanoke River samples were processed on site at the Edenton Fish Hatchery.

Genomic DNA isolation.

The protocol used for recovering genomic DNA was a modification of the "Genomic DNA Isolation System" from Gibco BRL (Gaithersburg, MD). Whole blood was collected from the caudal vein of live striped bass in a 1-5ml syringe containing 250mM EDTA solution as dead volume to inhibit clotting and prevent degradation of the DNA. The red blood cells were allowed to settle and then were separated from the serum using a pipette. One-hundred microliters of red blood cells (RBC) were aliquoted into 15ml centrifuge tubes (Nalge Company, Rochester, NY). An equal volume of 1X lysis buffer (640mM sucrose, 20mM Tris-HCl pH 7.8, 10mM MgCl₂, 2% w/v Triton X-100) (Gibco BRL, Gaithersburg, MD) was added and the tubes were incubated on ice for 5 minutes. The volume of each tube was brought to 13ml with 0.5X lysis

buffer and the RBCs gently resuspended. The tubes were then centrifuged for 30 minutes at 2000 RPM (Sorvall RC-5) to pellet the cellular nuclei. After centrifugation, the supernatant was decanted and discarded. The nuclear pellet was gently resuspended in 2ml Saline-EDTA (75mM NaCl, 24mM EDTA) (Gibco BRL, Gaithersburg, MD), then 100 μ l 20% (w/v) sodium dodecyl sulfate (SDS), and 250 μ l Proteinase K (20mg/ml) were added and the tubes incubated at 54°C overnight. Each sample was extracted one time with phenol, one time with phenol:chloroform:isoamyl alcohol (PCI 25:24:1), and one time with chloroform. To separate the phases between extractions, the tubes were centrifuged at 10,000 RPM (Sorvall RC-2 or RC-5) for 10 minutes. The aqueous phase was gently removed with a wide-bore disposable (Cat.No.13-711-7) pipette (Fisher Scientific, Pittsburgh, PA) and transferred to a clean 15ml centrifuge tube.

Precipitation of the DNA was accomplished by pouring the aqueous phase into 10ml of absolute ethanol and then pouring the solution between two 50ml centrifuge tubes to precipitate the DNA. The precipitated DNA was removed with a flame sealed Pasteur pipette (Fisher Scientific, Pittsburgh, PA) and resuspended in 2ml TES (10mM Tris pH 8, 5mM EDTA, 100mM NaCl). Every DNA sample was dialyzed against 4 liters of TE (10mM Tris pH 8, 1mM EDTA) overnight with a buffer change after one hour. Once the DNA samples were removed

from the dialysis tubing, they were stored at 4°C until needed (Gibco BRL, Gaithersburg, MD).

Total RNA isolation.

Spleen tissue was removed from freshly sacrificed striped bass. The tissue was placed immediately in a guanidinium isothiocyanate containing solution and homogenized in a hand-held 10ml Dounce homogenizer for RNA extraction by the recommended manufacturer's protocol (RNAzol™, Tel-Test, Inc., Friendswood, TX). The RNA pellets were resuspended in 0.1-1ml ddH₂O and stored at -20°C until needed.

Primer selection and construction.

To construct the striped bass *Mhc* class II β chain DNA primers for the PCR amplifications, DNA sequence from known *Mhc* class II genes of *Cyprinus carpio* were examined (Hashimoto et al, 1991). Highly conserved sequences were identified and PCR primers (TM215/216, see Table 1) were constructed based upon these sequences. The 5' end of selected PCR primers was modified to incorporate restriction endonuclease cleavage site, e.g. site for Eco RI, Bam HI, or Hind III. The incorporation of restriction endonuclease cleavage sites facilitated the subcloning of PCR amplified DNA. Table 1 shows the sequence of all the primers used in this study.

Table 1. Oligonucleotide primers and sequence used for polymerase chain reaction amplification.

Primer designation	Primer sequence (5'→3') ^{1,2}
TM215	TGY WGN GCC TAY GAY TTC TAY CC
TM216	SYN SYN GCR TGC TCC ACN ACR CA
TM221	ACG GAC AGG AAG TSA CCT C
TM222	GAG GTS ACT TCC TGT TCG T
TM223	CTA GGA ATT CGA GGT SAC TTC CTG TCC GT
TM232	GAG <u>GAA TTC</u> TGY WGN GCC TAY GAY TTC TAY CC
TM233	AGA <u>GGA ATT</u> CSW NSW NGC RTG CTC CAC NAC RCA
TM248	GAG <u>GGA TCC</u> <u>GTC</u> GAC ACG GAC AGG AAG TCA CCT CTG ATG
TM254	GAG <u>GGA TCC</u> <u>GAA TTC</u> TGT TTG TGA AGT GAA CAG AGG AAC
TM255	TGT <u>TTG TGA</u> <u>AGT GAA</u> CAG AGG AAC
TM256	GAG <u>GGA TCC</u> <u>GAA TTC</u> GTC AGG ACA GAC CAG ACC AGG ACA
TM259	GAG <u>AAG CTT</u> <u>TGT TTG</u> TGA AGT GAA CAG AGG ACC
TM260	GAG <u>TCT AGA</u> GTC AGG ACA GAC CAG ACC AGG ACA
TM262	GCC <u>TCC GGA</u> CTG ATC CCT GGG
TM263	GCA GTA CST CTC CTT CTS AGC
TM264	CTS AGA AGG AGA SGT ACT GCC
m13/pUC forward	GTA AAA CGA CGG CCA GT
m13/pUC reverse	CAG GAA ACA GCT ATG AC
(dT) ₁₇ -R ₁ -R ₀	AAG GAT CCG TCG ACA TCG ATA ATA CGA CTC ACT ATA GGG A(T) ₁₇

¹Restriction endonuclease cleavage sites are represented by the underlined region.

²Y = C/T , W = A/T , R = A/G , S = G/C , N = A/C/G/T.

The oligonucleotide primers in Table 1 were synthesized by the Biotechnology Core Facility at the East Carolina University School of Medicine.

Polymerase chain reaction (PCR) amplification

PCR amplification of DNA

The genomic DNA was used as a template with the oligonucleotide primers in PCR amplification of the *Mhc* class II β chain. The PCR amplification was accomplished by using the *Amplitaq* DNA polymerase kit (Perkin Elmer Cetus, Norwalk, CT). In independent PCR amplifications, two specific primers bordering the third exon of the *Mhc* class II region were used in each reaction mixture. The reaction mixtures were amplified in one of two thermal cyclers (Perkin Elmer Cetus, Norwalk, CT; MJ Research, Inc., Watertown, MA) that were checked for temperature uniformity in the reaction wells (Linz, 1990).

Each reaction was assembled on ice and consisted of 54 μ l ddH₂O, 10 μ l 10X PCR Reaction buffer II (1X, Perkin Elmer Cetus, Norwalk, CT), 16 μ l deoxynucleotide triphosphates (200 μ M, dNTPs), 5 μ l primer #1 (TM215, 263.5 μ M), 5 μ l primer #2 (TM216, 263.5 μ M), 8 μ l MgCl₂ (2 μ M), 1 μ l DNA, and 1 μ l *Taq* polymerase (5U/ μ l, Perkin-Elmer Cetus, Norwalk, CT) for a total volume of 100 μ l. Each reaction mixture was overlaid with 100 μ l mineral oil to prevent evaporation. The thermocycle used for amplification was 94°C for 1 minute to denature the DNA, 55°C for 1 minute to allow

annealing of the primers to the templates, 72°C for 1 minute to allow the *Taq* polymerase to catalyze extension of the primers. Reaction mixtures were subjected to 30-50 cycles to allow for sufficient amplification. These reactions were repeated with the more closely matched primers, TM232 and TM233.

The PCR products (5µl aliquots) were analyzed on 3% NuSieve® (FMC, Rockland, ME) by gel electrophoresis (20mA for 3 hours) and stained with ethidium bromide.

PCR amplification of RNA.

The protocol used for cDNA construction and PCR of the RNA was adapted for use by Dr. Thomas McConnell prior to the use of the RACE systems of Gibco BRL. The following reagents were placed into 0.5ml microfuge tubes: 48µl ddH₂O, 1µl of 100mM dithiothreitol (10µM, DTT), 10µl 10X PCR buffer II (1X, Perkin Elmer Cetus, Norwalk, CT), 16µl deoxynucleotide triphosphates (200µM), 1µl (from a 1:5 dilution) Human Placental Ribonuclease Inhibitor (HPRI) (Amersham), 5µl downstream primer Oligo-d(T) (108µM), TM216 (263.5µM), or TM233 (209µM), and 5µl RNA. The tubes were incubated at 65°C for 5 minutes. The tubes were cooled on ice then 8µl MgCl₂ (2µM), and 1µl (200 units) RNase H⁻ reverse transcriptase (Gibco BRL, Gaithersburg, MD) were added. The tubes were then incubated at 37°C for 1 hour. After this incubation, 5µl upstream primer (400ng/µl) and 1µl *Taq* polymerase (5U/µl) were added, while maintaining the tubes on ice. The

reaction mixtures were then overlaid with 100 μ l of mineral oil and cycled as the genomic DNA above.

Amplification of the 5' ends from RNA (5' RACE Protocol)

The manufacturer's protocol (Gibco BRL, Gaithersburg, MD) was followed using TM233 and TM223 as the anchoring and nested primers, respectively.

First strand cDNA synthesis

Five microliters of total RNA (approximately 1 μ g) and 9 μ l of ddH₂O were placed in a microfuge tube and incubated for 10 minutes at 70°C.

After the incubation, the following were added to the microfuge tube: 1 μ l gene-specific primer (TM233,), 1 μ l dNTPs (200 μ M each), 2 μ l 10X synthesis buffer (200mM Tris-HCl pH 8.4, 500mM KCl, 25mM MgCl₂, 1mg/ml BSA), and 2 μ l 0.1M DTT (10mM). The reaction mixture was equilibrated to 42°C for 2 minutes then 1 μ l RNase H⁻ (BRL) Reverse Transcriptase (200 units) was added and the mixture was incubated for 30 minutes at 42°C. The mixture was equilibrated to 55°C, 2 units of RNase H were added and then it was incubated for 10 minutes at 55°C.

Each reaction mixture was purified using BRL's GlassMax spin cartridges following the manufacturer's protocol modified by adding a final wash of 400 μ l 70% ethanol. Each cDNA elutate was vacuum dried and resuspended in 5 μ l ddH₂O.

Tailing Reaction

The 5 μ l of cDNA was denatured at 70°C for 5 minutes, chilled on ice, and centrifuged to collect contents into the bottom of tube. The remaining reagents were then added to the tube: 1 μ l synthesis buffer, 4 μ l dATP (200 μ M, 1mM stock), 1 μ l TdT (10 units), and suspended to a final volume of 20 μ l total volume with ddH₂O. This reaction mixture was then incubated 5 minutes at 37°C and then for 10 minutes at 65°C.

Amplification of tailed cDNA

The following reagents were added to a microfuge tube on ice: 5 μ l of tailed cDNA, 5 μ l 10x synthesis buffer, 8 μ l dNTP mix (10mM stock), 2 μ l Oligo d(T) (4 pmoles), 2 μ l nested gene specific primer #2 (TM222 or TM 223) (4 pmoles), 32 μ l ddH₂O, and 1 μ l Taq polymerase for a total volume of 50 μ l. This mixture was thermocycled for 35 cycles at 94°C for 45 seconds to denature, 55°C for 30 seconds to anneal, and 72°C for 2 minutes for extension. An initial denaturation step of 94°C for 6 minutes and a final extension step of 72°C for 10 minutes were also performed.

Amplification of the 3' ends from RNA (3' RACE Protocol)

Each 3' RACE reaction followed the manufacturer's protocol (Gibco BRL, Gaithersburg, MD) using TM232 as the gene-specific primer. The thermocycle used was 94°C for 45 seconds, 55°C for 1 minute, and 72°C for 1 minute for 30 cycles.

Amplification of the *Mhc* class II genes (PCR+1 protocol)

Construction of a cDNA copy from RNA was the same as in the 5' RACE protocol. Each PCR+1 (Borriello et al., 1990) reaction was performed by combining the following in a microfuge tube: 82µl ddH₂O, 10µl 10x buffer (20mM Tris-HCl pH 8.4 50mM KCl, 2.5mM MgCl₂ 100µg/ml BSA), 2µl 10mM dNTPs, 1µl (40ng) TM255, 1µl (400ng) TM256, 3µl cDNA, and 1µl Taq DNA polymerase. This mixture was heated to 94°C for 6 minutes and thermocycled at 94°C for 2 minutes, 68°C for 2 minutes, and 72°C for 3 minutes for 30 cycles. The reactions were immediately cooled to 4°C.

