

**Variation at Major Histocompatibility Complex Class I Loci  
In Two Killifish Species with Reduced  
Genetic Variance**

Michael Todd Fisher

Dissertation submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
in  
Biology

Bruce J. Turner, Chair  
Brenda S. Winkel-Shirley  
Stephen A. Smith  
Klaus D. Elgert  
Prakash S. Nagarkatti

October 23, 2000  
Blacksburg, Virginia

Keywords: Conservation, Inbreeding, Self-fertilization, *Rivulus marmoratus*, *Cyprinodon diabolis*, Immunogenetics

# Variation at Major Histocompatibility Complex Class I Loci In Two Killifish Species with Reduced Genetic Variance

Michael T. Fisher

(Abstract)

The ability of natural selection to promote and preserve genetic variation at Major Histocompatibility Complex (MHC) class I loci was examined in two fish species known to have low genetic variation either as a consequence of their breeding system or population structure. The tempo and modes of molecular evolution acting on these loci has also been discussed. The marine killifish *Rivulus marmoratus* is the only vertebrate known to exist in nature in homozygous form. The findings of this study suggest that MHC class I genes at one locus in *R. marmoratus* have rapidly accumulated variation, particularly in region of the gene encoding functional important domains. This rapid accumulation of variation is likely due to the increased intensity of natural selection acting on these genes resulting from the homozygosity of the species. The contention that the variation characterized at one MHC locus is of recent origin is further supported by an analysis of the mitochondrial control region of *R. marmoratus*, which suggests that the individuals included in this study are recently diverged. The variation characterized here is the first evidence of genetic variation at coding loci in this species, and may be evidence of the cost of homozygosity for *R. marmoratus*.

Variation at a single MHC class I locus was also characterized in several members of the Death Valley pupfish species complex. Two of these species, *Cyprinodon diabolis* and *Cyprinodon salinus* have undergone intense historical or ongoing population

bottlenecks. Despite the action of random genetic drift acting on these species, both maintain multiple MHC class I alleles at one locus. The alleles characterized in *C. diabolis* and *C. salinus* were also recovered from other members of the Death Valley complex. This suggests that overdominant natural selection has acted to maintain ancestral genetic variation at this locus in both species. This result has important implications for the design of conservation programs for these endangered species, and may have broader implications for the design of captive breeding programs of species with highly reduced population sizes in general.

## **Acknowledgments**

I would like to thank my advisor, Bruce Turner, for his years of help, and most importantly, for the opportunity he has given me. For this, I owe a debt of gratitude. I would also like to thank my committee, Drs. Prakash Nagarkatti, Stephen Smith, Brenda Winkel-Shirley, and Klaus Elgert, for their support and guidance. Felipe Figueroa and Jan Klein have given me invaluable training, without which this project would not have been possible.

I owe a debt to the many people who have collected the samples I have used in this study, and in particular Scott Taylor and Will Davis. I also thank those who have helped me through numerous discussions on either conceptual or technical matters, including Tom Bunt, Jason Bond, Lee Weight, Alicia Schulteis, and Liz Tucker.

## **Granting Agencies and Collecting Permits**

This work was supported by grants from the Jeffress Memorial Trust. Collections of Death Valley pupfish were under the auspices of the following permits or special licenses issued by units of the U.S. Department of the Interior or by state agencies: Endangered Species Permit PRT-769851 and Special Use Permit 56034, U.S. Fish and Wildlife Service; Collecting Permit A9103, Death Valley National Monument, National Park Service; 6840, California Desert District Office, Bureau of Land Management; S9494, Nevada Department of Wildlife, and a Memorandum of Understanding with the California Department of Fish and Game.

## Table of Contents

Abstract.....	i
Acknowledgments.....	iii
Granting Agencies and Collecting Permits.....	iv
Introduction.....	1
Chapter 1: Variation at MHC class I loci in a naturally homozygous organism, the killifish, <i>Rivulus marmoratus</i>	
Introduction.....	9
Materials and Methods.....	13
Results.....	18
Discussion.....	34
Chapter 2: Limited mitochondrial control region variation in Central American and Caribbean populations of the self-fertilizing, hermaphroditic killifish, <i>Rivulus marmoratus</i>	
Introduction.....	41
Materials and Methods.....	44
Results and Discussion.....	48
Chapter 3: Variation in a single Major Histocompatibility Complex class I locus in the Death Valley Pupfish complex ( <i>Cyprinodon spp.</i> )	
Introduction.....	51
Materials and Methods.....	55

Results.....	59
Discussion.....	68
Conclusions.....	73
Literature Cited.....	76
Curriculum Vitae.....	82

## List of Figures

Figure 1.1 MHC I Haplotype polymorphism among siblings in <i>R. marmoratus</i> .....	20
Figure 1.2 Phylogenetic trees of exon and intron sequences.....	21-25
Figure 1.3 Alignment of locus MHC <i>UA-Db</i> $\alpha$ 1 region amino acid sequences.....	29
Figure 1.4 Alignment of <i>Rima Mhc UA</i> $\alpha$ 1 sequences.....	32
Figure 2.1 TCS network of <i>R. marmoratus</i> mtDNA haplotypes.....	47
Figure 3.1 Resolution of heterologous PCR products.....	60
Figure 3.2 Alignment of <i>Cyne UA</i> $\alpha$ 1 region sequences.....	63
Figure 3.3 Alignment of <i>Cyne UA</i> $\alpha$ 1 amino acid sequences.....	64



## List of Tables

Table 1.1 Collection locations and samples sizes of <i>R. marmoratus</i> .....	15
Table 1.2 Comparison of the ration of synonymous and non-synonymous substitutions at MHC I loci between <i>R. marmoratus</i> and <i>O. tshawytscha</i> .....	30
Table 2.1 Collection locations and samples sizes of <i>R. marmoratus</i> .....	45
Table 3.1 Collection locations and sample sizes of Death Valley pupfishes.....	56

## List of Abbreviations

ARS	Antigen recognition site
$\alpha_1$	$\beta_2$ -microglobin binding region of the class I MHC molecule
$\alpha_2$	ARS region of the class I MHC molecule
$\alpha_3$	ARS region of the class I MHC molecule
CTL	Cytotoxic T lymphocyte (cytotoxic T-cell)
$d_N$	Rate of non-synonymous nucleotide substitution per non-synonymous site
$d_S$	Rate of synonymous nucleotide substitution per synonymous site
HLA	Human leukocyte antigen, the human MHC
LINE	Long interspersed nuclear element
MHC	Major Histocompatibility Complex
PCR	Polymerase chain reaction
RAPD	Random amplification of polymorphic DNA
TCR	T-cell receptor
TCS	Templeton, Crandall, and Singh (authors of an algorithm for haplotype network construction)
VNTR	Variable number tandem repeat

## Introduction

The genes of the Major Histocompatibility Complex (MHC) encode cell-surface glycoproteins that bind fragments of peptides derived from for antigens and present them to the T-cells of the immune system (Klein 1986). The interaction between the T-cell receptor (TCR) and the MHC/peptide complex is the initiating step of an immune response against the pathogens from which the peptides were derived. The MHC family of genes is divided into two major subfamilies, class I and class II. Class I MHC molecules present peptides derived from intra-cellular antigens and class II genes present peptides derived from extra-cellular antigens. These genes are characterized by extraordinary levels of polymorphism and heterozygosity, with many alleles present at intermediate frequencies in natural populations (Klein 1986). This pattern of polymorphism is consistent with the action of balancing selection, or overdominance (Hedrick and Thompson 1983). We have characterized variation at MHC class I loci in two fish species that are known to possess extremely low levels of genetic variation and whose breeding system or population structure promote intense inbreeding. These conditions test the limits of natural selection to promote and maintain genetic variation at MHC loci.

### *Structure and Function of MHC class I molecules*

MHC class I molecules are present on the surface of all somatic nucleated cells. The molecule is a heterodimer composed of the MHC encoded chain and the  $\beta_2$  microglobin molecule, which is not encoded in the MHC gene complex. The class I

molecule consists of several domains, each roughly corresponding to individual exons. These domains are the extra-cellular domains (designated  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ ), the transmembrane domain, and the cytoplasmic tail. The  $\alpha_1$  and  $\alpha_2$  domains form the antigen recognition site (ARS) which is the region of the molecule that binds foreign antigen derived peptides. This region is encoded by exons 2 and 3 of the MHC class I gene (Bjorkman *et al.* 1987). The  $\alpha_3$  region (encoded by exon 4) is non-covalently associated with  $\beta_2$  microglobulin.

Absent infection, class I genes bind and present peptides derived from the cellular proteins, often proteins encoded within the MHC. However, cytotoxic T-cells (CTL), which constantly survey cell the surface, do not recognize these MHC class I/self-peptide complexes because the process of T-cell education in the thymus does not allow the release of T-cells bearing receptors that recognize (bind to) self-peptides in complex with MHC class I molecules. During infection, self-peptides are replaced by foreign antigen derived peptides in the ARS of class I molecules (Bjorkman *et al.* 1987). This complex is recognized by CTLs, which initiate a cytotoxic response resulting in the death of the cell presenting foreign peptide in complex with and MHC class I molecule. CTLs recognize foreign peptides only when presented by a MHC class I molecule, a process known as MHC class I restriction. This process provides a targeted, intensive response against cells that harbor foreign antigen.

#### *Overdominant selection on MHC genes and MHC polymorphism*

Doherty and Zinkernagel (1974) first advanced the overdominance hypothesis for MHC polymorphism. They demonstrated that MHC molecules differed in the array of

peptides that they could present. Therefore, they hypothesized that individuals heterozygous at MHC loci would be able to bind and present a broader array of antigen-derived peptides than homozygous individuals. These heterozygous individuals would then be able to mount an immune response against a broader array of pathogens, or a more intense immune response against pathogens. However, this hypothesis is difficult to test directly because almost all individuals in natural populations are heterozygous at MHC loci, so the performance of homozygotes in natural populations cannot easily be measured.

The overdominance hypothesis has been tested by examining the pattern of nucleotide substitution at MHC loci (Hughes and Nei 1988, 1988). In most genes, the rate of synonymous nucleotide substitution ( $d_S$ ) exceeds the rate of non-synonymous nucleotide substitution ( $d_N$ ), reflecting the fact that genes generally evolve under purifying selection, and non-synonymous substitution resulting in changes in the amino acid sequence of a molecule are often deleterious. Hughes and Nei (1988) proposed that overdominant selection should enhance the rate of non-synonymous nucleotide substitution in regions of the MHC molecule that were involved in peptide binding. In an analysis of human MHC class I (HLA) genes, Hughes and Nei demonstrated convincingly that  $d_N$  was significantly greater than  $d_S$  in the regions of the class I gene encoding the ARS. Also, the rates of  $d_S$  compared among different regions of the gene were not different, evidence that MHC polymorphism was not due to an enhanced rate of nucleotide substitution.

It has also been demonstrated that amino acid substitutions in the ARS of the MHC molecule tend to be radical rather than conservative. Radical amino acid

substitutions are those in which involve changes in characteristics including charge, size, and polarity (Zuckerkindl and Pauling 1965; Zhang 1999). As the function of the ARS is to bind peptides, a process mediated by the side chains of the amino acids bordering the ARS, radical amino acid substitutions are likely to alter the peptide binding spectra of the molecule. The high rate of non-synonymous nucleotide substitution and the preponderance of radical amino acid substitutions in the ARS are strong evidence that MHC genes have evolved under overdominant natural selection (Hughes and Nei 1988, 1989).

#### *Trans-species polymorphism at MHC loci*

Additional support for the overdominance hypothesis as an explanation for MHC polymorphism is a phenomenon termed trans-species polymorphism (Gyllensten and Erlich 1989; Lawlor *et al.* 1988). Many allelic lineages are ancient, so much so that their rise predates speciation events. Humans and chimpanzees share allelic lineages, which suggests that these lineages have persisted since the divergence of the species. Under neutral selection, polymorphisms are not expected to persist for long periods. However, selection such as overdominance can cause polymorphism to persist for a period much longer than predicted under neutrality (Takahata and Nei 1990)

#### *Gene conversion at MHC loci*

There is evidence from human (HLA) MHC genes that inter-locus and intra-locus gene recombination (gene conversion) events have acted to generate new allelic variants (Kuhner *et al.* 1990; Parham and Ohta 1996; Ohta 1991). These putative recombination

events typically involve the exchange of short segments of DNA. The evidence for the action of gene conversion is the finding of alleles that differ from other alleles by short segments present in third alleles (Allsopp *et al.* 1991). Such allelic variants have been detected in populations of South American Amerindians. The alleles detected in this population are novel, and not found in Asian populations from which Amerindians are derived, and therefore must have evolved since the divergence of the two populations, or within the last 10 – 20,000 years. These novel allelic variants could not have arisen by simple accumulation of point mutation as insufficient time has passed for such mutations to accumulate (Belich *et al.* 1992; Watkins *et al.* 1992; Parham and Ohta 1996). There is also evidence that the rate of gene conversion at MHC loci may be quite high. A study of MHC class I loci in human sperm from individuals heterozygous at six HLA loci suggested that the rate of intra allelic gene conversion may be as high as  $0.81 \times 10^{-3}$  gametes per generation (Zangenberg *et al.* 1995). A similar study in mice suggests that the rate of inter allelic gene conversion may be as high as  $0.5 \times 10^{-6}$  gametes per generation (Hongstrand and Bohme 1994). Both of these rates are high enough to impact the level of variation in these genes. Although controversial, largely because an adequate mechanism through which gene conversion might act has yet to be advanced, it seems likely that gene conversion has acted in the evolution of MHC genes.

#### *Characterization of MHC genes in teleost species*

Much of the characterization of the processes involved in the evolution and maintenance of polymorphism has been done in primates and rodents, especially in the human and murine systems. However, the MHC of numerous fish species has been

characterized. These include several cichlid species from the African Rift lakes (Ono, O'hUigin, Tichy, and Klein 1993; Klein, Ono, and O'hUigin 1993; Sato *et al.* 1997), the zebrafish, *Brachydanio rerio* (Takeuchi, Figueroa, O'hUigin, and Klein 1995), the guppy, *Poecilia reticulata* (Sato *et al.* 1995), many salmonid species (Miller, Withler, and Beacham 1997; Glamann 1995; Grimholt, Hordvick, and Fosse 1993; Miller and Withler 1996; Kim, Parker and Hedrick 1999), the Gila topminnow, *Poeciliopsis o. occidentalis* (Hedrick and Parker 1998) and the coelacanth (Betz, Mayer, and Klein 1994).

### *Rivulus marmoratus*

The marine killifish, *Rivulus marmoratus*, is the only vertebrate known to reproduce by virtually obligate internal self-fertilization (Harrington 1961). Self-fertilization is the ultimate form of inbreeding, and results in homozygosity in a small number of generations (Maynard Smith 1978). Once homozygosity is reached parents and progeny are genetically identical, and the breeding system is effectively clonal. In fact, almost all individuals collected from natural populations have proven to be homozygous at highly variable VNTR loci by DNA fingerprint/progeny testing (Turner *et al.* 1992).

In general, homozygosity is thought to be deleterious, either because of the loss of overdominant effects of multiple alleles or due to the appearance of deleterious effects of recessive alleles appearing in homozygous form. However, it has proven difficult to detect any deleterious effects that homozygosity may have on *R. marmoratus*. We hypothesize that the effects of homozygosity should be evident in the MHC of *R. marmoratus*, or, put another way, that the homozygosity of *R. marmoratus* will increase



the intensity of natural selection on their MHC loci. The increased intensity of this selection will be evident in high numbers of non-synonymous substitutions at amino acid positions interacting with bound peptide (ARS). We have characterized MHC class I variation at several loci in *R. marmoratus* collected over a wide geographic range and determined that the pattern of nucleotide substitution at one locus are consistent with the hypothesis that natural selection acting on this locus is intense. We have also characterized variation in the mitochondrial control region of the individuals examined in this study to determine time frame over which the variation at MHC loci may have evolved (Awise 1994; Kocher and Carleton 1997; Lee, Conroy, Howell and Kocher 1995). The individuals in this study have extremely limited variation in the mitochondrial control region, suggesting that the divergence time between these individuals is short, and therefore, that the variation characterized at MHC class I loci has evolved recently.

#### *Cyprinodon diabolis and Cyprinodon salinus*

The Death Valley pupfishes are a diverse group of species stranded in various refugia by the drying of Lake Manly at the end of the Pleistocene (Miller 1948; Soltz and Naiman 1978; Soltz and Hirshfield 1981; Echelle and Dowling 1992; Echelle and Echelle 1993). Two species, *Cyprinodon diabolis* and *Cyprinodon salinus* have apparently undergone intense historical or ongoing population bottlenecks (Turner 1974; Duvernell and Turner 1998). Studies of neutral genetic marker loci indicate that these two species both have extremely limited genetic variation. The population structure of *C. diabolis* in particular has contributed to intense inbreeding (Chernoff 1985). The species is the most

highly localized of any vertebrate, inhabiting a narrow shelf in an isolated sink (Devil's Hole) in Ash Meadows, Nye County, Nevada, USA. Also, the population size has ranged from 250-500 individuals since annual censuses began in the 1970s. It is likely that the population has not been significantly larger since the species was stranded in Devil's Hole 10-15,000 years ago (Soltz and Naiman 1978). *C. salinus* inhabits a large spring at the head of Salt Creek, also in Ash Meadows, and currently has a census size in the thousands, but the effects of past population bottlenecks are still evident in the low genetic variation of this population (Duvernell and Turner 1998).

Other populations of Death Valley pupfish, in particular *Cyprinodon nevadensis* and *Cyprinodon amargosae* do not exhibit evidence of historical population bottlenecks (Echelle and Echelle 1993; Duvernell and Turner 1998). Therefore, these populations will likely maintain genetic variation similar to the ancestral population of pupfish that inhabited Lake Manly, and gave rise to *C. diabolis* and *C. salinus*. By comparing the levels of variation at MHC loci in *C. nevadensis* and *C. amargosae* with the variation in *C. diabolis* and *C. salinus*, it will be possible to determine how much ancestral MHC variation, if any, remains in these two species. We hypothesize that overdominant natural selection has acted to preserve MHC variation in these species. As both of these species are currently listed as endangered, and efforts are ongoing to preserve the species through propagation in artificial refugium, it is critical to determine if these species do maintain adaptive genetic variation, and if so, to ensure that conservation efforts are targeted toward the preservation of this variation (Hughes 1991, Turner 1984).

**Chapter 1: Variation at MHC Class I loci in a naturally homozygous organism,  
the killifish, *Rivulus marmoratus***

**Introduction**

*Rivulus marmoratus*

The marine killifish *Rivulus marmoratus* inhabits coastal mangrove swamps from Southern Florida, USA to at least as far south as Rio de Janeiro, Brazil, as well as throughout the Caribbean and probably Central America (Taphorn 1980). *Rivulus marmoratus* is the only vertebrate known to reproduce by virtually obligate internal self-fertilization (Kallman and Harrington 1961, Turner *et al.* 1990). Populations are composed almost exclusively of fully functional hermaphrodites. Rare males have been described in a few natural populations, but they generally comprise less than 1% of the population (Turner, Davis, and Taylor 1991). These males do not routinely participate in matings (but see Lubinski *et al.* 1995), and no mechanism has been described whereby hermaphrodites are able to curtail spermatogenesis and ovulate viable but unfertilized eggs. True females are entirely unknown.

Iterative self-fertilization is the ultimate form of inbreeding, and quickly leads to homozygosity. Transplantation studies using fins, scales, and hearts suggests that grafts are readily accepted between parental hermaphrodites and their progeny, and between siblings (Kallman and Harrington 1964). The absence of a typical graft rejection response is good evidence for homozygosity at loci involved in this response. More recently, progeny testing using multi-locus VNTR (Variable Number Tandem Repeat)

DNA fingerprinting suggested that almost all *R. marmoratus* in natural populations are homozygous (Turner *et al.* 1992). Because parents and progeny are genetically identical, this mode of reproduction is effectively clonal. However, individuals collected from natural populations do vary in their DNA fingerprint patterns, so that at least at VNTR loci, there is population level genetic variation. Therefore, the population structure is described as “arrays of homozygous clones”.

It is generally accepted that inbreeding leads to lowered fitness. This reduction in fitness is due to the appearance of deleterious recessive genes in homozygous form. Also, self-fertilization is thought to be a long-term evolutionary dead end because sex cannot act as a mechanism through which favorable combinations of new mutations can arise in a population through recombination (Smith 1978). However, it is not clear that *R. marmoratus* suffers a significant reduction in fitness due to its homozygosity. At most, it may experience increased developmental instability during embryogenesis (unpublished data). Also, *R. marmoratus* appears to have almost no identifiable genetic variation at the population level beyond that at hyper-variable VNTR. Studies of 31 allozyme loci found no variation among individuals collected from Florida and the Turks and Caicos islands (Vrijenhoek 1985). Also, this species has extremely limited variation in the mitochondrial control region (unpublished data). Therefore, *R. marmoratus* inhabits a challenging environment, the coastal mangroves, over an enormous range despite being homozygous and apparently lacking adaptively significant genetic variation at the population level.

### *The Major Histocompatibility Complex*

The genes of the Major Histocompatibility Complex (MHC) are loci at which overdominant selection acts to maintain levels of genetic variation and heterozygosity much higher than predicted by neutral theory (Hedrick and Thomson 1983). In fact, MHC genes, which have been found in all vertebrate species examined, generally have the highest levels of polymorphism and heterozygosity of any protein coding sequences yet characterized (Klein 1986). MHC genes encode cell surface glycoproteins that bind and present peptides derived from foreign antigens to the T-cells of the immune system as the initiating step of the innate immune response. There are two classes of MHC genes. Class I MHC genes present peptides derived from intracellularly processed antigens (viruses) and are present on all cell types, while Class II genes present peptides derived from extracellular antigens. The spectrum of peptides that can be bound by a specific MHC molecule, and therefore the pathogens against which an immune response can be mounted, are determined by the amino acid motifs present in the antigen recognition site (ARS) of the molecule (Klein 1986). Because MHC genes are co-dominantly expressed, heterozygotes with alleles that differ in their peptide binding spectra presumably are able to bind and present peptides derived from a broader array of antigens than are MHC homozygotes (Doherty and Zinkernagel 1975).

Variation at MHC loci is typified by a high number of non-synonymous substitutions in the region encoding the ARS of the molecule (Hughes and Nei 1988). Alleles at a single locus can differ by up to 30% of their amino acid sequence. Arden and Klein (1982) formulated the *trans-species theory* which proposes that newly evolved species arise with a complete suite of MHC alleles carried through the speciation process

from the ancestral species. A corollary of this theory is that many MHC alleles are older than species, so much so that some class II allelic lineages are shared between humans and apes (Lawlor *et al.* 1988).

It is of interest to examine any effect that the homozygosity of *R. marmoratus* might have had on the evolution of their MHC genes. Here we report on variation at MHC Class I loci in individuals of *R. marmoratus* collected over a wide geographic range. We characterize the variation at these loci and discuss the likely effects that homozygosity has had on the rate of evolution of genetic variation at MHC Class I genes.

## Materials and Methods

### *Samples and DNA preparation*

Specimens of *R. marmoratus* were collected using dip nets, miniature hook-and-line, or traps by Bruce Turner, William Davis, and D. Scott Taylor. Individuals were either shipped live in moist cotton or preserved in the field in 100% ethanol. A complete list of collection locations and dates is found in Table 1.1. Genomic DNA was extracted from the ethanol-fixed specimens using either a standard phenol:chloroform extraction procedure or the Puragene DNA Extraction Kit (Gentra Systems, Minneapolis, MN).

### *MHC Primers, PCR amplification, and automated sequencing*

Degenerate PCR primers were designed based on conserved regions of exons 2 and 3 of MHC class I genes from a variety of teleost fishes. These exons encode the  $\alpha 1$  and  $\alpha 2$  regions of the MHC molecule, which form the peptide-binding domain. This is the portion of the molecule upon which diversifying selection operates. The primer sequences were: mhc 1 G5 (sense) 5'-GCW GTG ACT CAC WCT CTG AAG-3' and mhc 1 3R (antisense) 5'-CCT CCC ATA GTT CAC ATA CTT CTT-3'. PCR reactions were prepared with MasterAmp PCR premix F and MasterAmp *Taq* DNA polymerase (Epicentre, Madisons, WI) according to the manufacturer's protocol. Standard PCR amplifications without the use of commercial additives were far less successful, possibly due to secondary structure within the sequence. Reactions were cycled on a PTC-100 thermocycler (MJ Research, Incline Village, NV) as follows: 94°C (3:00 m), 6 cycles of 94°C (30 s), 46°C (30 s), and 72°C (75 s), followed by 24 cycles of 94°C (30 s), 48°C (30

s), and 72°C (75 s). This reaction produced multiple bands, reflecting the fact that the primers



Table 1.1. Collection locations and sample sizes of *R. marmoratus* specimens.

Collection Location	N
Florida	
Brevard County	10
Charlotte City	6
Marco Island	6
Key West	1
Mate Cumbe	8
No Name Key	4
Everglades	8
Lost Man River	5
Dolphin Smile	3
Belize	
Dangriga (mainland)	10
Twin Cays (heterozygous)	11
Honduras	8
Panama	1
Bahamas	5
Total	86

were not specific to particular loci. The bands were excised from 1-% agarose gels and prepared for direct sequencing using Qiagen's Gel Extraction Kit (QIAGEN, Valencia, CA). In this case, it is possible to sequence PCR products directly because these individuals are homozygous. Sequencing was done using DigDye terminator chemistry (PE Biosystems, Foster City, CA), purified using Qiagen's DyeEX spin columns (QIAGEN, Valencia, CA) and the reactions were run on an ABI PRISM 310 Genetic Analyzer (PE Biosystems, Foster City, CA).

#### *Determination of Orthology/Paralogy, Phylogenetic Analysis*

One of the pitfalls of working with a multigene system is that it can be difficult to determine if two sequences from different individuals are orthologous without direct genetic information. We have used several lines of evidence to assign sequences to specific loci so that orthologous comparisons could be made. Initially, PCR products were assigned to a specific locus based on size, on the assumption that the size of intron 2 would be conserved within loci, but not across loci. Once these PCR fragments were sequenced, phylogenetic tools were used to resolve the relationships of exon 2 sequences and, separately, exon 3 sequences. Finally, the sequences of a portion of intron 2 were subjected to similar analysis. Sequences were aligned using CLUSTAL W (Higgins *et al.* 1996). The limits of the exons were determined by alignment with known MHC sequences from other teleost species. Dendrograms of exon 2, exon 3, and intron 2 sequences were constructed in PAUP\* 4.02b (Swofford 1999), using Felsenstein's 1981 distance model. The congruence between the topology of the three trees and the initial locus assignments allows us to be confident that the comparisons that we have made are

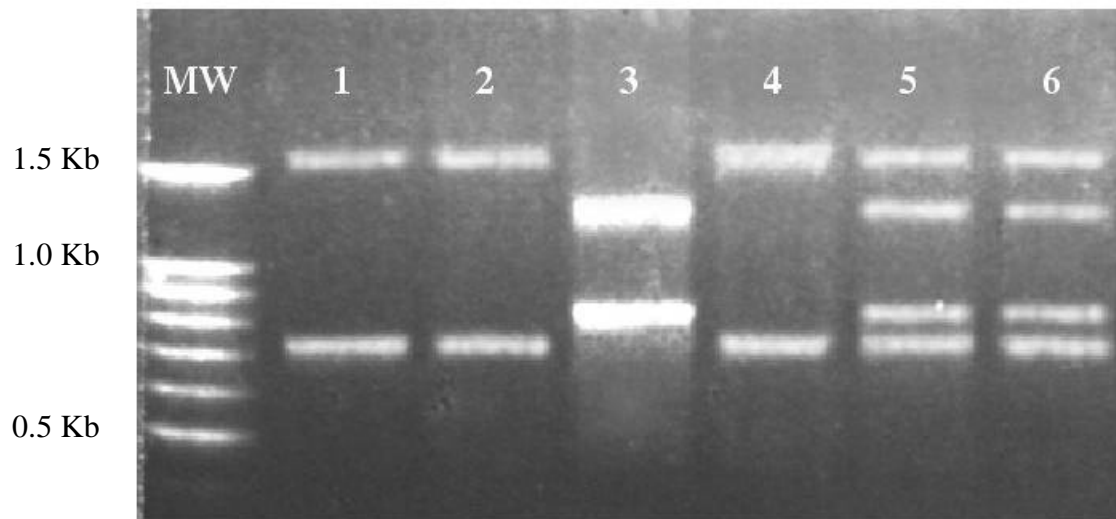
valid. This is because the congruence between the topologies of phylogenetic trees derived from intron and exon data provide independent diagnosis of orthology, which can be used to confirm each other. Independent phylogenetic analysis of intron and exon sequences may then be used to identify a gene as belonging to a particular locus. We recognize that the two data sets, intron and exon, are not entirely independent as they are linked, but the selection pressures driving the evolution of the two regions are so different as to at least partially decouple the evolution of one from the other. There are situations where the approach we have chosen could give misleading results. This method could lead to a underestimate of the number of loci if the genes had very recently undergone a duplication event. These duplicated, and therefore paralogous, genes might not have diverged sufficiently to be identified as separate loci. However, gene duplication events are rare and this is therefore unlikely.

## Results

### *Haplotype Polymorphism*

PCR amplification using MHC Class I-specific primers in six progeny of a known heterozygous individual suggest that at least two distinct MHC Class I haplotypes exist in *R. marmoratus*. The presumptive intron sequences of amplicons separated by gel electrophoresis differed markedly, and exon sequences appeared in different clades when subjected to phylogenetic analysis. The two haplotypes found in the parental heterozygous hermaphrodite appear to assort among progeny (Figure 1.1). It is therefore likely that each amplicon represents a distinct locus. Haplotype polymorphism was subsequently noted in 2 other broods of heterozygous hermaphrodites.

Haplotype polymorphism might explain why we were unable to amplify all loci from all individuals. Generally, we were only able to amplify one to three loci from each individual. We presume loci that were not recovered from individuals did not exist in the MHC haplotype possessed by those individuals. Therefore, it is unlikely that loci were missed due to polymorphism in priming sites, so that some loci amplified more readily than others and the missing loci were out-competed during PCR. The degenerate primers used here were designed based on conserved regions of exons 2 and 3 in a number of divergent fish taxa, including the guppy, stickleback, and several cichlid species. These primers have successfully amplified multiple loci in several Cyprinodont species, and one Poecileid.



**Figure 1.1.** Haplotype polymorphism demonstrated in the progeny of a known heterozygous individual. Each PCR band represents a single putative MHC Class I locus. Two distinct haplotypes are present, with lanes 1,2, and 4 showing amplification from individuals homozygous for one haplotype, lane 3 showing amplification from an individual heterozygous for another haplotype, and lanes 5 and 6 showing amplification from heterozygous individuals. The first lane is a molecular weight marker.