A second reaction was prepared using 25µl of the above asymmetric PCR reaction, 1µl (32µM) TM254, and 1µl Taq polymerase (5U). This second reaction was thermocycled at 94°C for 2 minutes, 68°C for 2 minutes, and 72°C for 3 minutes for 1 cycle. Again, this reaction was cooled to 4°C. A 5µl aliquot of each reaction was electrophoresed in 0.8% SeaPlaque® agarose (FMC, Rockland, ME) and the fragments visualized by ethidium bromide staining.

Extraction of PCR fragments from NuSieve agarose

The entire remaining PCR product from a single reaction (45-95µl) was loaded onto a 3% NuSieve® (FMC, Rockland, ME) agarose gel in 1X TBE (89mM Tris, 89mM Boric acid, 10mM EDTA). The 7cm gels were electrophoresed at 20mA for 3 hours. The bands were visualized by ethidium bromide staining and UV fluorescence. The DNA was recovered by

excising the bands from the agarose with a scalpel and placing the gel slices into a 1.5ml microfuge tube. The microfuge tubes containing the gel slices were placed at 68°C for 15 minutes. Once the NuSieve® had melted, 300µl of 68°C TE (10mM Tris pH 8, 1mM EDTA) and 500µl of 68°C phenol were added. The tubes were then mixed vigorously, centrifuged (Eppendorf 5415C) at 14,000 RPM for 5 minutes and the aqueous layers removed to new microfuge tubes. Another 500µl of 68°C phenol extraction was performed, followed by an extraction with 500µl chloroform and then 500µl of isobutanol. After the isobutanol extraction, the bottom layers (aqueous) were removed to new microfuge tubes. The DNA samples were precipitated by adding 1/10th the volume of 3M NaOAc, two volumes of ethanol, and incubating at -70°C overnight. The DNA samples were recovered by centrifugation at 14,000 RPM (Eppendorf 5415C), 4°C, for 20 minutes. The pellets were washed with 100µl of 70% EtOH and air dried. The DNA pellets were resuspended in 10-30µl ddH₂O and stored at -20°C. These samples were then used for subcloning into a suitable plasmid vector for double stranded sequencing.

Preparation of PCR fragments for ligation

The isolated PCR DNA fragments, containing restriction endonuclease cleavage sites within the primers, were digested with the appropriate restriction endonucleases (New England Biolabs, Beverly, MA) (typically *Eco* RI, *Bam* HI, or

Hind III). The DNA recovered from the NuSieve® preparation (30µl) was combined with 4µl of the appropriate enzyme buffer, 4µl ddH₂O, and 1µl of restriction enzyme for a total volume of 40µl. These were incubated overnight at 37°C. The digestion mixtures were then filtered to remove the termini extending from the restriction endonuclease cleavage site using a Millipore Ultra-MC spin filter. Briefly, the reaction volume was brought to 300µl with ddH₂O and transferred to the spin filter. The filters were centrifuged at 8,500 RPM for 3 minutes. Again, the volume was brought to 300µl with ddH₂O and the filters centrifuged as above. This procedure was repeated three times. After the final spin, the volume was brought to 300µl with ddH₂O and the complete volume was transferred to a new microfuge tube. The PCR fragments were precipitated by adding 30µl 3M NaOAc, two volumes of ethanol and incubating the tubes at -70°C overnight. The DNA was then pelleted by centrifugation (14,000 RPM, 4°C), washed once with 70% ethanol, air dried, and resuspended in 20-30µl ddH₂O.

Ligation of PCR fragments into a plasmid vector

pUC19 plasmid

Restriction endonuclease treated PCR fragments were ligated into the plasmid vector pUC19 (Gibco BRL, Gaithersburg, MD) by "cohesive end" or blunt-end ligation. For the "cohesive end" ligations, 5µl DNA was combined with 2µl T4 DNA ligase buffer, 1µl ddH₂O, 1µl T4 DNA ligase

(Gibco BRL, Gaithersburg, MD), and 1 μ l pUC19 (previously digested and phosphatase treated) for a total volume of 10 μ l. The ligation mixtures were incubated overnight at 12°C. After ligation was completed, the reaction volumes were diluted with 40 μ l ddH₂O prior to transfection into competent cells (Maniatis et al, 1990).

Blunt-end ligations were performed by first removing the single stranded overhanging ends of the fragments with Mung Bean (MB) Nuclease (Gibco BRL, Gaithersburg, MD). On ice, 15 μ l DNA, 3 μ l MB nuclease buffer (300mM NaOAc pH 4.5, 500mM NaCl, 10mM ZnCl₂, and 50% glycerol), 3 μ l (45 units) MB Nuclease, and 9 μ l ddH₂O were added into a microfuge tube and incubated at 37°C for 1 hour. The volume was brought to 300 μ l with ddH₂O and extracted one time with PCI and one time with chloroform. To separate the phases between extractions the tubes were centrifuged at 14,000 RPM (Eppendorf 5415C) for 5 minutes. The aqueous layer was transferred to a new microfuge tube, and 30 μ l 3M NaOAc and two volumes of ethanol were added. The tubes were then incubated at -70°C overnight. The PCR fragments were pelleted by centrifugation at 12,000 RPM (Jouan MR14.11), 4°C, for 20 minutes. The pellet was washed gently with 70% ethanol, air dried and resuspended in 5 μ l TE (10mM Tris pH 8, 1mM EDTA). The blunt-end ligation protocol was the same as for the "cohesive end" ligations (Maniatis et al, 1990).

TA cloning vector (Invitrogen)

Ligations using the Invitrogen's TA cloning vector system followed the manufacturer's protocol with only one exception. The DNA used was added undiluted directly from the completed PCR+1 reactions.

Cloning and plasmid minipreparations of PCR fragments

Using 5 μ l of the diluted ligation mix, the plasmid constructs were transformed into XL-1 Blue competent cells according to the manufacturer's protocol (Stratagene, La Jolla, CA). The transformed cells were plated onto media containing Blueo-gal (1.25 μ g/ml) and ampicillin (50 μ g/ml). The plates were incubated overnight at 37°C and then screened for recombinant white colonies.

The recombinant white colonies were placed into individual 50ml tubes with 10ml Luria broth (10g/L tryptone, 5g/L yeast extract, 5g/L NaCl, and 1ml/L 1M NaOH) containing 50 μ g/ml ampicillin and allowed to grow 16-18 hours at 37°C with agitation (200 RPM). The bacterial cells were pelleted by centrifugation (Sorvall RC2) 2,000 RPM for 10 minutes. The supernatant was poured off, the pellet resuspended in 90 μ l solution I (50mM glucose, 10mM EDTA, 25mM Tris pH 8) and transferred to a clean microfuge tube. Lysozyme (10 μ l) of 50mg/ml was added and the tubes incubated at room temperature for 10 minutes. Following this incubation, 200 μ l of fresh solution II (200mM NaOH, 1% SDS) and 150 μ l of solution III (300mM KOAc) were added. The tubes were mixed

gently and incubated on ice for 15 minutes. Cell debris was pelleted by centrifugation (Eppendorf 5415C) at 14,000 RPM, 4°C, for 15 minutes. The supernatant from each tube was transferred to a new microfuge tube and extracted twice with PCI (25:24:1). Two volumes of ethanol and one-tenth the volume of 3M NaOAc were added to the aqueous layer and mixed gently. The tubes were incubated at room temperature for 5 minutes and then centrifuged (Eppendorf 5415C) at 14,000 RPM for 15 minutes. The pellets were air dried and resuspended in 100µl TE (10mM Tris pH 8, 1mM EDTA). Two and one-half microliters of RNase A (10mg/ml stock) was added and the tubes were incubated for 10 minutes at room temperature. Each volume was increased with 260µl ddH₂O, 40µl 3M NaOAc and then extracted once with an equal volume of PCI and once with an equal volume of chloroform. After the aqueous layers were transferred to new microfuge tubes, two volumes of ethanol was added, and the tubes incubated at -70°C for 30 minutes to overnight. The DNA (plasmid) was collected by centrifugation (Eppendorf 5415C) 14,000 RPM, 4°C, for 20 minutes. The plasmid pellet was washed once with 100µl 70% ethanol, air dried, and resuspended in 30µl ddH₂O.

Sequencing of PCR fragments

³⁵S Sanger Method

Standard Sanger dideoxy sequencing was used, following the manufacturer's protocol (Sequenase 2.0, US Biochemical, Cleveland, OH), except for the primer annealing step. In

individual tubes, 2-4 μ l (approximately 2-3 μ g) aliquots of the plasmid DNA were treated with NaOH, for 5 minutes at room temperature, to denature the DNA. To the DNA mixture, 3 μ l of 3M NaOAc, 1 μ l pUC/M13 universal sequencing primer (100ng/ μ l), and 5 μ l ddH₂O were added. This mixture was precipitated by adding 75 μ l ethanol, mixing well, and incubating at -70°C for 30 minutes. The labeling isotope used for the sequencing reactions was α -³⁵S. All sequencing gels were 0.4mm thick, 6% polyacrylamide (National Diagnostics, Atlanta, GA) and were run at 65-70 watts for approximately 6 hours. After the sequencing run, the gels were transferred to 3mm filter paper and rinsed with 500ml of 10% acetic acid/12% methanol solution and vacuum dried. The dried gels were placed directly against Kodak XAR-5 x-ray film and exposed for 2-3 days at room temperature (Maniatis et al, 1990).

Fluorescence-based dideoxy sequencing

These sequencing reactions were performed with a battery of sequencing primers (TM221, TM222, TM233, TM259, TM260, TM262, TM263, TM264, m13/pUC forward, and m13/pUC reverse). Approximately 1 μ g of plasmid DNA was used for sequencing with each primer following the manufacturer's protocol in the PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, CA). The sequencing reactions were analyzed on the Applied

Biosystems Model 373A DNA Sequencing System according to the manufacturer's protocols.

Results

Isolation of DNA encoding *Mhc* class II β chain.

PCR primers based on those described by Hashimoto et al. (1990) were used to isolate a putative *Mhc* class II β fragment from exon 3 of fish representative of three different populations of *Morone saxatilis*. From the nine resulting 145 base pair (bp) fragments, new primers (TM221, TM222, and TM223) were designed to facilitate the use of 5' and 3' RACE technology (Gibco BRL, Gaithersburg, MD).

Isolation of cDNA encoding the 5' and 3' ends of the *Mhc* class II β chain.

A single fragment derived from the 5' RACE reactions was isolated from each population, subcloned into pUC19, and sequenced. Only one 3' RACE fragment from the Chesapeake population was isolated and sequenced. Together the DNA sequence information from 5' and 3' RACE amplified fragments was used for comparison to the carp TlaII gene (Hashimoto et al. 1990). Flanking PCR primers (TM255, TM259, and TM260) encompassing the entire *Mhc* class II β chain, excluding the 3' untranslated region, were constructed for amplification using the PCR+1 technique (see below).

Borriello et al. (1990) described the potential for generation of artifactual DNA sequences resulting from mismatch repair during PCR amplification and subcloning of regions within multigene families. Since the *Mhc* genetic locus has been demonstrated to contain multiple alleles, it

was necessary to adopt the Borriello et al. (1990) PCR+1 methodology to prevent the incorporation and cloning of any "repaired" heterologous PCR fragments, thus avoiding any chimeric artifacts. PCR primers flanking the *Mhc* region (TM255, TM259, and TM260) were constructed to address this potential technical problem.

Seven PCR+1 fragments were subcloned into pGEM-T (Promega, Madison, WI), and sequenced. Of the seven clones, six were from independent PCR+1 reactions. Three clones were from the Santee-Cooper population (Mosa-S-1, Mosa-S-2, Mosa-S-3), three clones from the Chesapeake Bay population (Mosa-C-1, Mosa-C-2, Mosa-C22), and one clone from the Roanoke River population (Mosa-R41). Mosa-C-2 and Mosa-C22 were derived from the same PCR reaction, but different subclones and after sequencing were found to be identical. Additionally, Mosa-S-2 and Mosa-S-3 were derived from different PCR reactions and after sequencing were also identical, therefore, Mosa-C22 and Mosa-S-3 were not used in any comparisons because of their redundancy.

The nomenclature for the clones follows a variation of the proposal by Klein et al. (1990), in which the acronym, Mosa refers to Morone saxatilis, the single capital letter following the four letter acronym refers to the population (C = Chesapeake Bay, R = Roanoke River, and S = Santee-Cooper Reservoir), the number following the acronym identifies the individual PCR reaction and/or clone.

Sequence Analysis

Figure 1 shows the comparison of nucleotide sequences from all seven *Mhc* class II β clones. All the clones contain complete coding regions for the β_1 and β_2 domains. In addition to 50 base pairs (bp) of leader sequence, each clone is truncated within the 3' untranslated region approximately 9-35bp after the stop codon. Exon junctions (indicated by the vertical lines in Figure 1) have been deduced from comparisons with *Mhc* class II β chain DNA sequences of other teleosts, the cichlid fish sequence of D. Klein et al. (1993), the carp sequence of Hashimoto et al. (1990), the salmon sequence of Hordvik and co-workers (1993), and the zebrafish sequences of Ono et al. (1992). Exon 1 encodes the leader peptide and four amino acids of the β_1 domain (residues 1-4). Assuming that the ATG at the 5' end of each cDNA clone is the initiation codon, the putative leader sequence would be 15 amino acid residues long. Exon 2 encodes the remainder of the β_1 domain (residues 5-95) and exon 3 encodes the entire β_2 domain (residues 96-189). Exon 4 encodes a short connecting peptide (residues 191-200), followed by the amino acids of the transmembrane region (residues 201-221), and the cytoplasmic tail (residues 222-236).

For the Santee-Cooper and Chesapeake Bay fish, two different *Mhc* class II β chain clones were isolated. In the Roanoke River fish only one *Mhc* class II β chain clone was

	1				50
Consensus	ATGGCTTCAT	CCTTTCTCAG	CTTCTCCCTC	CTCTTCATCA	CCAGCCTCTA
Mosa-S-2	-----	-----	-----a---	-----	t-----
Mosa-S-3	-----	-----	-----a---	-----	t-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----
	51	exon 2			100
Consensus	CACAGCAGAT	GGATTTCTGA	ATTATG-AAT	AGATCGCTGT	GTGTTTAACT
Mosa-S-2	-----	-----	g----ac--	-a-----	-----
Mosa-S-3	-----	-----	g----ac--	-a-----	-----
Mosa-S-1	-----	-----c-g-t	---t-tgg-c	-----	-----
Mosa-C-2	-----	-----	-----c-g-	-g-----	-----
Mosa-C22	-----	-----	-----c-g-	-g-----	-----
Mosa-C-1	-----g-	-----	-----c-g-	-a-----	-----
Mosa-R41	-----	-----c-g-t	---t--ag-c	-----	-----
	101				150
Consensus	CCACTGATCC	GAAGAACATC	GAGTACATCT	ACTCT-A-TA	TTACAACAAG
Mosa-S-2	-----	-----	-----	t--gt-c--	-----
Mosa-S-3	-----	-----	-----	t--gt-c--	-----
Mosa-S-1	-----	--g-----	-----a	--gt-c--	-----
Mosa-C-2	-----	-----	-----	-----g-a-	-----
Mosa-C22	-----	-----	-----	-----g-a-	-----
Mosa-C-1	-----	-----	-----	-----c-t-	-----
Mosa-R41	-----t	---g-----	a-----a	gg--g-a-	-----
	151				200
Consensus	TTGGAGATCG	CCAGGTTTCAG	CAGCAGTGTG	GGAGAGTTTG	TTGGTTACAC
Mosa-S-2	-----t-c	t-----	-----a-	-----	-----
Mosa-S-3	-----t-c	t-----	-----a-	-----	-----
Mosa-S-1	-----ta-	-----	-----a-	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----a-	-----
Mosa-R41	-----	-----	-----	-----	-----
	201				250
Consensus	TGAGCTGGGA	GTGAAA-AGG	CTAAATACTG	GAACAATGAT	CCTTCATATC
Mosa-S-2	---a-----	----ga-c-	--g-cg--t	t----a---	---a-----
Mosa-S-3	---a-----	----ga-c-	--g-cg--t	t----a---	---a-----
Mosa-S-1	---a-----	----ga-c-	--g-cg--t	t----a---	-----
Mosa-C-2	---tac--t	c----c--	-----	----g----	-----g-a-
Mosa-C22	---tac--t	c----c--	-----	----g----	-----g-a-
Mosa-C-1	c---t-c--	-----c--	-----	----g----	-----g-a-
Mosa-R41	-----	c---g-tg--	-----	-----a-	-----

Figure 1. (Figure continues on the next page)

	251				300
Consensus	TGGCTCAGAG	GAGAGCTGAG	AAGGAGAGGT	ACTGCCAACA	CAACATTAAC
Mosa-S-2	-----	-----	-----	-----tgac	-----
Mosa-S-3	-----	-----	-----	-----tgac	-----
Mosa-S-1	-----	-----	-----	---t-tgac	-----
Mosa-C-2	-----t	-----c	-----	-----	-----
Mosa-C22	-----t	-----c	-----	---c-----	-----
Mosa-C-1	-----g--	---t---c-	-----cag	t-----	-----
Mosa-R41	-----t	---g---	-----	-----	-----ggg
	301			exon 3	350
Consensus	ATTGACTACC	AGAATGCTCT	GGATAAATCA	GTT-AGCCCT	CTGTGGTCTCT
Mosa-S-2	-----	-----	-----	---a-a---	-----
Mosa-S-3	-----	-----	-----	---a-a---	-----
Mosa-S-1	--a-----	-----	-----	---a-a---	-----
Mosa-C-2	-----	--gc-----	-----	---c-----	-----
Mosa-C22	-----	--gc-----	-----	---c-----	-----
Mosa-C-1	-----	--gt--t--	-----	---g-----	-----
Mosa-R41	-actgg--a	gc--cat--	-----	---c--t--	a---caga--
	351				400
Consensus	GCACTCGGTG	GCGCCCCCTG	CTGGTAAACA	TCCCTCCATG	TTGGTCTGCA
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----
	401				450
Consensus	GCGTCTACGA	CTTCTTCCCC	AAACACATCA	GAGTGAGCTG	GCTCAGAGAC
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	---g-----	-----	-----
Mosa-R41	-----g--	-----	-----	-----	-----
	451				500
Consensus	GGACAGGAAG	TCACCTCTGA	TGTCATTCC	ACTGATGAGC	TGGCAGACGC
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----c--	-----	-----
Mosa-C22	-----	-----	-----c--	-----	-----
Mosa-C-1	-----	-----	---c-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----

Figure 1. (Figure continues on the next page)

	501				550
Consensus	TGATTGGTTC	TACCAGATCC	ACTCTCACCT	GGAGTACACG	CCCAGGTCTG
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----
	551				600
Consensus	GAGAGAAGAT	CTCCTGTGTG	GTGGAGCACG	CCAGCCTGAA	AGAACCTCTG
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----
	601		exon 4		650
Consensus	GTTACTGACT	GGGACCCGTC	CATGCCTGAG	TCTGAGAGAA	ACAAGATCGC
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----
	651				700
Consensus	CATCGGAGCC	TCCGGACTGA	TCCTGGGTCT	GATCTTGTCT	CTGGCTGGAT
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----	-----
Mosa-C-2	-----	-----	-----	-----	-----
Mosa-C22	-----	-----	-----	-----	-----
Mosa-C-1	-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----
	701				750
Consensus	TCATCTACTA	CAAGAGGAAG	GCCCGAGGGA	GGATTCTGGT	TCCCAGTAAC
Mosa-S-2	-----	-----	-----	-----	-----
Mosa-S-3	-----	-----	-----	-----	-----
Mosa-S-1	-----	-----	-----	-----c-----	-----
Mosa-C-2	-----	-----	-----a-----	-----	-----
Mosa-C22	-----	-----	-----a-----	-----	-----
Mosa-C-1	-----t-----	-----	-----	-----	-----
Mosa-R41	-----	-----	-----	-----	-----

Figure 1. (Figure continues on the next page)

	751				792
Consensus	TGATCCTGTC	CCTGTCCCTA	CCCCTGTACC	CTGGTCCT--	--
Mosa-S-2	-----	-----	-----	-----..	..
Mosa-S-3	-----	-----	-----	-----..	..
Mosa-S-1	-----	-----	-----	-----..	..
Mosa-C-2	-----gt	--.....			..
Mosa-C22	-----gt	--.....			..
Mosa-C-1	-----	-----g	t----acc--		..
Mosa-R41	-----gt	-----g	gt---gt--	tg-tc-tgtc	cc

Figure 1. Nucleotide sequences of *Mhc* class II β chains from striped bass, *Morone saxatilis*, as deduced the PCR+1 clones. Identical residues with consensus are denoted by *dashes*. Gaps were introduced for better alignment and are denoted by *asterisks*. The coding regions (domain borders) are indicated by *vertical lines*.

Table 2. Nucleotide and amino acid percent similarity matrices of the Mosa clones containing the *Mhc* class II β chains from *Morone saxatilis*.

Percent of nucleotide similarity ¹					
	Mosa-C-1	Mosa-C-2	Mosa-R41	Mosa-S-1	Mosa-S-2
Mosa-C-1	--				
Mosa-C-2	96	--			
Mosa-R41	91	93	--		
Mosa-S-1	92	93	92	--	
Mosa-S-2	93	94	91	97	--

Percent of amino acid similarity ¹					
	Mosa-C-1	Mosa-C-2	Mosa-R41	Mosa-S-1	Mosa-S-2
Mosa-C-1	--				
Mosa-C-2	94	--			
Mosa-R41	86	90	--		
Mosa-S-1	90	90	88	--	
Mosa-S-2	90	91	86	95	--

¹Dashes indicate 100% identity.

isolated. Considering striped bass appear to be diploid (Hassler 1988), a maximum of two alleles is expected to occur at each locus per individual. Additional study would be required to determine whether these clones originate as alleles at a locus or from genes at different loci. This distinction will be possible when the genomic organization of the striped bass *Mhc* is elucidated.

Alignment of the seven *Mosa* nucleotide sequences (Figure 1) revealed overall nucleotide conservation in three of the four domain-encoding regions. The region of exon 1 that codes for the leader peptide contains only three nucleotide substitutions, the DNA sequence coding for the connecting peptide, transmembrane region, and cytoplasmic tail (exon 4) contains four substitutions. The highly conserved exon 3, which codes for the β_2 domain, contains some variability within the first sixteen base pairs of the 5' end and only four nucleotide substitutions for the remainder of the sequence.

The most extensive variability within the striped bass class II *Mhc* was in exon 2, a region that encodes the β_1 domain of the *Mhc* class II molecule. This high degree of variability in the coding region for the β_1 domain was consistent with similar variability observed in class II β genes of mammals. The highly variable DNA sequences in the β_1 domain form the peptide-binding region (PBR) where variability has been proposed to accommodate the broad

peptide binding specificity of *Mhc* class II β gene products (Brown et al. 1993; Brown et al. 1988).

Figure 2 shows the alignment of the presumptive amino acid sequence based on the nucleotide sequence of the seven *Mosa* clones of *Mhc* class II β from *M. saxatilis* using DNASIS (National BioSciences, Plymouth, MN). Exon borders are indicated with vertical lines. Note the overall similarity between the sequences, which might be expected, even though they are from different populations within the same species. The putative glycosylation site at amino acid residues 18-20 is conserved. The variability within exon 2 is grouped into three regions (residues 9, 11, and 13; residues 28, 30, 32, and 38; residues 56-89). Seventy-one percent of the variability found in exon 2 is associated with amino acid residues of the PBR in the human (*HLA-DRB1*) *Mhc* class II β chain (Brown et al. 1993). The amino acid translations of the *Mosa* clones are consistent with that of functional human *Mhc* class II β chain genes (Brown et al. 1988). The concentration of variability in exon 2 (β_1 domain) is similar to sequences of mammalian *Mhc* class II β genes and supports the probability that these genes are functional.