### *Geographic Distribution of MHC Class I Loci*

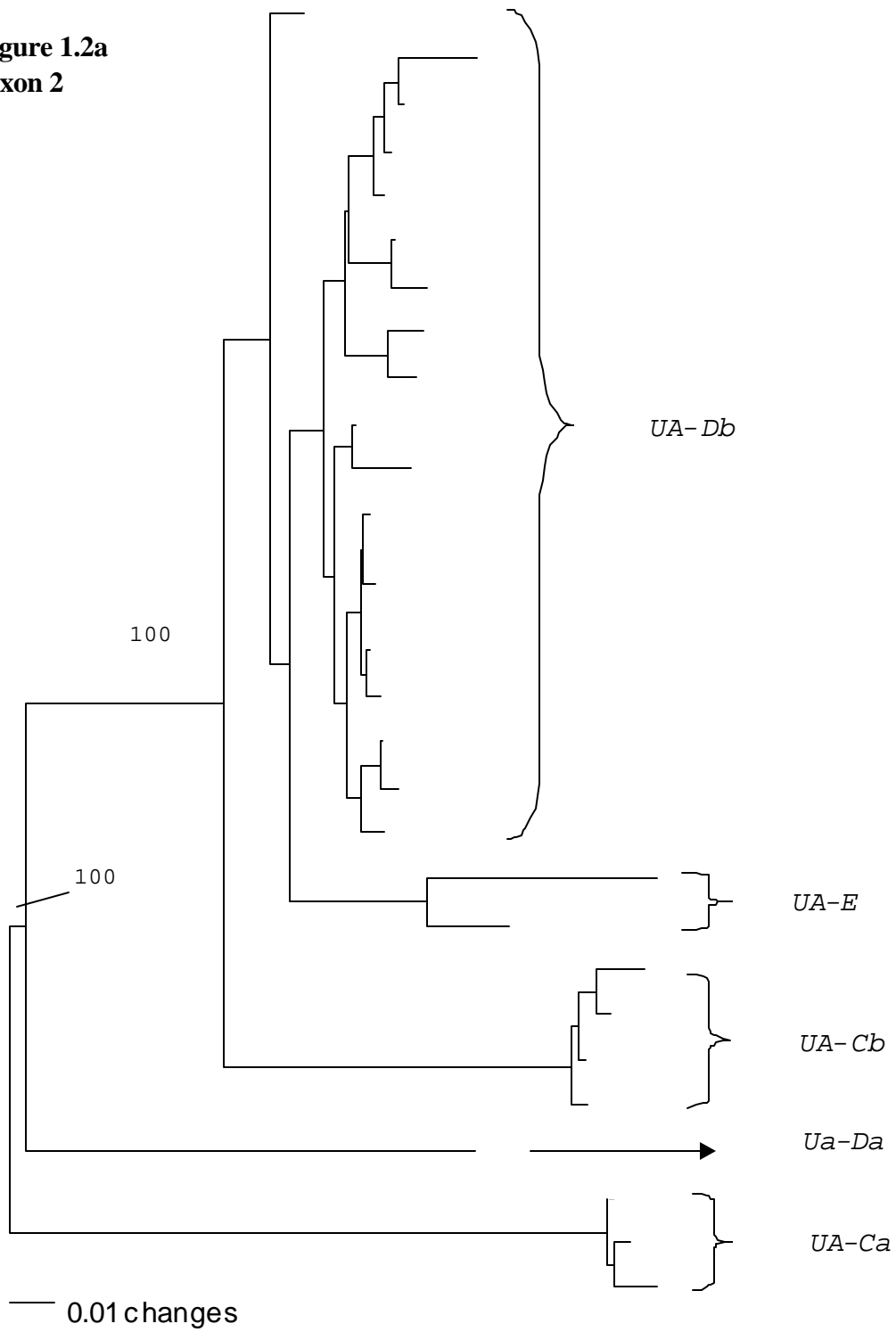
Not all loci were recovered from all geographic locations. Locus *Rima-UA-Db* (nomenclature after Klein *et al.* 1990) was found in Florida, Belize Cays, the Belize mainland, and Panama. Locus *Rima-UA-E* was recovered from Floridian and Honduran samples, locus *Rima-UA-Da* was recovered from Bahamian and Honduran samples. Locus *Rima-UA-Ca* was found in Belize mainland and Belize Cays samples, while locus *Rima-UA-Cb* (pseudogene) was found in Florida and mainland Belize.

### *Phylogenetic Analysis*

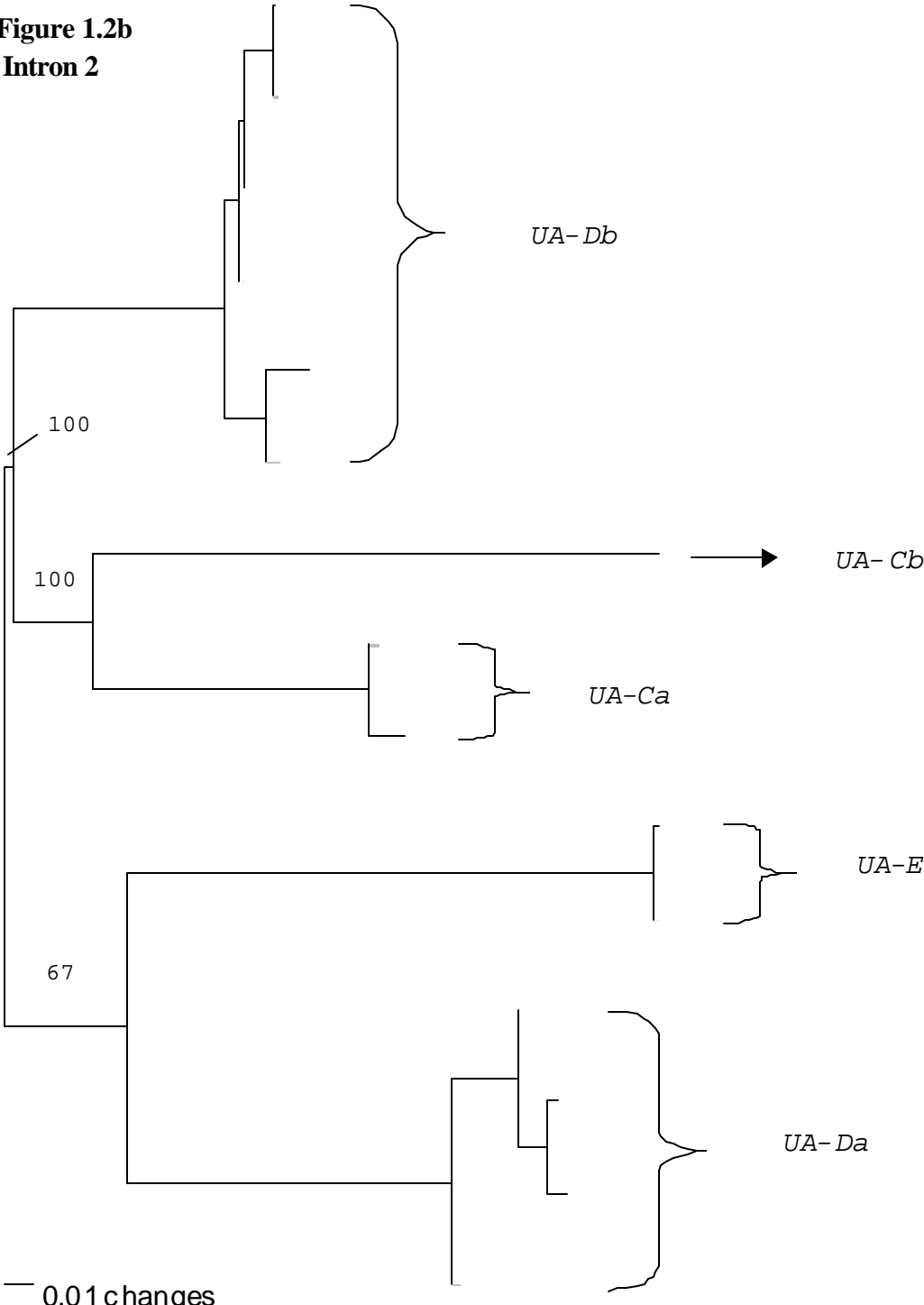
A total of 29 unique MHC Class I  $\alpha 1$  sequences were recovered from 87 *R. marmoratus* individuals. Also, one  $\alpha 1$  sequence was recovered from two *Rivulus caudomarginatus*, and one sequence was recovered from two *Rivulus tenuis*. These sequences formed five well-supported clades when subjected to phylogenetic analysis (Figure 1.2a). Sequences of exon 3 were analyzed separately in the same fashion (Figure 1.2b). Both phylogenies of exon data share the same general topology. Intron 2 sequences were also independently subjected to similar analysis (Figure 1.2c). All five clades in the tree of exon data were also present in the tree generated using intron data.

The topology of all three trees suggests that these sequences represent five distinct MHC class I loci, which we have named *MhcRima-UA-Ca*, *-Cb*, *-Da*, *Db*, and *E*. The sequences isolated from *R. caudomarginatus* group with the clade formed by alleles at the putative *Rima-UA-Cb* locus, while the *R. tenuis* sequence was highly divergent and serves as an

**Figure 1.2a**  
**Exon 2**

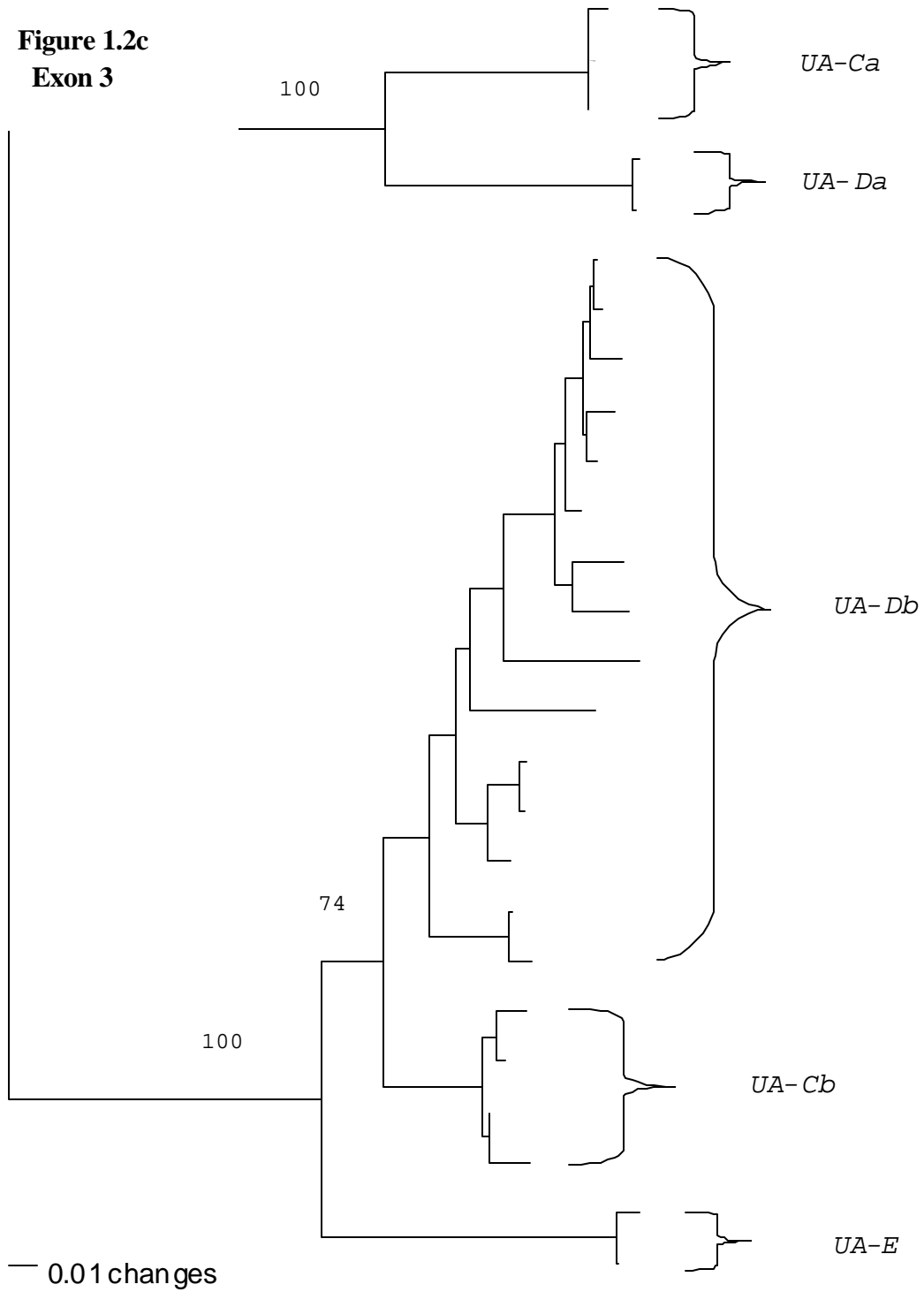


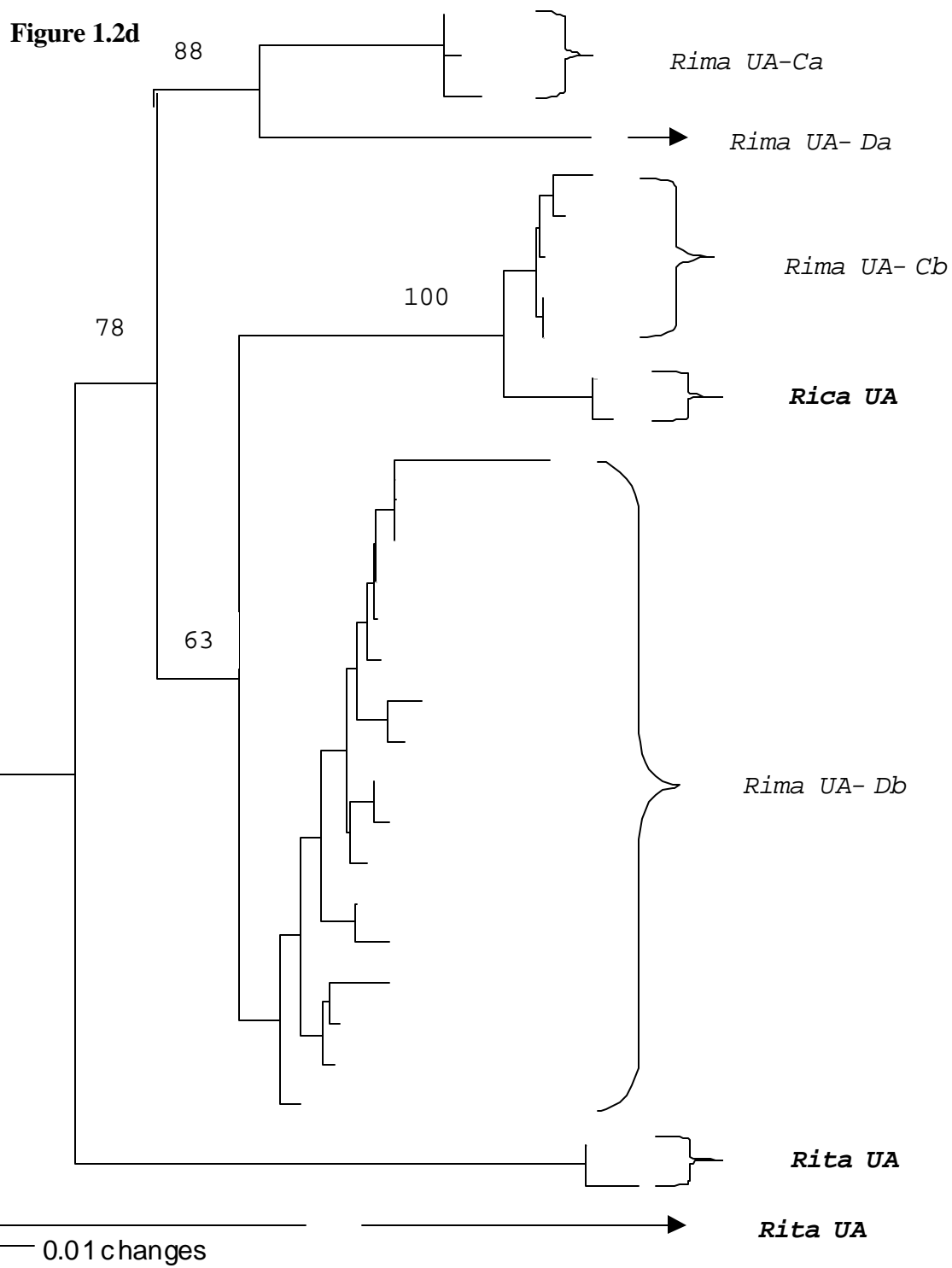
**Figure 1.2b**  
**Intron 2**





**Figure 1.2c**  
**Exon 3**





**Figure 1.2.** Phylogenetic analyses of MHC Class I exon 2 and 3 sequences (Figures 1.2a and c respectively), and intron 2 sequences (Figure 1.2b) recovered from *R. marmoratus*. Phylogenetic trees were produced for exon sequences using Felsenstein's 1984 Distance method, and using Felsenstein's 1981 Distance method for intron sequences. Trees produced for these data using Likelihood and Parsimony methods did not differ significantly in their topologies. Numbers above nodes indicate bootstrap support generated over 100 replications. Brackets indicate clades composed of alleles at different putative MHC Class I loci. Sequences recovered from *R. caudomarginatus* (*Rica UA*) and *R. tenuis* (*Rita UA*) are included in Figure 1.2d.

outgroup (Figure 1.2d) . The fact that sequences recovered from *R. caudomarginatus* form a clade close to one sequence from one putative locus in *R. marmoratus* supports our contention that each clade represents sequences at different loci. The branch lengths separating *R. caudomarginatus* sequences from *R. marmoratus* locus *UA-Cb* sequences are shorter than the branches separating the clades of putative different class I loci from *R. marmoratus* (Figure 1.2a). It is unlikely that homologous alleles within a species have diverged to a greater degree than alleles at homologous loci between species, especially considering that *R. marmoratus* and *R. caudomarginatus* likely diverged 20 mya. There is complete congruence between the clades present in the trees generated from exon 2, exon 3 and intron 2 sequences, so that the membership in a clade/locus suggested by exon data is supported by intron data. In all analyses, bootstrap support for the nodes separating putative loci is very high (Figures 1.2a-d).