Comparison of the *Mhc* class II β chains of *Morone saxatilis* to other teleosts

Comparison of the derived amino acid sequences of *M. saxatilis* *Mhc* class II β chains with similar sequences of other teleosts (cichlid, salmon, zebrafish, and carp;

	-15	1	exon 2	35
Consensus	MASSFLSFSL	LFITSLYTAD	GFL-Y-V-RC	VFNSTDPKNI EYIYS-YYNK
Mosa-S-1	-----	-----	--ry-wtd--	-----r-- --n-y---
Mosa-S-2	-----	---i-----	---s-d-n--	----- --y---
Mosa-C-1	-----	-----g	---n-a-n--	----- --h---
Mosa-C-2	-----	-----	---n-a-g--	----- --e---
Mosa-R41	-----	-----	--ry-etd--	-----l-d- k--r-e---
			S-S	
	36			85
Consensus	LEIARFSSSV	GEFVGYTELG	VK-AKYWN-D	PSYLAQRRAE KERYCQHNIN
Mosa-S-1	--y-----e	-----	--n-er--k-	----- --lt---
Mosa-S-2	--fl-----e	-----	--n-er--k-	-t----- --lt---
Mosa-C-1	-----	-----	--q-----s	--e--r-s-q --tv-----
Mosa-C-2	-----	-----	l-q-----s	--e--m--a -----
Mosa-R41	-----	-----	lrw-----nn	-----m-g- -----g
	86	exon 3		135
Consensus	IDYQNALDKS	V-PSVVLHSV	APPAGKHPSM	LVCSVYDFFP KHIRVSWLRD
Mosa-S-1	-----	-k-----	-----	-----
Mosa-S-2	-----	-k-----	-----	-----
Mosa-C-1	----vv-----	-e-----	-----	-----
Mosa-C-2	----a-----	-q-----	-----	-----
Mosa-R41	nw-s-i----	-qsy-r----	-----	-----
			S-S	
	136			185
Consensus	GQEVTSQVTS	TDELADADWF	YQIHSHLEYT	PRSGEKISCV VEHASLKEPL
Mosa-S-1	-----	-----	-----	-----
Mosa-S-2	-----	-----	-----	-----
Mosa-C-1	-----l--	-----	-----	-----
Mosa-C-2	-----p	-----	-----	-----
Mosa-R41	-----	-----	-----	-----
	186	exon 4		235
Consensus	VTDWDPSPME	SERNKIAIGA	SGLILGLILS	LAGFIYYKRK ARGRILVPSN -
Mosa-S-1	-----	-----	-----	-----p----- *
Mosa-S-2	-----	-----	-----	----- *
Mosa-C-1	-----	-----	-----	----- *
Mosa-C-2	-----	-----	-----	----- *
Mosa-R41	-----	-----	-----	----- *

Figure 2. Deduced amino acid sequence of the *Mhc* class II β chains from striped bass, *Morone saxatilis*. The sequence is given in international standard single-letter code. Dashes indicated identity with consensus sequence; Asterisks indicate stop codons; Exon borders are indicated by vertical lines. Note that at a limited number of sites, dashes represent highly conserved amino acids to the consensus sequence with the alignment based on PILEUP from GCG version 7.

	-16	1	exon 2	34
Consensus	-MASSFLSFS	LLFITSLYTA	TADGFL-Y-V	DRCVFNSTD L K-IEYIYSYY
Sasa.c157	...m-i-cv-	-t--l-i-sg	-#-----eqv-	rq-r-s-k-- qq----d--v
OnmyOM#55				
Mosa-S-1			##--ry-wt	-----p rn---n--
Mosa-S-2		---i-----	##--s-d	n-----p -n-----
Mosa-C-1		-----g	##--n-a	n-----p -n-----h-
Mosa-C-2		-----	##--n-a	g-----p -n-----e-
Mosa-R41		-----	##--ry-et	-----d-k--r-e-
Auha-M-231b	--c--	---#--s--	##--s-g	---q-----d--k--q-
BrreDAB4*01	-kp-l	vi-ml-t--g	---d-ik	qq-f-sts-- sdmv-la--s
CycaDAB*01	mslpkv-c-h	-i-ml-a--g	a-n---hsw	tk--hs-r-- sdmv--dn-i
			S-S	
	35			84
Consensus	YNKLE-ARFS	SSVGEFVGYT	ELGVKNAEYF	NKDPSYLAQR RAEKERYCQH
Sasa.c157	---a-yi--n	-t--k-----	-----a-	-s-aav--ve -g-l---k-
OnmyOM#55				
Mosa-S-1	----y---	--e-----	-----r-	-----lt
Mosa-S-2	----fl---	--e-----	-----r-	-----lt
Mosa-C-1	----i---	-----	---q-k--	-s---e--r- s-q--tv--
Mosa-C-2	----i---	-----	---l-q-k--	-s---e--m -a-----
Mosa-R41	----i---	-----	---lrw-k--	-nn---m -g-----
Auha-M-231b	----iy---	---k-----	---q-k--	-dq-avv-- -a-t---
BrreDAB4*01	---vvdqtq-n	---k-----	-q-lif--n-	---qa--h-l k-qv-t--r-
CycaDAB*01	---dvyiq-n	-t-----	a---y--r-	---ni-q-e --qv---k-
			exon 3	
	85			134
Consensus	NINI-YSNIL	DKSVQPSVVL	HSVAPPAGKH	P-MLVCSVYD F-PKHIRVSW
Sasa.c157	-adlh--t--	--t-e-h-r-	s-----s-r-	-a--m--a-- ---p--t-
OnmyOM#55				---a--m--a-- ---q--t-
Mosa-S-1	---d-q-a-	---k-----	-----	-s-----
Mosa-S-2	---d-q-a-	---k-----	-----	-s-----
Mosa-C-1	---d-qv--	---e-----	-----	-s-----
Mosa-C-2	---d-qaa-	-----	-----	-s-----
Mosa-R41	--gnw----	---sy-r-	-----	-s-----
Auha-M-231b	--g-w---p-	s---t-k-	--tt-lssq-	-a----- ---se-k---
BrreDAB4*01	-aq-wd-a-r	--a-l-e-ti	k--rqae-r-	-av-l--a-- ---k-km--
CycaDAB*01	-aelyq-a-r	--t-p-k-k-	s--k-ag-r-	-av-m--a-r ---hg-k---

Figure 3 (Figure continues on the next page)

	S-S										
	135									184	
Consensus	LRDGQEVTS	VTSTDELAD-	DW-YQIHS	HL	EYTPRS	GEKI	SCVVE	HAS	LK		
Sasa.c157	----r--k--	-----ng	-----	-----	-----	-----	--m--i--t				
OnmyOM#55	----r--k--	-----ng	-----	-----	----k-----	-----	--m--i--t				
Mosa-S-1	-----	-----a	-----	-----	-----	-----	-----				
Mosa-S-2	-----	-----a	-----	-----	-----	-----	-----				
Mosa-C-1	-----	l-----a	-----	-----	-----	-----	-----				
Mosa-C-2	-----	--p-----a	-----	-----	-----	-----	-----				
Mosa-R41	-----	-----a	-----	-----	-----	-----	-----				
Auha-M-231b	i-----	-----g	-----t-----	-----	-----	-----	-----				
BrreDAB4*01	----k-----	---m---g	-----	-----	----k-----	q-l-----t					
CycaDAB*01	--n-ev-kt-	-----png	-----e-	----k-----	--a---g-t						

	185	exon 4								238	
Consensus	EPLVTDW	DPS	MPESERN	KIA	IGASGLI	LGL	ILSLAG	FIYY	KRKARG	RILV	PSN-
Sasa.c157	---yh---	---a---	-----a	--a-----	-k-ss-vl..						
OnmyOM#55	---myh---	---a---	-----t	--a-----	-k-ss-vl**						
Mosa-S-1	-----	-----	-----	-----	-----p-	---					
Mosa-S-2	-----	-----	-----	-----	-----	---					
Mosa-C-1	-----	-----	-----	-----	-----	---					
Mosa-C-2	-----	-----	-----	-----	-----	---					
Mosa-R41	-----	-----	-----	-----	-----	---					
Auha-M-231b	-----s-	-----l-	-----	-----	-----	---					
BrreDAB4*01	q--tk--n-h	is-----f-	-----i	-iai-----	-k-st-----	-n..					
CycaDAB*01	k-i-i-i---	-----	-----i	-i-a-----	-k-ss-----	-t..					

Figure 3. Amino acid comparison of the *Mhc* class II β chain of Mosa, *Morone saxatilis* and other teleosts (carp (CycaDAB*01), Ono et al. 1993; salmon (Sasa.c157), Hordvik et al. 1993; rainbow trout (OnmyOM#55), Juul-Madsen et al. 1992; zebrafish (BrreDAB4*01), Ono et al. 1992; cichlid (Auha-M-132b), Ono et al. 1993). The sequence is given in international standard single-letter code. Dashes indicate identity with consensus sequence; Asterisks indicate stop codons; dots indicate the unavailability of sequence information; pound signs indicate gaps introduced for better alignment. Note that at a limited number of sites, dashes represent highly conserved amino acids to the consensus sequence with the alignment based on PILEUP from GCG version 7.

Table 3a. Amino acid similarity matrix of the *Mhc* class II β chain of *Mosa*, *Morone saxatilis* to other teleosts.

Percent of amino acid similarity (total) ¹						
	Mosa-C-1	Mosa-C-2	Mosa-R41	Mosa-S-1	Mosa-S-2	
Mosa-C-1	--					
Mosa-C-2	94	--				
Mosa-R41	86	90	--			
Mosa-S-1	90	90	88	--		
Mosa-S-2	90	91	86	95	--	
Auha-M-231b	82	83	82	80		82
Sasa.c157	69	69	69	69		70
<i>BrreDAB4*01</i>	60	60	60	60		61
<i>CycaDAB*01</i>	57	59	59	61		61

¹Dashes indicate 100% identity.

Table 3b. Amino acid similarity matrix of the *Mhc* class II β_1 domain of *Mosa*, *Morone saxatilis* to other teleosts.

	Percent of amino acid similarity (β_1)				
	Mosa-C-1	Mosa-C-2	Mosa-R41	Mosa-S-1	Mosa-S-2
Mosa-C-1	--				
Mosa-C-2	88	--			
Mosa-R41	68	76	--		
Mosa-S-1	76	77	73	--	
Mosa-S-2	78	80	67	89	--
Auha-M-231b	74	76	73	68	70
Sasa.c157	53	55	54	56	57
<i>BrreDAB4*01</i>	45	46	45	45	46
<i>CycaDAB*01</i>	42	44	45	51	48

¹Dashes indicate 100% identity.

Table 3c. Amino acid similarity matrix of the *Mhc* class II β_2 domain of *Mosa*, *Morone saxatilis* to other teleosts.

	Percent of amino acid similarity (β_2)				
	Mosa-C-1	Mosa-C-2	Mosa-R41	Mosa-S-1	Mosa-S-2
Mosa-C-1	--				
Mosa-C-2	97	--			
Mosa-R41	95	96	--		
Mosa-S-1	96	98	96	--	
Mosa-S-2	96	98	96	99	--
Auha-M-231b	82	83	83	83	83
Sasa.c157	79	78	79	79	79
<i>BrreDAB4*01</i>	70	70	70	71	71
<i>CycaDAB*01</i>	65	65	65	66	66

¹Dashes indicate 100% identity.

references given in Figure 3) is presented in Figure 3. The cichlid sequence, Auha-M-132b, demonstrated the highest degree of similarity with the *Mosa* clones with a percent of similarity ranging from 80-83% for the entire coding region (Table 3a). For salmon, zebrafish, and carp the percent similarity to the *Mosa* sequences for the entire coding region was 69-70%, 60-61%, and 57-61%, respectively. The cysteine residues in both exon 2 and 3 were conserved at sites 17, 82, 120, and 176. This result is extremely relevant in that these cysteines are essential for the formation of disulfide bridges within both the β_1 and β_2 domains and are absolutely essential for peptide binding activity.

A high degree of amino acid sequence variability exists in the β_1 domain among the teleosts (see Table 3b and Figure 3). The percent of similarity in the β_1 domain between the *Mosa* clones and cichlid fish is again the highest at 68-76%, followed by salmon (53-57%), zebrafish (45-46%), and carp (42-51%). This variability is consistent with the β_1 domain forming part of the peptide binding region (PBR). The entire PBR is formed from the α_1 and β_1 domain when the α and β chains non-covalently associate.

Although the amino acids composing the β_1 domain were shown to be highly variable among fishes, the β_2 domain tends to be very conserved. Despite some variability in the 5' end of the β_2 domain, comparison of the β_2 domain of the

Mosa clones with other teleosts reveals an overall conservation of the amino acid sequence (Figure 3). When compared to other teleosts, the range of variability is considerably lower in the β_2 domain versus the β_1 domain (18% vs. 34%).

Conservation of *Mhc* class II β chain sequences in different species

The phylogenetic relationship of striped bass to other species is reflected in the comparison of the *Mhc* class II β genes from striped bass to the corresponding sequences from human, mouse, amphibian, cichlid, salmon, zebrafish, and carp (Figure 4). Thirty-nine amino acid residues were shared between all the compared species. Within the β_1 domain, the structural cysteines in both the β_1 and β_2 domains are conserved among all the species compared. This was expected because of their structural significance in the formation of disulfide bonds.

The presumptive glycosylation site located at amino acid residues 18-20 is conserved in all the striped bass and the cichlid (Auha-M-231b), but the location of this site is different for salmon and zebrafish. Hordvik and co-workers' (1993) suggested that the difference in location of the presumptive glycosylation site among teleosts and other species was due to a difference in size of the mature class II β molecule among these species. Such a size difference was not apparent (Figure 4) in the striped bass studied. As

for human, mouse, and amphibian, the glycosylation site occurs only two amino acid residues downstream of the site in striped bass. Brown *et al.* (1988) reported that the glycosylation site was conserved at the same location in five different isotypes of human *Mhc* class II β molecules. Although gaps were introduced in the teleost sequences to enhance the alignment, these gaps occur at the junction between exon 1 and exon 2 which implies either a structurally larger class II β molecule in human, mouse, and amphibian, a longer leader sequence in these species, or insertions/deletions during exon-intron splicing.

	-30		1	exon 2	20
Mosa-S-1			MASSF	LSFSLLFITS	LYTA##DGFR YFWTDRQVFN
Mosa-S-2			MASSF	LSFSLLFITS	LYTA##DGFL SYDINRCVFN
Mosa-C-1			MASSF	LSFSLLFITS	LYTG##DGFL NYAVNRCVFN
Mosa-C-2			MASSF	LSFSLLFITS	LYTA##DGFL NYAVGRCVFN
Mosa-R41			MASSF	LSFSLLFITS	LYTA##DGFR YFETDRQVFN
Auha-M-231b			SF	LCLSLLFIS	LSTA##DGFL SYGVDRQVFN
Sasa.c157			MSI	FCVSLTLVLS	IFSGT#DGYF EQVVRQCRYF
BrreDAB4*01			L	KPFLVILMLS	TFTGTADGYF DYIKQQCFYS
CycaDAB*01			MSLPKV	LCFHLILMLS	AFTGAANGYF HSWWTKCIHS
Xelmhcbp1			SGVSVRV	VSVLLTSLVC	LCSSLPEYDV FQW#KSQCYF
Hb23_Mouse					WF LEYSTSECHF
Hummhdqzb1	MSWKKALRIP	GGLRAATVTL	MLSMLSTPVA	EGRDSPEDFV	YQF#KGMQCYF
					++
			S-S		
	21				70
Mosa-S-1	STDPKNIEYI	NSYYYNKLEY	ARFSSSEGEF	VGYTELGVKV	AERFNKDPSY
Mosa-S-2	STDPKNIEYI	FSYYYNKLEF	LRFSSSEGEF	VGYTELGVKV	AERFNKDPTY
Mosa-C-1	STDPKNIEYI	YSHYYNKLEI	ARFSSSVGEY	VGYTEFGVKQ	AKYWNNDPSE
Mosa-C-2	STDPKNIEYI	YSEYYNKLEI	ARFSSSVGEF	VGYTEYGLKQ	AKYWNNDPSE
Mosa-R41	STDLKDIKYI	RSEYYNKLEI	ARFSSSVGEF	VGYTELGVRW	AKYWNNDPSY
Auha-M-231b	STELKDIEYI	YSQYYNKLEI	YRFSSSVGKF	VGYTEYGVKQ	AKYFNDQPAV
Sasa.c157	SKDLQGIEFI	DSYVFNAEY	IRFNSTVGKF	VGYTELGVKV	AEAWNSDAAV
BrreDAB4*01	TSDYSDMVYL	ASYSFNKVVD	TQFNSSVGKF	VGYTEQGLIF	AENFNKDQAY
CycaDAB*01	SRDFSDMVYI	DNYIFNKDVY	IQFNSTVGEY	VGYTALGVYN	AERFNKDPNI
Xelmhcbp1	RNGTDNRYL	EGYANNQEEY	AYFSDVGEY	KAKNDFGEVQ	AKYWNSQKDL
Hb23_Mouse	YNGTQVRLL	ERYFYNLEEN	LRFDSDVGEF	RAVTELGRPD	AENWNSQPEF
Hummhdqzb1	TNGTERVRLV	SRSIYNREEI	VRFDSDVGEF	RAVTVLGLPA	AEYWNSQKDI
		+	+	+	+
	71			exon 3	120
Mosa-S-1	LAQRRAEKER	YCLTNINIDY	QNALDKSVKP	SVVLHSVAPP	AGKHPSMLVC
Mosa-S-2	LAQRRAEKER	YCLTNINIDY	QNALDKSVKP	SVVLHSVAPP	AGKHPSMLVC
Mosa-C-1	LARRSAQKET	VQOHNINIDY	QVVLDSVVEP	SVVLHSVAPP	AGKHPSMLVC
Mosa-C-2	LAQMRAAKER	YQOHNINIDY	QAALDKSVQP	SVVLHSVAPP	AGKHPSMLVC
Mosa-R41	LAQMRGEKER	YQOHNIGNWY	SNILDKSVQS	YVRLHSVAPP	AGKHPSMLVC
Auha-M-231b	VAQRRAAKET	YQOHNIGIY	SNPLSKSVQP	TVKLHSTTPL	SSQHPAMLVC
Sasa.c157	LAVERGELER	YCKHNADLHY	STILDKTVEP	HVRLSSVAPP	SGRHPAMLMC
BrreDAB4*01	LHQLKAQVDT	FQRHNAQIWD	SAVRDKAVLP	EVTIKSVRQA	EGRHPAVLLC
CycaDAB*01	LQQERAQVER	YCKHNAELYQ	SAIRDKTVP	KVKLSSVKPA	GGRHPAVLMC
Xelmhcbp1	LDQKRAVVD	VCRHNYQLYK	PYTVDRKSQP	DVKIVNTKTL	DLEHENLITC
Hb23_Mouse	LEQRRAAVDT	YCRHNYEILD	KFLVPRFVEP	TVTVPYPTKTQ	PLEHHNLLVC
Hummhdqzb1	LERKRAAVDR	VCRHNYQLEL	RTTLQRRVEP	TVTISPSRTE	ALNHHNLLVC
	+	+	+	+	+

Figure 4. (Figure continues on the next page)

S-S									
	121							170	
Mosa-S-1	SVYDFFPKHI	RVSWLRDQGE	VTSDVTSTDE	LADADWIFYQI	HSHLEYTPRS				
Mosa-S-2	SVYDFFPKHI	RVSWLRDQGE	VTSDVTSTDE	LADADWIFYQI	HSHLEYTPRS				
Mosa-C-1	SVYDFFPKHI	RVSWLRDQGE	VTSDLTSTDE	LADADWIFYQI	HSHLEYTPRS				
Mosa-C-2	SVYDFFPKHI	RVSWLRDQGE	VTSDVTPTDE	LADADWIFYQI	HSHLEYTPRS				
Mosa-R41	SVCDFFPKHI	RVSWLRDQGE	VTSDVTSTDE	LADADWIFYQI	HSHLEYTPRS				
Auha-M-231b	SVYDFYPSEI	KVSWIRLDGE	ITSVDVTSTEE	MADGDWYYQT	HSHLEYTPRS				
Sasa.c157	SAYDFYPKPI	RVTWLRDGRE	VKSDVTSTEE	LANGDWYYQI	HSHLEYTPRS				
BrreDAB4*01	SAYEFYPKKI	KMSWLRDQGE	VTSDVTSTME	MADGDWYYQI	HSHLEYTPKS				
CycaDAB*01	SAYRFYPHGI	KVSWMRNGEV	VKTDVTSTEE	MPNGDWYYQI	HSELEYTPKS				
Xelmhcp1	IVDGGFFPMI	KVTWLKNGIE	EREQVTSSAL	LKNGDWTFEI	HVMLETTIKH				
Hb23_Mouse	SVSDFYPGNI	EVWRFRNGKE	EKTGIVSTGL	VRNGDWTFQT	LVMLETVPQS				
Hummhdqzb1	SVTDFYPAQI	KVRWFRNGQE	ETAGVVSTPL	IRNGDWTFQI	LVMLEMTPQR				
	+	+	+	+	+	+	+	+	+
	171		exon 4					220	
Mosa-S-1	GEKISCVVEH	ASLKEPLVTD	WDPSMPESER	NKIAIGASGL	ILGLILSLAG				
Mosa-S-2	GEKISCVVEH	ASLKEPLVTD	WDPSMPESER	NKIAIGASGL	ILGLILSLAG				
Mosa-C-1	GEKISCVVEH	ASLKEPLVTD	WDPSMPESER	NKIAIGASGL	ILGLILSLAG				
Mosa-C-2	GEKISCVVEH	ASLKEPLVTD	WDPSMPESER	NKIAIGASGL	ILGLILSLAG				
Mosa-R41	GEKISCVVEH	ASLKEPLVTD	WDPSMPESER	NKIAIGASGL	ILGLILSLAG				
Auha-M-231b	GEKISCVVEH	ASLKEPLVTD	WDSSLPESES	NKLAIGASGL	ILGLILSLAG				
Sasa.c157	GEKISCMVEH	ISLTEPMVYH	WDPSLPEAER	NKIAIGASGL	VLGAILALAG				
BrreDAB4*01	GEKIQCLVEH	ASLTQPLTKD	WNPHISESDR	NKFAIGASGL	VLGIIIIAAG				
CycaDAB*01	GEKISCQVDH	AGLTKPIIID	WDPSLPESER	NKIAIGASGL	VLGIIIIAAG				
Xelmhcp1	GDTFTCQVEH	SSLQQPVSVN	WEPDVSESAR	NKMLTGIVGF	VLGSIFIIVG				
Hb23_Mouse	GEVYTCQVEH	PSLTDPVTVE	WKAQ#STSAQ	NKMLSGVGGF	VLGLLFLGAG				
Hummhdqzb1	GDVYTCQVEH	PSLQSPITVE	WRAQ#SESAQ	SKMLSGIGGF	VLGLIFLGLG				
	+	+	+	+	+	+	+	+	+
	221		244						
Mosa-S-1	FIYYKRKARG	RIPVPSN*..							
Mosa-S-2	FIYYKRKARG	RILVPSN*..							
Mosa-C-1	FIYYKRKARG	RILVPSN*..							
Mosa-C-2	FIYYKRKARG	RILVPSN*..							
Mosa-R41	FIYYKRKARG	RILVPSN*..							
Auha-M-231b	FIYYKRKARG	RILVPSN...							
Sasa.c157	LIYYKKKSSG	VL.....							
BrreDAB4*01	LIYYKKKSTG	RILVPN....							
CycaDAB*01	LIYYKKKSSG	RILVPT....							
Xelmhcp1	LVVYLRSKKT	MTRFSVVQNE	NLMS						
Hb23_Mouse	LFIYFRNQKG	QSGLOPTGLL	S...						
Hummhdqzb1	LIHHRSQKG	LLH.....							
		+							

Figure 4. (Figure continues on the next page)

Figure 4. Amino acid comparison of the *Mhc* class II β chain of *Mosa*, *Morone saxatilis*, other teleosts, and higher vertebrates (carp (CycaDAB*01), Ono et al. 1993; cichlid (Auha-M-231b), Ono et al. 1993; salmon, (Sasa.c157), Hordvik et al. 1993; zebrafish, (BrreDAB4*01), Ono et al. 1992; amphibian, (Xelmhcp1), Sato et al. 1992; mouse, (Hb23_Mouse), Mengle et al. 1985; human, (Hummhdqbz1), Hall, M.A. 1992). The sequence is given in international standard single-letter code. The *plus signs* indicate residue conservation in all species or all but one species; *Asterisks* indicate stop codons; *dots* indicate the unavailability of sequence information; *pound signs* indicate gaps inserted for better alignment. The structural cysteines and glycosylation sites are underlined.