All the sequences hypothesized to be alleles at the locus *Rima-UA-Cb* have a stop codon at position 56. However, these alleles do not appear to have accumulated any synonymous substitutions. This indicates that these are almost certainly recently inactivated pseudogenes.

#### *Polymorphism at locus Rima-UA-Db*

##### Exon 2 $\alpha$ 1 Region-locus *UA-Db*

Analysis of variation within loci is confined to a single locus, *Rima-UA-Db*, as the other loci have not been recovered from a sufficiently large sample size to support detailed analysis. Locus *UA-Db* was sequenced in 46 individuals. Exon 2 contains 16

polymorphic nucleotide sites, with alleles differing by an average of 5.6 +/- 2.6 nucleotides. Translation to amino acid sequences results in 15 unique alleles at the amino acid level with eleven polymorphic amino acid sites. The ratio of non-synonymous to synonymous substitutions ( $d_N:d_S$ ) is 10.4 (see Table 1.2 for a comparison of  $d_N:d_S$  at MHC Class I loci between *R. marmoratus* and *O. tshawytscha*). Of these eleven polymorphic sites, ten are located at positions associated with the ARS in mammalian MHC molecules (Figure 1.3). There are thirteen amino acid positions in the fragment of the exon 2  $\alpha 1$  region that was sequenced that point into the ARS. Of these thirteen positions, the amino acids at three positions are conserved in the mammalian MHC. These three conserved amino acids likely interact with constant features on processed antigens (Takeuchi, Figueroa, O'hUigin, and Klein 1995). These features (F22, G26, and Y62) are also conserved in the sequences reported here. Of the remaining ten internal ARS sites, seven are variable. In all, only 10% of the total amino acids in the  $\alpha 1$  region were variable, but 70% (seven of ten) of the non-conserved internal ARS sites were variable. Positions 43 and 53 are not known to be involved in the formation of the ARS in mammals, but the fact that they are variable here, as well as in other non-mammalian MHC systems, might suggest that these residues interact with the antigen.

The pattern of nucleotide substitution in the  $\alpha 1$  region is consistent with that reported in other species. There are no synonymous substitutions at positions encoding the mammalian ARS. Of the thirteen nonsynonymous substitutions, seven are first position, four are second position, and three are third position. All three of the synonymous substitutions are at third positions. One allele, *UA-Db\*14*, contains a three base substitution at the position encoding amino acid 85. No bases in this codon are

shared with any other *UA-Db* alleles. Allele *UA-Db\*15* has five substitutions in the first 27 bases of exon 2 that are shared with all *UA-E* alleles. Allele *UA-E\*2* shares sequence motifs through all of exon 2 with locus *UA-Db*. However, the intron 2 and exon 3 sequences place this allele at the *UA-E* locus (Figure 1.4), suggesting that this allele is the result of an inter-locus recombination event between alleles from locus *UA-Db* and *UA-E*.

### *Exon 3 $\alpha$ 2 Region-locus UA-Db*

The portion of the exon 3  $\alpha$ 2 region that was sequenced spanned codons 91 through 164. There were 20 polymorphic nucleotides in this region, 18 of which were non-synonymous substitutions. The ratio of non-synonymous to synonymous substitutions was 5.3:1.  $\alpha$ 2 region sequences differed by an average of 10.1 +/- 6.7 nucleotides. In contrast with the  $\alpha$ 1 region data, polymorphisms in the  $\alpha$ 2 region were not as strongly associated with codons involved in peptide binding in the mammalian MHC. Excluded from the determination of synonymous and non-synonymous substitutions were sequences that appeared to have evolved through inter-locus recombination events leading to motifs from one locus being present in sequences from another (Figure 1.4). The  $\alpha$ 2 encoding region of alleles *UA-Db\*11* and *UA-Db\*12* contained a motif including the last base of codon 93, the first two nucleotides of codon 94, and the first nucleotide of codon 97 that was found in all sequences at locus *UA-Da*. Alleles *UA-Db\*14*, *Db\*15*, and *Db\*16* contained a four nucleotide motif including the last nucleotide of codon 98, and codon 99 found in all alleles at locus *UA-E*. Allele *UA-Db\*15* also contained sequence motifs from alleles at locus *UA-E* in the 5' portion of



Table 1.2. Comparison of the ratio of synonymous ( $d_S$ ) to non-synonymous ( $d_N$ ) substitutions at a single MHC Class I locus between *R. marmoratus* and *O. tshawytscha*.  $d_N$  and  $d_S$  in *O. tshawytscha* are typical for MHC class I genes.

Species	Region	Codons	$d_S$ (SE)	$d_N$ (SE)	$d_N/d_S$
<i>R. marmoratus</i>	Non-ARS	50	0.0027 (0.0002)	0.00 (0.00)	$\infty$
	ARS	20	0.00 (0.00)	0.028 (0.002)	0
	Total	70			9.78
<i>O. tshawytscha</i>	Non-ARS	49	0.021 (0.019)	0.029 (0.010)	1.38
	ARS	27	0.035 (0.034)	0.086 (0.021)	2.46
	Total	76	0.026 (0.17)	0.050 (0.010)	1.92



exon 2. Alleles *UA-Db\*17* and *UA-Db\*20* contained a different motif from locus E that included the last nucleotide of codon 64, codon 65, and codon 66. Alleles *UA-Db\*19* and *UA-Db\*20* also contained elements from locus E in this region, but the motif was not reproduced exactly. Amino acid residues conserved in other teleost species also found here include T143, K 146, and Y160. These are residues thought to interact with conserved portions of processed antigen (Takeuchi, Figueroa, O'hUigin, and Klein 1995).

### *Polymorphism in Intron 2*

Approximately 350 bases of intron 2 immediately 5' of exon 3 were compared between all loci. There were eight unique locus *UA-Db* intron sequences, with a total of seven polymorphic nucleotides and an average pairwise divergence of 2.6 +/- 1.1 nucleotides. The intron sequences compared between loci were so divergent in this region as to be essentially unalignable. Intron sequences from locus *UA-E* are very similar to those at locus *UA-Db* across most of the alignment. However, the alignment breaks down dramatically at position 230.

### *Microsatellite Present in Intron 2 in Multiple Loci*

All five putative class I loci characterized have a CA dinucleotide repeat located approximately 50 bases downstream of the 5' end of exon 2. In three of the loci, this repeat is only five to nine repeat units in length, and is not polymorphic. In locus *UA-Db*, the repeat is at least 25 repeat units long and does appear to be polymorphic. In locus *UA-Da*, the CA repeat varies between five and eight repeat units, and is immediately followed by a GATA repeat that is at least nine units in length. Variation in these repeats

```

UA-Db*01 TGATGATCAGGTATAGCTCAACAACCTATCGGATCTTTAGCCAGGAGAAATACAGCACGCCACCAG----GGGGCAATGCTGCCCTTCGGATGATTCTACTATTG
UA-Db*02 .....
UA-Db*03 .....
UA-Db*04 .....
UA-Db*05 .....
UA-Db*06 .....T.....
UA-Db*07 .....T.G.T.....G.....
UA-Db*08 .....G.....
UA-Db*09 .....C.....C.....C.....
UA-Db*10 .....C.....C.....C.....
UA-Db*11 ..G.TTA.T.....
UA-Db*12 ..G.TTA.T.....
UA-Db*13 .....TAGTA...A.....G.TT...C.....C.....C.AC.....A.A.G.....
UA-Db*14 .....TAGTA...A.....G.TT...C.....C.....C.AC.....A.A.G.....
UA-Db*15 .....TAGTA...A.....G.TT...C.....C.....C.AC.....A.A.G.....
UA-Db*16 .....TAGTA...A.....G.TT...C.....C.....C.AC.....A.A.G.....
UA-Db*17 .....T...T...T.A.....G.TT.....T.....A...C.G...A.TACAT.....C...
UA-Db*18 .....T.....GG.....C.AC.....A.A.G.....A.T.C.A...T.C.....
UA-Db*19 .....T.....GG.....C.AC.....A.A.G...T.TA.T.C.A...T.C.....
UA-Db*20 .....T...C.....C.....C.....A.TACAT.....C...C...

UA-Ca*01 G...CTT.....CGAT.TGGAGGTAATAGAT..GGACTTA.GA.GCGTGGGTT..G...ACG...ACTAAAATTG..AA.TTATACA.A.A.TC.TACAG.A..
UA-Ca*02 G...CTT.....CGAT.TGGAGGTAATAGAT..GGACTTA.GA..CGTGGGTT..G...ACG...ACTAAAATTG..AA.TTATACA.A.A.TC.TACAG.A..
UA-Ca*03 G...CTT.....CGAT.TGGAGGTAATAGAT..GGACTTA.GA..CGTGGGTT..G...ACG...ACTAAAATTG..AA.TTATACA.A.A.TC.TACAG.A..

UA-Da*01 GG...TTA.T...CG.T.TGGAGGTAATAT...GGA.TTATGA..CGTGGGTT...T.AC.A.AGCTGAAA.GT..AAGCTAGC.ATA.-A.CCTACAGGAG.
UA-Da*02 GG...TTA.T...CG.T.TGGAGGTAATAT...GGA.TTATGA..CGTGGGTT...T.AC.A.AGCTGAAA.GT..AAGCTAGC.ATA..T.CCTACAGGAG.

UA-Cb*01 .A...G..TT.....TA.....TT.G..C.....C..GG.....C...ACAAA..ATT..A...A.AGC.....T..C.....
UA-Cb*02 .....TT.....TA.....T...C.....C..GG.....C...ACAAA..ATT..A...A.AGC.....T..C.....
UA-Cb*03 .....C..TT.....TA.....T...C.....C..GG...AG.....C...ACAAA..ATT..A...A.AGC.....T..C.....

UA-E*01 ..G..T...TAGTA..T.....A.AG.GC...T.C.C...G.....TTC..C.AGTAA...TT.AT.AA.TACAT.T.C..CT..GC...T
UA-E*02 ..G..T...TAGTA..T.....A.AG.GC...T.C.C...G.....TTC..C.AGTAA...TT.AT.AA.TACAT.T.C..CT..GC...T
UA-E*03 ..G..T...TAGTA..T.....A.AG.GC...T...C..GG.....TTC..C.AGTAA...TT.AT.AA.TACAT.T.C..CT..GC...T

```

**Figure 1.4.** Nucleotide alignment of exon 2 alleles from 5 *Rivulus marmoratus* MHC loci. Invariant sites across all loci have been removed. Nucleotide motifs hypothesized to have resulted from inter-locus recombination events in inter-locus recombination events are in bold .

elements are the only sites of intra-locus variation, save for two single nucleotide substitutions, one in locus *UA-Db* (T insertion) and one in locus *UA-Da* (T $\leftrightarrow$ A).

## Discussion

### *Variation at locus Rima UA-Db and the Rate of MHC Class I Evolution in R. marmoratus*

High levels of non-synonymous nucleotide substitution relative to synonymous substitution are hallmarks of classical MHC genes (Hughes and Nei 1988). However, the ratio of non-synonymous to synonymous substitution reported here is far higher than that reported for other species (see Table 1.2 for a comparison of  $d_N:d_S$  across species). We propose that this is due to a higher rate of non-synonymous nucleotide substitution in the ARS of MHC Class I loci in *R. marmoratus*. We are not proposing a higher intrinsic rate of nucleotide substitution at these loci, as this would require invoking a novel substitution mechanism. Non-synonymous substitutions are actually more common than synonymous ones, but the apparent rate of non-synonymous substitution is lower in general because non-synonymous substitutions tend to be deleterious and are eliminated by purifying selection. We propose that the homozygosity of *R. marmoratus* increases the intensity of diversifying selection at MHC loci, so that non-synonymous substitutions accumulate more rapidly than in an outcrossed, heterozygous species. Heterozygous individuals have approximately twice the variation at MHC loci of homozygous individuals, given the levels of polymorphism and heterozygosity that typify most MHC systems. Heterozygosity might serve to dampen pathogen-driven negative selection acting on certain alleles because other alleles present in homozygotes may provide immunologic protection. Therefore, if a particular allele is ineffective against the array of pathogens challenging a population, that allele may remain in the population in heterozygous form. This might allow heterozygous populations to maintain a pool of alleles that confer little

immediate fitness in the form of immunologic protection without a significant fitness deficit. However, the fitness of these alleles might increase should the pathogenic challenge facing a population change. There is abundant evidence of allele and haplotype specific resistance to specific pathogens as well as a number of associations between MHC haplotype and disease susceptibility (Hill 1991). A homozygous species such as *R. marmoratus* would be less able to maintain MHC alleles of low immediate fitness because these alleles would always appear in homozygous form. Low-fitness MHC alleles would therefore come under more intense selection in homozygous form than heterozygous. Individuals carrying these alleles would experience decreased fitness because they would be more susceptible to infection by pathogens. This might not be the case for heterozygotes, because the susceptibility conferred by certain alleles could be alleviated by the efficacy of other MHC alleles.

The rapidity with which alleles at locus *UA-Db* seem to have evolved is further supported by the low degree of divergence between the intron sequences recovered from this locus. Intron sequences differ by an average of 0.74 +/- 0.31 percent sequence divergence. This extremely low level of divergence is also comparable to the average non-synonymous divergence between exon 2 and exon 3 sequences (0.27 +/- 0.31 and 0.43 +/- 0.44 percent sequence divergence, respectively). Taken together, the low exon non-synonymous substitution rate and low intron substitution rate are good evidence that these alleles have evolved recently.

The high number of non-synonymous substitutions between alleles could appear to be elevated if our comparisons included paralogous loci. However, paralogous comparisons are not a good explanation for the low synonymous substitution rate because

paralogous alleles are expected to accumulate both synonymous and non-synonymous differences, not just non-synonymous differences.

MHC genes do not evolve at a faster rate than other nuclear genes (Klein, 1986). The extraordinary level of polymorphism described at MHC loci in many species is believed to be due to the retention of ancestral polymorphism, often across the species barrier. Therefore, much of the variation at these genes is likely quite old and new species evolve with a large suite of MHC alleles already in existence. This hypothesis may not be the best explanation for the pattern of variation in the MHC Class I genes in *R. marmoratus*. MHC alleles evolve and diversify by a variety of potential mechanisms, the most important probably being point mutation. Point mutation is a stochastic process. If these alleles are diverging at a rate similar to that described in other MHC systems, it is unlikely that a number of non-synonymous substitutions would have accumulated between alleles, while almost no synonymous substitutions occurred.

#### *Conservative and Radical Non-synonymous Substitution in Locus Rima UA-Db*

Amino acids are a diverse group of molecules that can be classified based on charge, polarity, and size. Non-synonymous substitutions are either conservative or radical, depending on whether the new amino acid is in the same group (conservative) or a different group (radical), as the original (Zuckerandl and Pauling 1965, Epstein 1966). The neutral theory predicts that most non-synonymous substitutions will be conservative, and a departure from this can be taken as evidence for positive Darwinian selection (Hughes and Nei 1988,1989). Among the non-synonymous substitutions in locus *UA-Db*, there are a large number of radical non-synonymous substitutions (Table 1.3). Of the

11 positions with amino acid polymorphisms, 9 of these polymorphisms involve amino acids that differ in charge and/or polarity. Two of these positions also have a radical substitution involving a shift in amino acid size. These radical substitutions almost certainly lead to a change in the peptide-binding spectra of the MHC molecules and are therefore immunologically significant.

#### *Inter-locus Recombination*

Several alleles have sequence motifs that may be evidence of inter-locus recombination (gene conversion) events. These include alleles *UA-Db*\*11, -\*12, -\*13, -\*14, -\*15, -\*16, -\*17, -\*20, and *UA-E*\*2 (see figure 4). Our assumption is that the sequences involved in such exchanges are at different loci. This assumption is based on several lines of evidence, the most compelling being that alleles from all loci involved have been recovered in some cases from the same homozygous individual. Because these different sequences were recovered from homozygotes, as defined by DNA fingerprinting/progeny testing (Laughlin *et al.* 1992), these loci are paralogous loci by definition.

#### *Conserved Microsatellite in Intron 2*

In all of the five MHC loci characterized in *R. marmoratus*, there is a microsatellite repeat element 1-200 bp downstream of exon 2. The presence of this repeat element in all loci is surprising as microsatellites are generally thought to be unstable on an evolutionary time scale (Eppelen *et al.* 1991). In locus *UA-Db*, *UA- UA-*

*Da*, and *UA-Cb*, the microsatellite is a CA dinucleotide repeat. However, it appears to be polymorphic only at the *UA-Db* locus. In the *UA-Ca* and *UA-Da* loci, the repeat is stabilized by a thymine nucleotide. Interestingly, in locus *UA-Ca*, the CA repeat is immediately followed by a polymorphic GATA repeat element. The presence of this element in three separate loci might be explained if the loci were of recent origin. This is not a completely satisfactory explanation given the high degree of divergence between the three loci in both exon 2 and intron 2 sequences. In fact, the intron divergence between the introns of these three loci is greater than the divergence between the introns of locus *UA-Cb* and the intron of the locus characterized in *R. caudomarginatus*. It appears that this microsatellite has been maintained over significant evolutionary time, a contention supported by the presence of a similar repeat element in the sequences isolated from *R. caudomarginatus*. If this repeat element is in fact orthologous between the two species, then it must be at least as old as the divergence time of the two species, some 20 million years. More surprising is the presence of a CA repeat in intron 2 of the guppy, *Poecilia reticulata*. A similarly persistent microsatellite has been more thoroughly characterized in the MHC-*DRB* genes of several artiodactyl species (Epplen 1991). In this case, the microsatellite appears to have been stable for at least  $7 \times 10^7$  years, as it is present in primates, including humans. While processes involved in the evolution of class I and class II MHC genes are not entirely analogous, classes of MHC gene are highly polymorphic and evolve under balancing selection. It has been suggested that these repeats are involved in gene conversion through inter-locus recombination events that are one of the mechanism proposed for the evolution of MHC genes (Epplen 1991).



These ancient microsatellites are be maintained over evolutionary time because of the role that they play as sites involved in inter-locus recombination.

*Note on Rare Outcrossing Events in Rivulus marmoratus*

We have described *R. marmoratus* as a virtually obligate self-fertilizing hermaphrodite. *R. marmoratus* is described as an ‘obligate’ selfing hermaphrodite because the ovarian and testicular tissues are undifferentiated and fertilization is effectively synonymous with ovulation. We began including ‘virtually’ in the description with the discovery of a single population on the Belize island of Twin Cays. This population was composed of fully functioning hermaphrodites and approximately 25% males, which are easily identified by their bright orange body wash (Soto and Noakes 1994). It is thought that matings must occur between males and hermaphrodites, though the mechanism through which this occurs is unknown. While males have been collected in other locations, they are generally rare (less than 1%) or non-existent, and there was no evidence of other episodes of outcrossing until recently. The discovery of a second outcrossed population in Honduras casts some doubt on the description of *R. marmoratus* as an obligate self-fertilizer. It may be more accurate to describe the species as engaging in episodic outcrossing. However, these episodes appear to be rare and localized, and are not frequent enough to maintain significant levels of heterozygosity because self-fertilization reduces heterozygosity by  $1/2$  each generation, and leads to homozygosity in approximately 12 generations (Maynard Smith 1978). We include this discussion here to circumvent speculation that perhaps the species is not generally homozygous. Rare

outcrossing events are insufficient to prevent the species from existing in an almost exclusively homozygous state.

### *Conclusions*

Alleles at a single MHC class I locus in *R. marmoratus* exhibit an extremely high level of non-synonymous substitution in the exons encoding the  $\alpha 1$  and  $\alpha 2$  domains. These same individuals have little to no variation in the intron intervening between the two regions. We hypothesize that this variation is due to intense positive natural selection for immunologic diversity, and that this selection is more intense in this species because of its homozygous nature.

**Chapter 2: Low variation in the mitochondrial control region in Central American and Caribbean populations of the self-fertilizing, hermaphroditic killifish, *Rivulus marmoratus***

**Introduction**

The mangrove killifish, *Rivulus marmoratus*, inhabits coastal mangrove swamps from Vero Beach, Florida, USA, to as far south as Rio de Janeiro, Brazil, and is also found throughout the Caribbean and coastal Central America (Huber 1992). In fact, it has one of the most extensive ranges of any vertebrate, and its range is complementary to that of the entire *Rivulus* genus (Taphorn 1980). *R. marmoratus* is the only vertebrate known to reproduce by virtually obligate internal self-fertilization (Harrington 1961). Self-fertilization is the ultimate form of inbreeding, and results in homozygosity in relatively few generations. Once homozygosity is reached, parents and progeny are genetically identical, and this mode of reproduction becomes effectively clonal. DNA fingerprinting studies coupled with progeny testing of hermaphrodites confirm that almost all of the individuals collected from natural populations are homozygous (Turner *et al.* 1992). However, individual clonal lines differ in their DNA fingerprint patterns, and populations are typically comprised of a number of different clonal lines, so that the population structure has been described as ‘arrays of homozygous clones’.

This result has given rise to speculation that different clonal lines are niche specialists, a hypothesis termed the ‘frozen niche variation’ model, in which each clone is uniquely adapted to withstand certain environmental pressures (Vrijenhoek 1984). This model is usually applied to (hemi)clonal vertebrates of interspecific hybrid origin. Also,

some literature on *R. marmoratus* has suggested that clones vary in their response to various environmental factors, including work by R.W. Harrington, who pioneered study of this species. Turner et al. (1992) specifically advanced the ‘frozen niche’ hypothesis regarding *R. marmoratus*. A number of studies have attempted to elucidate these proposed differences, including studies of clonal variation under different environmental stresses, variation in embryological development under low-temperature incubation, and behavioral differences between clones.

However, the underlying assumption of these studies, including those performed by the authors, is that variation (presumably neutral) in DNA fingerprint patterns indexes adaptive variation among clonal lines. This suggests one of two hypothesis; either significant ancestral variation has been partitioned and retained between clonal lines, or that clonal lines diverged a long time ago, and accumulated a significant number of adaptive mutations since diverging. We reject the first hypothesis as unlikely because it would suggest the transition to self-fertilization was a large-scale event, involving a significant number of genetically distinct individuals. The second hypothesis suggests that the transition to self-fertilization was a rare event, and was successful because of the increased reproductive potential that self-fertilizing individuals enjoy. This hypothesis allows for the transition to selfing involving a relatively small number of founding individuals, which would appear to be more likely.

However, if clonal lines have been diverged for a significant period of time, then clones would be expected to accumulate neutral mutations in a clock-like fashion, particularly in the rapidly evolving mitochondrial control region. We report here on very limited mitochondrial control region variation characterized in 55 individual *R.*

*marmoratus* collected from Florida to Panama, and speculate on the time frame of the invasion of the Panamian isthmus and Caribbean by the species.

## Materials and Methods

### *Samples and DNA extraction*

Specimens were collected using dip nets, miniature hook-and-line, or traps. Individuals were either shipped live in moist cotton or preserved in the field in 100% ethanol. A complete list of collection locations is found in Table 2.1. Genomic DNA was extracted from the ethanol fixed specimens either using a standard phenol:chloroform extraction or using the Puragene DNA Extraction Kit (Gentra Systems).

### *PCR amplification of the mitochondrial control region and DNA sequencing*

Primers E and K, with homology to conserved regions of the mitochondrial control region (Lee *et al.* 1995), were used to amplify a 389 base pair fragment of the mitochondrial Control region. Reactions were prepared in 50  $\mu$ l volumes using Epicentre's PCR Premix F as per the manufacturer's protocol. Reactions were cycled on a PTC-100 thermocycler using the following program; 95°C – 30s, 50°C – 30s, and 72°C – 30s for 30 cycles. PCR products were excised from 1-% agarose gels and prepared for direct sequencing using the Qiagen Gel Extraction Kit. Sequencing reactions were performed using ABI PRISM BigDye terminator chemistry, and the sequences were run on an ABI PRISM 310 Genetic Analyzer.

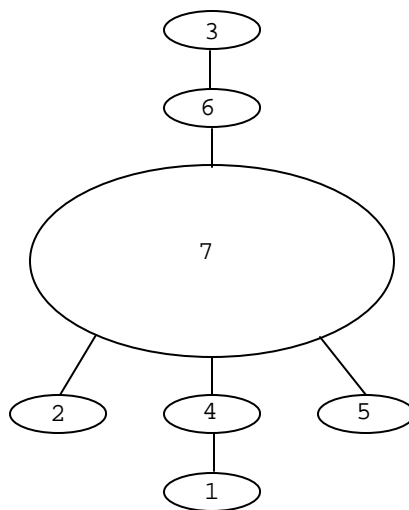
**Table 2.1.** Collection locations of *R. marmoratus* samples. The mitochondrial control region haplotypes recovered from each population are listed, with the occurrence of each haplotype in parenthesis.

Collection location	N	haplotypes (number)
Florida		
Brevard County	10	7(10)
Charlotte City	3	7(3)
Marco Island	3	7(3)
Mate Cumbe	1	7(1)
No Name Key	3	6(1)
Everglades	4	7(4)
Lost Man River	4	7(4)
Dolphin Smile	3	7(2), 5(1)
Belize		
Dangriga	10	3(1), 7(6), 5(3)
Twin Cays	9	7(6), 2(1), 3(1), 5(1)
Honduras	3	7(3)
Panama	1	1(1)
Bahamas	1	7(1)
Total		55

### *Haplotype network construction*

The haplotype network of mitochondrial control region sequences was constructed using TCS version 1.01, which develops a graphical representation of the likely evolutionary relationships between DNA sequences, with 95% confidence intervals. Results are shown in Figure 2.1.





**Figure 2.1.** Haplotype network of *Rivulus marmoratus* mitochondrial control region haplotypes constructed using TCS version 1.01, representing the likely evolutionary relationships of these sequences. Numbers of each haplotype correspond with the numbers in Table 1. The size of each oval roughly corresponds with the frequency of each haplotype. Connected haplotypes differ by single nucleotide substitutions

## Results and Discussion

PCR amplification using primers E and K yielded 371 bases of the mitochondrial control region. A total of 7 mitochondrial control region haplotypes were recovered from the 55 individuals. These haplotypes differed by no more than 4 mutational steps (Figure 2.1), with haplotypes differing by an average of 2.1 nucleotides, or 0.56%. The most common haplotype was recovered from 42 individuals, while the other haplotypes were recovered from 1 to 4 individuals. Table 2.1 summarizes the haplotypes recovered from each sampling location.

### *Lack of Mitochondrial Control Region Variation and the Rapid Colonization of Central America and the Caribbean by *R. marmoratus**

The almost complete lack of variation in the mitochondrial control region suggests 1) that Central American, Caribbean, and Floridian populations of *R. marmoratus* are very recently separated, and 2) the colonization of this region by *R. marmoratus* occurred very rapidly. The most common mitochondrial haplotype is shared (44 of 55 samples, see Table 2.1) by individuals collected from Honduras, mainland Belize, several Belize cays, and locations throughout Florida, and there is no apparent geographic pattern underlying the few substitutions in the control region. Given the geographic distances covered by our sampling, the paucity of variation in what is sometimes termed a “hyper-variable” region is startling. While we cannot invoke a molecular clock using this data set, the rate of divergence for the fish mitochondrial control region has been proposed to be ~3.6% per million years +/- 0.46% (Donaldson

and Wilson 1999). If the rate of molecular evolution in the control region of *R. marmoratus* approximates this, then the data presented here suggest a very recent divergence since the maximum divergence between samples is ~1%. It is important to note that many of the samples from Belize, Florida and Honduras locations have been DNA fingerprinted and exhibit different DNA fingerprint patterns. Therefore, the paucity of mitochondrial Control region sequence variation is not an artifact of population structure, or put more simply, we are not repeatedly sampling representatives of the same clonal line. We have also sequenced a portion the ND2 gene from a limited number of individuals and found no variation (data not shown). We conclude that the individuals in this study are very closely related.

It is unsurprising that the colonization of Central America and the Caribbean by *R. marmoratus* was so rapid because the species is an extraordinarily good colonist. *R. marmoratus* can colonize new habitat with a single individual. Also, these fish are tolerant of desiccation, poor water quality conditions (Davis, Taylor and Turner 1990), are able to respire aurally (Abel 1984), and are capable of traversing dry land for short distances (Turner pers. obs.), so much so that they actually spend a significant portion of time out of water in the laboratory. These are all characteristics likely to aid in rapid colonization and dispersal (<http://www.rivmar.bsi.vt.edu>).

Our hypothesis that *R. marmoratus* is a recent arrival in Central America and the Caribbean is consistent with the geologic history of the Panamanian Isthmus and the geographic origins of the species. The species arose in South America, probably Brazil, and migrated north (Costa 1998). Brazil is given as the likely ancestral range because that is the sole range of *R. marmoratus*' sister taxa, *Rivulus caudomarginatus*. Analysis

of the mitochondrial cytochrome *b* gene suggests that *R. marmoratus* diverged from *R. caudomarginatus* approximately 20 mya. The geologic history of the Panamanian isthmus is also consistent with a recent colonization of the Caribbean and Central America by *R. marmoratus*. The isthmus did not exist until approximately 3-3.5 mya (Coates *et al.* 1992). Open sea between North and South America may have been a significant barrier to the northern migration of the species.

The time frame of the origin of self-fertilization in *R. marmoratus* is unclear. While the divergence of *R. marmoratus* and *R. caudomarginatus* is dated to approximately 20 mya, *R. marmoratus* may have made the transition to self-fertilization much later. However, these data certainly suggest that *R. marmoratus* invaded the Panamanian Isthmus as a self-fertilizing hermaphrodite.

These data clearly suggest that the invasion of Central America and the Caribbean by *R. marmoratus* was rapid and relatively recent. Also, this invasion was likely accomplished by a group of closely related clones. It is therefore unlikely that the extant clonal lineages of Central America and the Caribbean harbor significant adaptive variation because such variation was probably not present in the founding population, and not enough time has passed since colonization for adaptive differences to occur.

## **Chapter 3: Variation in a Single Major Histocompatibility Complex Class I Locus in the Death Valley Pupfish Complex (*Cyprinodon spp.*)**

### **Introduction**

Conservation genetics focuses on the maintenance of genetic variation in threatened or endangered populations or species in an effort to preserve their viability and evolvability (Frankham, 1995; Frankel and Soule, 1981; Soule and Wilcox, 1980). Variation in species or populations under management is typically assessed through the use of neutral molecular markers, including, most commonly, mitochondrial DNA, allozymes, RAPD's, and microsatellites. The primary assumption underlying these approaches is that variation at neutral marker loci is correlated with or indexed by adaptively significant variation at other loci, so that the preservation of adaptively significant variation is a byproduct of preservation neutral variation (Hedrick and Miller, 1992). However, in practice this assumption is difficult to test.

Ideally, efforts at assessing genetic variation should be targeted directly toward loci at which this variation is adaptively significant. The advantages of this would be two-fold. First, because preservation of genetic variation is expensive, targeted efforts would likely be more cost effective. Second, these methods could be used to directly assess the success of conservation genetic efforts. The primary barrier to using this approach is that it can be difficult to determine what genetic variation is adaptively significant.

The Death Valley pupfish complex (*Cyprinodon spp.*) provide two natural experiments with which adaptively significant genetic variation may be identified. The Death Valley *Cyprinodon* complex is an assemblage of species isolated in various refugia following the drying of Lake Manly at the end of the Pleistocene (Soltz and Naiman, 1978; Miller, 1948). These forms are generally noted for their high degree of local endemism and their diversity of physiological adaptations (Soltz and Hirshfield, 1981), morphology (Miller, 1948; LaBounty and Deacon, 1972), mating behaviors (Liu, 1969), and habitats. The species and sub-species comprising the complex are *Cyprinodon salinus salinus*, *Cyprinodon salinus milleri*, *Cyprinodon nevadensis nevadensis*, *Cyprinodon nevadensis amargosae*, *Cyprinodon nevadensis shoshone*, *Cyprinodon nevadensis mionectes*, *Cyprinodon nevadensis pectoralis*, and *Cyprinodon diabolis* (Miller 1948, Turner 1974, Echelle and Dowling 1992, Duvernell and Turner 1998).