Discussion

Recombinant DNA cloning and DNA sequence analysis of PCR amplified RNA prepared from *Morone saxatilis* revealed the presence of five distinct *Mhc* class II β chain genes. All five of the striped bass *Mhc* class II β chain gene sequences contain the complete encoding sequence for the β chain protein. As striped bass are diploid, the possibility exists for multiple loci of *Mhc* class II β chain genes. The identification of at least five different *Mhc* class II β chain genes in *Morone saxatilis* indicates a high level of polymorphism in this species. The high degree of polymorphism at the striped bass *Mhc* class II β gene locus was expected based on examination of this multigene system in other vertebrates and teleosts (Ono et al. 1992; Ono et al. 1993c; Brown et al. 1993). In zebrafish, Ono et al. (1992) demonstrated the presence of four different loci for the *Mhc* class II β chain gene. Ono et al. (1993c) also demonstrated the existence of greater than twenty-three *Mhc* class II β genes in cichlids.

DNA sequence similarities (Table 2) among the striped bass *Mhc* class II β chain genes were remarkably high (97-91%) with most of the variability occurring in exon 2. The variability found in exon 2 of the DNA sequence was also evident in the deduced amino acid sequence of the striped bass *Mhc* class II β chain genes. The Mosa-R41 *Mhc* class II β

chain was the most divergent among the five striped bass *Mhc* class II β chains with a consistently lower similarity score (Table 3a,b,c) when compared to the clones from the Santee-Cooper Reservoir and Chesapeake Bay. However, note the similarity of amino acid residues 8-13 of the Mosa-R41 and Mosa-S-1 clones (Figure 2). This may be a pattern of gene conversion or the Mosa-R41 and Mosa-S-1 are *Mhc* class II β chain alleles from the same gene locus. Based on the amino acid sequence similarities, the striped bass *Mhc* class II β chain genes cluster into their respective populations. However, this may not be significant considering the small sample size and the possibility that the *Mhc* class II β chain genes isolated in this research may not have originated from the same gene locus.

The considerable variability among the striped bass *Mhc* class II β chain genes isolated in the present study may be useful in delineating between populations, however, the sample size is insufficient to characterize intrapopulation variation. Two factors should be addressed in increasing the sample size. First, the populations should have minimal migrational mixing and second, the populations sampled should have no known artificial stocking events. A PCR approach would still be appropriate. However, PCR amplifications of exon 2 and intron 1 of the *Mhc* class II β

chain gene would yield the better resolution as in Ono et al. 1993c and D. Klein et al. 1993.

Comparison of the amino acid sequence within different regions of the striped bass *Mhc* class II β chain gene revealed the β_1 domain of striped bass to be more variable than the β_2 domain. The variability within the β_1 domain of striped bass is grouped into three regions (residues 9, 11, and 13; residues 28, 30, 32, and 38; residues 56-89). Amino acid residues in the β_1 domain of striped bass (sites 9, 11, 13, 38, 60, 61, 65, 65, 68, 71, 74, 78, 81, 82, 85, 86, and 89) correspond to polymorphic amino acid residues in the β_1 domain of the human *Mhc* class II β genes (Brown et al. 1993). Grouping of the variable amino acid residues in the *Mhc* class II β_1 domain was also observed in zebrafish, salmon, and cichlids (Ono et al. 1992; Hordvik et al. 1993; D. Klein et al. 1993; Ono et al. 1993a, 1993b, 1993c). These short regions of variability correspond with regions of antigen binding within the PBR. After peptides resulting from digested native proteins are bound in this region, they are presented in association with the *Mhc* class II molecule to T_h cells. These activated T_h cells will subsequently initiate the immune response.

In the present study, seventy-one percent of the variability within the presumed PBR was found to correspond with that of the PBR from a human *Mhc* class II β chain gene.

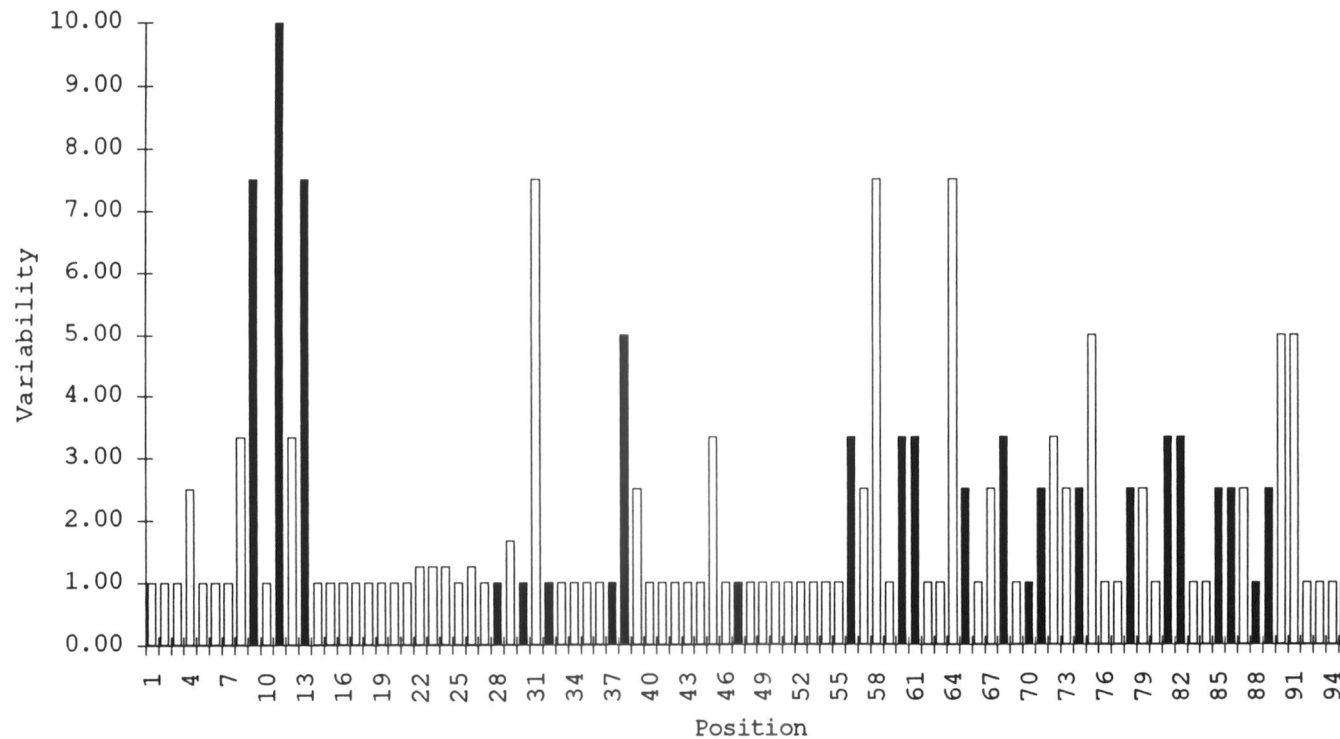


Figure 5. Wu-Kabat variability plot of Mosa *Mhc* class II β_1 domain amino acid sequences from Figure 2. Variability score (ordinate) is given by the number of different amino acid residues/frequency of the most common residue for each site (Wu and Kabat 1970). The horizontal axis shows the position of each site; the numbering corresponds to that used in Figure 2. The solid bars indicate residues corresponding to the PBR of the human (*HLA-DR1*) *Mhc* class II β_1 domain (Brown et al. 1993).

Ono et al. 1992 demonstrated that in zebrafish, sixty-five percent of the variability found in exon 2 was associated with the putative PBR, however, the comparison of variable regions within the PBR of Ono et al. 1992 were based on a putative model of the *Mhc* class II molecule extrapolated from a *Mhc* class I x-ray crystallograph (Brown et al. 1988). In this study, the comparison of variable regions within the PBR was based on the actual three-dimensional (determined by X-ray crystallography) structure of the complete *Mhc* class II molecule (Brown et al. 1993). The polymorphism in the β_1 domain supports the notion that this region forms part of the PBR of a complete *Mhc* molecule. Based on the conservation of peptide contact points within the PBR (see Figure 5) of the human *Mhc* class II β chain (*HLA-DR1*), the Mosa *Mhc* class II β chain clones most likely contain the DNA sequences necessary for synthesis of functional gene products (Brown et al. 1993). This concentration of polymorphism in the presumed PBR of the Mosa *Mhc* class II β_1 domain implies that the function of the striped bass genes is the same as corresponding genes of other vertebrate species.

The overall organization of the *Mhc* class II β chain genes in striped bass is remarkably similar to other *Mhc* class II β chain genes. The polymorphic amino acid residues of the striped bass *Mhc* class II β_1 domain occur in short

regions (presumably the PBR) a feature common to the *Mhc* class II β_1 domains of human, zebrafish, salmon, and cichlids (Brown *et al.* 1993; Ono *et al.* 1992; Hordvik *et al.* 1993; D. Klein *et al.* 1993; Ono *et al.* 1993a, 1993b, 1993c). Although the amino acid sequence similarity between the Mosa *Mhc* class II β chain genes and corresponding genes in amphibian, mouse, and human were low, (30-34%, 35-37%, and 36-38%, respectively), the general level of genomic organization appears to have been preserved evolutionarily.

The amino acid sequence similarity data of the *Mhc* class II β chain genes correlates well with the classical phylogenetic relationships as determined by morphological characteristics and the fossil record of these teleost species. Phylogenetic relationships among teleosts is reflected in the comparison of the amino acid sequences of the *Mhc* class II β chain genes. The cichlid (Auha-M-231b) *Mhc* class II β chain is most similar (80-82%) to striped bass for the entire *Mhc* class II β chain. Cichlid fishes and striped bass are grouped into the same order, Perciformes, which is considered a more advanced order for teleosts (Carroll 1988). From the similarity comparison of the *Mhc* class II β chain gene amino acid sequences, salmon (69-70%) are the next closely related species to striped bass after cichlids (80-82%) followed by zebrafish (60-61%) and carp (57-61%).

Classical phylogenetic relationships consider the Salmonids as a primitive species with zebrafish (and carp) branching from the main lineage just above or below Salmonoids followed by the relatively recent radiation of Perciforms (Carroll 1988, Moyle et al. 1982). Amino acid similarity data (Table 3a) in this study suggest that the Perciforms are more closely related to the Salmonids (69-70%) as opposed to the Cyprinids (57-61%), therefore, the molecular sequence data from the present study supports Carroll's (1988) suggestion that the Osteriophysii branched from the main lineage of teleost evolution prior to the appearance of the Salmonids. Conservation of the general structure within the *Mhc* class II β chain emphasizes the evolutionary conservation of function within these *Mhc* molecules which is the defense against microbial pathogens. Simultaneously we see extensive variability, particularly at the β_1 domain, providing a "trail" to act as a marker for following speciation which otherwise might be impossible to track using mitochondrial DNA analysis (Ono et al. 1993c). The diversity and function of the *Mhc* class II genes of teleosts may provide insight into not only immune defense in fishes but speciation as well.

References

- Abbas, Abdul K. et al. *Cellular and Molecular Immunology*. Philadelphia, PA: W.B. Saunders Company, 1991.
- Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- Borriello, Frank, and Kenneth S. Krauter. 1990. Reactive site polymorphism in the murine protease inhibitor gene family is delineated using a modification of the PCR reaction (PCR+1). *Nucleic Acids Research* 18(18): 5481-5487.
- Bourlet, Y., et al. 1989. Isolation of chicken major histocompatibility complex class II (B-L) beta chain sequences: comparison with mammalian beta chains and expression in lymphoid organs. *EMBO* 7:1031.
- Braschler, David W., Miller G. White, and Jeffrey W. Foltz. 1988. Movements and habitat selection of striped bass in the Santee-Cooper Reservoirs. *Proc. Annu. Conf. SEAFWA*. pgs.27-34.
- Brown, J.H., T. Jardetzky, M.A. Saper, B. Samraoui, P.J. Bjorkman, and D.C. Wiley 1988. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature (London)* 332: 845-850.
- Brown, Jerry H., T.S Jardetzky, J.C Gorga, L.J. Stern, R.G. Urban, J.L. Strominger, and D.C. Wiley. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 364: 33-39.
- Carroll, Robert L. *Vertebrate paleontology and evolution*. New York: W. H. Freeman and Company, 1988.
- Crouse, J. and Amorese, D. (1986) Gibco BRL Modified Genomic DNA Isolation System *Focus* 8:3.
- Deuel, Dave, Dave McDaniel, and Steve Taub. 1989. Atlantic coastal striped bass: road to recovery. Annual Congressional report available from the National Marine Fisheries Service or the U.S. Fish and Wildlife Service.
- Ezquerra, A., et al. 1990. T-cell receptor delta gene expression and diversity in the mouse spleen. *J. Immunol.* 145:1311.