The Devil's Hole pupfish, *Cyprinodon diabolis* (Wales, 1930), is of particular interest because it is the most highly endemic of any species known, and has had a census size of no more than approximately 500 individuals since 1971, when yearly census began (Deacon 1979; Chernoff 1985). The entire species inhabits a shallow shelf of 3.5 by 22 meters on the edge of Devil's Hole in Ash Meadows, Nye County, Nevada, USA (Chernoff 1985). It is estimated that *C. diabolis* has been isolated in this habitat for at least 10,000 years, and probably longer (Soltz and Naiman, 1979; Miller, 1981). Conservatively, this means it has survived *in situ* for over 20,000 generations. The Salt Creek/McClain's Spring pupfish (Miller, 1943), *C. s. salinus*, is also highly endemic, limited to a single drainage in Ash Meadows. The species currently numbers in the thousands, but molecular genetic evidence suggests that it has undergone a significant

population bottleneck (Duvernell and Turner, 1999; LaBounty and Deacon 1972). *C. s. salinus* has probably been isolated for several thousand years, much like *C. diabolis*.

The power of random genetic drift to fix genes in small populations is well known (Wright, 1969). Given the population size and the number of generations involved, *C. diabolis* would be expected to be genetically monomorphic, even at highly polymorphic neutral loci. In fact, allozyme loci (Turner, 1974), mitochondrial control region sequences (Duvernell and Turner, 1998), highly polymorphic (in other species) LINE element loci (Duvernell and Turner, 1998), and a number of microsatellites (Wilcox, pers. comm.), have shown no variation in *C. diabolis*. This is not surprising given the number of generations that have likely passed since its isolation in Devil's Hole. Only intense natural selection would be sufficient to preserve genetic variation against loss through genetic drift. Therefore, any genetic variation at coding loci in *C. diabolis* is almost certainly of adaptive significance.

The genes of the Major Histocompatibility Complex (MHC) are sequences at which overdominant selection has acted to maintain extreme levels of polymorphism and heterozygosity in most species (reviewed in Klein, 1986). Rates of heterozygosity in some species are upwards of 90% and in well-characterized MHC systems such as the human HLA, as many as 100 alleles have been described at a single locus. MHC genes encode cell surface glycoproteins involved in the binding and presentation of foreign antigen derived peptides to the T-cells of the immune system as the initiating step of the innate immune response. The spectra of peptides that can be bound by the product of an allele are determined by the amino acid motifs present in the peptide-binding domain (ARS) of the molecule (Doherty and Zinkernagel, 1975). Individuals that are

heterozygous at MHC loci may enjoy enhanced immunologic protection compared to MHC homozygotes (Hughes and Nei, 1988; Hughes et al. 1994; Hughes, 1999).

We report here on variation at a single MHC class I locus in *C. diabolis*, *C. s. salinus*, as well as other members of the Death Valley Cyprinodon complex. The goal of this project is to assess adaptively significant genetic variation in *C. diabolis* and *C. s. salinus*, and compare these results with assessments of genetic variation at neutral loci. These results will potentially be useful for assessing the efficacy of captive breeding populations maintained in artificial refugia in preserving adaptively significant genetic variation in these species.



## Materials and Methods

### *Samples*

All samples used in this study were collected in 1994 under permits from the California Department of Fish and Game by B. J. Turner. A complete list of samples is found in Table 3.1. Genomic DNA was extracted from the ethanol fixed specimens either using a standard phenol:chloroform extraction or using the Puragene DNA Extraction Kit (Gentra Systems).

### *Initial Primers, PCR amplification, and cloning, and automated sequencing*

Degenerate PCR primers were designed based on conserved regions of exons 2 and 3 of MHC class I genes from a variety of teleost fishes. The primer sequences were: mhc 1 G5 (sense) 5'-GCW GTG ACT CAC WCT CTG AAG-3' and mhc 1 3R (antisense) 5'-CCT CCC ATA GTT CAC ATA CTT CTT-3'. PCR reactions were prepared with Epicentre's MasterAmp PCR premix F and MasterAmp *Taq* DNA polymerase according to the manufacturer's protocol (Epicentre Technologies, Madison, WI). Reactions were cycled on an MJ Research PTC-100 as follows: 94°C (3:00 m), 6 cycles of 94°C (30 s), 46°C (30 s), and 72°C (75 s), followed by 24 cycles of 94°C (30 s), 48°C (30 s), and 72°C (75 s). These reactions produced a complex banding pattern, which was difficult to fully resolve using agarose gel electrophoresis. Therefore, the PCR products of several individuals were cloned into a pUC18 vector using the SureClone Ligation Kit (Amersham Pharmacia Biotech, Piscataway, NJ). These plasmids were used to transform competent *E. coli* (strain

**Table 2.1.** Collection locations of Death Valley pupfish samples with observed heterozygosities at locus *Mhc Cyne UA*.

Population Collection location	N	H <sub>obs</sub>
<i>C. n. amargosae</i> Amargosa River (Shoshone)	6	0.83
<i>C. n. amargosae</i> Amargosa River (Tecopa)	4	1.0
<i>C. n. amargosae</i> Amargosa River (China Ranch)	4	0.75
<i>C. n. amargosae</i> Cairn Spring	6	0.33
<i>C. n. amargosae</i> Amargosa River (Valley Springs)	7	0.29
<i>C. n. shoshone</i> Shoshone Headsprings Pool	4	1.0
<i>C. n. nevadensis</i> Saratoga Springs	12	0.75
<i>C. diabolis</i> Devil's Hole	10	0.5
<i>C. s. salinus</i> Salt Creek / McClaine's Spring	10	0.4

DH10B). Positive colonies were allowed to grow overnight in selective media.

Plasmids were isolated using the QiaPrep Spin Miniprep Kit (Qiagen, Valencia, CA).

The inserts in these plasmids were sequenced using T3 and T7 primers and PE

Biosystems BigDye chemistry. Reactions purified using the DyeEx spin column system

(Qiagen, Valencia, CA) were run on an ABI 310 Genetic Analyzer (PE Biosystems,

Foster City, CA).

#### *Locus specific intron primer design*

Alignment of the above sequences suggested that at least three MHC I loci had been amplified in the initial reaction. The intron sequences from these different loci were so divergent as to be unalignable. However, it was possible to design a locus-specific antisense primer 3' of exon 2 in intron 2. Amplification using this primer (mhc 1 int2R1 5'-TCC TTT AAT TTC TCC CCA CTG AT-3') in conjunction with primer mhc 1 G5, using a 1 µl aliquot of the original PCR reaction as template, and using the same PCR conditions as the initial reaction, produced a single band. PCR products were excised from 1% agarose gels and prepared for sequencing using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA)). Sequencing was done using BigDye chemistry and the reactions were run on an ABI 310 Genetic Analyzer.

#### *Sequence Analysis*

Sequences were aligned using CLUSTAL X (Higgins *et al.* 1996). Identification of the intron/exon boundary was made by comparison with MHC sequences from other

teleost fishes, acquired using the Genbank BLAST program (Altschul 1990). These sequences were also used to identify conserved amino acid residues .

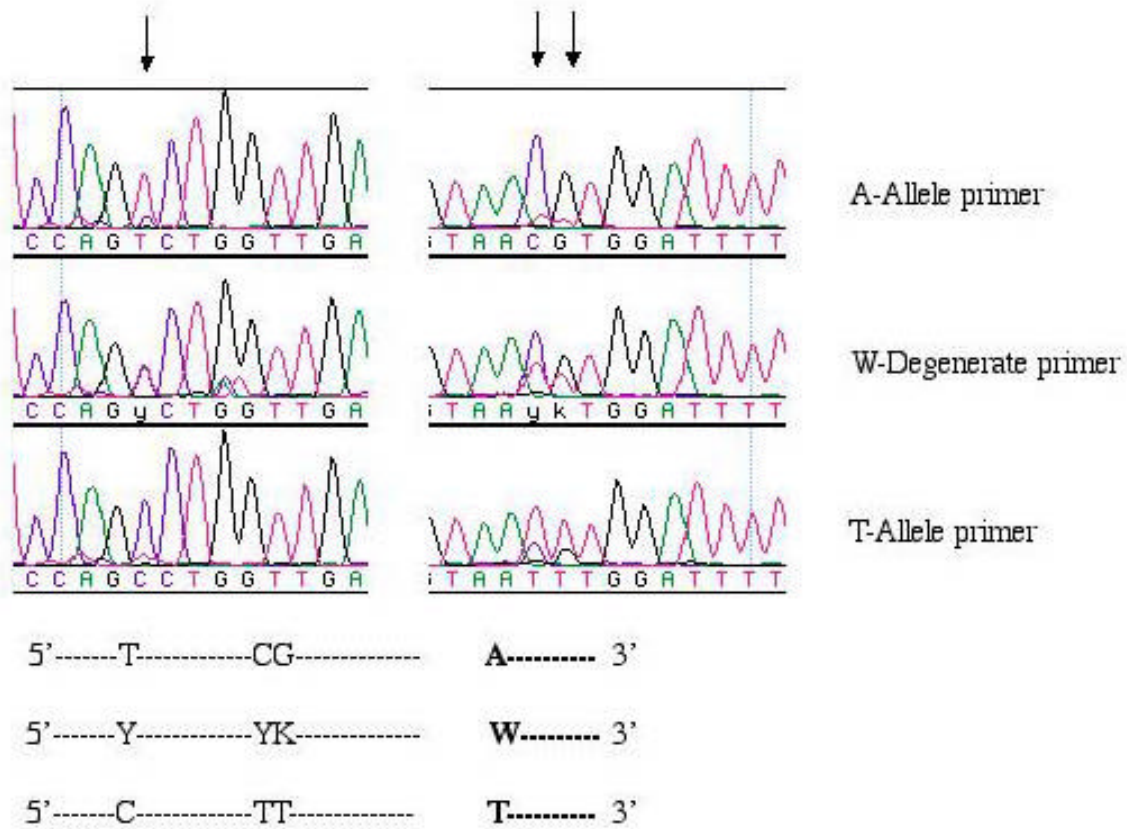
## Results

### *Amplification of class I sequences*

PCR amplification of genomic DNA using the locus specific intron primer (mhc 1 int2R1) and the degenerate exon 2 primer (mhc 1 G5) produced a fragment of 450 bp. This fragment included the nucleotides encoding amino acids at positions 5 through 85 in the  $\alpha$ 1 region of the Mhc class I molecule, based on alignment with sequences from other teleosts.

### *Identification of sequences from heterozygous individuals*

Assessment of these sequences revealed that many of the individuals were likely heterozygous at the MHC I locus under study. Polymorphic sites appear in the electropherogram produced by automated fluorescent dye-terminator sequencing as two peaks at the same position, with many individuals being polymorphic at multiple sites. To confirm that we could reliably identify polymorphisms present in the sequences from heterozygotes, primers were designed with the 3' end at a polymorphic site. Three sets of primers were designed; one ending with one component of the polymorphic site, another ending with the other component, and the third being degenerate with both components of the polymorphic site. These were used as sequencing primers for single locus PCR products from putative heterozygotes. Sequencing from a polymorphic site in the 3' end of the PCR product allowed us to resolve the phase of other polymorphic positions further downstream (Figure 3.1). Sequences were analyzed in the same manner as above. The results of this experiment suggest that polymorphisms present in an



**Figure 3.1.** Sample results of sequencing approach using reverse primers position so that the 5' most nucleotide corresponds to a polymorphic site. Arrows indicate polymorphic sites. The alignment indicates the polymorphic nucleotides present in the heterogeneous sequences derived from heterozygous individuals resolved into their allelic components.

electropherogram can reliably diagnose heterozygotes and that these polymorphisms can be resolved into alleles by direct sequencing using allele specific primers.

### *Sequence analysis*

A total of 14 unique Mhc class I alleles were found in the *Cyprinodon* Death Valley complex (Figure 3.2). These alleles, 193 bp in length, varied at 11 nucleotide positions (5.7%) and had a mean nucleotide similarity of 94.5%, with a range of 96.3-99.4% similarity. The ratio of non-synonymous to synonymous substitutions was 1.75, typical of regions of the Mhc molecule involved in peptide binding and indicative of past or ongoing selection for immunologic diversity. Consistent with this, the four synonymous substitutions were at third positions and occurred in regions that do not encode amino acids involved in peptide binding, while non-synonymous substitutions all occurred at either first or second positions. Surprisingly, the non-synonymous substitutions did not occur at positions that are hypothesized to be involved in peptide binding based on homology to positions in the mammalian MHC.

The deduced amino acid sequences of these alleles contained eleven unique amino acid residues. Motifs conserved across many Mhc systems also present in these sequences included F22 and G26, residues that may be involved in binding constant portions of processed antigen (Bjorkman *et al.* 1987). There were a total of 7 (8%) variable amino acid positions. The pattern of amino acid substitution was consistent with that in other teleost Mhc systems, with variation concentrated in regions of the  $\alpha 1$  portion of the molecule likely involved in peptide binding (Figure 3.3).

### *Heterozygosity and polymorphism*

Individuals heterozygous at the Mhc class I locus under study were identified by the presence of polymorphism in the automated sequencing electropherogram. No heterozygous sites were found that were not also found in homozygous form, with the exception of a single sequence from *C. diabolis*. Given this, coupled with the results of sequencing using allele specific sequencing primers and a limited number of sequences cloned from heterozygotes, we are confident that heterozygotes can reliably be diagnosed by multiple overlapping peaks present in sequence electropherograms. However, it is difficult to accurately deduce the allelic components present in heterozygous sequences. This is because the linkage between two polymorphisms did not appear to be consistent across all sequences, likely due to inter-allelic recombination. To allow us to analyze the data more fully, we have assumed that alleles found in homozygous form in a population are likely to be components of heterozygous genotypes. Using this assumption, we deduced the allelic components by essentially subtracting common alleles recovered in homozygous form from heterozygous genotypes. In most cases, the allelic components hypothesized to comprise heterozygous genotypes were not unique alleles. It is important to note that this method is a conservative approach to estimating allelic variation at this locus, so that it is unlikely that we have over-estimated allelic diversity in this system. Also, it is expected that rare alleles are most likely to be recovered from heterozygotes, as was the case for all of the unique alleles that are described.





	10	20	30	40	50	60	70	80	90	
<i>Cyne UA 6-2b</i>	V	VTHSLKYFYT	ASSGVPDFPS	YVSVGLVDDV	QISYCDSTRD	RNVPKQEWMK	EVSLENPTYW	EEETATCLGK	QQVFKANMEI	AKQRFN
<i>Cyne UA 6-2a</i>	.	.....	.....F	.....	.....	.....	.....R..	.....	.....	.....
<i>Cyne UA 6-4</i>	.	.....	.....L	.....	.....	.....	.....R..	.....	.....	.....
<i>Cyne UA 4-28</i>	.	.....	.....	.....	.....	.....	.....R..	.....	.....	.....
<i>Cyne UA 6-1</i>	.	.....	.....	.....	.....	K.....	.....N..	.....	.....	.....
<i>Cyne UA 4-1</i>	.	.....	.....	.....	.....	.....M.....	.....N..	.....	.....	.....
<i>Cyne UA 4-2</i>	.	.....	.....	.....	.....	.....M.....	.....N..	.....	.....	.....E.....
<i>Cyne UA 8-8</i>	.	.....	.....	.....	.....	.....	.....N..	.....	.....	.....
<i>Cyne UA 8-18</i>	.	.....	.....F	.....	.....	.....	.....	.....	.....	.....
<i>Cyne UA 5-9</i>	.	.....	.....	.....	.....I.....	.....	.....	.....	.....	.....*..
<i>Cyne UA 3-12</i>	.	.....	.....	.....	.....	.....M.....	.....R..	.....	.....	.....

**Figure 3.3.** Alignment of putative *Mhc cyne UA* amino acid sequences translated from genomic sequences. Numbers above the alignment indicate codon numbers determined by alignment with mammalian MHC genes. (.) indicates unity with the first sequence and (\*) indicates a stop codon.

Heterozygosity at the population level ranged from 1.0 to 0.29, with an average heterozygosity across all populations of 0.65. Individual population heterozygosities and sample sizes can be found in Table 3.1. Six of the ten *C. diabolis* individuals were heterozygous at the MHC locus under study. A total of five alleles were recovered. Allele *Cyne UA 1-7* was the most common allele overall, and had a frequency of 0.4 in *C. diabolis*. Allele *Cyne UA 3-10* was also common, and had a frequency of 0.35 in *C. diabolis*. The other three alleles were recovered from single individuals, one in homozygous form. Only allele *Cydi UA 13-8* was unique to *C. diabolis*. The other four alleles present in the *C. diabolis* sample were recovered from other members of the Death Valley complex.

#### *Distribution of Alleles Among Populations and Drainages*

None of the populations had complete overlap in their allelic composition. Allele *Cyne UA 1-7* was the most common, and was absent only in population 3, which included only three individuals. *Cyne UA 3-10* was found in all populations except 1, 6 and 7. All other alleles were more restricted, being found in four or fewer populations. Populations 5, 8, 15, 4, and 13 all had unique alleles that were recovered in heterozygous form. There is no apparent pattern of allelic distribution among drainages.

#### *Estimation of the time required to fix genetic variation at Mhc Cyne UA in C. diabolis*

The time required for a single allele to become fixed ( $T_{(q)}$ ) in *C. diabolis* can be calculated using the following equation;

$$T_{(q)} = (4N(1-q) \ln(1-q))/q$$

where the initial allele frequency is  $q$  and the population size is  $N$ . We have assumed that the population size has remained static at 500 individuals since the isolation of *C. diabolis* in Devil's Hole. Also, we have assumed that the initial population contained only *Mhc Cyne UA* alleles found in homozygous form and that these alleles were in equal frequency in the founding population. Therefore,  $q = 0.09$ , assuming 11 alleles in the initial population. Given these assumptions, the time to fixation for a single allele in *C. diabolis* is estimated to be ~12,000 years. Doubling the number of alleles to 22 and halving the frequency of each allele in the founding population results in an estimated time to fixation of ~25,000 years. While our assumptions probably underestimate the number of alleles present in the founding population, it is also likely that the initial allele frequencies were not equal, and some alleles may have had significantly higher frequencies than our estimate. Therefore, this is a conservative estimate of the time to fixation for *C. diabolis*. As *C. diabolis* has likely been isolated in Devil's Hole for at least 10,000 years, and probably closer to 15,000 years, the population has maintained a high level of MHC polymorphism despite random genetic drift.

*Estimation of the Selection Co-efficient Required to Maintain Putative Level of Heterozygosity Reported in C. diabolis*

The loss of heterozygosity each generation due to the effects of random genetic drift can be calculated by the following equation.

$$H_t = (1 - (1/2N)) \times H_{t-1}$$

where  $H_{t-1}$  is the initial population heterozygosity. Maintenance of rates of heterozygosity higher than expected over time is accounted for by the persistence,  $d$ , of heterozygosity, which ranges from 0 to 1. The equation then becomes

$$H_t = ((1 - (1/2N) \times d) \times H_{t-1})$$

Calculating this iteratively over 2000 generations indicates that a persistence of at least  $d = 0.5$  is required to maintain heterozygosity at the levels we have reported in *C. diabolis*. To explain the retention of heterozygosity over the 20,000 generations that *C. diabolis* has been isolated requires an even greater persistence. The persistence of these alleles is hypothesized to be mediated by overdominant natural selection.

## Discussion

### *Maintenance of MHC Heterozygosity and Polymorphism in Cyprinodon diabolis and Cyprinodon salinus salinus*

The data reported here provide the first evidence of any genetic variation in *C. diabolis*. As discussed previously, this variation is certainly of adaptive significance because of the general nature of MHC variation and its role in innate immunity (Klein 1986), and because this variation has been maintained against strong random genetic drift. Overdominant selection for immunologic diversity is the agent proposed to maintain variation at MHC loci in most well characterized systems (Hughes and Nei 1988, 1989; Parham and Ohta 1990), and is probably acting here as well. Frequency dependent selection regimes, in which the fitness of specific alleles fluctuate as the pathogenic threat facing a population changes, have also been proposed as a mechanism for maintaining MHC polymorphism. However, a frequency dependent selection scheme in a population such as *C. diabolis* is not adequate to explain the preservation of polymorphism reported over 20,000 generations. Such a selection regime would almost certainly result in the rapid fixation of favored alleles after a relatively few rounds of selection. Given that *C. diabolis* has no means beyond mutation to counter such selection, this hypothesis can be rejected on the basis that the variation recovered in *C. diabolis* is not novel, but is shared with other species in the Death Valley pupfish complex, and is therefore probably ancestral. The quantitative analysis presented here support the hypothesis that natural selection is likely responsible for the maintenance of MHC variation in *C. diabolis* and *C. s. salinus*.

*Source of MHC Variation in Cyprinodon diabolis*

All of the alleles recovered from *C. diabolis* were also found in other members of the Death Valley complex. This is consistent with the trans-species polymorphism hypothesis, which proposes that new species arise with an extant suite of MHC alleles inherited from their ancestral species (Klein 1986). Under this hypothesis, recently diverged species are expected to share MHC alleles, which is the pattern of variation described here. It is unlikely that convergent evolution could account for the high degree of similarity among all the alleles recovered from *C. diabolis*, as MHC genes are not thought to evolve rapidly (Klien 1986). Also, stochastic sorting of ancestral polymorphism has been hypothesized to explain patterns of variation characterized at other loci in the Death Valley complex, including mitochondrial DNA haplotypes and retro-element variation (Duvernell and Turner 1998). It is therefore likely that such stochastic events have shaped the distribution of MHC alleles as well.

*Comparison of levels of polymorphism at neutral marker loci and locus Mhc cyne UA among species of the Death Valley Complex*

A comparison between the levels of variation in LINE element loci with allelic variation at MHC I loci illustrates a problem in using neutral genetic markers to assess overall genetic variance in small populations or populations that have had intense historical bottlenecks. A rank order comparison of LINE element variation indicates that

*C. n. nevadensis* from the Amargosa River at Valley Spring and from nearby Saratoga Springs retain high levels of variation (Duvernell and Turner 1998). It is likely that neither of these populations has experienced a severe population bottleneck as both these water sources have historically been copious (Soltz and Naiman 1978). The same is true of populations of *C. n. nevadensis* from Shoshone Spring, Tecopa Bore, and the confluence of Shoshone Spring and the Amargosa River (Soltz and Naiman 1978). These three populations had somewhat lower LINE element diversity than the previous two, but still retain significant ancestral variation (Duvernell and Turner 1998). However, populations of *C. s. salinus* from McLain Spring (2 collection locations) are fixed for a single LINE element banding pattern. *C. diabolis* is also fixed for a single banding pattern (Duvernell and Turner 1998). These three populations have experienced intense historical bottlenecks. Though populations of *C. s. salinus* currently number in the thousands, Salt Creek intermittently dries and the population is reduced to a few hundred individuals. *C. diabolis* has likely not had a census size greater than 500 for at least 10,000 years. Despite this, both species retain significant allelic variation at one MHC locus. *C. s. salinus* has at least 12 alleles and *C. diabolis* retains at least five alleles. This disparity between neutral marker diversity and MHC I allelic variation in *C. diabolis* and *C. salinus* demonstrates the difficulty in using neutral marker variation to assess overall genetic variation, is likely due to the different effects that population bottlenecks have on neutral genetic variation and variation at loci under selection.



### *Management Implications*

Standard conservation genetic management tools (assessing population genetic variation using neutral molecular markers) have failed to detect variation in *C. diabolis* and *C. s. salinus*. Both species have been shown to be monomorphic for several microsatellites, mitochondrial DNA sequences, allozymes, and LINE element polymorphisms. The sample sizes used in these studies (N=10 for *C. diabolis*, and slightly higher for *C. s. salinus*), were not sufficiently large to have a high likelihood of detecting rare alleles, and so variation at these marker loci may have been underestimated (Sjoren and Wyoni, 1994). However, this does not mitigate the fact that levels of variation at neutral marker loci are far lower than the levels of variation described here at one MHC locus. At least in these species, variation at neutral markers is not indexed to adaptively significant variation. This presents an opportunity to undertake a more comprehensive approach toward the preservation of genetic variation in these endangered species. As discussed above, the variation maintained in these species is almost certainly of adaptive significance. Conservation genetic efforts are designed to preserve just this sort of variation, and therefore MHC variation should be monitored in a truly comprehensive management program of these species.

Hughes (1991) proposed a more radical approach, specifically, that conservation genetic efforts should be targeted solely toward the preservation of MHC variation. This suggestion was criticized for the following reasons (Vrijenhoek and Leberg 1991). Breeding programs designed to preserve variation at a single locus increase the overall inbreeding of a population, particularly when breeding programs designed to increase the frequency of rare alleles are instituted. Such programs result in an increased rate of loss

of genetic variation at other loci as compared to a traditional management approach designed to maintain overall genetic variation. These arguments are not compelling in regard to *C. diabolis* and *C. s. salinus* as these species do not have any identified genetic variation except at the MHC I locus characterized here. However, Hughes' approach may warrant a second look for its application to other management programs. Many of the assumptions regarding genetic variation in managed populations are based on data from studies of large, healthy populations. It is possible that assumptions drawn from such studies are not applicable to small populations or populations that have experienced intense historical populations bottlenecks. Specifically, the data presented here call into question that validity of assessing genetic variation in such populations using neutral marker loci. We suggest that efforts at assessing variation in these types of populations should be targeted toward loci at which variation is known to be of adaptive significance, such as MHC loci.

## Conclusions

### *Rivulus marmoratus*

The results presented in chapters 1 and 2 suggest that MHC class I genes in *R. marmoratus* have evolved rapidly under intense positive natural selection. The intensity of natural selection acting on these genes in *R. marmoratus*, evidenced by high numbers of non-synonymous nucleotide substitutions and a preponderance of radical amino acid substitutions, is almost certainly a result of the homozygosity of the species, and may be one of the costs that the species bears as a result of its self-fertilizing mode of reproduction. The inference that the variation characterized at MHC class I loci has evolved recently is based on evidence from the mtDNA control region and MHC intron sequences. However, this evidence provides insight into the divergence between extant clonal lines, but does not provide evidence regarding the length of time that the species has been self-fertilizing. It is critical to determine how long *R. marmoratus* has been self-fertilizing and the population size involved in the shift to a self-fertilizing mode of reproduction. Such determinations will allow the following questions to be addressed;

1. How much genetic variation in *R. marmoratus* is ancestral (i.e. evolved prior to the shift to self-fertilization), and how much has evolved since self-fertilization began?
2. How much genetic variation do different clones share?

These two questions are central to many ongoing studies of *R. marmoratus*. Much current work is focusing on elucidation of adaptive differences between clones and

populations. These studies would be strengthened by genetic evidence that such adaptive differences might exist. Also, only by determining if the genetic variation present in the species is ancestral or has evolved since the shift to self-fertilization will it be possible to hypothesize as to whether different clones have adapted to fill different micro-niches. The results described in chapters 1 and 2 essentially provide two answers to these questions. The MHC class I data demonstrate that there is significant genetic variation present in the species, however, the mtDNA control region data suggest that the species is young and recently diverged, limiting the likelihood of significant genetic variation either among clones or populations. Future studies should attempt to resolve this issue.

#### *Cyprinodon diabolis and Cyprinodon salinus*

The results presented in chapter 3 demonstrate that two species of Death Valley pupfish, *C. diabolis* and *C. salinus*, both maintain genetic variation and heterozygosity in at least one MHC class I locus. This is despite both species having undergone intense population genetic bottlenecks that have rendered them monomorphic at highly variable neutral marker loci. Both species have maintained MHC variation for many thousands of generations, and calculations suggest that natural selection has prevented the loss of this variation. *C. diabolis* and *C. salinus* are both endangered species and artificial propagation programs are in place. These results suggest that monitoring the genetic composition of refugia populations is an important aspect of the conservation biology of these species. Also, this preliminary study must be extended to other MHC loci, both class I and class II, so that conservation efforts will include the preservation of all MHC variation in these species. These results also have broad implications for the practice of

conservation genetics. Current management principles involve the use of highly variable neutral markers to characterize population genetic variation, and then focus on the preservation of this neutral variation, presuming that preservation of adaptive variation will be a byproduct. This study demonstrates that neutral variation may not always index adaptive variation. Therefore, assessment of MHC variation and conservation genetic programs designed to preserve MHC variation may be appropriate for some endangered species with highly reduced populations.

## Literature Cited

- Abel, D.C. 1984. Terrestriality in the cyprinodontid fish *Rivulus marmoratus*: potential utility in H<sub>2</sub>S and other bioassays. *Bulletin of Environmental Contamination and Toxicology* 33:33-39.
- Alsopp, C. E. M., A. V. S. Hill, D. Kwiatkowski, A. Hughes, M. Bunce, C. J. Taylor, L. Pazmany, D. Brewster, A. J. McMichael, and B. M. Greenwood. 1991. Sequence analysis of HLA-Bw35, a common West African allele, suggests and origin by gene conversion of HLA-B35. *Human Immunology* 30:104-9.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410.
- Arden, B. and J. Klein. 1982. Biochemical comparison of major histocompatibility complex molecules from different subspecies of *Mus musculus*: evidence for trans-species evolution of alleles. *Proceedings of the National Academy of Sciences, USA* 79:2342-2346.
- Avise, J. C. 1994. *Molecular Markers, Natural History, and Evolution*. Chapman and Hall, New York.
- Belich, M. P., J. A. Madrigal, W. H. Hildebrand, J. Zemmour, R. C. Williams, R. Luz, M. L. Petzl-Erler, and P. Parham. 1992. Unusual *HLA-B* alleles in two tribes of Brazilian Indians. *Nature (London)* 357:326-9.
- Betz, U. A. K., W. E. Mayer, and J. Klein. 1994. Major histocompatibility complex class I genes of the coelacanth *Latimeria chalumnae*. *Proceeding of the National Academy of Sciences, USA* 91:11065-69.
- Bjorkman, P.J., M.A. Saper, B. Samaroui, W.S. Bennet, J.L. Stominger, and D.C. Wiley. 1987. Structure of the human histocompatibility antigen, HLA-A2. *Nature* 329:506-512.
- Chernoff, B. 1985. Population dynamics of the Devil's Hole pupfish. *Environmental Biology of Fishes* 13(2):139-47.
- Costa, W.J.E.M. 1998. Phylogeny and classification of Rivulidae revisited: origin and evolution of annualism and miniaturization in rivulid fishes (Cyprinodontiformes: Aplochelidae). *Journal of Computational Biology* 3:33-92.
- Davis, W.P., D.S. Taylor, and B.J. Turner. 1990. Field observations on the ecology and habitats of the mangrove rivulus (*Rivulus marmoratus*) in Belize and Florida (Teleostei: Cyprinodontiformes: Rivulidae). *Ichthyological Explorations of Freshwaters* 1:123-134.

- Deacon, J. 1979. Endangered and threatened fishes of the West. *Great Basin Naturalist* 3:41-64.
- Doherty, P.C., and R.M. Zinkernagel. 1975. Enhanced immunologic surveillance in mice heterozygous at the H-2 gene complex. *Nature* 256:50-52.
- Donaldson, K.A., and R.R. Wilson, Jr. 1999. Amphi-panamic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Molecular Phylogenetics and Evolution* 13:208-213.
- Duvernell, D.D. and B.J. Turner. 1998. Evolutionary genetics of Death Valley pupfish populations: mitochondrial DNA sequence variation and population structure. *Molecular Ecology* 7:279-288.
- Echelle, A. A., and T. E. Dowling. 1992. Mitochondrial DNA variation and evolution of the Death Valley pupfishes (Cyprinodon: Cyprinodontidae). *Evolution* 46:193-206.
- Echelle, A. A., and A. F. Echelle. 1993. Allozyme perspective on mitochondrial DNA variation and evolution of the Death Valley pupfishes (Cyprinodontidae: Cyprinodon). *Copeia* 1993(2):275-87.
- Epstein, C. J. 1966. Role of the amino acid "code" and of selection for conformation in the evolution of proteins. *Nature* 215:355-59.
- Frankel, O.H. and M.E. Soulé. 1981. *Conservation and Evolution*. Cambridge: Cambridge University Press.
- Frankham, R. 1995. Conservation Genetics. *Annual Review of Genetics* 29:305-27
- Glamman, J. 1995. Complete coding sequence of rainbow trout Mhc II beta chain. *Scandinavian Journal of Immunology* 41:365-72.
- Grimholt, U., I. Hordvik, and V. M. Fosse. 1993. Molecular cloning of major histocompatibility complex class I cDNAs in Atlantic salmon (*Salmo salar*). *Immunogenetics* 37:469-73.
- Gyllensten, U. B., and H. A. Erlich. 1988. Ancient roots for polymorphism at the HLA-DQA locus in primates. *Proceedings of the National Academy of Sciences, USA* 86:9986-90.
- Harrington, R.W., Jr. 1961. Oviparous hermaphroditic fish with internal self-fertilization. *Science* 134:1749-1750.
- Hedrick, P.W., and G. Thomsom. 1983. Evidence for balancing selection at HLA. *Genetics* 104:449-456.