Genetics Computer Group, (1991), *GCG Program Manual*, Version 7, April 1991, 575 Science Drive, Madison, Wisconsin, USA 53711.

Gibbs, H.L., P.J. Weatherhead, P.T. Boag, B.N. White, L.M. Tabak, and D.J. Hoysak. 1990. realized reproductive success of polygynous red-winged blackbirds revealed by DNA markers. *Science* 250:1394.

Guillemot, et al. 1988. A molecular map of the chicken major histocompatibility complex: the class II beta genes are closely linked to the class I genes and the nucleolar organizer. *EMBO* 7:2775.

Hall, M.A. 1992. HLA-DQ beta second domain polymorphisms may explain enhanced risk of coeliac disease associated with heterozygous genotypes. Unpublished.

Hashimoto, Kehchiro, Teruyuki Nakanishi, and Yoshikazu Kurosawa. 1990. Isolation of carp genes encoding major histocompatibility complex antigens. *Proc. Natl. Acad. Sci. USA* 87: 6863-6867.

Hassler, T.J. 1988. Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (Pacific SouthWest)--striped bass. U.S. Fish Wildl. Serv. Biol. Rep. 82(11.82). U.S. Army Corps of Engineers, TR EL-82-4. 29 pp.

Helfrich, L.A., et al. 1989. North Carolina Aquaculture Development Report. NCBC Nov.1989:130 pp.

Hill, J., J.W. Evans, and M.J. Van Den Avyle. 1989. Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (South Atlantic)--striped bass. U.S. Fish Wildl. Serv. Biol. Rep. 82(11.118). U.S. Army Corps of Engineers, TR EL-82-4. 35 pp.

Hordvik, Ivar, Unni Grimholt, Viggo M. Fosse, Øystein Lie, and Curt Endresen. 1993. Cloning and sequence analysis of cDNAs encoding the MHC class II β chain in Atlantic salmon (*Salmo salar*). *Immunogenetics* 37: 437-441.

Juul-Madsen, H.R., J. Glamann, H.O. Madsen, and M. Simonsen. 1992. MHC class II beta-chain expression in the rainbow trout. *Scand. J. Immunol* 35: 687-694.

- Kaufman, Jim. 1990. Evolution of the MHC: Lessons from the nonmammalian vertebrates. *Immunol. Res.* 9: 123-134.
- Klein, Jan. *Natural History of the Major Histocompatibility Complex*. New York: John Wiley & Sons, Inc., 1986.
- Klein, J., R.E. Bontrop, R.L. Dawkins, H.A. Erlich, U.B. Gyllenstein, E.R. Heise, P.P. Jones, P. Parham, E.K. Wakeland, and D.I. Watkins. 1990. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics* 31: 217-219.
- Klein, Dagmar, Hideki Ono, Colm O'hUigin, Vladimir Vincek, Tijs Goldschmidt, and Jan Klein. 1993. Extensive MHC variability in cichlid fishes of Lake Malawi. *Nature* 364: 330-334.
- Kuby, Janis. *Immunology*. New York: W.H. Freeman and Company, 1992.
- Linz, U. 1990. Thermocycler temperature variation invalidates PCR results. *Biotechniques* 9:286.
- Medrano, J.F. et al. 1990. DNA extraction from nucleated red blood cells. *Biotechniques* 8:43.
- Miller, N.W., Sizemore, R.C., and Clem, L.W. 1985. Phylogeny of lymphocyte heterogeneity: the cellular requirements for in vitro antibody responses of channel catfish leucocytes. *J Immunology* 134:2884-2888.
- Ono, Hideki, Dagmar Klein, Vladimir Vincek, Felipe Figueroa, Colm O'hUigin, Herbert Tichy, and Jan Klein. 1992. Major histocompatibility complex class II genes of zebrafish. *Proc. Natl. Acad. Sci. USA* 89: 11886-11890.
- Ono, Hideki, Colm O'hUigin, Vladimir Vincek, René J.M. Stet, Felipe Figueroa, and Jan Klein. 1993a. New β chain-encoding *Mhc* class II genes in the carp. *Immunogenetics* 38: 146-149.
- Ono, Hideki, Colm O'hUigin, Vladimir Vincek, and Jan Klein. 1993b. Exon-intron organization of fish major histocompatibility complex class II β genes. *Immunogenetics* 38: 223-234.

Ono, Hideki, Colm O'hUigin, Herbert Tichy, and Jan Klein. 1993c. Major histocompatibility complex variation in two species of cichlid fishes from Lake Malawi. *Mol. Biol. Evol.* 10(5): 1060-1072.

Payne, Elmer Sturgis. 1990. Genetic heterogeneity of Roanoke River striped bass. Masters Thesis. East Carolina University.

Perbal, Bernard. *A Practical Guide to Molecular Cloning*. 2nd ed. New York: John Wiley & Sons, Inc., 1988.

Pillay, T.V.R. *Aquaculture: principles and practices*. Oxford: Fishing News Books, 1990.

Rago, Paul J., R. Anne Richards, Robert M. Dorazio, and David G. Deuel. 1989. Emergency Striped Bass Research Study: Report for 1987. U.S. Fish and Wildlife Service and National Marine Fisheries Service.

Roitt, I.M. et al. *Immunology*. 2nd ed. London, England: Gower Medical Publishing, 1989.

Sambrook, J. et al. 1990. *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. 74: 5463-5467.

Sato, K., Flajnik, M.F., Pasquier, L.D., Katagiri, M. and M. Kasahara. 1992. Evolution of the MHC: isolation of class II beta-chain cDNA clones from the Amphibian *Xenopus laevis*. *J. Immunol.* 150: 2831-2843.

Setzler, E.M., W.R. Boynton, K.V. Wood, H.H. Zion, L. Lubbers, N.K. Mountford, P. Frere, L. Tucker, and J.A. Mihursky. 1980. Synopsis of biological data on striped bass, *Morone saxatilis* (Walbaum). *FAO Synopsis No. 121. NOAA Technical Report NMFS Circular 433*.

Schwartz, R.M. and Dayhoff, M.O. 1979. *Atlas of Protein Sequence and Structure*. Dayhoff, M.O. Ed. National Biomedical Research Foundation, Washington D.C., pp. 353-358.

Stet, R.J.M., and E. Egberts. 1991. The histocompatibility system in teleostean fishes: from multiple histocompatibility loci to a major histocompatibility complex. *Fish & Shellfish Immunology* 1: 1-16.

Stet, René J.M., Saskia H.M. van Erp, Trudi Hermsen, Holger A. Sülthmann, and Egbert Egberts. 1993. Polymorphism and estimation of the number of *MhcCyca* class I and class II genes in laboratory strains of the common carp (*Cyprinus carpio* L.). *Developmental and Comparative Immunology* 17: 141-156.

Vallejo, A.N., Miller, N.W., and Clem, L.W. 1992. Antigen processing and presentation in teleost immune responses. *Annu Rev Fish Diseases*, pp. 73-89.

Waldman, J.R. et al. 1988. A review of stock discrimination techniques for striped bass. *N. Am. J. Fish. Manage.* 8:410.

Williams, J.G., and R.K. Patient. *Genetic Engineering*. Oxford, England: IRL Press Limited, 1988.

Appendix: Complete sequence data

Complete nucleotide and amino acid sequences of the *Mhc* class II β chain in the seven Mosa clones from striped bass, *Morone saxatilis*. Amino acid sequence is given in international standard single-letter code.

Mosa-C-1.DNA

1	M	A	S	S	F	L	S	F	S	L	L	F	I	T	S	L	16
1	ATG	GCT	TCA	TCC	TTT	CTC	AGC	TTC	TCC	CTC	CTC	TTC	ATC	ACC	AGC	CTC	48
17	Y	T	G	D	G	F	L	N	Y	A	V	N	R	C	V	F	32
49	TAC	ACA	GGA	GAT	GGA	TTT	CTG	AAT	TAT	GCA	GTA	AAT	CGC	TGT	GTG	TTT	96
33	N	S	T	D	P	K	N	I	E	Y	I	Y	S	H	Y	Y	48
97	AAC	TCC	ACT	GAT	CCG	AAG	AAC	ATC	GAG	TAC	ATC	TAC	TCT	CAT	TAT	TAC	144
49	N	K	L	E	I	A	R	F	S	S	S	V	G	E	Y	V	64
145	AAC	AAG	TTG	GAG	ATC	GCC	AGG	TTC	AGC	AGC	AGT	GTG	GGA	GAG	TAT	GTT	192
65	G	Y	T	E	F	G	V	K	Q	A	K	Y	W	N	S	D	80
193	GGT	TAC	ACC	GAG	TTC	GGA	GTG	AAA	CAG	GCT	AAA	TAC	TGG	AAC	AGT	GAT	240
81	P	S	E	L	A	R	R	S	A	Q	K	E	T	V	C	Q	96
241	CCT	TCA	GAA	CTG	GCT	CGG	AGG	AGT	GCT	CAG	AAG	GAG	ACA	GTC	TGC	CAA	288
97	H	N	I	N	I	D	Y	Q	V	V	L	D	K	S	V	E	112
289	CAC	AAC	ATT	AAC	ATT	GAC	TAC	CAG	GTT	GTT	CTG	GAT	AAA	TCA	GTT	GAG	336
113	P	S	V	V	L	H	S	V	A	P	P	A	G	K	H	P	128
337	CCC	TCT	GTG	GTC	CTG	CAC	TCG	GTG	GCG	CCC	CCT	GCT	GGT	AAA	CAT	CCC	384
129	S	M	L	V	C	S	V	Y	D	F	F	P	K	H	I	R	144
385	TCC	ATG	TTG	GTC	TGC	AGC	GTC	TAC	GAC	TTC	TTC	CCC	AAG	CAC	ATC	AGA	432
145	V	S	W	L	R	D	G	Q	E	V	T	S	D	L	T	S	160
433	GTG	AGC	TGG	CTC	AGA	GAC	GGA	CAG	GAA	GTC	ACC	TCT	GAT	CTC	ACT	TCC	480
161	T	D	E	L	A	D	A	D	W	F	Y	Q	I	H	S	H	176
481	ACT	GAT	GAG	CTG	GCA	GAC	GCT	GAT	TGG	TTC	TAC	CAG	ATC	CAC	TCT	CAC	528
177	L	E	Y	T	P	R	S	G	E	K	I	S	C	V	V	E	192
529	CTG	GAG	TAC	ACG	CCC	AGG	TCT	GGA	GAG	AAG	ATC	TCC	TGT	GTG	GTG	GAG	576
193	H	A	S	L	K	E	P	L	V	T	D	W	D	P	S	M	208
577	CAC	GCC	AGC	CTG	AAA	GAA	CCT	CTG	GTT	ACT	GAC	TGG	GAC	CCG	TCC	ATG	624
209	P	E	S	E	R	N	K	I	A	I	G	A	S	G	L	I	224
625	CCT	GAG	TCT	GAG	AGA	AAC	AAG	ATC	GCC	ATC	GGA	GCC	TCC	GGA	CTG	ATC	672
225	L	G	L	I	L	S	L	A	G	F	I	Y	Y	K	R	K	240
673	CTG	GGT	CTG	ATC	TTG	TCT	CTG	GCT	GGA	TTC	ATC	TAT	TAC	AAG	AGG	AAG	720
241	A	R	G	R	I	L	V	P	S	N	*	S	C	P	C	P	256
721	GCC	CGA	GGG	AGG	ATT	CTG	GTT	CCC	AGT	AAC	TGA	TCC	TGT	CCC	TGT	CCC	768
257	C	P	Y	P	260												
769	TGT	CCC	TAC	CCC	780												

Mosa-C-2

1	M	A	S	S	F	L	S	F	S	L	L	F	I	T	S	L	16
1	ATG	GCT	TCA	TCC	TTT	CTC	AGC	TTC	TCC	CTC	CTC	TTC	ATC	ACC	AGC	CTC	48
17	Y	T	A	D	G	F	L	N	Y	A	V	G	R	C	V	F	32
49	TAC	ACA	GCA	GAT	GGA	TTT	CTG	AAT	TAT	GCA	GTA	GGT	CGC	TGT	GTG	TTT	96
33	N	S	T	D	P	K	N	I	E	Y	I	Y	S	E	Y	Y	48
97	AAC	TCC	ACT	GAT	CCG	AAG	AAC	ATC	GAG	TAC	ATC	TAC	TCT	GAA	TAT	TAC	144
49	N	K	L	E	I	A	R	F	S	S	S	V	G	E	F	V	64
145	AAC	AAG	TTG	GAG	ATC	GCC	AGG	TTC	AGC	AGC	AGT	GTG	GGA	GAG	TTT	GTT	192
65	G	Y	T	E	Y	G	L	K	Q	A	K	Y	W	N	S	D	80
193	GGT	TAC	ACT	GAG	TAC	GGT	CTG	AAA	CAG	GCT	AAA	TAC	TGG	AAC	AGT	GAT	240
81	P	S	E	L	A	Q	M	R	A	A	K	E	R	Y	C	Q	96
241	CCT	TCA	GAA	CTG	GCT	CAG	ATG	AGA	GCT	GCG	AAG	GAG	AGG	TAC	TGC	CAA	288
97	H	N	I	N	I	D	Y	Q	A	A	L	D	K	S	V	Q	112
289	CAC	AAC	ATT	AAC	ATT	GAC	TAC	CAG	GCT	GCT	CTG	GAT	AAA	TCA	GTT	CAG	336
113	P	S	V	V	L	H	S	V	A	P	P	A	G	K	H	P	128
337	CCC	TCT	GTG	GTC	CTG	CAC	TCG	GTG	GCG	CCC	CCT	GCT	GGT	AAA	CAT	CCC	384
129	S	M	L	V	C	S	V	Y	D	F	F	P	K	H	I	R	144
385	TCC	ATG	TTG	GTC	TGC	AGC	GTC	TAC	GAC	TTC	TTC	CCC	AAA	CAC	ATC	AGA	432
145	V	S	W	L	R	D	G	Q	E	V	T	S	D	V	T	P	160
433	GTG	AGC	TGG	CTC	AGA	GAC	GGA	CAG	GAA	GTC	ACC	TCT	GAT	GTC	ACT	CCC	480
161	T	D	E	L	A	D	A	D	W	F	Y	Q	I	H	S	H	176
481	ACT	GAT	GAG	CTG	GCA	GAC	GCT	GAT	TGG	TTC	TAC	CAG	ATC	CAC	TCT	CAC	528
177	L	E	Y	T	P	R	S	G	E	K	I	S	C	V	V	E	192
529	CTG	GAG	TAC	ACG	CCC	AGG	TCT	GGA	GAG	AAG	ATC	TCC	TGT	GTG	GTG	GAG	576
193	H	A	S	L	K	E	P	L	V	T	D	W	D	P	S	M	208
577	CAC	GCC	AGC	CTG	AAA	GAA	CCT	CTG	GTT	ACT	GAC	TGG	GAC	CCG	TCC	ATG	624
209	P	E	S	E	R	N	K	I	A	I	G	A	S	G	L	I	224
625	CCT	GAG	TCT	GAG	AGA	AAC	AAG	ATC	GCC	ATC	GGA	GCC	TCC	GGA	CTG	ATC	672
225	L	G	L	I	L	S	L	A	G	F	I	Y	Y	K	R	K	240
673	CTG	GGT	CTG	ATC	TTG	TCT	CTG	GCT	GGA	TTC	ATC	TAC	TAC	AAG	AGG	AAG	720
241	A	R	G	R	I	L	V	P	S	N	*	S	W	S			254
721	GCC	CGA	GGA	AGG	ATT	CTG	GTT	CCC	AGT	AAC	TGA	TCC	TGG	TCC			762

Mosa-C22

1	M	A	S	S	F	L	S	F	S	L	L	F	I	T	S	L	16
1	ATG	GCT	TCA	TCC	TTT	CTC	AGC	TTC	TCC	CTC	CTC	TTC	ATC	ACC	AGC	CTC	48
17	Y	T	A	D	G	F	L	N	Y	A	V	G	R	C	V	F	32
49	TAC	ACA	GCA	GAT	GGA	TTT	CTG	AAT	TAT	GCA	GTA	GGT	CGC	TGT	GTG	TTT	96
33	N	S	T	D	P	K	N	I	E	Y	I	Y	S	E	Y	Y	48
97	AAC	TCC	ACT	GAT	CCG	AAG	AAC	ATC	GAG	TAC	ATC	TAC	TCT	GAA	TAT	TAC	144
49	N	K	L	E	I	A	R	F	S	S	S	V	G	E	F	V	64
145	AAC	AAG	TTG	GAG	ATC	GCC	AGG	TTC	AGC	AGC	AGT	GTG	GGA	GAG	TTT	GTT	192
65	G	Y	T	E	Y	G	L	K	Q	A	K	Y	W	N	S	D	80
193	GGT	TAC	ACT	GAG	TAC	GGT	CTG	AAA	CAG	GCT	AAA	TAC	TGG	AAC	AGT	GAT	240
81	P	S	E	L	A	Q	M	R	A	A	K	E	R	Y	S	Q	96
241	CCT	TCA	GAA	CTG	GCT	CAG	ATG	AGA	GCT	GCG	AAG	GAG	AGG	TAC	TCC	CAA	288
97	H	N	I	N	I	D	Y	Q	A	A	L	D	K	S	V	Q	112
289	CAC	AAC	ATT	AAC	ATT	GAC	TAC	CAG	GCT	GCT	CTG	GAT	AAA	TCA	GTT	CAG	336
113	P	S	V	V	L	H	S	V	A	P	P	A	G	K	H	P	128
337	CCC	TCT	GTG	GTC	CTG	CAC	TCG	GTG	GCG	CCC	CCT	GCT	GGT	AAA	CAT	CCC	384
129	S	M	L	V	C	S	V	Y	D	F	F	P	K	H	I	R	144
385	TCC	ATG	TTG	GTC	TGC	AGC	GTC	TAC	GAC	TTC	TTC	CCC	AAA	CAC	ATC	AGA	432
145	V	S	W	L	R	D	G	Q	E	V	T	S	D	V	T	P	160
433	GTG	AGC	TGG	CTC	AGA	GAC	GGA	CAG	GAA	GTC	ACC	TCT	GAT	GTC	ACT	CCC	480
161	T	D	E	L	A	D	A	D	W	F	Y	Q	I	H	S	H	176
481	ACT	GAT	GAG	CTG	GCA	GAC	GCT	GAT	TGG	TTC	TAC	CAG	ATC	CAC	TCT	CAC	528
177	L	E	Y	T	P	R	S	G	E	K	I	S	C	V	V	E	192
529	CTG	GAG	TAC	ACG	CCC	AGG	TCT	GGA	GAG	AAG	ATC	TCC	TGT	GTG	GTG	GAG	576
193	H	A	S	L	K	E	P	L	V	T	D	W	D	P	S	M	208
577	CAC	GCC	AGC	CTG	AAA	GAA	CCT	CTG	GTT	ACT	GAC	TGG	GAC	CCG	TCC	ATG	624
209	P	E	S	E	R	N	K	I	A	I	G	A	S	G	L	I	224
625	CCT	GAG	TCT	GAG	AGA	AAC	AAG	ATC	GCC	ATC	GGA	GCC	TCC	GGA	CTG	ATC	672
225	L	G	L	I	L	S	L	A	G	F	I	Y	Y	K	R	K	240
673	CTG	GGT	CTG	ATC	TTG	TCT	CTG	GCT	GGA	TTC	ATC	TAC	TAC	AAG	AGG	AAG	720
241	A	R	G	R	I	L	V	P	S	N	*	S	W	S		254	
721	GCC	CGA	GGA	AGG	ATT	CTG	GTT	CCC	AGT	AAC	TGA	TCC	TGG	TCC		762	

Mosa-R41

1	M	A	S	S	F	L	S	F	S	L	L	F	I	T	S	L	16
1	ATG	GCT	TCA	TCC	TTT	CTC	AGC	TTC	TCC	CTC	CTC	TTC	ATC	ACC	AGC	CTC	48
17	Y	T	A	D	G	F	R	Y	F	E	T	D	R	C	V	F	32
49	TAC	ACA	GCA	GAT	GGA	TTC	CGG	TAT	TTT	GAG	ACA	GAT	CGC	TGT	GTG	TTT	96
33	N	S	T	D	L	K	D	I	K	Y	I	R	S	E	Y	Y	48
97	AAC	TCC	ACT	GAT	CTG	AAG	GAC	ATC	AAG	TAC	ATC	AGG	TCT	GAA	TAT	TAC	144
49	N	K	L	E	I	A	R	F	S	S	S	V	G	E	F	V	64
145	AAC	AAG	TTG	GAG	ATC	GCC	AGG	TTC	AGC	AGC	AGT	GTG	GGA	GAG	TTT	GTT	192
65	G	Y	T	E	L	G	L	R	W	A	K	Y	W	N	N	N	80
193	GGT	TAC	ACT	GAG	CTG	GGA	CTG	AGA	TGG	GCT	AAA	TAC	TGG	AAC	AAT	AAT	240
81	P	S	Y	L	A	Q	M	R	G	E	K	E	R	Y	C	Q	96
241	CCT	TCA	TAT	CTG	GCT	CAG	ATG	AGA	GGT	GAG	AAG	GAG	AGG	TAC	TGC	CAA	288
97	H	N	I	G	N	W	Y	S	N	I	L	D	K	S	V	Q	112
289	CAC	AAC	ATT	GGG	AAC	TGG	TAC	AGC	AAC	ATT	CTG	GAT	AAA	TCA	GTT	CAG	336
113	S	Y	V	R	L	H	S	V	A	P	P	A	G	K	H	P	128
337	TCC	TAT	GTC	AGA	CTG	CAC	TCG	GTG	GCG	CCC	CCT	GCT	GGT	AAA	CAT	CCC	384
129	S	M	L	V	C	S	V	C	D	F	F	P	K	H	I	R	144
385	TCC	ATG	TTG	GTC	TGC	AGC	GTC	TGC	GAC	TTC	TTC	CCC	AAA	CAC	ATC	AGA	432
145	V	S	W	L	R	D	G	Q	E	V	T	S	D	V	T	S	160
433	GTG	AGC	TGG	CTC	AGA	GAC	GGA	CAG	GAA	GTC	ACC	TCT	GAT	GTC	ACT	TCC	480
161	T	D	E	L	A	D	A	D	W	F	Y	Q	I	H	S	H	176
481	ACT	GAT	GAG	CTG	GCA	GAC	GCT	GAT	TGG	TTC	TAC	CAG	ATC	CAC	TCT	CAC	528
177	L	E	Y	T	P	R	S	G	E	K	I	S	C	V	V	E	192
529	CTG	GAG	TAC	ACG	CCC	AGG	TCT	GGA	GAG	AAG	ATC	TCC	TGT	GTG	GTG	GAG	576
193	H	A	S	L	K	E	P	L	V	T	D	W	D	P	S	M	208
577	CAC	GCC	AGC	CTG	AAA	GAA	CCT	CTG	GTT	ACT	GAC	TGG	GAC	CCG	TCC	ATG	624
209	P	E	S	E	R	N	K	I	A	I	G	A	S	G	L	I	224
625	CCT	GAG	TCT	GAG	AGA	AAC	AAG	ATC	GCC	ATC	GGA	GCC	TCC	GGA	CTG	ATC	672
225	L	G	L	I	L	S	L	A	G	F	I	Y	Y	K	R	K	240
673	CTG	GGT	CTG	ATC	TTG	TCT	CTG	GCT	GGA	TTC	ATC	TAC	TAC	AAG	AGG	AAG	720
241	A	R	G	R	I	L	V	P	S	N	*	S	W	S	C	P	256
721	GCC	CGA	GGG	AGG	ATT	CTG	GTT	CCC	AGT	AAC	TGA	TCC	TGG	TCC	TGT	CCC	768
257	W	S	W	S	W	S	C	P	264								
769	TGG	TCC	TGG	TCC	TGG	TCC	TGT	CCC	792								

Mosa-S-1

[illegible]

Mosa-S-2

[illegible]