Hedrick P.W. and P.S. Miller. 1992. Conservation genetics: techniques and fundamentals. *Ecological Applications* 2:30-46.

Hedrick, P. W., and K. M. Parker. 1998. MHC variation in the endangered Gila topminnow. *Evolution* 52(1):194-99.

Higgins, D.G., A.J. Bleasby, and R. Fuchs. 1996. CLUSAL V: Improved software for multiple sequence alignment. *Computer Applied Bioscience* 8:189-191.

Hogstrand, K., and J. Bohme. 1994. A determination of the frequency of gene conversion in unmanipulated mouse sperm. *Proceedings of the National Academy of Sciences, USA* 91:9921-25.

Hughes, A.L. 1991. MHC polymorphism and the design of captive breeding programs. *Conservation Biology* 5:249-251.

Hughes, A.L., and M. Nei. 1988. Pattern of nucleotide substitution at MHC class I loci reveals overdominant selection. *Nature* 335:167-170.

Hughes, A. L., and M. Nei. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proceeding of the National Academy of Sciences, USA* 86:958-62.

Kallman, K.D., and R.W. Harrington. 1964. Evidence for the existence of homozygous clones in the self-fertilizing hermaphroditic teleost *Rivulus marmoratus* (Poey). *Biological Bulletin* 126:101-114.

Kim, T. J., K. M. Parker, and P. W. Hedrick. 1999. Major histocompatibility complex differentiation in Sacramento River Chinook salmon. *Genetics* 151:1115-22.

Klein, J. 1986. *The Natural History of the Major Histocompatibility Complex*. John Wiley and Sons, New York.

Klein, J., R.E. Bontrop, R.L. Dawkins, H.A. Erlich, U.A. Gyllensten, 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, J., H. Ono, and C. O'hUgin. 1993. Extensive MHC variability in cichlid fishes of Lake Malawi. *Nature* 364:330-34.

Kocher, T. D., and K. L. Carleton. 1997. Base substitution in fish mitochondrial DNA: Pattern and rates. In: *Molecular Systematics of Fishes*, T. D. Kocher and C. A. Stepien eds. pp 13-24. Academic Press, San Diego.



- Kuhner, M., S. Watts, W. Klitz, G. Thomson, and R. S. Goodenow. 1990. Gene conversion in the evolution of both the *H-2* and the *Qa* class I genes in the murine major histocompatibility complex. *Genetics* 126:1115-26.
- Labounty, J.F. and J.E. Deacon. 1972. *Cyprinodon milleri*, a new species of pupfish (family Cyprinodontidae) from Death Valley, California. *Copeia* 1972:769-780.
- Lawlor, D.A., J. Zemmour, P.D. Ennis, and P. Parham. 1988. HLA-A, -B polymorphisms predate the divergence of humans and chimpanzees. *Nature* 335:268-271.
- Lee, W. J., J. Conroy, W. H. Howell, and T. D. Kocher. 1995. Structure and evolution of teleost mitochondrial control regions. *Journal of Molecular Evolution* 41:54-66.
- Liu, R.K. 1969. The comparative behavior of allopatric species (Teleostei: Cyprinodontidae: *Cyprinodon*). PhD thesis. University of California, Los Angeles.
- Lubinski, B.A., W.P. Davis, D.S. Taylor, and B.J. Turner. 1995. Outcrossing in a natural population of self-fertilizing hermaphroditic fish. *Journal of Heredity* 86:469-473.
- Maynard Smith, J. 1978. *The Evolution of Sex*. Cambridge University Press, Cambridge.
- Miller, K. M., and R. E. Withler. 1996. Sequence analysis of a *Mhc* polymorphic class II gene in Pacific salmon. *Immunogenetics* 43:337-51.
- Miller, K. M., R. E. Withler, and T. D. Beacham. 1997. Molecular evolution at *Mhc* genes in two populations of chinook salmon *Oncorhynchus tshawytscha*. *Molecular Ecology* 6:937-54.
- Miller, R.R. 1943. *Cyprinodon salinus*, a new species of fish from Death valley, California. *Copeia* 1943:70-78.
- Miller, R.R. 1948. The cyprinodont fishes of the Death Valley system of eastern California and southwestern Nevada. *Miscellaneous Publications, Museum of Zoology, University of Michigan* 68:1-155.
- Ohta, T. 1991. Role of diversifying selection and gene conversion in the evolution of major histocompatibility complex loci. *Proceedings of the National Academy of Sciences, USA* 88:6716-20.
- Ono, H., C. O'hUigin, H. Tichy, and J. Klein. 1993. Major histocompatibility complex variation in two species of cichlid fishes from Lake Malawi. *Molecular Biology and Evolution* 10:1060-72.
- Parham, P., and T. Ohta. 1990. Population biology of antigen presentation by MHC class I molecules. *Science* 272:67-74.

Sato, A., F. Figueroa, C. O'hUgin, D. N. Reznick, and J. Klein. 1995. Identification of the major histocompatibility complex genes in the guppy, *Poecilia reticulata*. *Immunogenetics* 43:38-49.

Soltz, D.L. and M.L. Hirshfield. 1981. Genetic differentiation of pupfishes (genus *Cyprinodon*) in the American southwest. Pp. 291-333 in R.J. Naiman and D.L. Soltz, eds. *Fishes in North American Deserts*. Wiley, New York.

Soltz, D.L. and R.J. Naiman. 1978. The natural history of the native fishes of the Death Valley system. *Natural History Museum of Los Angeles County, Science Series* 30:1-76.

Soulé, M.E. and B.A. Wilcox, eds. 1980. *Conservation Biology: An Evolutionary-Ecological Perspective*. Sunderland, MA: Sinauer.

Swofford, D.L. 1999. PAUP\*. *Phylogenetic Analysis Using Parsimony (and Other Methods)*. Version 4.0b4a. Sinauer Associates, Sunderland, Massachusetts.

Takahata, N., and M. Nei. 1990. Allelic genealogy under overdominant and frequency dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* 124:967-78.

Takeuchi, H., F. Figueroa, C. O'hUgin, and J. Klein. 1995. Cloning and characterization of class I Mhc genes of the zebrafish (*Brachydanio rerio*). *Genomics* 23:1-14.

Taphorn, D.C. 1980. The first record of *Rivulus marmoratus* Poey from the South American continent. *Zoologische mededelingen* 55:127-129.

Turner, B.J. 1974. Genetic divergence of Death Valley pupfish species: Biochemical versus morphological evidence. *Evolution* 28:281-94.

Turner, B. J. 1984. Evolutionary genetics of artificial refugium populations of an endangered species, the desert pupfish. *Copeia* 2:364-69.

Turner, B.J., J.F. Elder, Jr., T.F. Laughlin, W.P. Davis, and D.S. Taylor. 1992. Extreme clonal diversity and divergence in clonal vertebrates detected by simple sequence DNA fingerprinting. *Proceedings of the National Academy of Sciences, USA* 89:10643-10647.

Turner, B.J., W.P. Davis, and D.S. Taylor. 1992. Abundant males in populations of a self-fertilizing hermaphroditic fish, *Rivulus marmoratus*. *Journal of Fish Biology* 40:307-310.

Vrijenhoek, R.C. 1985. Homozygosity and interstrain variation in the self-fertilizing hermaphroditic fish, *Rivulus marmoratus*. *Journal of Heredity* 76:82-84.

Vrijenhoek, R. C., and P. L. Leberg. 1991. Let's not throw the baby out with the bathwater: A comment on management for MHC diversity in captive populations. *Conservation Biology* 5:252-54.

Wales, J. M. 1930. Biometrical studies of some races of cyprinodont fishes from the Death Valley region with description of *Cyprinodon diabolis* n. sp. *Copeia* 1930:61-70.

Watkins, D. I., S. N. McAdam, X. Liu, C. R. Strang, E. L. Milford, C. G. Levine, T. L. Garber, A. L. Dogon, C. I. Lord, S. H. Ghim, G. M. Troup, A. L. Hughes, and N. L. Letvin. 1992. New recombinant *HLA-B* alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature (London)* 357:329-33.

Weibel, A.C., T.E. Dowling, and B.J. Turner. 1999. Evidence that an outcrossing population is a derived lineage in a hermaphroditic fish (*Rivulus marmoratus*). *Evolution* 53:1217-1225.

Zangenberg, R. M., M.-M. Huang, N. Arnheim, and H. Erlich. 1995. New HLA-DPB1 alleles generated by intraallelic gene conversion detected by the analysis of sperm. *Nature Genetics* 10:407-14.

Zhang, Z. 1999. Rates of conservative and radical non-synonymous substitutions in mammalian nuclear genes. *Journal of Molecular Evolution*. 50:56-68.

Zuckerkandl, E., and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. In: Bryson, V. and H. J. Vogel (eds) *Evolving genes and proteins*. Academic Press, New York, pp 97-116.

## CIRRICULUM VITAE

*Michael Todd Fisher*

Department of Biology  
Virginia Polytechnic Institute and State University  
Blacksburg, Virginia 24061  
Phone: (540) 977-9768, e-mail: mifishe3@mail.vt.edu

### Personal Data

Born: 10 February 1969, Seattle, Washington. Citizenship: USA

Marital Status: married; Nathalie Molliet-Ribet, June 20, 1998

### Education

High School Diploma: June, 1987; Yorktown High School, Arlington, VA

Bachelor of Arts: December, 1994; Virginia Polytechnic Institute and State University, Blacksburg, VA. Major: History.

Master's of Science: January 1997. Virginia Polytechnic Institute and State University, Blacksburg, VA. Major Advisor: B.J. Turner.  
Thesis: Low temperature induction of males and other developmental anomalies in a self-fertilizing hermaphroditic fish species.

Doctoral Researcher: January 1997-September 1997. Max-Plank Institut für Biologie, abteilung Immunogenetic. Tübingen, Germany. Director: Jan Klein.

Doctoral Candidate: January 1997-November 2000. Virginia Polytechnic Institute and State University, Blacksburg, VA. Major Advisor: B.J. Turner.  
Dissertation: Effects of homozygosity and intense inbreeding on the evolution of Major Histocompatibility Class I genes.

Post-doctoral researcher: July 2000-present. Virginia Polytechnic Institute and State University, Blacksburg, VA. Advisor: M. Lederman  
Research focus: Development of transgenic plants using human adeno-associated viral vectors.

## Research

1. Variation in induction of males through low temperature incubation of embryos of a self-fertilizing vertebrate, the marine killifish *Rivulus marmoratus*, January 1995 - December, 1997
2. Characterization of the Major Histocompatibility Complex (Class I) in *Rivulus marmoratus* using PCR techniques, April, 1997 - present.
3. Characterization of variation at Major Histocompatibility Complex loci in the Death Valley pupfish complex, April 1997 – present
4. Phylogeography and population structure of *Cyprinodon variegatus*, October 1999 – present.

## Manuscripts

1. Variation at Major Histocompatibility Complex class I loci in the homozygous killifish, *Rivulus marmoratus*. M.T. Fisher and B.J. Turner. Submitted to: Molecular Biology and Evolution.
2. A sequence based approach to the resolution of alleles amplified from heterozygous individuals without sub-cloning. M.T. Fisher. Submitted to: Biotechniques.
3. Low variation in mitochondrial control region sequences from the self-fertilizing hermaphroditic fish, *Rivulus marmoratus*; evidence for a recent invasion of the Caribbean. M.T. Fisher and B.J. Turner. Submitted to: Molecular Ecology.
4. Outcrossing in a population of the generally self-fertilizing killifish, *Rivulus marmoratus*, from Honduras, with a reappraisal of the role of outcrossing in this species. D.S. Taylor, M.T. Fisher, and B.J. Turner. Submitted to: Environmental Biology of Fishes

## Manuscripts in Preparation

1. A species level phylogeny of *Cyprinodon variegatus* using mitochondrial markers. M.T. Fisher and B.J. Turner. for submission to: Molecular Ecology.

2. Low-temperature induction of males and other developmental anomalies in the self-fertilizing hermaphroditic fish, *Rivulus marmoratus*. M.T. Fisher and B.J. Turner. for submission to: Journal of Fish Biology.
3. Heterozygosity at a single MHC locus in the Death Valley pupfish, *Cyprinodon diabolis* and *Cyprinodon salinus*: implications for conservation genetics. M.T. Fisher and B.J. Turner. for submission to: Conservation Biology.
4. An issues based approach to teaching Biology to non-science majors. M.T. Fisher and A. Buikema. for submission to: The American Biology Teacher.

### **Presentations at Professional Meetings**

*American Society of Ichthyologists and Herpetologists Annual Meeting (ASIH)*, Seattle, Washington, 1997. *A possible explanation for the appearance of males in populations of self-fertilizing hermaphroditic fish*. M.T. Fisher (pres.) and B.J. Turner.

*Southeastern Ecology and Evolution Conference*, Mountain Lake, Virginia, 1999, *Rapid evolution of variation at MHC Class I loci in the self-fertilizing hermaphroditic fish, Rivulus marmoratus*. M.T. Fisher (pres.) and B.J. Turner.

*Annual Meeting of the Desert Fishes Council*, 1999, Ciudad Victoria, Tamaulipas, Mexico. *Intense overdominant selection acting to maintain variation at a single MHC class I locus in the Devil's Hole pupfish, Cyprinodon diabolis*. M.T. Fisher (pres.) and B.J. Turner.

*Ethology, Evolutionary Biology, and Conservation Ecology of Fishes (EEEF)*, Athens, GA, 2000. Symposium: *The Biology of Unisexual Fishes* (D. Bechler, org.) Introductory Keynote Address: *"Rivulus marmoratus: Does homozygosity matter? – Evidence for intense natural selection on mhc1 loci*. B.J. Turner (pres.) and M.T. Fisher.

*Society for the Study of Evolution*, Bloomington, IN, 2000. *Evidence for intense natural selection at mhc1 loci in the self-fertilizing and homozygous vertebrate, Rivulus marmoratus*. B.J. Turner (pres.) and M.T. Fisher.

### **Teaching Experience**

Graduate teaching assistant, Virginia Polytechnic Institute and State University.  
 Courses: General Biology Laboratory, Spring, 1996-Spring 1997  
 Principles of Biology Laboratory for Majors, Fall 1997.  
 Honors Biology Laboratory, Fall 1999

Course development:

1. Molecular Ecology - Spring 1999, development of an upper-level field and lab based course using molecular tools to explore population structure local flora.
2. Honors biology – Fall 1998, development of a molecular biology approach to teaching basic science concepts.
3. General Biology laboratory development – Fall 1999, an issues-based approach to teaching science to non-majors, using current issues in science to increase student awareness and understanding of basic science concepts and their social implications.

Guest Lecturer – Courses:

1. Evolutionary Genetics (Senior level)
2. Genetics (Junior level)
3. Evolutionary Biology (Sophomore level)
4. Molecular Biology (Senior Level)
5. History of Science (Graduate Seminar)

**Departmental Service**

Graduate student mentoring – Introductory seminar: *Whom do you go to when you're in trouble?* Fall, 1999.

Departmental DNA Sequencing Facility - Manager, 1998-present.

**University Service**

Virginia Science Teacher's Conference. Summer, 1999. Invited Speaker: *Teaching evolution for high school teachers; Web based exercises and real-world examples.*

**Military service**

United States Marine Corps Reserve, Company B, 4th Combat Engineer Battalion, 4th Marine Division, Roanoke, Virginia. October, 1990 - April, 1997. Separation rank: Sergeant (E-5).